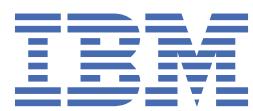


IBM SPSS Modeler 19.0 User's Guide



Nota

Antes de utilizar essas informações e o produto que elas suportam, leia as informações em “[Avisos](#)” na [página 267](#).

Informações sobre o produto

Esta edição se aplica à versão 19, liberação 0, modificação 0 do IBM® SPSS Modelador e a todas as liberações e modificações subsequentes até que seja indicado de outra forma em novas edições.

© Copyright International Business Machines Corporation .

US Government Users Restricted Rights – Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Contents

Chapter 1. SobreIBM SPSS Modelador.....	1
Produtos IBM SPSS Modelador.....	1
IBM SPSS Modelador.....	1
ServidorIBM SPSS Modeler.....	1
Console de administraçãoIBM SPSS Modeler.....	2
IBM SPSS Lote do modelador.....	2
Editor de soluçõesIBM SPSS Modeler.....	2
ServidorIBM SPSS Modeler Adaptadores para IBM SPSS Serviços de colaboração e implantação.....	2
Edições do IBM SPSS Modelador.....	2
Documentação.....	3
Documentação do SPSS Modeler Professional.....	3
SPSS Modeler Premium documentação.....	4
Exemplos de Aplicação.....	4
Pasta Demos.....	4
Rastreamento de Licença.....	4
Chapter 2. O que Há de Novo no IBM SPSS Modelador.....	5
Chapter 3. Visão geral do produto.....	7
Introdução.....	7
Iniciando o IBM SPSS Modelador.....	7
Ativando a partir da linha de comandos.....	7
Conectando-se ao ServidorIBM SPSS Modeler.....	8
Conectando-se ao Servidor analítico.....	10
Alterando o diretório temporário.....	11
Início Múltiplo IBM SPSS Modelador Sessões.....	11
IBM SPSS Modelador Interface em uma Glance.....	12
IBM SPSS Modelador tela de fluxo.....	12
Paleta de nós.....	13
IBM SPSS Modelador Gerenciadores.....	14
IBM SPSS ModeladorProjetos.....	15
IBM SPSS Modelador Barra de ferramentas.....	15
Customizando a barra de ferramentas.....	16
Como personalizar a janela IBM SPSS Modelador.....	17
Alterando o tamanho do ícone para um fluxo.....	18
Usando o Mouse em IBM SPSS Modelador.....	18
Usando teclas de atalho.....	18
Imprimindo.....	19
automatizandoIBM SPSS Modelador.....	20
Chapter 4. Entendendo a mineração de dados.....	21
Visão geral de mineração de dados.....	21
Avaliando dados.....	22
Uma estratégia para mineração de dados.....	23
O modelo de processo CRISP-DM.....	24
Tipos de Modelos.....	25
Exemplos de mineração de dados.....	30
Chapter 5. Construindo fluxos.....	31

Visão geral da construção de fluxo.....	31
Construindo fluxos de dados.....	31
Trabalhando com nós.....	31
Trabalhando com fluxos.....	36
Descrições de fluxo.....	49
Fluxos em execução.....	50
Trabalhando com modelos.....	51
Incluindo comentários e anotações em nós e fluxos.....	51
Salvando fluxos de dados.....	56
Carregando arquivos.....	58
Mapeando fluxos de dados.....	58
Dicas e atalhos.....	60
Chapter 6. Gráficos de construção.....	63
Catálogo de gráficos.....	63
Layout e termos.....	64
Construindo um gráfico a partir da galeria de tipos de gráficos.....	64
Tipos de gráficos.....	64
Gráficos 3D.....	64
Gráficos de barras.....	65
box plots.....	68
gráficos de bolhas.....	69
Gráficos Candlestick (candelabro).....	69
Gráficos de embalagem de círculo.....	70
Gráficos Customizados.....	71
Gráficos de Dendrogram.....	72
Gráficos de eixos duplos Y.....	72
Gráficos de barras de erro.....	74
Gráficos de avaliação.....	75
Gráficos de mapas de utilização.....	77
Gráficos de histograma.....	78
Gráficos de linhas.....	80
Gráficos de mapas.....	81
Gráficos de curva de matemática.....	84
Gráficos de multigráficos.....	85
Vários gráficos de séries.....	86
Gráficos paralelos.....	87
Gráficos de Pareto.....	88
Gráficos de setores circulares.....	88
Gráficos de pirâmide populacional.....	90
Gráficos Q-Q.....	90
Gráfico de radar.....	92
Gráficos de relacionamento.....	93
Tramas de dispersão e tramas de pontos.....	94
Gráficos de matriz de dispersão.....	96
Gráficos de matriz de série.....	97
Gráficos de explosão solar.....	97
gráficos t-SNE.....	98
Tramas de tempo.....	99
Gráficos do Rio Tema.....	100
Gráficos de árvore.....	101
Gráficos de treemap.....	103
Gráficos de nuvem de palavras.....	103
Preferências de visualização global.....	104
Chapter 7. Painel.....	107

Chapter 8. Trabalhando com saída..... 109

Viewer.....	109
Mostrando e ocultando resultados.....	109
Mover, excluir e copiar saída.....	110
Alterando o alinhamento inicial.....	110
Alterando o alinhamento dos itens de saída.....	110
Viewer esboço.....	110
Editando e adicionando itens para o Viewer.....	111
Como localizar e substituir informações no Viewer.....	112
Copiando a saída em outros aplicativos.....	113
Resultado interativo.....	114
Exportar saída.....	115
Opções de HTML.....	116
Opções de relatório da web.....	116
Opções de Palavra.....	117
Opções do Excel.....	118
Opções do PowerPoint.....	118
Opções de PDF.....	119
Opções de texto.....	120
Imagens apenas opções.....	120
Opções de formato de gráficos.....	121
Impressão do visualizador.....	122
Para imprimir saída e gráficos.....	122
Visualização da impressão.....	122
Atributos de Página: Cabeçalhos e Rodapés.....	123
Atributos da página: opções.....	123
Salvar saída.....	124
Salvando um Viewer documento.....	124
Tabelas dinâmicas.....	125
Tabelas dinâmicas.....	125
Manipulando uma tabela dinâmica.....	125
Trabalhando com estratos.....	128
Mostrando e ocultando itens.....	129
TableLooks.....	130
Propriedades da tabela.....	130
Propriedades de célula.....	133
Notas de rodapé e legendas.....	133
Larguras da célula de dados.....	136
Alterando a largura da coluna.....	136
Exibindo bordas ocultas em uma tabela dinâmica.....	136
Selecionando linhas, colunas e células em uma tabela dinâmica.....	136
Imprimindo tabelas dinâmicas.....	136
Criando um gráfico a partir de uma tabela dinâmica.....	137
Tabelas anteriores.....	138
Opções.....	138
Opções.....	138
Opções gerais.....	138
Opções do Visualizador.....	138
Opções da tabela dinâmica.....	139
Opções de saída.....	140

Chapter 9. Tratamento de valores ausentes..... 141

Visão geral de valores omissos.....	141
Manipulando valores omissos.....	142
Manipulando registros com valores omissos.....	142
Manipulando campos com valores omissos.....	142

Tratamento de registros com valores omissos do sistema.....	143
Imputando ou preenchendo valores omissos.....	145
Funções do CLEM para valores omissos.....	145
Chapter 10. Construindo expressões CLEM.....	147
SobreCLEM.....	147
Exemplos de CLEM.....	147
Valores e tipos de dados.....	149
Expressões e condições.....	150
Parâmetros de Fluxo, Sessão e SuperNode.....	150
Trabalhando com sequências de caracteres.....	151
Manipulando valores omissos e em branco.....	152
Trabalhando com números.....	152
Trabalhando com datas e horas.....	152
Resumindo diversos campos.....	153
Trabalhando com dados de múltiplas respostas.....	154
O construtor de expressões.....	155
Acessando o construtor de expressões.....	155
Criando expressões.....	155
Selecionando funções.....	156
Selecionando campos, parâmetros e variáveis globais.....	159
Visualizando ou selecionando valores.....	159
Verificando expressões CLEM.....	160
Localizar e substituir.....	160
Chapter 11. CLEM referência de idioma.....	163
CLEM Visão Geral de Referência.....	163
CLEM Datatypes.....	163
Números Inteiros.....	163
Reais.....	163
Caracteres.....	164
Sequências.....	164
Listas.....	164
Campos.....	165
Datas.....	165
Hora.....	166
CLEM operadores.....	166
Referência de funções.....	168
Convenções em descrições de função.....	170
Funções de informações.....	170
Funções de conversão.....	171
Funções de comparação.....	173
Funções Lógicas.....	175
Funções numéricas.....	176
Funções trigonométricas.....	177
Funções de probabilidade.....	178
Funções espaciais.....	178
Operações de número inteiro bit a bit.....	179
Funções aleatórias.....	180
Funções de sequência de caracteres.....	181
Funções SoundEx.....	187
Funções de Data e Hora.....	187
Funções de sequência.....	194
Funções globais.....	200
Funções manipulando valores em branco e nulos.....	201
Campos Especiais.....	202

Chapter 12. Usando IBM SPSS Modelador com um repositório.....	205
Sobre o IBM SPSS Reppositório de serviços de colaboração e implantação.....	205
Armazenando e implementando objetos do repositório.....	206
Conectando ao repositório.....	206
Inserindo credenciais para o repositório.....	206
Navegação pelas credenciais do repositório.....	207
Navegando no conteúdo do repositório.....	207
Armazenando objetos no repositório.....	207
Configurando propriedades do objeto.....	207
Armazenando fluxos.....	209
Armazenando projetos.....	210
Armazenando nós.....	210
Armazenando objetos de saída.....	210
Armazenando modelos e paletas de modelos.....	211
Recuperando objetos do repositório.....	211
Escolhendo um objeto para recuperar.....	212
Selecionando versão de um objeto.....	212
Procurando objetos no repositório.....	212
Modificando objetos do repositório.....	213
Criando, renomeando e excluindo pastas.....	214
Bloqueando e desbloqueando objetos do repositório.....	214
Excluindo objetos do repositório.....	214
Gerenciando propriedades dos objetos do repositório.....	214
Visualizando propriedades da pasta.....	215
Visualizando e editando propriedades do objeto.....	215
Gerenciando rótulos de versão do objeto.....	216
Implementando fluxos.....	217
Opções de Implementação do Fluxo.....	217
A ramificação de escoragem.....	219
Chapter 13. Exportando para aplicativos externos.....	223
Sobre Exportando para aplicativos externos.....	223
Abrindo um Fluxo em VantagensIBM SPSS Modeler.....	223
Importando e exportando modelos como PMML.....	223
Tipos de modelo que suportam PMML.....	224
Chapter 14. Projetos e relatórios.....	227
Introdução aos Projetos.....	227
Visualização CRISP-DM.....	227
Visualização Classes.....	228
Construindo um projeto.....	228
Criando um novo projeto.....	228
Incluindo em um projeto.....	228
Transferindo Projetos para o IBM SPSS Reppositório de serviços de colaboração e implantação...	229
Configurando propriedades do projeto.....	229
Anotando um projeto.....	230
Propriedades do Objeto.....	231
Encerrando um projeto.....	231
Gerando um relatório.....	231
Salvando e exportando relatórios gerados.....	232
Chapter 15. Customizando o IBM SPSS Modelador.....	235
Customizando as opções IBM SPSS Modelador.....	235
Configurando IBM SPSS Modelador opções.....	235
Opções de Sistema.....	235
Configurando diretórios padrão.....	236

Configurando opções do usuário.....	236
Customizando a Paleta de Nós.....	244
Customizando o Gerenciador de Paleta.....	244
Alterando uma visualização da guia paleta.....	246
Chapter 16. Considerações de desempenho para fluxos e nós.....	249
Ordem dos nós.....	249
Caches de nó.....	250
Desempenho: Nós de Processamento.....	251
Desempenho: Nós de Modelagem.....	252
Desempenho: CLEM Expressões.....	252
Chapter 17. Acessibilidade em IBM SPSS Modelador.....	253
Visão geral da Acessibilidade em IBM SPSS Modelador.....	253
Tipos de suporte à acessibilidade.....	253
Acessibilidade para deficiência visual.....	253
Acessibilidade para usuários cegos.....	254
Teclado de acessibilidade.....	254
Usando um leitor de tela.....	262
Dicas de uso.....	263
Interferência de outro software.....	264
JAWS e Java.....	264
Usando Gráficos em IBM SPSS Modelador.....	264
Chapter 18. Suporte a Unicode.....	265
Suporte a Unicode em IBM SPSS Modelador.....	265
Avisos.....	267
Marcas comerciais.....	268
Termos e condições para documentação do produto.....	268
Index.....	271

Capítulo 1. Sobre IBM SPSS Modelador

O IBM SPSS Modelador é um conjunto de ferramentas de mineração de dados que permite desenvolver rapidamente modelos preditivos usando o conhecimento de negócios, e implementá-los em operações de negócios para melhorar a tomada de decisão. Projetado em torno do modelo CRISP-DM padrão de mercado, o IBM SPSS Modelador suporta todo o processo de mineração de dados, a partir dos dados para melhores resultados de negócios.

O IBM SPSS Modelador oferece uma variedade de métodos de modelagem tomados do aprendizado de máquina, inteligência artificial e estatística. Os métodos disponíveis na paleta Modelagem permitem derivar informações novas a partir dos dados, e desenvolver modelos preditivos. Cada método possui certas forças e é mais adequado para certos tipos de problemas.

O ModeladorSPSS pode ser comprado como um produto independente, ou usado como um cliente na combinação com o Servidor do SPSS Modeler. Várias opções adicionais também estão disponíveis, conforme resumidas nas seções a seguir. Para mais informações, consulte <https://www.ibm.com/analytics/us/en/technology/spss/>.

Produtos IBM SPSS Modelador

A família de produtos IBM SPSS Modelador e o software associado abrangem o seguinte.

- IBM SPSS Modelador
- ServidorIBM SPSS Modeler
- Console de administraçãoIBM SPSS Modeler (incluído com IBM SPSS Gerente de implantação)
- IBM SPSS Lote do modelador
- Editor de soluçõesIBM SPSS Modeler
- ServidorIBM SPSS Modeler adaptadores para IBM SPSS Serviços de colaboração e implantação

IBM SPSS Modelador

ModeladorSPSS é uma versão funcionalmente completa do produto que você instala e executa em seu computador pessoal. É possível executar o ModeladorSPSS no modo local como um produto independente ou usá-lo no modo distribuído com ServidorIBM SPSS Modeler para melhorar o desempenho em conjuntos de dados grandes.

Com o ModeladorSPSS, é possível construir modelos preditivos exatos de maneira rápida e intuitiva, sem programação. Usando a interface visual exclusiva, é possível visualizar facilmente o processo de mineração de dados. Com o suporte da análise avançada integrada ao produto, é possível descobrir tendências e padrões ocultos anteriormente em seus dados. É possível modelar resultados e entender os fatores que os influenciam, permitindo que você aproveite as vantagens das oportunidades de negócios e diminua os riscos.

ModeladorSPSS está disponível em duas edições: SPSS Modeler Professional e SPSS Modeler Premium. Consulte o tópico “[Edições do IBM SPSS Modelador](#)” na página 2 para obter mais informações.

ServidorIBM SPSS Modeler

ModeladorSPSS usa uma arquitetura de cliente/servidor para distribuir solicitações para operações cheias de recursos para poderosos softwares de servidor, resultando em desempenho mais rápido em conjuntos de dados maiores.

Servidor do SPSS Modeler é um produto licenciado separadamente que é executado de forma contínua no modo de análise distribuído em um host do servidor com uma ou mais instalações do IBM SPSS Modelador. Dessa maneira, o Servidor do SPSS Modeler fornece desempenho superior em conjuntos de dados grandes, pois operações com uso intensivo de memória podem ser executadas no servidor

sem fazer download dos dados no computador cliente. ServidorIBM SPSS Modeler também fornece suporte para otimização de SQL e capacidades de modelagem dentro da base de dados, entregando mais benefícios para o desempenho e a automação.

Console de administraçãoIBM SPSS Modeler

O Console de administração do Modeler é uma interface gráfica de usuário para o gerenciamento de muitas das opções de configuração Servidor doSPSS Modeler , que também são configuráveis por meio de um arquivo de opções. O console é incluído em IBM SPSS Gerente de implantação, pode ser usado para monitorar e configurar suas instalações Servidor doSPSS Modeler , e está disponível gratuitamente para os clientes atuais Servidor doSPSS Modeler . O aplicativo pode ser instalado somente em computadores Windows; no entanto, ele pode administrar um servidor instalado em qualquer plataforma suportada.

IBM SPSS Lote do modelador

Embora geralmente a mineração de dados seja um processo interativo, também é possível executar o ModeladorSPSS a partir de uma linha de comandos, sem a necessidade de uma interface gráfica com o usuário. Por exemplo, você pode ter tarefas repetidas ou de longa execução que deseja executar sem intervenção do usuário. SPSS Lote do modelador é uma versão especial do produto que fornece suporte para capacidades de análise completa do ModeladorSPSS sem acessar a interface com o usuário regular. Servidor doSPSS Modeler é necessário para usar o SPSS Lote do modelador.

Editor de soluçõesIBM SPSS Modeler

Editor de soluçõesSPSS Modeler é uma ferramenta que permite criar uma versão do pacote de um fluxo do ModeladorSPSS que pode ser executado por um mecanismo de tempo de execução externo ou integrado a um aplicativo externo. Dessa maneira, é possível publicar e implementar fluxos completos do ModeladorSPSS para uso em ambientes que não têm o ModeladorSPSS instalado. Editor de soluçõesSPSS Modeler é distribuído como parte do serviço IBM SPSS Serviços de colaboração e implantação - Pontuação, para o qual uma licença separada é necessária. Com essa licença, você recebe o Tempo de execuçãoSPSS Modeler Solution Publisher, que permite executar os fluxos publicados.

Para obter mais informações sobre Editor de soluçõesSPSS Modeler, consulte a documentação do IBM SPSS Serviços de colaboração e implantação. A documentação IBM SPSS Serviços de colaboração e implantação IBM contém seções chamadas " IBM SPSS Modeler Solution Publisher " e " IBM SPSS Analytics Toolkit "

ServidorIBM SPSS Modeler Adaptadores para IBM SPSS Serviços de colaboração e implantação

Inúmeros adaptadores para o IBM SPSS Serviços de colaboração e implantação estão disponíveis para permitir que o ModeladorSPSS e o Servidor doSPSS Modeler interajam com um repositório do IBM SPSS Serviços de colaboração e implantação. Dessa forma, um fluxo do ModeladorSPSS implementado no repositório pode ser compartilhado por diversos usuários ou acessado a partir do aplicativo thin client VantagensIBM SPSS Modeler. Você instala o adaptador no sistema que hospeda o repositório.

Edições do IBM SPSS Modelador

ModeladorSPSS está disponível nas seguintes edições.

SPSS Modeler Professional

SPSS Modeler Professional fornece todas as ferramentas necessárias para você trabalhar com a maioria dos tipos de dados estruturados, como comportamentos e interações controlados em sistemas CRM, demográficos, comportamento de compra e dados de vendas.

SPSS Modeler Premium

SPSS Modeler Premium é um produto licenciado separadamente que se estende SPSS Modeler Professional para trabalhar com dados especializados e com dados de texto não estruturados. SPSS Modeler Premium inclui IBM SPSS Modelador de análise de texto:

IBM SPSS Modelador de análise de texto usa tecnologias de linguística avançada e processamento de linguagem natural (NLP) para processar rapidamente uma grande variedade de dados de texto não estruturados, extrair e organizar conceitos chave e agrupar esses conceitos em categorias. Categorias e conceitos extraídos podem ser combinados com dados estruturados existentes, como demográficos, e aplicados à modelagem usando o conjunto completo de ferramentas de mineração de dados do IBM SPSS Modelador para gerar decisões melhores e mais focadas.

AssinaturaIBM SPSS Modeler

AssinaturaIBM SPSS Modeler fornece todas as mesmas capacidades de analítica preditiva que o cliente IBM SPSS Modelador tradicional. Com a edição de Assinaturas, é possível fazer o download de atualizações do produto regularmente.

Documentação

A documentação está disponível no menu **Ajuda** em ' ModeladorSPSS. Isso abre a documentação on-line IBM, que está sempre disponível fora do produto.

A documentação completa de cada produto (incluindo instruções de instalação) também está disponível em formato PDF. Consulte a seguinte página de suporte: [**SPSS Modeler 19.0 documentation**](#).

Documentação do SPSS Modeler Professional

O conjunto de documentações do SPSS Modeler Professional (excluindo instruções de instalação) é o seguinte.

- **IBM SPSS Modelador User's Guide.** Introdução geral para usar ModeladorSPSS, incluindo como construir fluxos de dados, manipular valores ausentes, construir expressões CLEM , trabalhar com projetos e relatórios, e streams de pacotes para implementação em IBM SPSS Serviços de colaboração e implantação ou VantagensIBM SPSS Modeler.
- **Nós de Origem, de Processo e de Saída do IBM SPSS Modelador.** Descrições de todos os nós usados para ler, processar e emitir dados em diferentes formatos. Efetivamente, isso significa todos os nós além dos de modelagem.
- **Nós de Modelagem do IBM SPSS Modelador.** Descrições de todos os nós usados para criar modelos de mineração de dados. OIBM SPSS Modelador oferece uma variedade de métodos de modelagem tomados do aprendizado de máquina, inteligência artificial e estatística.
- **Guia de Aplicativos do IBM SPSS Modelador.** Os exemplos neste guia fornecem introduções sintetizadas e direcionadas para técnicas e métodos de modelagem específicos. Uma versão online deste guia também está disponível no menu Ajuda. Veja o tópico ["Exemplos de Aplicação"](#) na página 4 para obter mais informações.
- **Script e Automação Python do IBM SPSS Modelador.** Informações sobre como automatizar o sistema por meio de script Python, incluindo as propriedades que podem ser usadas para manipular nós e fluxos.
- **Guia de Implementação do IBM SPSS Modelador.** Informações sobre a execução de fluxos IBM SPSS Modelador como etapas de processamento de tarefas sob IBM SPSS Gerente de implantação.
- **Guia de Mineração Dentro do Banco de Dados do IBM SPSS Modelador.** Informações sobre como usar o poder do seu banco de dados para melhorar o desempenho e ampliar o intervalo de capacidades analíticas por meio de algoritmos de terceiros.
- **Guia de Desempenho e de Administração do ServidorIBM SPSS Modeler.** Informações sobre como configurar e administrar o ServidorIBM SPSS Modeler.

- **Guia do Usuário do IBM SPSS Gerente de implantação.** Informações sobre o uso da interface de usuário do console de administração incluídas no aplicativo Gerente de implantação para monitoramento e configuração ServidorIBM SPSS Modeler.
- **IBM SPSS Modelador Guia CRISP-DM.** Guia passo a passo para o uso da metodologia CRISP-DM para mineração de dados com ModeladorSPSS.
- **IBM SPSS Lote do modelador User's Guide.** Guia completo para o uso do IBM SPSS Modelador no modo em lote, incluindo detalhes da execução do modo em lote e argumentos de linha de comandos. Este guia está disponível somente em formato PDF.

SPSS Modeler Premium documentação

O conjunto de documentações do SPSS Modeler Premium (excluindo instruções de instalação) é o seguinte.

- **SPSS Modelador de análise de texto User's Guide.** Informações sobre o uso de analítica de texto com ModeladorSPSS, cobrindo os nós de mineração de texto, ambiente de trabalho interativo, modelos e outros recursos.

Exemplos de Aplicação

Enquanto as ferramentas de mineração de dados no ModeladorSPSS podem ajudar a resolver uma ampla variedade de negócios e problemas organizacionais, os exemplos de aplicativos fornecem introduções breves e destinadas aos métodos e técnicas de modelagem específicos. Os conjuntos de dados utilizados aqui são muito menores do que as enormes lojas de dados gerenciadas por alguns mineiros de dados, mas os conceitos e métodos que estão envolvidos são escaláveis para aplicações do mundo real.

Para acessar os exemplos, clique em **Exemplos de aplicativos** no menu Ajuda em ModeladorSPSS.

Os arquivos de dados e os fluxos de amostra são instalados na pasta Demos no diretório de instalação do produto. Para obter mais informações, consulte “[Pasta Demos](#)” na [página 4](#).

Exemplos de modelagem da base de dados. Consulte os exemplos no *Guia de Mineração dentro do Banco de Dados do IBM SPSS Modelador*.

Exemplos de script. Consulte os exemplos no *Guia de Script e Automação do IBM SPSS Modelador*.

Pasta Demos

Os arquivos de dados e fluxos de amostra que são utilizados com os exemplos de aplicação são instalados na pasta Demos sob o diretório de instalação do produto (por exemplo: C :\Program Files\IBM\SPSS\Modeler\<version>\Demos). Esta pasta também pode ser acessada a partir do grupo de programas IBM ModeladorSPSS no menu Iniciar do Windows, ou clicando em Demos na lista de diretórios recentes na caixa de diálogo **Arquivo > Open Stream** .

Rastreamento de Licença

Quando você usa o ModeladorSPSS, o uso sob licença é controlado e registrado em intervalos regulares. As métricas de licença que são registradas são *AUTHORIZED_USER* e *CONCURRENT_USER* e o tipo de métrica que é registrado depende do tipo de licença que você possui para o ModeladorSPSS.

Os arquivos de log que são produzidos podem ser processados pelo IBM License Metric Tool, do qual é possível gerar relatórios de uso sob licença.

Os arquivos de log de licença são criados no mesmo diretório onde os arquivos de log do Client log do ModeladorSPSS são registrados (por padrão, %ALLUSERSPROFILE%\IBM\SPSS\Modeler\<version>\log).

Capítulo 2. O que Há de Novo no IBM SPSS Modelador

IBM SPSS Modelador adiciona os seguintes novos recursos e aprimoramentos nesta versão.

Versão 19.0.0

Lançado em: 28 de outubro de 2025

Azure Suporte da Databricks

IBM SPSS Modelador agora oferece suporte ao Azure Databricks como fonte de dados, com pushback completo de SQL e integração de funções do Modeler.

Novos nós do Spark

Os nós Logistic Regression e Bisecting K-Means Clustering são adicionados. Eles são desenvolvidos com base em algoritmos nativos do Spark.

SQL IntelliSense e formatação dinâmica

IntelliSense e a formatação dinâmica para consultas SQL estão disponíveis no nó Database Source. Isso melhora a experiência de edição de consultas.

Nó OLAP

IBM SPSS Modelador agora inclui o nó OLAP. Com o novo nó, você pode fazer análises avançadas com funções de janela para classificação, numeração de linhas e cálculos agregados sem recolher as linhas.

Nó OFFSET

Com o nó OFFSET, você pode recuperar valores anteriores ou posteriores em um conjunto de dados. É possível agrupar, classificar e personalizar o controle do intervalo de abrangência.

Teradata Autenticação da carteira

A autenticação da carteira Teradata aprimora as opções de conexão segura.

Integrado ModeladorSPSS com CPLEX

Suporte aprimorado para a passagem de dados adicionais para o nó CPLEX para fluxos de trabalho de otimização mais flexíveis e robustos.

Suporte ao idioma chinês para análise de texto

O Text Analytics agora é suportado em chinês, o que amplia os recursos multilíngues do produto.

Otimização do desempenho do Spark

As configurações padrão do Spark são revisadas para aumentar o desempenho dos nós do Spark.

Apache Spark 3.5.4

ModeladorSPSS agora suporta Apache Spark 3.5.4. Essa alteração torna o produto compatível com os recursos e aprimoramentos mais recentes do Spark.

Compilação de base de código modernizada

IBM SPSS Modelador agora é construído usando compiladores C++ atualizados, melhorando o desempenho do tempo de execução, aprimorando a estabilidade do código e garantindo a compatibilidade de longo prazo com plataformas modernas.

Suporte expandido à plataforma

IBM SPSS Modelador agora oferece suporte a uma ampla variedade dos mais recentes sistemas operacionais e bancos de dados nas plataformas Linux, macOS, e Windows, conforme detalhado no relatório de compatibilidade.

Integração com IBM SPSS Estatísticas 31

IBM SPSS Modelador agora está integrado ao site IBM SPSS Estatísticas 31. A integração permite fluxos de trabalho mais suaves e maior interoperabilidade entre a modelagem preditiva e a análise estatística.

Aperfeiçoamentos de Segurança

Segurança e conformidade reforçadas com a correção de CVEs críticos e atualização de componentes principais e de código aberto.

Consulte [as notas de versão](#) para obter os requisitos do sistema e os problemas e limitações conhecidos para usar o [IBM SPSS Modelador](#).

- Conectividade com `watsonx.data` por meio do mecanismo de consulta Presto SQL com suporte a pushback de SQL.
- Suporte aos recursos do R 4.4.0 nos nós Extension Import e Extension Export.
- O tempo de execução do Java agora está atualizado para IBM Semeru JRE 17.
- Desempenho aprimorado para a conectividade Azure Synapse com pushback do SQL.
- As seguintes novas versões de sistema operacional são compatíveis: Ubuntu 24 e RHEL 9.4.
- O Text Analytics agora é compatível com o idioma japonês.

Notas sobre a liberação

Consulte [as notas de versão](#) para obter os requisitos do sistema e os problemas e limitações conhecidos para usar o [IBM SPSS Modelador](#).

Capítulo 3. Visão geral do produto

Introdução

Como um aplicativo de mineração de dados, o IBM SPSS Modelador oferece uma abordagem estratégica para localizar relacionamentos úteis em grandes conjuntos de dados. Em contraste com métodos estatísticos mais tradicionais, você não precisará necessariamente saber o que está procurando ao iniciar. Você pode explorar os dados, ajustando diferentes modelos e investigando diferentes relacionamentos até localizar informações úteis.

Iniciando o IBM SPSS Modelador

Para iniciar o aplicativo, clique em:

Iniciar > [Todos] Programas > IBM SPSS Modeler < version> > IBM SPSS Modeler < version>

A janela principal é exibida após alguns segundos.

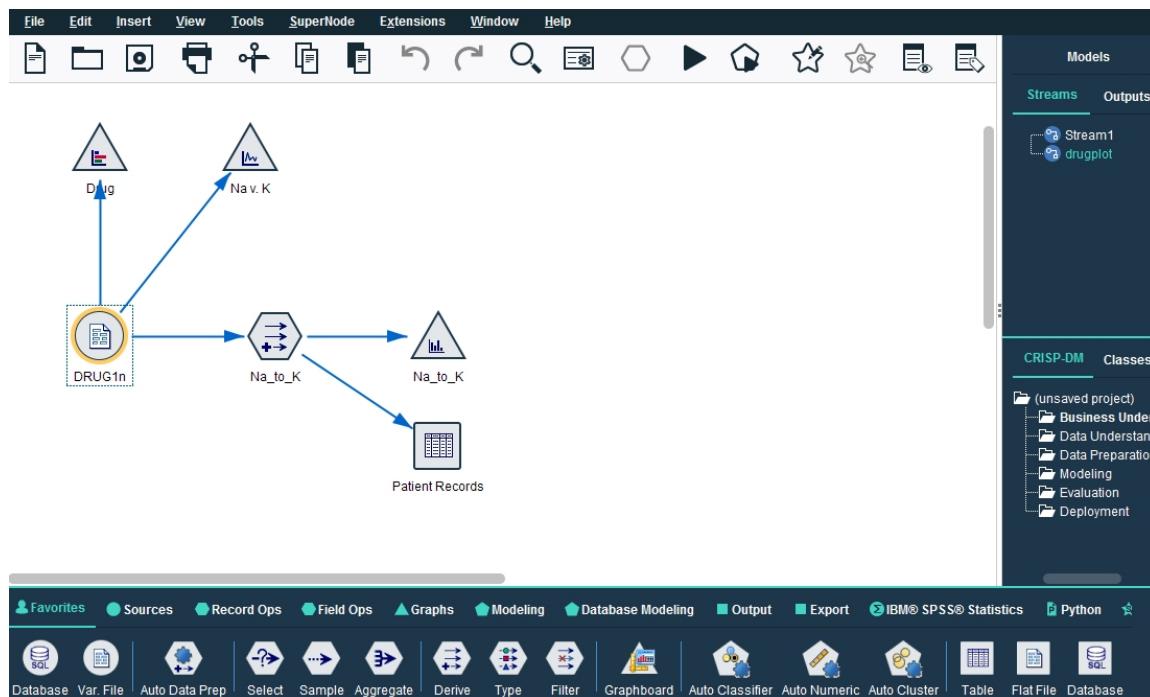


Figura 1. IBM SPSS Modelador janela principal do aplicativo

Ativando a partir da linha de comandos

Você pode usar a linha de comando do seu sistema operacional para iniciar o 'IBM SPSS Modelador da seguinte forma.

Microsoft Windows

1. Em um computador no qual o IBM SPSS Modelador está instalado, abra um DOS, um prompt de comandos ou uma janela.
2. Alterne para o caminho de instalação de IBM SPSS Modelador (por exemplo, [Installpath]\Program Files\IBM\SPSS\Modeler\19.0\bin).

3. Para iniciar a interface ' IBM SPSS Modelador no modo interativo, digite o comando ' modelerclient seguido dos argumentos necessários; por exemplo:

```
modelerclient -stream report.str -execute
```

Você pode usar os argumentos disponíveis (sinalizadores) para se conectar a um servidor, carregar fluxos, executar scripts ou especificar outros parâmetros, conforme necessário.

Mac OS

1. Localize o caminho do comando Mac OS para IBM SPSS Modelador (por exemplo, [Installpath]/Applications/IBM/SPSS/Modeler/19.0/IBM SPSS Modeler.app/Contents/MacOS).
2. Para iniciar a interface ' IBM SPSS Modelador no modo interativo, execute o comando modeler seguido dos argumentos necessários; por exemplo:

```
./modeler -stream report.str -execute
```

Conectando-se ao ServidorIBM SPSS Modeler

IBM SPSS Modelador pode ser executado como um aplicativo independente, ou como um cliente conectado diretamente ao ServidorIBM SPSS Modeler, ou a um ServidorIBM SPSS Modeler ou cluster de servidores por meio do plug-in Coordinator of Processes do IBM SPSS Serviços de colaboração e implantação. O status da conexão atual é exibido na parte inferior esquerda da janela do IBM SPSS Modelador.

Sempre que você desejar se conectar a um servidor, é possível inserir manualmente o nome do servidor ao qual deseja se conectar ou selecionar um nome que você definiu anteriormente. No entanto, se você tiver o IBM SPSS Serviços de colaboração e implantação, é possível procurar em uma lista de servidores ou clusters de servidores na caixa de diálogo Login do Servidor. A capacidade de navegar pelos serviços do Estatísticas em execução em uma rede é disponibilizada por meio do Coordinator of Processes.

Para se conectar a um servidor

1. No menu Ferramentas, clique em **Login do Servidor**. A caixa de diálogo Login do Servidor se abre. Alternativamente, dê um clique duplo na área de status da conexão da janela IBM SPSS Modelador.
2. Usando a caixa de diálogo, especifique opções para se conectar ao computador servidor local ou selecione uma conexão da tabela.
 - Clique em **Incluir** ou **Editar** para incluir ou editar uma conexão. Consulte o tópico [“Como adicionar e Editar a Conexão ServidorIBM SPSS Modeler”](#) na página 9 para obter mais informações.
 - Clique em **Procurar** para acessar um servidor ou cluster de servidores no Coordinator of Processes. Veja o tópico [“Procurando por Servidores em IBM SPSS Serviços de colaboração e implantação”](#) na página 9 para obter mais informações.

Tabela de servidores. Essa tabela contém o conjunto de conexões de servidor definidas. A tabela exibe a conexão padrão, o nome do servidor, a descrição e o número da porta. É possível incluir manualmente uma nova conexão, bem como selecionar ou procurar uma conexão existente. Para configurar um determinado servidor como conexão padrão, marque a caixa de seleção na coluna Padrão na tabela para a conexão.

Caminho de dados padrão. Especifique um caminho usado para dados no computador servidor. Clique no botão de reticências (...) para navegar para o local necessário.

Configurar credenciais. Deixe essa caixa desmarcada para ativar a variável **conexão única**, que tenta efetuar seu login no servidor usando detalhes de nome de usuário e senha do computador local. Se uma conexão única não for possível ou se você marcar essa caixa para desativar a conexão única (por exemplo, para efetuar login em uma conta do administrador), os campos a seguir serão ativados para você inserir suas credenciais.

ID do Usuário. Insira o nome de usuário com o qual efetuar logon no servidor.

Senha. Insira a senha associada ao nome de usuário especificado.

Domínio. Especifique o domínio usado para efetuar logon no servidor. Um nome de domínio só é necessário quando o computador servidor está em um domínio do Windows diferente daquele do computador cliente.

3. Clique em **OK** para concluir a conexão.

Para se desconectar de um servidor

1. No menu Ferramentas, clique em **Login do Servidor**. A caixa de diálogo Login do Servidor se abre. Alternativamente, dê um clique duplo na área de status da conexão da janela IBM SPSS Modelador.
2. Na caixa de diálogo, selecione o Servidor Local e clique em **OK**.

Como adicionar e Editar a Conexão ServidorIBM SPSS Modeler

É possível editar ou incluir manualmente uma conexão do servidor na caixa de diálogo Login do Servidor. Clicando em Incluir, é possível acessar uma caixa de diálogo Incluir/Editar Servidor vazia na qual é possível inserir detalhes da conexão do servidor. Com a seleção de uma conexão existente e um clique em Editar na caixa de diálogo Login do Servidor, a caixa de diálogo Incluir/Editar é aberta com os detalhes para essa conexão, de modo que seja possível fazer quaisquer mudanças.

Nota: Não é possível editar uma conexão do servidor que foi incluída do IBM SPSS Serviços de colaboração e implantação, já que o nome, a porta e outros detalhes são definidos no IBM SPSS Serviços de colaboração e implantação. A melhor prática determina que as mesmas portas devem ser usadas para comunicação com IBM SPSS Serviços de colaboração e implantação e ModeladorSPSS Client. Isso pode ser configurado como `max_server_port` e `min_server_port` no arquivo `options.cfg`.

Para incluir conexões do servidor

1. No menu Ferramentas, clique em **Login do Servidor**. A caixa de diálogo Login do Servidor se abre.
 2. Nesta caixa de diálogo, clique em **Incluir**. A caixa de diálogo Incluir/Editar Servidor de Login do Servidor é aberta.
 3. Insira os detalhes de conexão do servidor e clique em **OK** para salvar a conexão e retornar para a caixa de diálogo Login do Servidor.
- **Servidor.** Especifique um servidor disponível ou selecione um na lista. O computador servidor pode ser identificado por um nome alfanumérico (por exemplo, *myserver*) ou um endereço IP designado ao computador servidor (por exemplo, 202.123.456.78).
 - **Porta.** Forneça o número da porta no qual o servidor está atendendo. Se o padrão não funcionar, peça ao administrador do sistema o número da porta correto.
 - **Descrição.** Insira uma descrição opcional para essa conexão do servidor.
 - **Assegurar conexão segura (usar SSL).** Especifica se uma conexão SSL (**Secure Sockets Layer**) deve ser usada. O SSL é um protocolo usado comumente para proteger os dados enviados em uma rede. Para usar essa variável, o SSL deve ser ativado no servidor hospedando o ServidorIBM SPSS Modeler. Se necessário, entre em contato com o administrador local para obter detalhes.

Para editar conexões do servidor

1. No menu Ferramentas, clique em **Login do Servidor**. A caixa de diálogo Login do Servidor se abre.
2. Nessa caixa de diálogo, selecione a conexão que deseja editar e clique em **Editar**. A caixa de diálogo Incluir/Editar Servidor de Login do Servidor é aberta.
3. Altere os detalhes de conexão do servidor e clique em **OK** para salvar as mudanças e retornar para a caixa de diálogo Login do Servidor.

Procurando por Servidores em IBM SPSS Serviços de colaboração e implantação

Em vez de inserir uma conexão do servidor manualmente, é possível selecionar um servidor ou cluster de servidores disponíveis na rede por meio do Coordinator of Processes, disponível no IBM SPSS Serviços

de colaboração e implantação. Um cluster de servidores é um grupo de servidores a partir do qual o Coordinator of Processes determina o servidor mais adequado para responder a uma solicitação de processamento.

Embora seja possível incluir manualmente servidores na caixa de diálogo Login do Servidor, procurar por servidores disponíveis permite conectar-se aos servidores sem precisar saber o nome e o número da porta corretos do servidor. Estas informações são fornecidas automaticamente. Entretanto, você ainda precisará das informações de logon corretas, como nome de usuário, domínio e senha.

Nota: Se você não tiver acesso à capacidade Coordinator of Processes, ainda será possível inserir manualmente o nome do servidor ao qual deseja se conectar ou selecionar um nome que você definiu anteriormente. Veja o tópico [“Como adicionar e Editar a Conexão ServidorIBM SPSS Modeler”](#) na página 9 para obter mais informações.

Para procurar servidores e clusters

1. No menu Ferramentas, clique em **Login do Servidor**. A caixa de diálogo Login do Servidor se abre.
2. Nesta caixa de diálogo, clique em **Procurar** para abrir a caixa de diálogo Procurar por Servidores. Se você não tiver efetuado logon no IBM SPSS Serviços de colaboração e implantação quando tentou navegar no Coordinator of Processes, será solicitado que faça isso agora.
3. Selecione o servidor ou cluster de servidores da lista.
4. Clique em **OK** para fechar a caixa de diálogo e incluir essa conexão na tabela na caixa de diálogo Login do Servidor.

Conectando-se ao Servidor analítico

Se você tiver vários Servidor analíticos disponíveis, você pode usar o diálogo Conexão do Servidor Analítico para definir mais de um servidor para uso em IBM SPSS Modelador. O seu administrador pode já ter configurado um padrão Servidor analítico no arquivo <Modeler_install_path>/config/options.cfg. Mas você também pode usar outros servidores disponíveis após defini-los. Por exemplo, ao usar os nós Servidor analítico Source e Export, você pode querer usar diferentes Servidor analítico conexões em diferentes ramos de um fluxo para que quando cada ramificação execute ele use seus próprios Servidor analítico e nenhum dado será puxado para o ServidorIBM SPSS Modeler. Note que se uma ramificação contém mais de uma conexão Servidor analítico , os dados serão retirados dos Servidor analíticos para o ServidorIBM SPSS Modeler.

Para criar uma nova conexão Servidor analítico , vá em **Ferramentas > Conexões do Servidor Analítico** e forneça as informações necessárias nas seguintes seções do diálogo.

Conexão

URL. Digite URL para o Servidor analítico no formato `https://hostname:port/contextroot`, em que `hostname` é o endereço IP ou o nome do host do Servidor analítico, `port` é o número da porta e `contextroot` é a raiz do contexto do Servidor analítico.

Inquilino. Digite o nome do locatário que o ServidorIBM SPSS Modeler é um membro. Entre em contato com o administrador se você não conhece o inquilino.

Autenticação

Modo. Selecione a partir dos seguintes modos de autenticação.

- **Nome de usuário e senha** requer que você digite o nome de usuário e a senha.
- **Credencial Armazenada** requer que você selecione uma credencial a partir do IBM SPSS Repositório de serviços de colaboração e implantação.
- **Kerberos** requer que você digite o nome principal do serviço e o caminho de arquivo config. Entre em contato com o administrador se você não souber essas informações.

Nome de usuário. Digite o nome de usuário Servidor analítico .

Resmas. Selecione o reino a ser usado para a conexão Servidor analítico .

Senha. Digite a senha do Servidor analítico .

Conectar. Clique em **Conectar** para testar a nova conexão.

Conexões

Depois de especificar as informações acima e clicar em **Connect**, a conexão será adicionada a esta tabela Conexões. Se você precisar remover uma conexão, selecione-a e clique em **Remover**.

Se o seu administrador definiu uma conexão padrão do Servidor analítico no arquivo `options.cfg`, você pode clicar em **Adicionar conexão padrão** para adicioná-lo às suas conexões disponíveis também. Você será solicitado o nome de usuário e a senha.

Alterando o diretório temporário

Algumas operações que são feitas pelo ServidorIBM SPSS Modeler podem requerer que arquivos temporários sejam criados Por padrão, o IBM SPSS Modelador usa o diretório temporário do sistema para criar arquivos temp. É possível alterar a localização do diretório temporário usando os passos a seguir.

1. Crie um novo diretório chamado `spss` e um subdiretório chamado `servertemp`.
2. Edite `options.cfg`, localizado no diretório `/config` do diretório de instalação do IBM SPSS Modelador. Edite o parâmetro `temp_directory` neste arquivo para ler: `temp_directory, "C:/spss/servertemp"`.
3. Reinicie o serviço ServidorIBM SPSS Modeler. Isso pode ser feito clicando na guia **Serviços** do Painel de Controle do Windows. Pare o serviço e, em seguida, inicie-o para ativar as mudanças feitas. Reiniciar a máquina também reinicia o serviço.

Todos os arquivos temp agora são gravados neste novo diretório

Nota: Barras devem ser usadas.

Diretório temporário para visualização de dados

Para o serviço de visualização de dados, conclua as etapas a seguir para configurar o diretório temporário

1. Crie o diretório `temp D:/SPSSTemp`
2. Os arquivos temp da visualização de dados são controlados por `-Djava.io.tmpdir=D:/SPSSTemp`
Inclua essa opção no script de início para visualização de dados no seguinte arquivo: `${ModelerInstallDir}/dataview/start_graph_micro_service.sh`.
3. Inicie o IBM SPSS Modelador.

Arquivos agora são gravados em `D:/SPSSTemp`.

Início Múltiplo IBM SPSS Modelador Sessões

Se você precisar ativar mais de uma sessão do IBM SPSS Modelador por vez, deve-se fazer algumas mudanças nas configurações do IBM SPSS Modelador e do Windows. Por exemplo, talvez seja necessário fazer isso se você tiver duas licenças de servidor separadas e desejar executar dois fluxos com relação a dois servidores diferentes a partir do mesmo computador cliente.

Para ativar várias sessões do IBM SPSS Modelador:

1. Clique em:
Iniciar > [Todos] Programas > IBM SPSS Modeler
2. No atalho IBM SPSS Modelador (aquele com o ícone), clique com o botão direito do mouse e selecione **Propriedades**.
3. Na caixa de texto **Resposta**, inclua `-noshare` no final da sequência de caracteres.

IBM SPSS Modelador Interface em uma Glance

Em cada ponto no processo de mineração de dados, a interface do IBM SPSS Modelador fácil de usar convida seus conhecimentos de negócios específicos. Os algoritmos de modelagem, como predição, classificação, segmentação e detecção de associação, asseguram modelos poderosos e exatos. Os resultados do modelo podem ser facilmente implementados e lidos nos bancos de dados, IBM SPSS Estatísticas e em uma grande variedade de outros aplicativos.

O trabalho com o IBM SPSS Modelador é um processo de três passos de trabalho com dados.

- Primeiro, você lê dados no IBM SPSS Modelador.
- Depois, você executa os dados por meio de uma série de manipulações.
- Por fim, você envia os dados para um destino.

Essa sequência de operações é conhecida como **fluxo de dados**, pois os dados fluem registro por registro, desde a origem, passando por cada manipulação e, por fim, até seu destino—um modelo ou tipo de saída de dados.

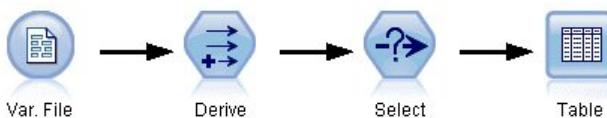


Figura 2. Um fluxo simples

IBM SPSS Modelador tela de fluxo

A tela de fluxo é a maior área da janela do IBM SPSS Modelador e é onde você construirá e manipulará fluxos de dados.

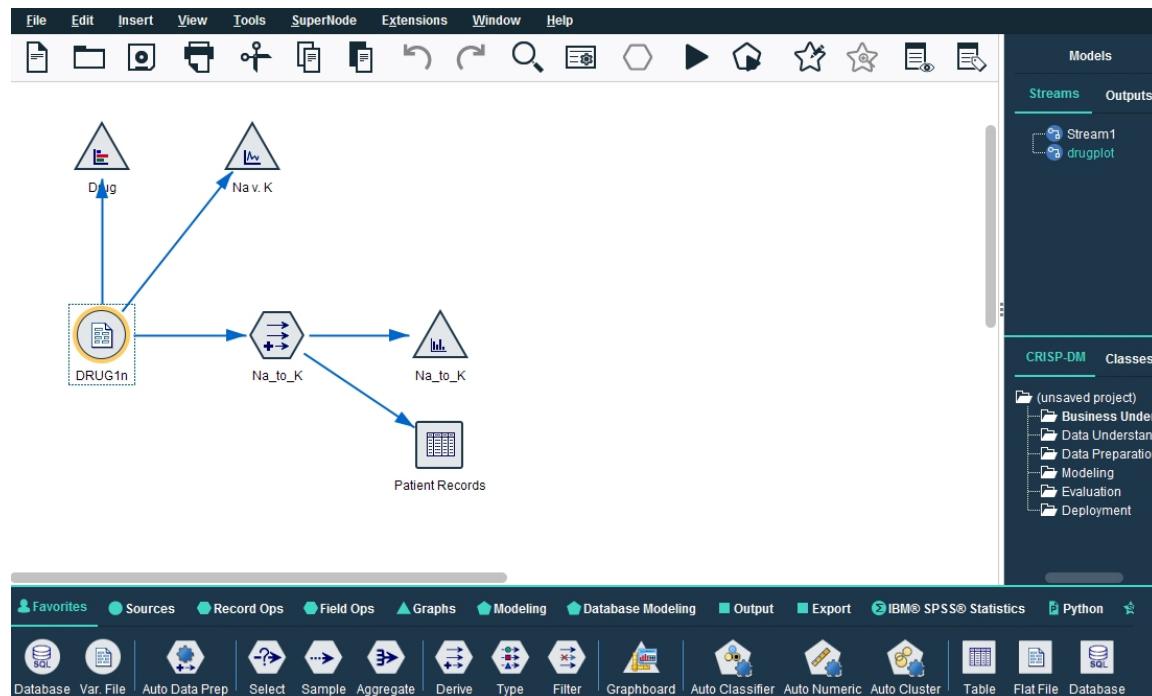


Figura 3. Área de trabalho do IBM SPSS Modelador (visualização padrão)

Fluxos são criados desenhando diagramas de operações de dados relevantes para seus negócios na tela principal da interface. Cada operação é representada por um ícone ou **nó**, e os nós são vinculados em um **fluxo** representando o fluxo de dados em cada operação.

É possível trabalhar com vários fluxos de uma vez no IBM SPSS Modelador, ou na mesma tela de fluxo ou abrindo uma nova. Durante uma sessão, fluxos são armazenados no gerenciador de Fluxos, no lado superior direito da janela do IBM SPSS Modelador.

Nota: Se estiver usando um MacBook com a configuração **Force Click and haptic feedback** do trackpad integrado ativada, arrastar e soltar nós da paleta de nós para a tela de fluxo pode resultar na inclusão de nós duplicados na tela. Para evitar essa questão, recomendamos a desativação da preferência do sistema **Force Click and haptic feedback** trackpad.

Paleta de nós

A maioria das ferramentas de dados e modelagem em ModeladorSPSS estão disponíveis a partir da *Paleta de Nodes*, através da parte inferior da janela abaixo da tela do fluxo.

Por exemplo, a guia paleta **Registrar Ops** contém nós que você pode usar para executar operações nos dados *registros*, como selecionar, mesclar e anexar.

Para adicionar nós à tela, dê um duplo clique em ícones da Paleta de Nós ou arraste-os para a tela. Depois conecte-os para criar um *fluxo*, representando o fluxo de dados.



Figura 4. Guia Operações de Registro na paleta de nós

Cada guia da paleta contém uma coleção de nós relacionados usados para diferentes fases de operações de fluxo, como:

- **Fontes** nós trazem dados em ModeladorSPSS.
- **Record Ops** nós executam operações em dados *records*, como selecionar, mesclar e anexar.
- **Ops de campo** nós executam operações em dados *campos*, como filtragem, derivando novos campos e determinando o nível de medição para determinados campos.
- **Gráficos** nós exibem graficamente os dados antes e depois da modelagem. Gráficos incluem plots, histogramas, nós da web e gráficos de avaliação.
- **Modelagem** nós utilizamos os algoritmos de modelagem disponíveis em ModeladorSPSS, tais como redes neurais, árvores de decisão, algoritmos de clustering e sequenciamento de dados.
- **Modelagem de Banco de Dados** nós usamos os algoritmos de modelagem disponíveis em Microsoft SQL Server, IBM Db2, e bancos de dados Oracle e Netezza .
- **Nós de saída** produzem várias saídas para dados, gráficos e resultados de modelos que podem ser visualizados em ModeladorSPSS.
- **Os nós de exportação** produzem várias saídas que podem ser visualizadas em aplicativos externos, como IBM SPSS Coleta de dados ou Excel.
- **IBM SPSS Estatísticas** nós importar dados de, ou exportar dados para, IBM SPSS Estatísticas , bem como executar procedimentos IBM SPSS Estatísticas .
- Os nós **Python** podem ser usados para executar algoritmos Python .
- Os nós **Spark** podem ser usados para executar algoritmos de Spark.

Conforme você se familiariza com ModeladorSPSS, é possível customizar o conteúdo da paleta para seu próprio uso.

No lado esquerdo da Paleta de Nós, é possível filtrar os nós que são exibidos selecionando Supervised, Association ou Segmentação.

Localizada abaixo da Paleta de Nós, uma área de janela de relatório fornece feedback sobre o progresso de várias operações, como quando os dados estão sendo lidos no fluxo de dados. Também localizada abaixo da Paleta de Nós, uma área de janela de status fornece informações sobre o que o aplicativo está fazendo no momento, bem como indicações de quando um feedback do usuário é necessário.

Nota: Se estiver usando um MacBook com a configuração **Force Click and haptic feedback** do trackpad integrado ativada, arrastar e soltar nós da paleta de nós para a tela de fluxo pode resultar na inclusão de nós duplicados na tela. Para evitar essa questão, recomendamos a desativação da preferência do sistema **Force Click and haptic feedback** trackpad.

IBM SPSS Modelador Gerenciadores

Na parte superior direita da janela está a área de janela de gerenciadores. Ela possui três guias, que são usadas para gerenciar fluxos, saída e modelos.

É possível usar a guia Fluxos para abrir, renomear, salvar e excluir os fluxos criados em uma sessão.



Figura 5. Guia Fluxos



Figura 6. Guia Saídas

A guia Saídas contém uma variedade de arquivos, como gráficos e tabelas, produzidos por operações de fluxo no IBM SPSS Modelador. É possível exibir, salvar, renomear e fechar as tabelas, gráficos e relatórios listados nessa guia.

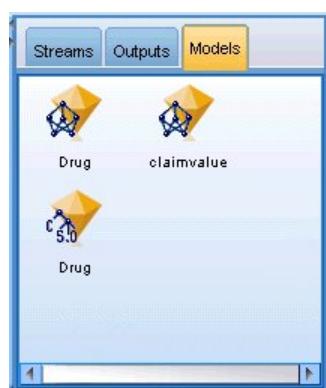


Figura 7. Guia Modelos contendo nuggets de modelo

A guia Modelos é a mais poderosa das guias de gerenciadores. Essa guia contém todos os **nuggets** de modelo, que contêm os modelos gerados no IBM SPSS Modelador, para a sessão atual. Esses modelos podem ser navegados diretamente a partir da guia Modelos ou incluídos no fluxo na tela.

IBM SPSS Modelador Projetos

No lado inferior direito da janela está a área de janela do projeto, usada para criar e gerenciar **projetos** de mineração de dados (grupos de arquivos relacionados a uma tarefa de mineração de dados). Há duas maneiras de se visualizar projetos que você cria no IBM SPSS Modelador—na visualização Classes e na visualização CRISP-DM.

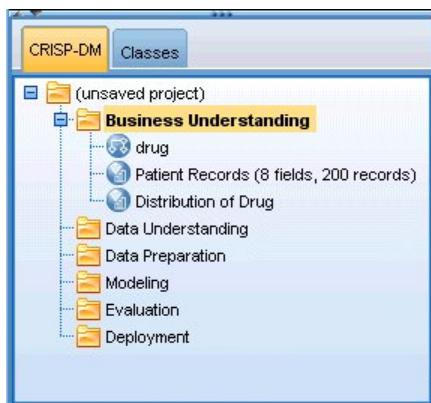


Figura 8. Visualização CRISP-DM

A guia CRISP-DM fornece uma maneira de organizar projetos de acordo com o processo padrão de vários segmentos de mercados para mineração de dados, uma metodologia não proprietária comprovada pela indústria. Tanto para mineradores de dados iniciantes quanto para os experientes, o uso da ferramenta CRISP-DM ajudará você a organizar e comunicar melhor seus esforços.

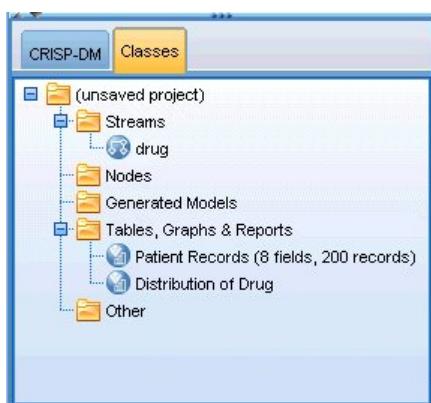


Figura 9. Visualização Classes

A guia Classes fornece uma maneira de organizar seu trabalho no IBM SPSS Modelador categoricamente—por tipos de objetos criados. Essa visualização é útil durante a realização de inventário de dados, fluxos e modelos.

IBM SPSS Modelador Barra de ferramentas

Na parte superior da janela do IBM SPSS Modelador, você localizará uma barra de ferramentas de ícones que fornecem inúmeras funções úteis. A seguir estão os botões da barra de ferramentas e suas funções.



Criar novo fluxo



Abrir fluxo

	Salvar fluxo		Imprimir fluxo atual
	Cortar e mover para a área de transferência		Copiar para área de transferência
	Colar seleção		Desfazer última ação
	Refazer		Procurar nós
	Edite propriedades do fluxo		Visualizar geração de SQL
	Executar fluxo atual		Executar seleção de fluxo
	Parar fluxo (Ativo somente enquanto o fluxo está em execução)		Incluir Supernó
	Aumentar zoom (Somente SuperNodes)		Diminuir zoom (Somente SuperNodes)
	Nenhuma marcação no fluxo		Inserir comentário
	Ocultar marcação de fluxo (se houver)		Mostrar marcação de fluxo oculta
	Abrir fluxo em VantagensIBM SPSS Modeler		

Marcação de fluxo consiste em comentários do fluxo, ligações de modelo e indicações de ramificação de escoragem.

Ligações de modelo são descritas no guia *Nós de Modelagem do IBM SPSS*.

Customizando a barra de ferramentas

É possível alterar vários aspectos da barra de ferramentas, como:

- Se ela é exibida
- Se os ícones têm dicas de ferramentas disponíveis
- Se ela usa ícones grandes ou pequenos

Para ativar e desativar a exibição da barra de ferramentas:

1. No menu principal, clique em:

Visualizar > Barra de Ferramentas > Display

Para alterar as configurações de dica de ferramenta ou tamanho do ícone:

1. No menu principal, clique em:

View > Barra de Ferramentas > customizar

Clique em **Mostrar Dicas de Ferramenta** ou **Botões Grandes** conforme necessário.

Como personalizar a janela IBM SPSS Modelador

Usando as divisórias entre várias partes da interface do ModeladorSPSS, é possível redimensionar ou fechar ferramentas para atender às suas preferências. Por exemplo, se você estiver trabalhando com um grande fluxo, é possível usar as pequenas setas localizadas em cada divisória para fechar a paleta do nó, a área de janela de gerenciadores e a área de janela de projeto. Isso maximiza a tela de fluxo, fornecendo um espaço de trabalho suficiente para fluxos grandes ou para vários fluxos.

Alternativamente, no menu Visualizar, clique em **Paleta de Nós, Gerenciadores** ou **Projeto** para ativar ou desativar a exibição desses itens.

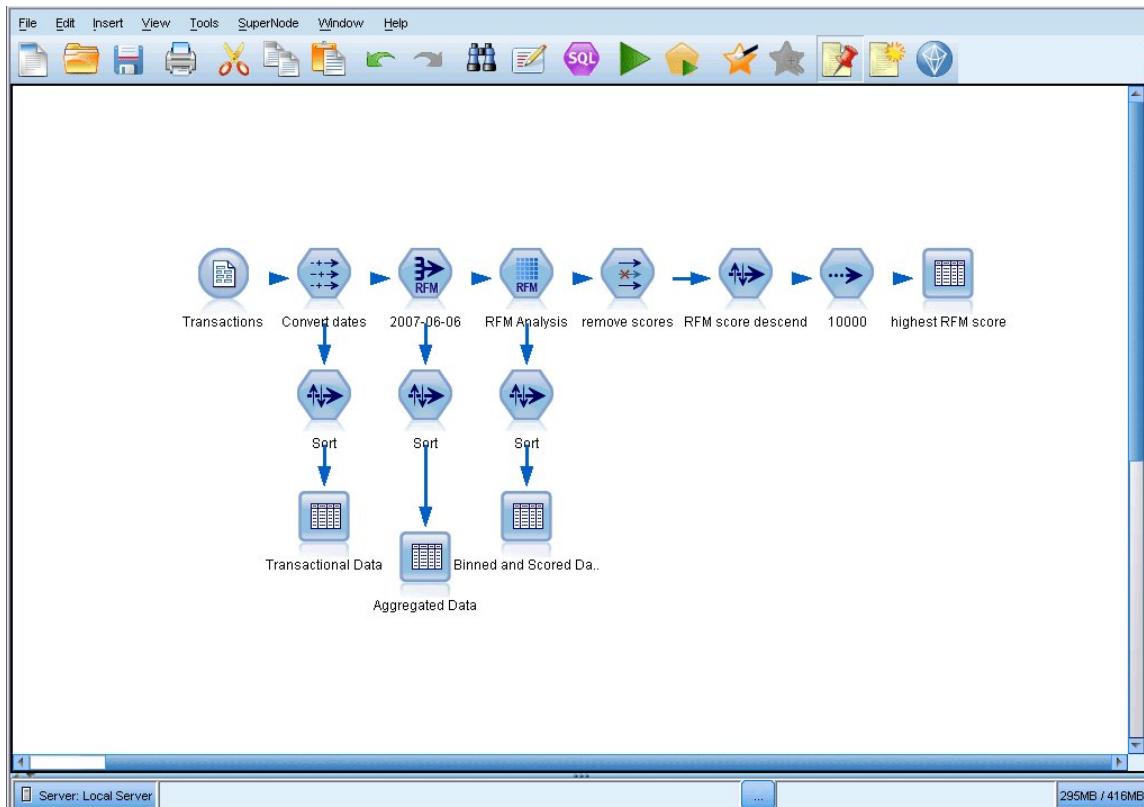


Figura 10. Tela de fluxo maximizada

Como uma alternativa para o fechamento da paleta de nós e para as áreas de janela de gerenciador e projeto, é possível usar a tela de fluxo como uma página rolável se movendo vertical e horizontalmente com as barras de rolagem ao lado e na parte inferior da janela do ModeladorSPSS.

Também é possível controlar a exibição da marcação da tela, o que consiste em comentários do fluxo, ligações de modelo e indicações de ramificação de escoragem. Para ativar ou desativar essa exibição, clique em:

Exibir > marcação de fluxo

Alterando o tamanho do ícone para um fluxo

É possível alterar o tamanho dos ícones de fluxo das seguintes formas.

- Por meio de uma configuração de propriedades de fluxo
- Por meio de um menu pop-up no fluxo
- Usando o teclado

É possível escalar a visualização de fluxo inteira para um de vários tamanhos entre 8% e 200% do tamanho do ícone padrão.

Para escalar o fluxo inteiro (método propriedades do fluxo)

1. No menu principal, escolha

Ferramentas > Propriedades do Fluxo > Opções > Layout.

2. Escolha o tamanho desejado no menu Tamanho do Ícone.
3. Clique em **Aplicar** para ver o resultado.
4. Clique em **OK** para salvar a mudança.

Para escalar o fluxo inteiro (método menu)

1. Clique com o botão direito no segundo plano do fluxo na tela.
2. Escolha **Tamanho do Ícone** e selecione o tamanho desejado.

Para escalar o fluxo inteiro (método teclado)

1. Pressione Ctrl + [-] no teclado principal para diminuir zoom para o próximo menor tamanho.
2. Pressione Ctrl + Shift + [+] no teclado principal para aumentar zoom para o próximo maior tamanho.

Observe que este método de zoom em pode não funcionar dependendo do seu sistema operacional e teclado utilizado.

Essa variável é particularmente útil para se ter uma visualização geral de um fluxo complexo. Também é possível usá-la para minimizar o número de páginas necessárias para se imprimir um fluxo.

Usando o Mouse em IBM SPSS Modelador

Os usos mais comuns do mouse no IBM SPSS Modelador incluem o seguinte:

- **Clique único.** Use o botão direito ou esquerdo do mouse para selecionar opções dos menus, abrir menus pop-up e acessar vários outros controles e opções padrão. Clique e segure o botão para mover e arrastar nós.
- **Clique duplo.** Dê um clique duplo usando o botão esquerdo do mouse para colocar nós na tela de fluxo e editar nós existentes.
- **Clique do meio.** Clique com o botão do meio do mouse e arraste o cursor para conectar nós na tela de fluxo. Dê um clique duplo no botão do meio do mouse para desconectar um nó. Se você não tiver um mouse com três botões, é possível simular essa variável pressionando a tecla Alt ao clicar e arrastar o mouse.

Usando teclas de atalho

Muitas operações de programação visual no IBM SPSS Modelador têm teclas de atalho associadas a elas. Por exemplo, é possível excluir um nó clicando nele e pressionando a tecla Delete em seu teclado. Da mesma forma, é possível salvar rapidamente um fluxo pressionando a tecla S enquanto você mantém pressionada a tecla Ctrl. Comandos de controle como esse são indicados por uma combinação de Ctrl e outra tecla -- por exemplo, Ctrl+S.

Há inúmeras teclas de atalho usadas em operações padrão do Windows, como Ctrl+X para cortar. Esses atalhos são suportados no IBM SPSS Modelador com os seguintes atalhos específicos do aplicativo.

Nota: Em alguns casos, antigas teclas de atalho usadas em IBM SPSS Modelador entram em conflito com teclas de atalho padrão do Windows. Esses atalhos antigos são suportados com a adição da tecla Alt. Por exemplo, Ctrl+Alt+C pode ser usado para ativar e desativar o cache.

Tabela 1. Teclas de atalho suportadas

Tecla de Atalho	Função
Ctrl+A	Selecionar todos
Ctrl+X	Cortar
Ctrl+N	Novo fluxo
Ctrl+O	Abrir fluxo
Ctrl+P	Imprimir
Ctrl+C	Copiar
Ctrl+V	Colar
Ctrl+Z	Desfazer
Ctrl+Q	Selecionar todos os nós de recebimentos de dados do nó selecionado
Ctrl+W	Cancelar todos os nós de recebimento de dados (alterna-se com Ctrl+Q)
Ctrl+E	Executar a partir do nó selecionado
Ctrl+S	Salvar fluxo atual
Alt+Teclas de Seta	Move os nós selecionados na tela de fluxo na direção da tecla usada
Shift+F10	Abre o menu pop-up para o nó selecionado

Tabela 2. Atalhos suportados para antigas teclas de atalho

Tecla de Atalho	Função
Ctrl+Alt+D	Duplicar nó
Ctrl+Alt+L	Carregar nó
Ctrl+Alt+R	Renomear nó
Ctrl+Alt+U	Criar nó de Entrada do Usuário
Ctrl+Alt+C	Ativar/desativar cache
Ctrl+Alt+F	Limpar o cache
Ctrl+Alt+X	Expandir Supernó
Ctrl+Alt+Z	Aumentar zoom/diminuir aumentar zoom
Excluir	Excluir nó ou conexão

Imprimindo

Os objetos a seguir podem ser impressos no IBM SPSS Modelador:

- Diagramas de fluxo
- Gráficos

- Tabelas
- Relatórios (no nó Relatório e Relatórios do Projeto)
- Scripts (das caixas de diálogo de propriedades do fluxo, Script Independente ou Script SuperNode)
- Modelos (navegadores Modelo, guias de caixa de diálogo com foco atual, visualizadores de árvore)
- Anotações (usando a guia Anotações para saída)

Para imprimir um objeto:

- Para imprimir sem visualizar, clique no botão Imprimir na barra de ferramentas.
- Para configurar a página antes da impressão, selecione **Configuração de Página** no menu Arquivo.
- Para visualizar antes de imprimir, selecione **Visualização de Impressão** no menu Arquivo.
- Para visualizar a caixa de diálogo de impressão padrão com opções para selecionar impressoras e especificar opções de aparência, selecione **Imprimir** no menu Arquivo.

automatizandoIBM SPSS Modelador

Como a mineração de dados avançada pode ser um processo complexo e, às vezes, longo, o IBM SPSS Modelador inclui vários tipos de codificação e o suporte de automação.

- **Control Language for Expression Manipulation** (CLEM) é uma linguagem para analisar e manipular os dados que fluem ao longo dos fluxos do IBM SPSS Modelador. Os mineradores de dados usam CLEM extensivamente em operações de fluxo para executar tarefas tão simples quanto derivar lucros dos dados de custo e renda ou tão complexas quanto transformar dados de log da web em um conjunto de campos e registros com informações utilizáveis.
- **Script** é uma ferramenta poderosa para automatizar processos na interface com o usuário. Scripts podem executar os mesmos tipos de ações que os usuários executam com um mouse ou teclado. Também é possível especificar saída e manipular modelos gerados.

Capítulo 4. Entendendo a mineração de dados

Visão geral de mineração de dados

Por meio de uma variedade de técnicas, **mineração de dados** identifica nuggets de informações em corpos de dados. A mineração de dados extrai informações de uma maneira que elas possam ser usadas em áreas como suporte à decisão, predição, previsões e estimativa. Geralmente os dados são volumosos, mas têm baixo valor e pouca utilidade direta em sua forma bruta. São as informações ocultas nos dados que têm valor.

Na mineração de dados, o sucesso vem da combinação do seu conhecimento (ou do especialista) em dados com técnicas de análise ativa avançadas, nas quais o computador identifica as variáveis e os relacionamentos subjacentes nos dados. O processo de mineração de dados gera modelos a partir de dados históricos que, posteriormente, são usados para previsões, detecção de padrões e mais. A técnica para a construção desses modelos é chamada **aprendizado por máquina** ou **modelagem**.

Técnicas de Modelagem

IBM SPSS Modelador inclui inúmeras tecnologias de modelagem e aprendizado por máquina, que podem ser grosseiramente agrupadas de acordo com os tipos de problemas que pretendem resolver.

- Métodos de modelagem preditiva incluem árvores de decisão, redes neurais e modelos estatísticos.
- Os modelos de armazenamento em cluster focam na identificação de grupos de registros semelhantes e na rotulagem dos registros de acordo com o grupo ao qual eles pertencem. Os métodos de armazenamento em cluster incluem Kohonen, *k*-médias e TwoStep.
- As regras de associação associam uma determinada conclusão (como a compra de um certo produto) a um conjunto de condições (a compra de vários outros produtos).
- Os modelos de rastreamento podem ser usados para rastrear dados para localizar campos e registros com maior probabilidade de serem interessantes para modelagem e identificar valores discrepantes que podem não se ajustar aos padrões conhecidos. Os métodos disponíveis incluem seleção de variável e detecção de anomalias.

Manipulação de dados e descoberta

IBM SPSS Modelador também inclui vários recursos que lhe permitem aplicar seu conhecimento aos dados:

- **Manipulação de dados.** Constrói novos itens de dados derivados dos existentes e divide os dados em subconjuntos significativos. Dados de várias origens podem ser mesclados e filtrados.
- **Navegação e visualização.** Exibe aspectos dos dados usando o nó de Auditoria de Dados para executar uma auditoria inicial, incluindo gráficos e estatísticas. A visualização avançada inclui gráficos interativos, que podem ser exportados para inclusão em relatórios do projeto.
- **Estatísticas.** Confirma relacionamentos suspeitos entre as variáveis nos dados. As estatísticas do IBM SPSS Estatísticas também podem ser usadas dentro do IBM SPSS Modelador.
- **Testes de hipótese.** Constrói modelos de como os dados se comportam e verifica esses modelos.

Normalmente você usará essas variáveis para identificar um conjunto promissor de atributos nos dados. Esses atributos podem ser alimentados para as técnicas de modelagem, que tentarão identificar relacionamentos e regras subjacentes.

Aplicativos típicos

Aplicativos típicos de técnicas de mineração de dados incluem o seguinte:

Correio direto. Determina quais grupos demográficos têm a taxa de resposta mais alta. Use essas informações para maximizar a resposta para futuras correspondências.

Escoragem de crédito. Use um histórico de crédito individual para tomar decisões sobre crédito.

Recursos humanos. Entenda antigas práticas de contratação e crie regras de decisão para aperfeiçoar seu processo de contratação.

Pesquisa médica. Crie regras de decisão que sugiram procedimentos apropriados com base em evidências médicas.

Análise de mercado. Determine quais variáveis, como geografia, preço e características do cliente, estão associadas às vendas.

Controle de qualidade. Analise dados de manufatura do produto e identifique variáveis determinando defeitos do produto.

Estudos de políticas. Use dados de pesquisas de opinião para formular políticas aplicando regras de decisão para selecionar as variáveis mais importantes.

Assistência médica. Os dados clínicos e as pesquisas de opinião do usuário podem ser combinados para descobrir variáveis que contribuam com a saúde.

Terminologia

Os termos **atributo**, **campo** e **variável** se referem a um único item de dados comum para todos os casos sendo considerados. Uma coleção de valores de atributo que se refere a um caso específico é chamada **registro**, **exemplo** ou **caso**.

Avaliando dados

Não é provável que uma mineração de dados seja produtiva, a menos que os dados que você deseja usar atendam a determinados critérios. As seções a seguir apresentam alguns dos aspectos dos dados e sua aplicação que deve ser considerada.

Assegure que os dados estejam disponíveis

Isso pode parecer óbvio, mas saiba que, embora os dados possam estar disponíveis, talvez eles não estejam em um formato que possa ser usado facilmente. IBM SPSS Modelador pode importar dados de bancos de dados (por meio de ODBC) ou de arquivos. Os dados, no entanto, podem ser retidos de alguma outra forma em uma máquina que não pode ser acessada diretamente. Eles precisarão ser transferidos por download ou submetidos a um dump de uma maneira adequada antes de poderem ser usados. Eles podem ser dispersados para diferentes bancos de dados e origens e precisam colaborar. Eles podem até não estar online. Se existir somente no papel, a entrada de dados será necessária para que seja possível iniciar a mineração de dados.

Verifique se os dados cobrem os atributos relevantes

O objetivo da mineração de dados é identificar atributos relevantes, portanto, a inclusão dessa verificação pode parecer estranha a princípio. No entanto, é muito útil examinar quais dados estão disponíveis e tentar identificar os prováveis fatores relevantes que não estão registrados. Na tentativa de prever as vendas de sorvetes, por exemplo, talvez você tenha inúmeras informações sobre estabelecimentos varejistas ou histórico de vendas, mas não tenha informações sobre o tempo e a temperatura, o que provavelmente tem um papel significativo. Atributos omissos não significam necessariamente que a mineração de dados não produzirá resultados úteis, mas podem limitar a precisão das previsões resultantes.

Uma maneira rápida de avaliar a situação é executar uma auditoria abrangente de seus dados. Antes de continuar, considere anexar um nó de Auditoria de Dados a sua origem de dados e executá-lo para gerar um relatório completo.

Tome cuidado com dados ruidosos

Dados costumam conter erros ou podem conter julgamentos subjetivos e, portanto, variáveis. Esses fenômenos são chamados coletivamente de **ruído**. Às vezes um ruído nos dados é normal. Pode haver regras subjacentes, mas não em 100% dos casos.

Normalmente, quanto mais ruído há nos dados, mais difícil fica a obtenção de resultados precisos. Porém, os métodos de aprendizado por máquina no IBM SPSS Modelador podem manipular dados ruidosos e vêm sendo usados com sucesso em conjuntos de dados contendo quase 50% de ruído.

Assegure que haja dados suficientes

Na mineração de dados, não é necessariamente o tamanho do conjunto de dados que importa. A representatividade do conjunto de dados é muito mais significativa, com sua cobertura de possíveis resultados e combinações de variáveis.

Normalmente, quanto mais atributos são considerados, mais registros serão necessários para fornecer uma cobertura representante.

Se os dados forem representantes e houver regras gerais subjacentes, talvez uma amostra de dados de alguns milhares (ou centenas) de registros forneçam resultados igualmente bons como milhões, e você terá esses resultados mais rapidamente.

Busque especialistas em dados

Em muitos casos, você estará trabalhando com seus próprios dados e, portanto, estará totalmente familiarizado com seu conteúdo e significado. No entanto, se você estiver trabalhando com dados de outro departamento de sua organização ou de seu cliente, talvez seja desejável que você tenha acesso aos especialistas que conhecem os dados. Eles podem guiá-lo na identificação de atributos relevantes e ajudá-lo a interpretar os resultados da mineração de dados, distinguindo nuggets verdadeiros de informações "baratas" ou artefatos causados por anomalias nos conjuntos de dados.

Uma estratégia para mineração de dados

Assim como acontece com a maioria dos esforços de negócios, a mineração de dados é muito mais efetiva se for feita de uma maneira sistemática e planejada. Mesmo com ferramentas de mineração de dados de ponta, como IBM SPSS Modelador, a maioria do trabalho na mineração de dados requer que um analista de negócios instruído mantenha o processo atualizado. Para guiar seu planejamento, responda as questões a seguir:

- Qual problema substantivo você deseja resolver?
- Quais origens de dados estão disponíveis e quais partes dos dados são relevantes para o problema atual?
- Que tipo de pré-processamento e limpeza de dados você precisa fazer antes de iniciar a mineração de dados?
- Qual(is) técnica(s) de mineração de dados você usará?
- Como você avaliará os resultados da análise de mineração de dados?
- Como você irá tirar o máximo proveito das informações obtidas da mineração de dados?

O processo de mineração de dados típico pode se tornar complicado muito rapidamente. Há muito o que acompanhar -- problemas de negócios complexos, várias origens de dados, qualidade de dados variada nas origens de dados, matriz de técnicas de mineração de dados, diferentes maneiras de medir o sucesso da mineração de dados, entre outras coisas.

Para você se manter atualizado, é útil ter um modelo de processo definido explicitamente para mineração de dados. O modelo de processo ajuda você a responder as questões listadas anteriormente nesta seção e se certificar de que os pontos importantes serão abordados. Isso serve como um roteiro de mineração de dados para que você não se perca ao se aprofundar nas complexidades de seus dados.

O processo de mineração de dados sugerido para uso com ModeladorSPSS é o processo padrão de vários segmentos de mercados para mineração de dados (CRISP-DM). Como o nome já diz, esse modelo é projetado como um modelo geral que pode ser aplicado a uma grande variedade de problemas de mercado e negócios.

O modelo de processo CRISP-DM

O modelo de processo CRISP-DM geral inclui seis fases que abordam os principais problemas da mineração de dados. As seis fases se ajustam em um processo cíclico projetado para incorporar mineração de dados a suas práticas de negócios maiores.

As seis fases incluem:

- **Entendimento de negócios.** Talvez esteja seja a fase mais importante da mineração de dados. Entendimento de negócios inclui determinar objetivos de negócios, avaliar a situação, determinar objetivos de mineração de dados e produzir um plano de projeto.
- **Entendimento de dados.** Dados fornecem os "materiais brutos" da mineração de dados. Essa fase aborda a necessidade de entender quais são seus recursos de dados e as características desses recursos. Isso inclui coletar dados iniciais, descrever dados, explorar dados e verificar a qualidade dos dados. O nó de Auditoria de Dados disponível na paleta de nós de Saída é uma ferramenta indispensável para o entendimento dos dados.
- **Preparação de dados.** Após catalogar seus recursos de dados, você precisará preparar seus dados para mineração. As preparações incluem selecionar, limpar, construir, integrar e formatar dados.
- **Modelagem.** Esta é, sem dúvida, a parte mais chamativa da mineração de dados, na qual métodos de análise sofisticados são usados para extrair informações dos dados. Essa fase envolve a seleção de técnicas de modelagem, geração de designs de teste e construção e avaliação de modelo.
- **Avaliação.** Após ter escolhido seus modelos, você estará pronto para avaliar como os resultados da mineração de dados podem ajudá-lo a atingir seus objetivos de negócios. Elementos dessa fase incluem avaliar resultados, revisar o processo de mineração de dados e determinar os próximos passos.
- **Implementação.** Agora que você investiu todo seu esforço, é hora de colher os benefícios. Essa fase foca na integração do seu conhecimento em processos de negócios do dia a dia para resolver seus problemas de negócios originais. Essa fase inclui implementação de plano, monitoramento e manutenção, produção de relatório final e revisão de projeto.

Há alguns pontos chave nesse modelo de processo. Primeiro, embora haja uma tendência geral para o processo fluir pelos passos na ordem esboçada nos parágrafos anteriores, também há inúmeros locais nos quais as fases influenciam umas às outras de maneira não linear. Por exemplo, geralmente a preparação de dados precede a modelagem. No entanto, geralmente as decisões tomadas e as informações reunidas durante a fase de modelagem podem levá-lo a repensar partes da fase de preparação de dados, o que pode então apresentar novos problemas de modelagem. As duas fases realimentam uma à outra até que ambas tenham sido resolvidas adequadamente. Da mesma forma, a fase de avaliação pode levá-lo a reavaliar seu entendimento de negócios original e você pode concluir que esteve tentando responder à questão errada. Nesse ponto, é possível revisar seu entendimento de negócios e continuar com o restante do processo novamente com uma resposta melhor em mente.

O segundo ponto chave é a natureza iterativa da mineração de dados. Raramente, ou nunca, você irá simplesmente planejar um projeto de mineração de dados, concluí-lo e depois empacotar seus dados e ir para casa. A mineração de dados para abordar demandas de clientes é um esforço constante. O conhecimento ganho com um ciclo de mineração de dados quase invariavelmente levará a novas questões, novos problemas e novas oportunidades para identificar e atender às necessidades do cliente. Essas novas questões, problemas e oportunidades geralmente podem ser abordados minerando os dados novamente. Esse processo de mineração e identificação de novas oportunidades deve fazer parte da forma como você pensa sobre negócios e da base de sua estratégia geral de negócios.

Essa introdução fornece apenas uma breve visão geral do modelo de processo CRISP-DM. Para obter detalhes completos sobre o modelo, consulte os seguintes recursos:

- O *Guia do CRISP-DM*, que pode ser acessado com outra documentação na pasta *|Documentation* no disco de instalação.
- Sistema de Ajuda CRISP-DM, disponível no menu Iniciar ou clicando em **Ajuda do CRISP-DM** no menu Ajuda no IBM SPSS Modelador.

Tipos de Modelos

O IBM SPSS Modelador oferece uma variedade de métodos de modelagem tomados do aprendizado de máquina, inteligência artificial e estatística. Os métodos disponíveis na paleta Modelagem permitem derivar informações novas a partir dos dados, e desenvolver modelos preditivos. Cada método possui certas forças e é mais adequado para certos tipos de problemas.

O *Guia de Aplicativos do IBM SPSS Modelador* fornece exemplos para vários desses métodos, com uma introdução geral ao processo de modelagem. Este guia está disponível como um tutorial online, e também em formato PDF. Consulte o tópico [“Exemplos de Aplicação”](#) na página 4 para obter mais informações.

Os métodos de modelagem são divididos nessas categorias:

- Supervisionados
- Associação
- Segmentação

modelos supervisionados

Modelos supervisionados usam os valores de um ou mais campos **input** para prever o valor de uma ou mais saída, ou **destino**, campos. Alguns exemplos dessas técnicas são: árvores de decisão (algoritmo Árvore C e R, algoritmo de árvore estatística eficiente, rápido e imparcial, algoritmo Detector de Interação Automática Chi-quadrado e C5.0), regressão (algoritmos linear, de logística, linear generalizado e algoritmo de regressão de Cox), redes neurais, Support Vector Machines e redes bayesianas.

Os modelos supervisionados ajudam as organizações a prever um resultado conhecido, como se um cliente irá comprar ou sair ou se uma transação se encaixa em um padrão conhecido de fraude. As técnicas de modelagem incluem aprendizado por máquina, indução de regra, identificação de subgrupo, métodos estatísticos e geração de modelos múltiplos.

Nós Supervisionados



O nó Previsor Categórico Automático cria e compara inúmeros modelos diferentes para resultados binários (sim ou não, perda de clientes ou não, e assim por diante), permitindo escolher a melhor abordagem para uma determinada análise. Vários algoritmos de modelagem são suportados, possibilitando a seleção dos métodos que você deseja usar, as opções específicas para cada e os critérios para comparar os resultados. O nó gera um conjunto de modelos com base nas opções especificadas e classifica os melhores candidatos de acordo com os critérios que você especificar.



O nó Previsor Contínuo Automático estima e compara modelos para resultados de intervalos numéricos contínuos usando vários métodos diferentes. O nó trabalha da mesma maneira que o nó Previsor Categórico Automático, permitindo escolher os algoritmos para usar e experimentar com várias combinações de opções em uma única passagem de modelagem. Os algoritmos suportados incluem redes neurais, Árvore C e R, algoritmo Detector de Interação Automática Chi-quadrado, regressão linear, regressão linear generalizada e Support Vector Machines (SVM). Os modelos podem ser comparados com base em correlação, erro relativo ou número de variáveis utilizadas.



O nó Classificação e Regressão (C e R) gera uma árvore de decisão que permite prever ou classificar futuras observações. O método usa particionamento recursivo para dividir os registros de treinamento em segmentos, minimizando a impureza em cada etapa, em que um nó na árvore será considerado “puro” se 100% dos casos no nó caírem em uma categoria específica do campo de destino. Campos de destino e entrada podem ser intervalos numéricos ou categóricos (nominal, ordinal ou sinalizadores); todos os splits são binários (apenas dois subgrupos).



O nó QUEST fornece um método de classificação binário para construir árvores de decisão, projetadas para reduzir o tempo de processamento necessário para grandes análises de C e R, enquanto também reduz a tendência localizada nos métodos de árvore de classificação para favorecer entradas que permitam mais divisões. Campos de entrada podem ser intervalos numéricos (contínuos), mas o campo de resposta deve ser categórico. Todas as splits são binárias.



O nó CHAID gera árvores de decisão usando estatísticas qui-quadrado para identificar as divisões ideais. Ao contrário dos nós Árvore C e R e QUEST, o CHAID pode gerar árvores não binárias, o que significa que algumas divisões têm mais de duas ramificações. Os campos de destino e de entrada podem ser um intervalo numérico (contínuo) ou categóricos. Exaustivo CHAID é uma modificação de CHAID que faz um trabalho mais minucioso de examinar todas as divisões possíveis mas demora mais tempo para computar.



O nó C5.0 constrói uma árvore de decisão ou um conjunto de regras. O modelo funciona dividindo a amostra com base no campo que fornece o ganho máximo de informações em cada nível. O campo de destino deve ser categórico. São permitidas várias divisões em mais de dois subgrupos.



O nó de Lista de Decisão identifica subgrupos, ou segmentos, que mostram uma probabilidade maior ou menor de um determinado resultado binário relativo à população geral. Por exemplo, você pode procurar clientes que provavelmente não irão migrar para o concorrente ou têm maior probabilidade de responder favoravelmente a uma campanha. É possível incorporar seu conhecimento de negócios no modelo, incluindo seus próprios segmentos personalizados e visualizando modelos alternativos lado a lado para comparar os resultados. Os modelos de Lista de Decisão consistem em uma lista de regras na qual cada regra possui uma condição e um resultado. As regras são aplicadas na ordem, e a primeira regra que corresponder determina o resultado.



Os modelos de regressão linear preveem um alvo contínuo baseado em relações lineares entre o alvo e um ou mais preditores.



O nó PCA/Fator fornece técnicas poderosas de redução de dados para reduzir a complexidade de seus dados. A análise de componentes principais (PCA) localiza combinações lineares dos campos de entrada que realizam as melhores tarefas de captura de variância no conjunto de campos inteiro, no qual os componentes são ortogonais (perpendiculares) uns aos outros. A análise fatorial tenta identificar os fatores subjacentes que explicam o padrão de correlações dentro de um conjunto observado de campos. Para ambas as abordagens, o objetivo é localizar um pequeno número de campos derivados que sumarizam efetivamente as informações no conjunto de campos original.



O nó Seleção de Variável verifica campos de entrada para remoção com base em um conjunto de critérios (como a porcentagem de valores omissos); em seguida, classifica a importância de entradas restantes com relação a um destino especificado. Por exemplo, dado um conjunto de dados com centenas de entradas potenciais, quais têm maior probabilidade de serem úteis na modelagem de resultados do paciente?



A análise discriminante faz pressupostos mais rigorosos do que a regressão logística mas pode ser uma alternativa valiosa ou complementar a uma análise de regressão logística quando essas suposições forem atendidas.



Regressão logística é uma técnica estatística para classificar registros com base em valores de campos de entrada. É análoga à regressão linear mas leva um campo de destino categórico em vez de um intervalo numérico.



O Modelo Linear Generalizado expande o modelo linear geral para que a variável dependente esteja linearmente relacionada aos fatores e às covariáveis por meio de uma função de ligação especificada. Além do mais, o modelo permite que a variável dependente tenha uma distribuição não normal. Ele cobre a funcionalidade de um grande número de modelos estatísticos, incluindo regressão linear, regressão logística, modelos de loglinear para dados de contagem e modelos de sobrevivência censurados de intervalo.



Um modelo linear generalizado misto (GLMM) estende o modelo linear para que a resposta possa ter uma distribuição não normal, esteja linearmente relacionada aos fatores e às covariáveis via uma função de ligação especificada e para que as observações possam ser correlacionadas. Os modelos lineares generalizados mistos abrangem uma ampla variedade de modelos, desde regressão linear simples até modelos multiníveis complexos para dados longitudinais não normais.



O nó de Regressão de Cox permite construir um modelo de sobrevivência para dados de sobrevivência na presença de registros censurados. O modelo produz uma função de sobrevivência que prevê a probabilidade de que o evento de interesse tenha ocorrido em um determinado momento (t) para valores dados das variáveis de entrada.



O nó Support Vector Machine (SVM) permite ordenar dados em dois grupos sem super ajuste. SVM trabalha bem com conjuntos de dados grandes, como aqueles com um número muito grande de campos de entrada.



O nó de Rede Bayesiana permite construir um modelo de probabilidade combinando evidências observadas e registradas com conhecimento do mundo real para estabelecer a probabilidade de ocorrências. O nó focaliza as redes Tree Augmented Naïve Bayes (TAN) e Markov Blanket que são usadas principalmente para classificação.



O nó Self-Learning Response Model (SLRM) permite construir um modelo no qual um único novo caso, ou um pequeno número de novos casos, pode ser usado para estimar novamente o modelo sem precisar treinar o modelo outra vez usando todos os dados.



O nó Séries Temporais estima modelos de suavização exponencial, Média Móvel Integrada AutoRegressiva (ARIMA) univariada e ARIMA multivariada (ou função de transferência) para dados de séries temporais e produz previsões do desempenho futuro. Este nó do Time Series é semelhante ao nó do Time Series anterior que foi descontinuado em ModeladorSPSS versão 18. No entanto, este nó do Time Series mais recente foi projetado para aproveitar a potência de IBM SPSS Servidor analítico para processar big data, e exibir o modelo resultante no visualizador de saída que foi adicionado em ModeladorSPSS versão 17.



O nó *k*-Nearest Neighbor (KNN) associa um novo caso à categoria ou valor dos objetos *k* mais próximos dele no espaço do preditor, em que *k* é um número inteiro. Casos semelhantes ficam próximos uns dos outros e os casos diferentes ficam distantes uns dos outros.



O nó Spatio-Temporal Prediction (STP) usa dados que contêm dados de localização, campos de entrada para previsão (preditores), um campo de tempo e um campo de resposta. Cada local possui várias linhas nos dados que representam os valores de cada preditor em cada momento da medição. Após os dados serem analisados, eles podem ser usados para prever valores de destino em qualquer local dentro dos dados de forma que são usados na análise.

Modelos da Associação

Modelos de associação localizam padrões em seus dados nos quais uma ou mais entidades (como eventos, compras ou atributos) são associadas a uma ou mais entidades. Os modelos constroem conjuntos de regras que definem esses relacionamentos. Aqui os campos dentro dos dados podem agir como entradas e respostas. Você poderia localizar essas associações manualmente, mas os algoritmos de regra de associação fazem isso com muito mais rapidez e podem explorar padrões mais complexos. Os modelos A Priori e Carma são exemplos do uso desses algoritmos. Um outro tipo de modelo de associação é um modelo de detecção de sequência, que localiza padrões sequenciais em dados estruturados por tempo.

Os modelos de associação são mais úteis ao prever múltiplos resultados—por exemplo, os clientes que compraram o produto X também compraram Y e Z. Os modelos de associação associam uma conclusão particular (como a decisão de comprar algo) com um conjunto de condições. A vantagem dos algoritmos de regra de associação sobre os algoritmos de árvore de decisão mais padrão (C5.0 and C&RT) é que associações podem existir entre qualquer um dos atributos. Um algoritmo de árvore de decisão construirá regras com apenas uma única conclusão, ao passo que os algoritmos de associação tentam localizar muitas regras, cada qual podendo ter uma conclusão diferente.

Nós de associação



O nó A Priori extrai um conjunto de regras dos dados, removendo as regras com o conteúdo de informações mais alto. A Priori oferece cinco métodos diferentes de seleção de regras e usa um esquema de indexação sofisticado para processar conjuntos de dados grandes com eficiência. Para grandes problemas, o A Priori geralmente é mais rápido para treinar; ele não tem um limite arbitrário no número de regras que podem ser retidas e pode manipular regras com até 32 condições prévias. O A Priori requer que os campos de entrada e saída sejam todos categóricos, mas entrega melhor desempenho por ser otimizado para esse tipo de dado.



O modelo CARMA extrai um conjunto de regras dos dados sem requerer a especificação de campos de entrada ou saída. Em contraste com Apriori o nó CARMA oferece configurações de construção para suporte de regra (suporte para tanto antecedentes e consequentes) em vez de apenas suporte antecedente. Isso significa que as regras geradas podem ser usadas para uma variedade maior de aplicativos—por exemplo, para localizar uma lista de produtos ou serviços (antecedentes) cujo consequente é o item que você deseja promover nesta temporada de férias.



O nó de Sequência descobre regras de associação em dados sequenciais ou orientados por tempo. Uma sequência é uma lista de conjuntos de itens que tende a ocorrer em uma ordem previsível. Por exemplo, um cliente que compra um aparelho de barbear e uma loção pós-barba pode comprar um creme de barbear na próxima compra. O nó Sequência é baseado no algoritmo de regras de associação CARMA, que usa um método eficiente de duas passagens para localizar sequências.



O nó de Regras de Associação é semelhante ao nó A Priori; no entanto, ao contrário do A Priori, o nó de Regras de Associação pode processar dados de lista. Além disso, o nó de Regras de Associação pode ser usado com IBM SPSS Servidor analítico para processar Big Data e aproveitar a vantagem do processamento paralelo mais rápido.

Modelos de Segmentação

Modelos de segmentação dividem os dados em segmentos, ou clusters, de registros com padrões semelhantes de campos de entrada. Como eles estão interessados somente nos campos de entrada, os modelos de segmentação não têm um conceito de campos de saída ou resposta. Os exemplos de modelos de segmentação são redes Kohonen, armazenamento em cluster de k-médias, clusterização em dois passos e detecção de anomalias.

Segmentação de modelos (também conhecidos como "modelos de armazenamento em cluster") são úteis nos casos em que o resultado específico é desconhecido (por exemplo, durante a identificação de novos padrões de fraude ou identificação de grupos de interesse em sua base de cliente). Os modelos de armazenamento em cluster focam na identificação de grupos de registros semelhantes e na rotulagem dos registros de acordo com o grupo ao qual eles pertencem. Isso é feito sem o benefício do conhecimento prévio sobre os grupos e suas características, e isso distingue os modelos de armazenamento em cluster de outras técnicas de modelagem porque não há saída predefinida ou campo de resposta para o modelo prever. Não há respostas certas ou erradas para esses modelos. Seu valor é determinado pela sua capacidade de capturar agrupamentos interessantes nos dados e fornecer descrições úteis desses agrupamentos. Os modelos de armazenamento em cluster geralmente são usados para criar clusters ou segmentos que são usados como entradas em análises subsequentes (por exemplo, segmentando possíveis clientes em subgrupos homogêneos).

Nós de segmentação



O nó Cluster Automático estima e compara modelos de armazenamento em cluster, que identificam grupos de registros com características semelhantes. O nó trabalha da mesma maneira que outros nós de modelagem automatizados, permitindo experimentar várias combinações de opções em uma única passagem de modelagem. Os modelos podem ser comparados usando medidas básicas com as quais tentar filtrar e classificar a utilidade dos modelos de cluster, além de fornecer uma medida com base na importância de campos particulares.



O nó K-médias armazena em cluster os dados configurados em grupos distintos (ou clusters). O método define um número fixo de clusters, designa registros aos clusters iterativamente e ajusta os centros de cluster até que os refinamentos adicionais não possam mais melhorar o modelo. Em vez de tentar prever um resultado, o k-médias usa um processo conhecido como aprendizado não supervisionado para descobrir padrões no conjunto de campos de entrada.



O nó Kohonen gera um tipo de rede neural que pode ser usada para agrupar o conjunto de dados em grupos distintos. Quando a rede é totalmente treinada, os registros que são semelhantes devem estar próximos no mapa de saída, enquanto os registros que forem diferentes estarão distantes. É possível verificar o número de observações capturadas por cada unidade no nugget do modelo para identificar as unidades fortes. Isso pode dar uma noção do número apropriado de clusters.



O nó TwoStep usa um método de clusterização em dois passos. O primeiro passo faz uma simples passagem pelos dados para compactar os dados de entrada brutos em um conjunto gerenciável de subclusters. O segundo passo usa um método de armazenamento em cluster hierárquico para mesclar progressivamente os subclusters em clusters cada vez maiores. TwoStep tem a vantagem de estimar automaticamente o número ideal de clusters para dados de treinamento. Ele pode manipular tipos mistos de campos e conjuntos grandes de dados de forma eficiente.



O nó de Detecção de Anomalias identifica casos incomuns, ou valores discrepantes, que não estão em conformidade com os padrões de dados “normais”. Com este nó, é possível identificar outliers mesmo que eles não se encaixem em nenhum padrão anteriormente conhecido e mesmo que você não esteja exatamente certo do que você está procurando.

Modelos de Mineração dentro da Base de Dados

IBM SPSS Modelador suporta integração com ferramentas de mineração de dados e modelagem que estão disponíveis a partir de fornecedores de banco de dados, incluindo Oracle Data Miner e Microsoft Analysis Services. É possível construir, escorar e armazenar modelos dentro do banco de dados-tudo isso de dentro do aplicativo IBM SPSS Modelador. Para obter detalhes completos, consulte o *IBM SPSS Modelador In-Database Mining Guide*.

modelos do IBM SPSS Statistics

Se você tiver uma cópia do IBM SPSS Estatísticas instalada e licenciada em seu computador, é possível acessar e executar certas rotinas do IBM SPSS Estatísticas de dentro do IBM SPSS Modelador para construir e escorar modelos.

Exemplos de mineração de dados

A melhor maneira de se aprender sobre mineração de dados na prática é começar com um exemplo. Inúmeros exemplos de aplicativo estão disponíveis no *Guia de Aplicativos do IBM SPSS Modelador*, que fornece introduções direcionadas sintetizadas para técnicos e métodos de modelagem específicos. Consulte o tópico [“Exemplos de Aplicação”](#) na página 4 para obter mais informações.

Capítulo 5. Construindo fluxos

Visão geral da construção de fluxo

A mineração de dados usando IBM SPSS Modelador foca no processo de execução de dados através de uma série de nós, referida como um fluxo de . Essa série de nós representa operações a serem executadas nos dados, enquanto ligações entre os nós indicam a direção do fluxo de dados. Normalmente você usa um fluxo de dados para ler dados no IBM SPSS Modelador, executá-lo por meio de uma série de manipulações e depois enviá-lo para um destino, como uma tabela ou um visualizador.

Por exemplo, suponha que você deseja abrir uma origem de dados, incluir um novo campo, selecionar registros com base em valores no novo campo e depois exibir os resultados em uma tabela. Nesse caso, seu fluxo de dados consistiria em quatro nós:



Um nó de Arquivo Variável, que você configura para ler os dados da origem de dados.



Um nó de Derivação, que você usa para incluir o campo novo calculado no conjunto de dados.



Um nó de Seleção, que você usa para configurar critérios de seleção para excluir registros do fluxo de dados.



Um nó de Tabela, que você usa para exibir os resultados de sua manipulações na tela.

Construindo fluxos de dados

A interface do ModeladorSPSS exclusiva permite minar seus dados visualmente trabalhando com diagramas de fluxos de dados. No nível mais básico, é possível construir um fluxo de dados usando os passos a seguir:

- Inclua nós na tela de fluxo.
- Conecte os nós para formar um fluxo.
- Especifique as opções de nó ou de fluxo.
- Execute o fluxo.

Esta seção contém informações mais detalhadas sobre como trabalhar com nós para criar fluxos de dados mais complexos. Discute também as opções e configurações para nós e fluxos. Para exemplos passo a passo de construção de fluxo usando os dados enviados com o ModeladorSPSS (na pasta Demos da instalação do programa, consulte [“Exemplos de Aplicação”](#) na página 4).

Trabalhando com nós

Nós são usados no IBM SPSS Modelador para ajudar a explorar dados. Vários nós na área de trabalho representam diferentes objetos e ações. A paleta na parte inferior da janela do IBM SPSS Modelador contém todos os nós possíveis usados na construção do fluxo.

Há inúmeros tipos de nós. **Nós de origem** reúnem dados no fluxo e estão localizados na guia Origens da paleta de nós. **Nós de processo** executam operações em campos e registros de dados individuais

e podem ser localizados nas guias Operações de Registro e Operações de Campo da paleta. **Nós de saída** produzem uma variedade de saída de para dados, gráficos e resultados de modelo, além de serem incluídos nas guias Gráficos, Saída e Exportação da paleta de nós. **Nós de modelagem** usam algoritmos estatísticos para criar nuggets de modelos e estão localizados na guia Modelagem e (se ativada) na guia Modelagem da Base de Dados da paleta de nós. Veja o tópico [“Paleta de nós”](#) na página 13 para obter mais informações.

Você conecta os nós para formar fluxos que, quando executados, permitem visualizar relacionamentos e tirar conclusões. Fluxos são como scripts--é possível salvá-los e reutilizá-los com diferentes arquivos de dados.

Um nó executável que processa dados de fluxo é conhecido como **nó terminal**. Um nó de saída ou modelagem será um nó terminal se estiver localizado no final do fluxo ou ramificação do fluxo. Não é possível conectar nós adicionais a um nó terminal.

Nota: Você pode customizar a paleta Nodes. Consulte o tópico [“Customizando a Paleta de Nós”](#) na página 244 para obter informações adicionais.

Incluindo nós em um fluxo

Há várias maneiras de incluir nós em um fluxo a partir da paleta de nós:

- Dê um clique duplo em um nó na paleta. *Nota:* Um clique duplo em um nó o conecta automaticamente ao fluxo atual. Veja o tópico [“Conectando nós em um fluxo”](#) na página 32 para obter mais informações.
- Arraste e solte um nó da paleta para a tela de fluxo.
- Clique em um nó na paleta e, em seguida, clique na tela de fluxo.
- Selecione uma opção apropriada no menu Inserir do IBM SPSS Modelador.

Após ter incluído um nó na tela de fluxo, dê um clique duplo no nó para exibir sua caixa de diálogo. As opções disponíveis dependem do tipo de nó sendo incluído. Para obter informações sobre controles específicos dentro da caixa de diálogo, clique em seu botão **Ajuda**.

Removendo Nós

Para remover um nó do fluxo de dados, clique nele e pressione a tecla Delete ou clique com o botão direito e selecione **Excluir** no menu.

Conectando nós em um fluxo

Nós incluídos na tela de fluxo não formam um fluxo de dados até que eles tenham sido conectados. As conexões entre os nós indicam a direção dos dados conforme eles fluem de uma operação para a seguinte. Há inúmeras maneiras de se conectar nós para formar um fluxo: dando um clique duplo, usando o botão do meio do mouse ou manualmente.

Para incluir e conectar nós dando um clique duplo

A maneira mais simples de formar um fluxo é dando um clique duplo nos nós na paleta. Esse método conecta automaticamente o novo nó ao nó selecionado na tela de fluxo. Por exemplo, se a tela contiver um nó de Banco de Dados, é possível selecionar esse nó e depois dar um clique duplo no próximo nó da paleta, como um nó de Derivação. Essa ação conecta automaticamente o nó de Derivação ao nó de Banco de Dados existente. É possível repetir esse processo até atingir um nó terminal, como um nó de Histograma ou Tabela, em cujo ponto quaisquer novos nós serão conectados ao último nó de envio de dados não terminal.

Para conectar nós usando o botão do meio do mouse

Na tela de fluxo, é possível clicar e arrastar de um nó para outro usando o botão do meio do mouse. (Se seu mouse não tiver um botão do meio, é possível simular isso pressionando a tecla Alt enquanto você arrasta com o mouse de um nó para outro.)

Para conectar nós manualmente

Se você não tiver um botão do meio do mouse e preferir conectar nós manualmente, é possível usar o menu pop-up para um nó para conectá-lo a outro nó já na tela.

1. Clique com o botão direito no nó do qual deseja iniciar a conexão. Isso abre o menu do nó.
2. No menu, clique em **Conectar**.
3. Um ícone de conexão é exibido no nó inicial e cursor. Clique em um segundo nó na tela para conectar os dois nós.

Durante a conexão dos nós, há várias diretrizes para serem seguidas. Você receberá uma mensagem de erro se tentar estabelecer algum dos tipos de conexão a seguir:

- Uma conexão levando a um nó de origem
- Uma conexão a partir de um nó terminal
- Um nó com mais do que seu número máximo de conexões de entrada
- Conexão de dois nós já conectados
- Circularidade (dados retornam para um nó do qual já fluíram)

Efetuando bypass nos nós em um fluxo

Quando você efetua bypass em um nó no fluxo de dados, todas as suas conexões de entrada e saída são substituídas pelas conexões que levaram diretamente dos seus nós de entrada para seus nós de saída. Se o nó não tiver conexões de entrada e saída, todas as suas conexões serão excluídas em vez de roteadas novamente.

Por exemplo, você pode ter um fluxo que deriva um novo campo, filtra campos e depois explora os resultados em um histograma e em uma tabela. Se também desejar visualizar o mesmo gráfico e tabela para os dados *antes* de os campos serem filtrados, é possível incluir novos nós de Histograma e Tabela ou efetuar bypass no nó de Filtro. Quando você efetua bypass no nó de Filtro, as conexões com o gráfico e a tabela passam diretamente do nó de Derivação. O nó de Filtro é desconectado do fluxo.

Para efetuar bypass em um nó

1. Na tela de fluxo, use o botão do meio do mouse para dar um clique duplo no nó no qual deseja efetuar bypass. Como alternativa, é possível usar Alt+clique duplo.

Nota: É possível desfazer essa ação clicando em **Desfazer** no menu Editar ou pressionando Ctrl+Z.

Desativando nós em um fluxo

Os nós de processamento com uma única entrada dentro dos fluxos podem ser desativados, cujo resultado é o nó ser ignorado durante a execução do fluxo. Isso evita que você precise remover ou efetuar bypass do nó e significa que é possível deixá-lo conectado aos nós restantes. Ainda é possível abrir e editar as configurações do nó; no entanto, nenhuma mudança entrará em vigor até você ativar o nó novamente.

Por exemplo, você pode ter um fluxo que filtre vários campos e depois construa modelos com o conjunto de dados reduzido. Se também desejar construir os mesmos modelos *sem* os campos serem filtrados, para ver se eles melhoraram os resultados do modelo, é possível desativar o nó de Filtro. Quando você desativa o nó de Filtro, as conexões com o nó de modelagem passam diretamente pelo nó de Derivação para o nó Tipo.

Para desativar um nó

1. Na tela de fluxo, clique com o botão direito no nó que deseja desativar.
2. Clique em **Desativar Nó** no menu pop-up.

Alternativamente, você pode clicar em **Nó > Nó Desativar** no menu Editar. Quando quiser incluir o nó de volta no fluxo, clique em **Ativar Nó** da mesma maneira.

Nota: É possível desfazer essa ação clicando em **Desfazer** no menu Editar ou pressionando Ctrl+Z.

Incluindo nós em conexões existentes

É possível incluir um novo nó entre dois nós conectados arrastando a seta que conecta os dois nós.

1. Com o botão do meio do mouse, clique e arraste a seta de conexão na qual deseja inserir o nó. Alternativamente, é possível manter a tecla ALT pressionada enquanto você clica e arrasta para simular o botão do meio do mouse.
2. Arraste a conexão para o nó que deseja incluir e solte o botão do mouse.

Nota: É possível remover novas conexões do nó e restaurar o original **efetuando bypass** no nó.

Excluindo conexões entre nós

Para excluir a conexão entre dois nós:

1. Clique com o botão direito na seta de conexão.
2. No menu, clique em **Excluir Conexão**.

Para excluir todas as conexões para e de um nó, execute um dos seguintes procedimentos:

- Selecione o nó e pressione F3.
- Selecione o nó e, no menu principal, clique em:

Editar > Nó > Desconectar

Configurando opções para nós

Uma vez que creat e conecte nós, há várias opções para personalizar nós. Clique com o botão direito em um nó e selecione uma das opções de menu.

- Clique em **Editar** para abrir a caixa de diálogo para o nó selecionado.
- Clique em **Conectar** para conectar manualmente um nó ao outro.
- Clique em **Desconectar** para excluir todos as ligações para e do nó.
- Clique em **Renomear e Anotar** para abrir a guia Anotações da caixa de diálogo de edição.
- Clique em **Novo Comentário** para incluir um comentário relacionado ao nó. Veja o tópico “[Incluindo comentários e anotações em nós e fluxos](#)” na [página 51](#) para obter mais informações.
- Clique em **Desativar Nó** para “ocultar” o nó durante o processamento. Para deixar o nó visível novamente para processamento, clique em **Ativar Nó**. Consulte o tópico “[Desativando nós em um fluxo](#)” na [página 33](#) para obter mais informações.
- Clique em **Cortar** ou **Excluir** para remover o(s) nó(s) selecionado(s) da tela de fluxo. *Nota:* Um clique em **Cortar** permite colar nós, enquanto **Excluir** não.
- Clique em **Copiar Nó** para fazer uma cópia do nó sem conexões. Isso pode ser incluído em um fluxo novo ou existente.
- Clique em **Carregar Nó** para abrir um nó salvo anteriormente e carregar suas opções no nó atualmente selecionado. Os nós devem ser de tipos idênticos.
- Clique em **Recuperar Nó** para recuperar um nó de um IBM SPSS Reppositório de serviços de colaboração e implantação conectado.
- Clique em **Salvar Nó** para salvar os detalhes do nó em um arquivo. É possível carregar detalhes do nó somente em outro nó do mesmo tipo.
- Clique em **Armazenar Nó** para armazenar o nó selecionado em um IBM SPSS Reppositório de serviços de colaboração e implantação conectado.
- Clique em **Cache** para expandir o menu com opções para armazenar em cache o nó selecionado.
- Clique em **Mapeamento de Dados** para expandir o menu com opções para mapear dados para uma nova origem ou especificar campos obrigatórios.
- Clique em **Criar SuperNode** para expandir o menu com opções para criar um SuperNode no fluxo atual.

- Clique em **Gerar Nô de Entrada do Usuário** para substituir o nó selecionado. Os exemplos gerados por esse nó terão os mesmos campos que o nó atual.
- Clique em **Executar Daqui** para executar o recebimento de dados de todos os nós terminais a partir do nó selecionado.

Armazenando opções em cache para nós

Para otimizar a execução do fluxo, é possível configurar um *cache* em um nó não terminal. Quando você configura um cache em um nó, o cache é preenchido com os dados que são passados pelo nó na próxima vez que você executa o fluxo de dados. Daí em diante, os dados são lidos do cache (que é armazenado em disco em um diretório temporário) e não da origem de dados.

O armazenamento em cache é mais útil após uma operação demorada, como uma ordenação, mesclagem ou agregação. Por exemplo, suponha que você tenha um nó de origem configurado para ler dados de vendas de um banco de dados e um nó agregado que resuma as vendas por localização. É possível configurar um cache no nó agregado em vez de no nó de origem, pois você deseja que o cache armazene os nós agregados, e não o conjunto de dados inteiro.

Nota: O armazenamento em cache nos nós de origem, que simplesmente armazena uma cópia dos dados originais como ele é lido em IBM SPSS Modelador, não irá melhorar o desempenho na maioria das circunstâncias.

Nós com o armazenamento em cache ativado são exibidos com um ícone de documento pequeno no canto superior direito. Quando os dados são armazenados em cache no nó, o ícone de documento fica verde.

Para ativar um cache

1. Na tela de fluxo, clique com o botão direito no nó e clique em **Cache** no menu.
2. No submenu de armazenamento em cache, clique em **Ativar**.
3. É possível desativar o cache clicando com o botão direito no nó e clicando em **Desativar** no submenu de armazenamento em cache.

Nós de armazenamento em cache em um banco de dados

Para fluxos executados em um banco de dados, dados podem ser armazenados em cache no midstream para uma tabela temporária no banco de dados em vez de para o sistema de arquivo. Quando combinado com otimização de SQL, isso pode resultar em ganhos significativos no desempenho. Por exemplo, a saída de um fluxo que mescla várias tabelas para criar uma visualização de mineração de dados pode ser armazenada em cache e reutilizada conforme necessário. Ao gerar SQL automaticamente para todos os nós de recebimento de dados, o desempenho pode ser melhorado.

Para aproveitar a vantagem do armazenamento em cache do banco de dados, a otimização de SQL e o armazenamento em cache do banco de dados devem estar ativados. Observe que as configurações de otimização do Servidor substituem as do Cliente. Consulte o tópico [“Configurando opções de otimização para fluxos”](#) na página 40 para obter informações adicionais.

Com o armazenamento em cache do banco de dados ativado, basta clicar com o botão direito em qualquer nó não terminal para armazenar dados em cache nesse ponto e, automaticamente, o cache será criado diretamente no banco de dados na próxima vez que o fluxo for executado. Se o armazenamento em cache do banco de dados ou a otimização de SQL não estiverem ativados, o cache será gravado no sistema de arquivos.

Nota: Os bancos de dados a seguir suportam tabelas temporárias para fins de armazenamento em cache: Db2, Oracle, SQL Server e Teradata. Outros bancos de dados, como o Netezza, usarão uma tabela normal para armazenamento em cache do banco de dados. O código SQL pode ser customizado para bancos de dados específicos - entre em contato com os Serviços para obter assistência.

Para limpar um cache

Um ícone de documento em branco em um nó indica que seu cache está vazio. Quando o cache estiver cheio, o ícone de documento ficará verde sólido. Se desejar substituir o conteúdo do cache, primeiro você deve limpar o cache e depois executar o fluxo de dados novamente para voltar a preenchê-lo.

1. Na tela de fluxo, clique com o botão direito no nó e clique em **Cache** no menu.
2. No submenu de armazenamento em cache, clique em **Limpeza**.

Para salvar um cache

Você pode salvar o conteúdo de um cache como um arquivo de dados IBM SPSS Estatísticas (*.sav). Você pode então recarregar o arquivo como um cache, ou pode configurar um nó que usa o arquivo de cache como sua fonte de dados. Também é possível carregar um cache que você salvou de outro projeto.

1. Na tela de fluxo, clique com o botão direito no nó e clique em **Cache** no menu.
2. No submenu de armazenamento em cache, clique em **Salvar Cache**.
3. Na caixa de diálogo Salvar Cache, navegue para a localização onde deseja salvar o arquivo de cache.
4. Insira um nome na caixa de texto Nome do Arquivo.
5. Certifique-se de que ***.sav** esteja selecionado na lista Arquivos do Tipo e clique em **Salvar**.

Para carregar um cache

Se você tiver salvado um arquivo de cache antes de removê-lo do nó, é possível recarregá-lo.

1. Na tela de fluxo, clique com o botão direito no nó e clique em **Cache** no menu.
2. No submenu de armazenamento em cache, clique em **Carregar Cache**.
3. Na caixa de diálogo Carregar Cache, navegue para a localização onde deseja salvar o arquivo de cache, selecione-o e clique em **Carregar**.

Visualizando dados em nós

Para assegurar que os dados estejam sendo alterados da maneira esperada conforme você constrói um fluxo, seria possível executar seus dados por meio de um nó de Tabela em cada passo significativo. Para salvá-lo de ter que fazer isso, é possível gerar uma pré-visualização a partir de cada nó que exibe uma amostra dos dados que serão criados, reduzindo assim o tempo que leva para construir cada nó.

Para o envio de dados de nós de um nugget do modelo, a visualização mostra os campos de entrada; para um nugget do modelo ou recebimento de dados de nós do nugget (exceto nós terminais), a visualização mostra campos de entrada e gerados.

O número padrão de linhas exibido é 10; no entanto, é possível alterar isso nas propriedades do fluxo. Consulte o tópico [“Configurando opções gerais para fluxos”](#) na página 38 para obter informações adicionais.

A partir do menu **Gerar**, é possível criar vários tipos de nós.

Nota: Ao visualizar dados gerados por este nó, todas as alterações de propriedade serão aplicadas a este nó e não poderão ser canceladas (o mesmo comportamento do clicar em **Aplicar**).

Bloqueando nós

Para evitar que outros usuários aditem as configurações de um ou mais nós em um fluxo, é possível conter o nó ou nós em um tipo especial de nó chamado SuperNode e depois bloquear o SuperNode aplicando proteção de senha.

Trabalhando com fluxos

Após ter conectado os nós de origem, processo e terminal na tela de fluxo, você criou um fluxo. Como uma coleção de nós, fluxos podem ser salvos, anotados e incluídos em projetos. Também é possível

configurar inúmeras opções para fluxos, como otimização, configurações de data e hora, parâmetros e scripts. Essas propriedades serão discutidas nos tópicos a seguir.

No IBM SPSS Modelador, é possível usar e modificar mais de um fluxo de dados na mesma sessão do IBM SPSS Modelador. O lado direito da janela principal contém o painel de gerentes, que ajuda você a navegar pelos fluxos, saídas e modelos que estão atualmente abertos. Se não for possível ver a área de janela de gerenciadores, clique em **Gerenciadores** no menu Visualizar e, em seguida, clique na guia **Fluxos**.

Nesta guia, você pode:

- Acessar fluxos.
- Salvar fluxos.
- Salvar fluxos no projeto atual.
- Fechar fluxos.
- Abrir novos fluxos.
- Armazenar e recuperar fluxos a partir de um repositório do IBM SPSS Serviços de colaboração e implantação (se disponível em seu site). Veja o tópico [“Sobre o IBM SPSS Repositório de serviços de colaboração e implantação”](#) na página 205 para obter mais informações.

Clique com o botão direito no fluxo na guia Fluxos para acessar essas opções.

Configurando opções para fluxos

É possível especificar inúmeras opções para aplicar ao fluxo atual. Também é possível salvar essas opções como padrões para aplicar a todos os fluxos. As opções são as seguintes.

- **Geral.** Opções variadas, como símbolos e codificação de texto, para usar no fluxo. Consulte o [“Configurando opções gerais para fluxos”](#) na página 38 para obter mais informações.
- **Data/Hora.** Opções relacionadas ao formato das expressões de data e hora. Consulte o [“Configurando opções de data e hora para fluxos”](#) na página 39 para obter mais informações.
- **Os formatos de número.** Opções controlando o formato de expressões numéricas. Consulte o [“Configurando opções de formato numérico para fluxos”](#) na página 40 para obter mais informações.
- **Otimização.** Opções para otimizar o desempenho do fluxo. Consulte o [“Configurando opções de otimização para fluxos”](#) na página 40 para obter mais informações.
- **Logging e Status.** Opções controlando criação de log SQL e status de registro. Consulte o [“Configurando opções de status de registro e criação de log SQL para fluxos”](#) na página 42 para obter mais informações.
- **Layout.** Opções relacionadas ao layout do fluxo na tela. Consulte o [“Configurando opções de layout para fluxos”](#) na página 42 para obter mais informações.
- **Servidor analítico.** Opções relacionadas ao uso de Servidor analítico com ModeladorSPSS. Consulte [“Propriedades de fluxo do Servidor analítico”](#) na página 43 para obter mais informações.
- **Geoespacial.** Opções relacionadas à formatação de dados geoespaciais para uso no fluxo. Consulte o [“Configurando opções geoespaciais para fluxos”](#) na página 43 para obter mais informações.

Para configurar opções de fluxo

1. No menu Arquivo, clique em **Propriedades do Fluxo** (ou selecione o fluxo na guia Fluxos na área de janela dos gerenciadores, clique com o botão direito e clique em **Propriedades do Fluxo** no menu pop-up).
2. Clique na guia **Opções**.

Alternativamente, no menu Ferramentas, clique em:

Propriedades do Fluxo > Opções

Configurando opções gerais para fluxos

As opções gerais são um conjunto de opções diversas que se aplicam a vários aspectos do fluxo atual.

A seção **Configurações Básicas** inclui as opções básicas a seguir:

- **Símbolo decimal.** Selecione uma vírgula (,) ou um período (.) como separador decimal.
- **Símbolo de Agrupamento.** Para formatos de exibição de número, selecione o símbolo usado para agrupar valores (por exemplo, a vírgula em 3.000,00). As opções incluem nenhum, ponto final, vírgula, espaço e definido por código de idioma (em cujo caso o padrão para o código de idioma atual é usado).
- **Codificação.** Especifique o método padrão de fluxo para codificação de texto. (Nota: Aplica-se a Var. Somente nó de origem Arquivo e nó de exportação Arquivo Simples. Nenhum outro nó usa essa configuração; a maioria dos arquivos de dados tem informações de codificação integradas.) É possível escolher o padrão do sistema ou UTF-8. O padrão do sistema é especificado no Painel de Controle do Windows ou, se estiver em execução no modo distribuído, no computador servidor. Consulte o tópico “[Suporte a Unicode em IBM SPSS Modelador](#)” na página 265 para obter mais informações.
- **Avaliação de Conjunto de Regras.** Determina como modelos de conjunto de regras são avaliados. Por padrão, conjuntos de regras usam **Voto** para combinar previsões de regras individuais e determinar a previsão final. Para assegurar que os conjuntos de regras usem a primeira regra de ocorrência por padrão, selecione **Primeira Ocorrência**. Observe que essa opção não se aplica aos modelos Lista de Decisão, que sempre usam a primeira ocorrência conforme definido pelo algoritmo.

Número máximo de linhas para mostrar na Visualização de Dados. Especifique o número de linhas a ser mostrado quando uma visualização de dados for solicitada para um nó. Veja o tópico “[Visualizando dados em nós](#)” na página 36 para obter mais informações.

Máximo de membros para campos nominais. Selecione para especificar um número máximo de membros para campos nominais (configurados) após o qual o tipo de dado do campo se torna **Sem Tipo**. Essa opção é útil quando se trabalha com grandes campos nominais. Nota: Quando o nível de medição de um campo é configurado para **Sem Tipo**, sua função é automaticamente configurada para **Nenhum**. Isso significa que os campos não estão disponíveis para modelagem.

Limitar tamanho da configuração para modelagem Kohonen e k-médias. Selecione para especificar um número máximo de membros para campos nominais usados em redes Kohonen e modelagem K-médias. O tamanho de configuração padrão é 20; depois disso, o campo é ignorado e um aviso é levantado, fornecendo informações sobre o campo em questão.

Observe que, para compatibilidade, essa opção também se aplica ao antigo nó Rede Neural que foi substituído na versão 14 do IBM SPSS Modelador; alguns fluxos anteriores ainda podem conter esse nó.

Atualizar nós de origem na execução. Selecione para atualizar automaticamente todos os nós de origem durante a execução do fluxo atual. Essa ação é análoga a um clique no botão **Atualizar** em um nó de origem, exceto que atualiza automaticamente todos os nós de origem (exceto nós Entrada do Usuário) para o fluxo atual.

Nota: A seleção dessa opção limpa os caches dos nós de recebimento de dados, mesmo se os dados não tiverem mudado. Se você usar a opção **Executar o fluxo atual** a partir da barra de ferramentas, o flushing ocorre apenas uma vez por execução do fluxo, no entanto, o que significa que você ainda pode usar caches de downstream como armazenamento temporário para uma única execução. Por exemplo, digamos que você tenha configurado um midstream de cache após uma operação de derivação complexa e tenha um recebimento de dados com vários gráficos e relatórios anexados desse nó Derivar. Durante a execução do fluxo, o cache no nó Derivar será limpo e preenchido novamente, mas somente para o primeiro gráfico ou relatório. Os nós terminais subsequentes lerão dados do cache do nó Derivar. Note que se você optar por executar cada nó do terminal individualmente (quando tiver mais de um nó do terminal), em vez de usar a opção **Executar o fluxo atual**, o cache flushing ocorre sempre que você executar um nó terminal.

Rótulos de valor e campo de exibição na saída. Rótulos de valor e campo de exibição em tabelas, gráficos e outra saída. Se os rótulos não existirem, os nomes de campo e os valores de dados serão exibidos. Os rótulos ficam desativados por padrão; no entanto, é possível alternar os rótulos

individualmente em qualquer lugar no IBM SPSS Modelador. Também é possível escolher a exibição de rótulos na janela de saída usando um botão de alternância disponível na barra de ferramentas.



Figura 11. Ícone da barra de ferramentas usado para alternar rótulos de valor e campo

Exibir tempos de execução. Exibe tempos de execução individuais para nós de fluxo na guia Tempos de Execução após o fluxo ser executado. Veja o tópico [“Visualizando tempos de execução de nós”](#) na página 44 para obter mais informações.

A seção **Criação de Nós Automática** inclui as seguintes opções para criar nós automaticamente em fluxos individuais. Essas opções controlam se deve-se ou não inserir os nuggets de modelagem na tela de fluxo durante a geração de novos nuggets. Por padrão, essas opções se aplicam somente aos fluxos criados na versão 16 ou posterior. No IBM SPSS Modelador 16 ou posterior, se você abrir um fluxo criado na versão 15 ou anterior e executar um nó de modelagem, o nugget não será colocado sobre a tela do fluxo como costumava ser em releases anteriores. Se você criar um novo fluxo usando IBM SPSS Modelador 16 ou posterior e executar um nó de modelagem, o nugget gerado será colocado sobre a tela do fluxo. Isso é tão projetado porque, por exemplo, a opção **Criar nós de modelo de aplicação para nova saída de modelo** provavelmente quebraria pre-16 fluxos que são executados em batch, em IBM SPSS Serviços de colaboração e implantação, e em outros ambientes em que a interface do usuário do cliente IBM SPSS Servidor Modelador não está presente.

- **Criar nós de aplicação de modelo para nova saída do modelo.** Cria automaticamente nós de aplicação de modelo para a nova saída do modelo. Se você selecionar essa opção, também é possível escolher em **Criar ligações de atualização de modelo** se você irá configurar as ligações como ativadas, desativadas ou não irá criá-las.

Quando um aplicador de novo modelo ou nó de origem é criado, as opções de ligação nos menus suspensos controlam se as ligações de atualização entre o nó do construtor e o novo nó são criadas e, se sim, em que modo elas estão. Se as ligações forem criadas, é provável que você as queira ativadas, mas essas opções fornecem controle completo ao usuário.

- **Criar nós de origem a partir de construtores de origem.** Cria automaticamente nós de origem a partir de construtores de origem. Semelhante à opção anterior, se você selecionar essa opção, também é possível escolher no menu suspenso **Criar ligações de atualização de origem** se você irá configurar ligações de atualização como ativadas, desativadas ou não irá criá-las.

Salvar Como Padrão. As opções especificadas se aplicam somente ao fluxo atual. Clique nesse botão para configurar essas opções como padrão para todos os fluxos.

Configurando opções de data e hora para fluxos

Essas opções especificam o formato a ser usado para várias expressões de data e hora no fluxo atual.

Importar data/hora como Selecione entre usar armazenamento de data/hora para campos de data/hora ou importá-los como variáveis de sequência.

Formato de data Selecione um formato de data para ser usado para campos de armazenamento de data ou quando sequências forem interpretadas como datas pelas funções de data do CLEM.

Formato de hora Selecione um formato de hora para ser usado para campos de armazenamento de hora ou quando sequências forem interpretadas como hora pelas funções de hora do CLEM.

Dias/minutos de rolagem Para formatos de hora, selecione se diferenças de tempo negativo são interpretadas como referência a um dia ou hora anterior.

Linha de base de data (1º Jan) Selecione os anos de linha de base (sempre 1º de janeiro) para serem usados pelas funções de data do CLEM que trabalham com data única.

Início de datas de 2 dígitos Especifica o ano de corte para incluir dígitos de século para anos que estão denotados com apenas 2 dígitos. Por exemplo, especificando 1930 como ano de corte significa que 11/05/02 é no ano 2002. A mesma configuração usará o século 20 para datas após o ano 30; assim, 05/11/73 será assumido como 1973.

Fuso horário Seleciona como o fuso horário é escolhido para ser usado com a expressão `datetime_now` de CLEM.

- Se **Servidor** for selecionado, o fuso horário dependerá dos seguintes itens:
 - Se o fluxo atual usa uma origem de dados de Servidor analítico, a expressão `datetime_now` usará a hora do Servidor analítico; por padrão, o servidor usa a Hora Universal Coordenada.
 - Se o fluxo atual usa um nó de origem de Banco de Dados, os bancos de dados suportados usarão o SQL pushback e a expressão `datetime_now` usará a hora do bando de dados.
 - Para todos os outros fluxos, o fuso horário usa a hora do Servidor do SPSS Modeler.
- Se **Cliente do Modelador** for selecionado, o fuso horário refletirá os detalhes de fuso horário da máquina em que o Modelador SPSS está instalado.
- Alternativamente, é possível selecionar qualquer valor de Hora Universal Coordenada para o fuso horário.

Salvar como Padrão. As opções especificadas são aplicadas apenas ao fluxo atual. Clique nesse botão para configurar essas opções como padrão para todos os fluxos.

Configurando opções de formato numérico para fluxos

Estas opções especificam o formato a ser usado para várias expressões numéricas no fluxo atual.

Formato de exibição numérico. É possível escolher entre os formatos padrão (####.###), científicos (#.###E+##) ou de moeda (\$###.##).

Casas decimais (padrão, científico, moeda). Para formatos de exibição numéricos, especifica o número de casas decimais a serem usadas durante a exibição ou impressão de números reais. Essa opção é especificada separadamente para cada formato de exibição.

Cálculos em. Selecione **Radianos** ou **Graus** como unidade de medida a ser usada em expressões trigonométricas do CLEM. Veja o tópico “[Funções trigonométricas](#)” na página 177 para obter mais informações.

Salvar Como Padrão. As opções especificadas se aplicam somente ao fluxo atual. Clique nesse botão para configurar essas opções como padrão para todos os fluxos.

Configurando opções de otimização para fluxos

É possível usar configurações de Otimização para otimizar o desempenho do fluxo. Observe que as configurações de desempenho e otimização no Servidor IBM SPSS Modeler (se usado) substituem quaisquer configurações equivalentes no cliente. Se essas configurações estiverem desativadas no servidor, o cliente não poderá ativá-las. Mas se elas estiverem ativadas no servidor, o cliente poderá escolher desativá-las.

Nota: A modelagem da base de dados e a otimização de SQL requerem que a conectividade do Servidor IBM SPSS Modeler esteja ativada no computador do IBM SPSS Modelador. Com essa configuração ativada, é possível acessar os algoritmos de banco de dados, realizar SQL pushback diretamente do IBM SPSS Modelador e acessar o Servidor IBM SPSS Modeler. Para verificar o atual status da licença, escolha o seguinte no menu do IBM SPSS Modelador.

Ajuda > Sobre > Detalhes Adicionais

Se a conectividade estiver ativada, você verá a opção **Ativação do Servidor** na guia Status da Licença.

Veja “[Conectando-se ao Servidor IBM SPSS Modeler](#)” na página 8 para obter mais informações.

Nota: Se a pushback SQL e a otimização são suportadas depende do tipo de banco de dados em uso. Para obter as informações mais recentes sobre quais bancos de dados e drivers do ODBC são suportados e testados para uso com IBM SPSS Modelador, consulte o site de Suporte corporativo em <http://www.ibm.com/support>.

Ativar regravação de fluxo. Selecione esta opção para ativar a regravação de fluxo no IBM SPSS Modelador. Há quatro tipos de regravação disponíveis, e é possível selecionar um ou mais deles. A

regravação de fluxo reordena os nós em um fluxo por trás dos cenários para uma operação mais eficiente, sem alterar a semântica do fluxo.

- **Otimizar geração de SQL.** Esta opção permite que os nós sejam reordenados dentro do fluxo, de modo que mais operações possam retroceder usando geração de SQL para execução no banco de dados. Ao localizar um nó que não pode ser renderizado em SQL, o otimizador aguardará para ver se há algum nó de recebimento de dados que pode ser renderizado em SQL e movido com segurança para a frente do nó problemático sem afetar a semântica do fluxo. Não só o banco de dados pode executar operações com mais eficiência que o IBM SPSS Modelador, mas esses retroprocessos agem para reduzir o tamanho do conjunto de dados que é retornado para o IBM SPSS Modelador para processamento. Isso, por sua vez, pode reduzir o tráfego de rede e acelerar as operações do fluxo. Observe que a caixa de seleção **Gerar SQL** deve ser selecionada para que a otimização de SQL tenha algum efeito.
- **Otimizar expressão do CLEM.** Esta opção permite que o otimizador procure expressões do CLEM que possam ser pré-processadas antes de o fluxo ser executado para aumentar a velocidade do processamento. Como um exemplo simples, se você tivesse uma expressão como *log(salário)*, o otimizador calcularia o valor do salário real e o passaria para processamento. Isso pode ser usado para melhorar o SQL pushback e o desempenho do Servidor IBM SPSS Modeler.
- **Otimizar a execução de sintaxe.** Esse método de regravação de fluxo aumenta a eficiência das operações que incorporam mais de um nó contendo sintaxe do IBM SPSS Estatísticas. A otimização é obtida combinando os comandos de sintaxe em uma única operação, em vez de executar cada um como uma operação separada.
- **Otimizar outra execução.** Este método de regravação de fluxo aumenta a eficiência das operações que não podem ser delegadas ao banco de dados. A otimização é obtida reduzindo a quantidade de dados no fluxo o mais cedo possível. Enquanto mantém a integridade dos dados, o fluxo é regravado para enviar as operações por push para mais perto da origem de dados, reduzindo assim o recebimento de dados para as operações mais caras, como junções.

Ativar o processamento paralelo. Durante a execução em um computador com vários processadores, esta opção permite que o sistema equilibre a carga nesses processadores, o que pode resultar em um desempenho mais rápido. O uso de diversos nós ou o uso dos seguintes nós individuais pode ser beneficiado pelo processamento paralelo: C5.0, Mesclagem (por chave), Ordem, Categorização (métodos de ranqueamento e ladrilho) e Agregado (usando um ou mais campos-chave).

Gerar SQL. Selecione esta opção para ativar geração de SQL, permitindo que operações de fluxo retrocedam para o banco de dados usando código SQL para gerar processos de execução, o que pode melhorar o desempenho. Para melhorar ainda mais o desempenho, **Otimizar Geração de SQL** também pode ser selecionada para maximizar o número de operações que retrocedem para o banco de dados. Quando operações para um nó retrocederem para o banco de dados, o nó será destacado em púrpura quando o fluxo for executado.

- **Armazenamento em cache de banco de dados.** Para que fluxos que geram SQL sejam executados no banco de dados, dados podem ser armazenados em cache no midstream para uma tabela temporária no banco de dados em vez de para o sistema de arquivo. Quando combinado com otimização de SQL, isso pode resultar em ganhos significativos no desempenho. Por exemplo, a saída de um fluxo que mescla várias tabelas para criar uma visualização de mineração de dados pode ser armazenada em cache e reutilizada conforme necessário. Com o armazenamento em cache do banco de dados ativado, basta clicar com o botão direito em qualquer nó não terminal para armazenar dados em cache nesse ponto e, automaticamente, o cache será criado diretamente no banco de dados na próxima vez que o fluxo for executado. Isso permite que o SQL seja gerado para nós de recebimento de dados, melhorando ainda mais o desempenho. Alternativamente, essa opção pode ser desativada se for necessário, como quando políticas ou permissões impedem que dados sejam gravados no banco de dados. Se o armazenamento em cache do banco de dados ou a otimização de SQL não estiverem ativados, o cache será gravado no sistema de arquivos. Consulte o tópico [“Armazenando opções em cache para nós”](#) na página 35 para obter mais informações.
- **Usar conversão afrouxada.** Esta opção permite a conversão de dados de sequências de caracteres em números, ou de números em sequências de caracteres, se eles estiverem armazenados em um formato adequado. Por exemplo, se os dados forem mantidos no banco de dados como uma sequência

de caracteres, mas de fato contiverem um número significativo, os dados poderão ser convertidos para uso quando ocorrer um retrocesso.

Nota: Devido a pequenas diferenças na implementação SQL, os fluxos executados em um banco de dados podem retornar resultados um pouco diferentes daqueles retornados quando executados em IBM SPSS Modelador. Por motivos semelhantes, essas diferenças também podem variar, dependendo do fornecedor do banco de dados.

Salvar Como Padrão. As opções especificadas se aplicam somente ao fluxo atual. Clique nesse botão para configurar essas opções como padrão para todos os fluxos.

Configurando opções de status de registro e criação de log SQL para fluxos

Essas configurações incluem várias opções controlando a exibição de instruções SQL geradas pelo fluxo, além da exibição do número de registros processados pelo fluxo.

Exibir SQL no log de mensagens durante a execução de fluxo. Especifica se a SQL gerada durante a execução do fluxo é passada para o log de mensagens.

Exibir detalhes de geração de SQL no log de mensagens durante a preparação do fluxo. Durante a visualização do fluxo, especifica se uma visualização da SQL que seria gerada é passada para o log de mensagens.

Exibir SQL. Especifica se alguma SQL que é exibida no log deve conter funções de SQL nativo ou funções de ODBC padrão na forma `{fn FUNC(...)}`, conforme gerado por ModeladorSPSS. O último conta com a funcionalidade do driver ODBC que pode não estar implementada.

Reformatar SQL para melhorar a capacidade de leitura. Especifica se a SQL exibida no log deve ser formatada para capacidade de leitura.

Mostrar status para registros. Especifica quando os registros devem ser relatados conforme eles chegam aos nós terminais. Especifique um número que seja usado para atualizar o status a cada N registros.

Salvar Como Padrão. As opções especificadas se aplicam somente ao fluxo atual. Clique nesse botão para configurar essas opções como padrão para todos os fluxos.

Configurando opções de layout para fluxos

Estas configurações fornecem inúmeras opções relacionadas à exibição e ao uso da tela de fluxo.

Largura mínima da tela de fluxo. Especifique a largura mínima da tela de fluxo em pixels.

Altura mínima da tela de fluxo. Especifique a altura mínima da tela de fluxo em pixels.

Taxa de rolagem do fluxo. Especifique a taxa de rolagem para a tela de fluxo para controlar a rapidez com que a área de janela da tela de fluxo rola quando um nó está sendo arrastado de um local para outro na tela. Os números mais altos especificam uma taxa de rolagem mais rápida.

Nome máximo do ícone. Especifique um limite nos caracteres para os nomes dos nós na tela de fluxo.

Tamanho do ícone. Selecione uma opção para escalar a visualização de fluxo inteira para um de vários tamanhos entre 8% e 200% do tamanho do ícone padrão.

Tamanho da célula de grade. Selecione um tamanho de célula de grade da lista. Esse número é usado para alinhar nós na tela de fluxo usando uma grade invisível. O tamanho da célula de grade padrão é 0,25.

Ajustar à grade. Selecione para alinhar ícones para um padrão de grade invisível (selecionado por padrão).

Localização do ícone gerado. Escolha onde na tela colocar ícones para nós gerados a partir dos nuggets de modelo. O padrão é superior esquerdo.

Salvar Como Padrão. As opções especificadas se aplicam somente ao fluxo atual. Clique nesse botão para configurar essas opções como padrão para todos os fluxos.

Propriedades de fluxo do Servidor analítico

Estas configurações fornecem diversas opções para trabalhar com o Servidor analítico.

Número máximo de registros a processar fora do Servidor analítico

Especifique o número máximo de registros a ser importado no servidor ModeladorSPSS a partir de uma origem de dados do Servidor analítico.

Notificação quando um nó não pode ser processado no Servidor analítico

Esta configuração determina o que acontece quando um fluxo que seria enviado para o Servidor analítico contém um nó que não pode ser processado no Servidor analítico. Especifique se um aviso será emitido e continue processando o fluxo, ou lance um erro e pare o processamento.

Configurações de Armazenamento de Modelo de Divisão

Armazene os modelos de divisão por referência no Servidor analítico quando o tamanho do modelo (MB) for excedido

Nuggets do modelo geralmente são armazenados como parte do fluxo. Modelos de divisão com muitas divisões podem produzir nuggets grandes, e mover o nugget para frente e para trás entre o fluxo e o Servidor analítico pode causar impacto no desempenho. Como uma solução, quando um modelo de divisão exceder o tamanho especificado, ele será armazenado no Servidor analítico, e o nugget no ModeladorSPSS conterá uma referência ao modelo.

Pasta padrão para armazenar modelos por referência no Servidor analítico quando a execução for concluída

Especifique o caminho padrão no qual deseja armazenar modelos de divisão no Servidor analítico. O caminho deve começar com um nome do projeto válido do Servidor analítico.

Pasta para armazenar modelos promovidos

Especifique o caminho padrão no qual deseja armazenar modelos "promovidos". Um modelo promovido não é limpo quando a sessão do ModeladorSPSS termina.

Configurando opções geoespaciais para fluxos

Qualquer campo geoespacial, independentemente de ser uma forma, uma coordenada ou um valor de eixo único (como x ou y, ou latitude e longitude) tem um sistema de coordenadas associado. Esse sistema de coordenadas configura atributos como o ponto de origem (0,0) e as unidades associadas aos valores.

Há inúmeros sistemas de coordenadas e dois tipos: Geográfico e Projetado. Todas as funções espaciais no ModeladorSPSS podem ser usadas somente com um sistema de coordenadas projetadas.

Devido à natureza dos sistemas de coordenadas, a mesclagem ou anexação de dados de duas origens de dados geoespaciais separadas requer que as origens usem o mesmo sistema de coordenadas. Por isso, você deve especificar uma configuração de coordenadas para quaisquer dados geoespaciais usados no fluxo.

Os dados são reprojetados automaticamente para usar o sistema de coordenadas de fluxo escolhido nas seguintes situações:

- Para funções espaciais (como area, closeto, within), o parâmetro passado para a função é automaticamente reprojetado; no entanto, os dados de linha originais são mantidos inalterados.
- Durante o uso dos nós de construção ou escoragem (nugget) no Spatio-Temporal Prediction (STP), o campo de localização é automaticamente reprojetado. Durante a escoragem, a localização que é revelada do nugget é a localização original.
- Durante o uso do nó Visualização de Mapa.

Sistema de coordenadas do fluxo. Disponível somente se você marcar a caixa de seleção. Clique em **Mudança** para exibir uma lista de Sistemas de Coordenadas Projetadas disponíveis e selecionar o desejado para usar para o fluxo atual.

Salvar Como Padrão. O sistema de coordenadas selecionado se aplica somente ao fluxo atual. Para selecionar o sistema como padrão para todos os fluxos, clique nesse botão.

Selecionando sistemas de coordenadas geoespaciais

Todas as funções espaciais no ModeladorSPSS podem ser usadas somente com um sistema de coordenadas projetadas.

A caixa de diálogo **Selecionar Sistema de Coordenadas de Fluxo** contém uma lista de todos os sistemas de coordenadas projetadas que podem ser selecionados para quaisquer dados geoespaciais usados em um fluxo.

As informações a seguir são listadas para cada sistema de coordenadas.

- **WKID** O Well Known ID que é exclusivo para cada sistema de coordenadas.
- **Nome** O nome do sistema de coordenadas.
- **Unidades** A unidade de medição associada ao sistema de coordenadas.

Além da lista de todos os sistemas de coordenadas, a caixa de diálogo tem um controle de **Filtragem**. Se você souber todo ou parte do nome do sistema de coordenadas requerido, digite-o no campo **Nome** na parte inferior da caixa de diálogo. A lista de sistemas de coordenadas da qual é possível escolher é filtrada automaticamente para mostrar somente os sistemas com nomes que contêm o texto que você inseriu.

Visualizando mensagens de operação de fluxo

Mensagens referentes a operações de fluxo, como execução, otimização e tempo decorrido para construção e avaliação de modelo, podem ser facilmente visualizadas usando a guia **Mensagens** na caixa de diálogo de propriedades do fluxo. Mensagens de erro também são relatadas nessa tabela.

Para visualizar mensagens de fluxo

1. No menu Arquivo, clique em **Propriedades do Fluxo** (ou selecione o fluxo na guia Fluxos na área de janela dos gerenciadores, clique com o botão direito e clique em **Propriedades do Fluxo** no menu pop-up).
2. Clique na guia **Mensagens**.

Alternativamente, no menu Ferramentas, clique em:

Propriedades do Fluxo > Mensagens

Além das mensagens referentes às operações de fluxo, mensagens de erro são relatadas aqui. Quando o fluxo em execução é finalizado devido a um erro, essa caixa de diálogo é aberta para a guia **Mensagens** com a mensagem de erro visível. Além disso, o nó com erros é destacado em vermelho na tela de fluxo.

Se a otimização de SQL e as opções de criação de log estiverem ativadas na caixa de diálogo **Opções do Usuário**, as informações no SQL gerado também serão exibidas. Consulte o tópico [“Configurando opções de otimização para fluxos”](#) na página 40 para obter informações adicionais.

É possível salvar mensagens relatadas aqui para um fluxo clicando em **Salvar Mensagens** na lista suspensa do botão **Salvar** (à esquerda, abaixo da guia **Mensagens**).

Você pode limpar todas as mensagens para um determinado fluxo selecionando **Limpar Todas as Mensagens** a partir do drop-down.

Observe que tempo de CPU é a quantidade de tempo que o processo do servidor está utilizando CPU. Tempo decorrido é o tempo total entre o início da execução e o fim da execução, portanto, ele também inclui coisas como transferência de arquivos e renderização de saídas. O tempo de CPU pode ser mais que o tempo decorrido quando um fluxo está alavancando várias CPUs (execução paralela). Quando um fluxo for totalmente enviado de volta para o banco de dados sendo usado como uma origem de dados, o tempo de CPU será zero.

Visualizando tempos de execução de nós

Na aba **Mensagens** você pode optar por exibir o **Execution Times**, em que é possível ver os tempos de execução individuais para nós no fluxo que executam no ServidorIBM SPSS Modeler. Observe que os

itens podem não ser exatos para fluxos executados em outras áreas, como R ou Servidor analítico. E o tempo de execução de alguns nós não pode ser calculado.

Nota: Para que este recurso funcione, a opção **Exibir tempos de execução** deve ser selecionada na configuração **Geral** da guia **Opções**.

Na tabela de tempos de execução de nó, as colunas são as seguintes. Clique em um título da coluna para ordenar as entradas em ordem crescente e decrescente (por exemplo, para ver quais nós têm os tempos de execução mais longos).

Nó Terminal. O identificador da ramificação à qual o nó pertence. O identificador é o nome do nó terminal no final da ramificação.

Rótulo do Nó. O nome do nó ao qual o tempo de execução se refere.

Id do Nó. O identificador exclusivo do nó ao qual o tempo de execução se refere. Esse identificador é gerado pelo sistema quando o nó é criado.

Tempo(s) de Execução. O tempo em segundos levado para a execução desse nó. Note que o tempo de execução muitas vezes pode variar a partir do momento em que você vê em mensagens gerais porque o tempo é gasto na preparação de dados e na recuperação de dados na saída ao executar um fluxo, e este tipo de tempo não pode ser calculado.

Configurando parâmetros de sessão e fluxo

Parâmetros podem ser definidos para uso em expressões do CLEM e em scripts. Na verdade, eles são variáveis definidas pelo usuário que são salvas e persistem no fluxo, sessão ou SuperNode atuais e que podem ser acessadas a partir da interface com o usuário, bem como por meio de scripts. Se você salvar um fluxo, por exemplo, quaisquer parâmetros configurados para esse fluxo também serão salvos. (Isso os distingue das variáveis do script local, que podem ser usadas somente no script no qual elas são declaradas.) Geralmente os parâmetros são usados em scripts para controlar o comportamento do script, fornecendo informações sobre campos e valores que não precisam ser codificados permanentemente no script.

O escopo do parâmetro depende de onde ele está configurado:

- Parâmetros de fluxo podem ser configurados em um script de fluxo ou na caixa de diálogo de propriedades do fluxo e estão disponíveis para todos os nós no fluxo. Eles são exibidos na lista Parâmetros no construtor de expressões.
- Parâmetros de sessão podem ser configurados em um script independente ou na caixa de diálogo de parâmetros de sessão. Eles estão disponíveis para todos os fluxos usados na sessão atual (todos os fluxos listados na guia Fluxos na área de janela de gerenciadores).

Parâmetros também podem ser configurados para SuperNodes, em cujo caso ficam visíveis somente para os nós contidos nesse SuperNode.

Para Configurar Parâmetros de Sessão e Fluxo por Meio da Interface com o Usuário

1. Para configurar parâmetros de fluxo, no menu principal, clique em:

Ferramentas > Propriedades do Fluxo > Parâmetros

2. Para configurar parâmetros de sessão, clique em **Configurar Parâmetros da Sessão** no menu Ferramentas.

Prompt? Marque esta caixa se desejar que o usuário receba um prompt no tempo de execução para inserir um valor para esse parâmetro.

Nome. Nomes de parâmetros são listados aqui. É possível criar um novo parâmetro inserindo um nome nesse campo. Por exemplo, para criar um parâmetro para temperatura mínima, você poderia digitar `minvalue`. Não inclua o prefixo `$P-` que denota um parâmetro nas expressões do CLEM. Esse nome também é usado para exibição no construtor de expressões do CLEM.

Nome longo. Lista o nome descritivo para cada parâmetro criado.

Armazenamento. Selecione um tipo de armazenamento da lista. Armazenamento indica como valores de dados são armazenados no parâmetro. Por exemplo, quando você trabalha com valores contendo zeros à esquerda que deseja preservar (como 008), é necessário selecionar **Sequência de Caracteres** como o tipo de armazenamento. Caso contrário, os zeros serão removidos do valor. Os tipos de armazenamento disponíveis são sequência de caracteres, número inteiro, real, hora, data e registro de data e hora. Para parâmetros de data, observe que os valores devem ser especificados usando uma notação padrão ISO, conforme mostrado no parágrafo seguinte.

Valor. Lista o valor atual de cada parâmetro. Ajuste o parâmetro conforme necessário. Observe que, para parâmetros de data, os valores devem ser especificados na notação padrão ISO (ou seja, YYYY-MM-DD). Datas especificadas em outros formatos não são aceitas.

Tipo (opcional). Se você pretende implementar o fluxo em um aplicativo externo, selecione um nível de medição da lista. Caso contrário, é recomendável deixar a coluna *Tipo* no estado em que se encontra. Se desejar especificar restrições de valor para o parâmetro, como limites mínimo e máximo para um intervalo numérico, selecione **Especificar** na lista.

Observe que as opções nome longo, armazenamento e tipo podem ser configuradas para parâmetros somente por meio da interface com o usuário. Essas opções não podem ser configuradas usando scripts.

Clique nas setas à direita para mover o parâmetro selecionado mais para cima ou para baixo na lista de parâmetros disponíveis. Use o botão Excluir (marcado com um X) para remover o parâmetro selecionado.

Especificando prompts de tempo de execução para valores de parâmetro

Se você tiver fluxos nos quais talvez seja necessário inserir valores diferentes para o mesmo parâmetro em ocasiões diferentes, é possível especificar prompts de tempo de execução para um ou mais valores de parâmetro de fluxo ou sessão.

Parâmetros. (Opcional) Insira um valor para o parâmetro ou deixe o valor padrão, se houver algum.

Desative esses prompts. Selecione essa caixa se não quiser que esses prompts sejam exibidos quando você executar o fluxo. Você pode fazer com que eles sejam reexibidos selecionando o **Prompt?** caixa de seleção na caixa de diálogo de propriedades de fluxo ou propriedades de sessão onde os parâmetros foram definidos. Consulte o tópico ["Configurando parâmetros de sessão e fluxo"](#) na página 45 para obter mais informações.

Especificando restrições de valor para um tipo de parâmetro

É possível criar restrições de valor para um parâmetro disponível durante a implementação do fluxo em um aplicativo externo que lê fluxos de modelagem de dados. Essa caixa de diálogo permite especificar valores disponíveis para um usuário externo executando o fluxo. Dependendo do tipo de dados, as restrições de valor variam dinamicamente na caixa de diálogo. As opções mostradas aqui são idênticas às opções disponíveis para valores do nó **Tipo**.

Tipo. Exibe o nível de medição atualmente selecionado. É possível alterar esse valor para refletir a maneira como você pretende usar o parâmetro no IBM SPSS Modelador.

Armazenamento. Exibe o tipo de armazenamento se ele for conhecido. Os tipos de armazenamento não são afetados pelo nível de medição (contínuo, nominal ou flag) que você escolher para trabalhar no IBM SPSS Modelador. É possível alterar o tipo de armazenamento na guia Parâmetros principal.

A metade inferior da caixa de diálogo muda dinamicamente, dependendo do nível de medição selecionado no campo **Tipo**.

Nível de Medição Contínuos

Inferior. Especifique um limite inferior para os valores de parâmetro.

Superior. Especifique um limite superior para os valores de parâmetro.

Rótulos. É possível especificar rótulos para qualquer valor de um campo de intervalo. Clique no botão **Rótulos** para abrir uma caixa de diálogo separada para especificar rótulos de valores.

Níveis de Medição Nominais

Valores. Esta opção permite especificar valores para um parâmetro que serão usados como um campo nominal. Valores não serão forçados no fluxo do IBM SPSS Modelador, mas serão usados em uma lista suspensa para aplicativos de implementação externos. Usando os botões de seta e exclusão, é possível modificar valores existentes, bem como reordenar ou excluir valores.

Nível de Medição Flag

True. Especifique um valor de flag para o parâmetro quando a condição for atendida.

False. Especifique um valor de flag para o parâmetro quando a condição não for atendida.

Rótulos. É possível especificar rótulos para os valores de um campo de flag.

Opções de Implementação do Fluxo

A guia Implantação da caixa de diálogo de propriedades do fluxo possibilita especificar opções para implantação do fluxo dentro de IBM SPSS Serviços de colaboração e implantação para fins de atualização de modelo ou agendamento de tarefas automatizado. Todos os fluxos requerem uma ramificação de pontuação designada antes que eles possam ser implementados. Consulte o “[Armazenando e implementando objetos do repositório](#)” na página 206 para obter mais informações.

Loop da execução de fluxos

Usando a guia Execução na caixa de diálogo de propriedades do fluxo, é possível configurar condições de loop para automatizar tarefas repetitivas no fluxo atual.

Após a configuração dessas condições, é possível usá-las como uma introdução aos scripts, já que elas preenchem a janela de scripts com scripts básicos para seu fluxo que, depois, pode ser modificado - talvez para você usar como base para a construção de scripts melhores. Consulte o tópico “[Funções globais](#)” na página 200 para obter mais informações.

Para configurar um loop para um fluxo

1. No menu Arquivo, clique em **Propriedades do Fluxo** (ou selecione o fluxo na guia Fluxos na área de janela dos gerenciadores, clique com o botão direito e clique em **Propriedades do Fluxo** no menu pop-up).
2. Clique na guia **Execução**.
3. Selecione o modo de execução **Loop / Execução Condicional**.
4. Clique na guia **Loop**.

Alternativamente, no menu Ferramentas, clique em:

Propriedades do Fluxo > Execução

Como uma alternativa adicional, clique com o botão direito no nó e, no menu de contexto, clique em:

Looping / Execução Condicional > Editar Configurações de Looping

Iteração. Não é possível editar esse valor de número da linha, mas é possível incluir, excluir ou mover uma iteração para cima ou para baixo usando os botões à direita da tabela.

Cabeçalhos da tabela. Refletem a chave de iteração e quaisquer variáveis de iteração criadas durante a configuração do loop.

Visualizar valores globais para fluxos

Usando a guia Globais na caixa de diálogo de propriedades do fluxo, é possível visualizar os valores globais configurados para o fluxo atual. Valores globais são criados usando um nó Configurar Globais para determinar estatísticas como média, soma ou desvio padrão para os campos selecionados.

Após o nó Configurar Globais ser executado, esses valores serão disponibilizados para uma variedade de usos em operações de fluxo. Consulte o tópico “[Funções globais](#)” na página 200 para obter informações adicionais.

Para visualizar valores globais para um fluxo

1. No menu Arquivo, clique em **Propriedades do Fluxo** (ou selecione o fluxo na guia Fluxos na área de janela dos gerenciadores, clique com o botão direito e clique em **Propriedades do Fluxo** no menu pop-up).
2. Clique na guia **Globais**.

Alternativamente, no menu Ferramentas, clique em:

Propriedades do Fluxo > Globais

Globais disponíveis. Globais disponíveis são listados nesta tabela. Não é possível editar valores globais aqui, mas é possível limpar todos os valores globais para um fluxo usando o botão Limpar Todos os Valores à direita da tabela.

Procurando nós em um fluxo

É possível procurar por nós em um fluxo especificando inúmeros critérios de procura, como nome do nó, categoria e identificador. Essa variável pode ser útil principalmente para fluxos complexos contendo um grande número de nós.

Para procurar nós em um fluxo

1. No menu Arquivo, clique em **Propriedades do Fluxo** (ou selecione o fluxo na guia Fluxos na área de janela dos gerenciadores, clique com o botão direito e clique em **Propriedades do Fluxo** no menu pop-up).
2. Clique na guia **Procurar**.

Alternativamente, no menu Ferramentas, clique em:

Propriedades do Fluxo > Pesquisar

É possível especificar mais de uma opção para limitar a procura, exceto que a procura por ID do nó (usando o campo **ID é igual a**) exclui as outras opções.

Rótulo do nó contém. Marque esta caixa e insira todo ou parte do rótulo de um nó para procurar um nó específico. As procura não fazem distinção entre maiúsculas e minúsculas, e várias palavras são tratadas como uma única parte de texto.

Categoria do nó. Marque esta caixa e clique em uma categoria na lista para procurar por um tipo de nó específico. **Nó de Processo** significa um nó da guia Operações de Registro ou Operações de Campo da paleta de nós; **Aplicar Nó de Modelo** se refere a um nugget de modelo.

Palavras-chave incluem. Marque esta caixa e insira uma ou mais palavras-chave completas para procurar campos com esse texto no campo Palavras-chave na guia Anotações da caixa de diálogo do nó. O texto da palavra-chave inserido deve ser uma correspondência exata. Separe palavras-chave múltiplas com pontos e vírgulas para procurar alternativas (por exemplo, a inserção de proton;neutron localizará todos os nós com qualquer uma dessas palavras-chave. Consulte o tópico [“Anotações” na página 56](#) para obter mais informações.

Anotação contém. Marque esta caixa e insira uma ou mais palavras para procurar por nós que contenham esse texto na área de texto principal na guia Anotações da caixa de diálogo do nó. As procura não fazem distinção entre maiúsculas e minúsculas, e várias palavras são tratadas como uma única parte de texto. Veja o tópico [“Anotações” na página 56](#) para obter mais informações.

Gera campo chamado. Marque esta caixa e insira o nome de um campo gerado (por exemplo, \$C-Drug). É possível usar essa opção para procurar nós de modelagem que gerem um campo específico. Insira somente um nome de campo, que deve ser uma correspondência exata.

ID é igual a. Marque esta caixa e insira um ID de nó para procurar um nó específico com esse identificador (a seleção dessa opção desativa todas as opções precedentes. IDs de nó são designados pelo sistema quando o nó é criado e podem ser usados para referenciar o nó para os propósitos de script ou automação. Insira somente um ID de nó, que deve ser uma correspondência exata. Veja o tópico [“Anotações” na página 56](#) para obter mais informações.

Procura em SuperNodes. Esta caixa é marcada por padrão, o que significa que a procura é executada nos nós dentro e fora de SuperNodes. Limpe essa caixa se desejar executar a procura somente em nós fora de SuperNodes, no nível superior do fluxo.

Localizar. Quando você tiver especificado todas as opções desejadas, clique neste botão para iniciar a procura.

Nós que correspondem às opções especificadas são listados na parte inferior da caixa de diálogo. Selecione um nó na lista para destacá-lo na tela de fluxo.

Renomeando fluxos

Usando a guia Anotações na caixa de diálogo de propriedades de fluxo, é possível incluir anotações descritivas para um fluxo e criar um nome customizado para o fluxo. Essas opções são úteis principalmente durante a geração de relatórios para fluxos incluídos na área de janela do projeto. Consulte o tópico “Anotações” na página 56 para obter informações adicionais.

Descrições de fluxo

Para cada fluxo criado, o IBM SPSS Modelador produz uma descrição do fluxo contendo informações sobre o conteúdo do fluxo. Isso pode ser útil se você estiver tentando ver o que um fluxo faz mas você não tem IBM SPSS Modelador instalado, por exemplo ao acessar um fluxo através de IBM SPSS Serviços de colaboração e implantação.

A descrição do fluxo é exibida em forma de um documento HTML consistindo em inúmeras seções.

Informações Gerais do Fluxo

Esta seção contém o nome do fluxo, com os detalhes de quando o fluxo foi criado e salvo pela última vez.

Descrição e Comentários

Esta seção inclui quaisquer:

- Anotações de fluxo (consulte “Anotações” na página 56)
- Comentários não conectados a nós específicos
- Comentários conectados a nós nas ramificações de modelagem e escoragem do fluxo

Informações de Escoragem

Esta seção contém informações sob vários títulos relacionados à ramificação de escoragem do fluxo.

- **Comentários.** Inclui comentários vinculados aos nós na ramificação de escoragem.
- **Entradas.** Lista os campos de entrada com seus tipos de armazenamento (por exemplo, sequência de caracteres, número inteiro, real, entre outros).
- **Saídas.** Lista os campos de saída, incluindo os campos adicionais geradas pelo nó de modelagem, juntamente com seus tipos de armazenamento.
- **Parâmetros.** Lista quaisquer parâmetros relativos à ramificação de escoragem do fluxo e que podem ser visualizados ou editados cada vez que o modelo é escorado. Esses parâmetros são identificados quando você clica no botão **Parâmetros de Escoragem** na guia **Implementação** da caixa de diálogo propriedades do fluxo.
- **Nó Modelo.** Mostra o nome e o tipo do modelo (por exemplo, Rede Neural, Árvore C&R e assim por diante). Esse é o nugget do modelo selecionado para o campo **Nó de Modelo** na guia **Implementação** da caixa de diálogo propriedades do fluxo.
- **Detalhes do Modelo.** Mostra detalhes do nugget do modelo identificado sob o título anterior. Onde for possível, os gráficos de avaliação e importância do preditor para o modelo serão incluídos.

Informações de modelagem

Contém informações relativas à ramificação modelagem do fluxo.

- **Comentários.** Lista quaisquer comentários ou anotações que estão conectados aos nós na ramificação de modelagem.
- **Entradas.** Lista os campos de entrada com seu papel na ramificação de modelagem (em forma de valor do papel do campo, por exemplo, Entrada, Resposta, Divisão, entre outros).
- **Parâmetros.** Lista quaisquer parâmetros relativos à ramificação de modelagem do fluxo e que podem ser visualizados ou editados cada vez que o modelo é atualizado. Esses parâmetros são identificados quando você clica no botão **Parâmetros de Construção de Modelo** na guia **Implementação** da caixa de diálogo propriedades do fluxo.
- **Modelagem do nó.** Mostra o nome e o tipo do nó de modelagem usado para gerar ou atualizar o modelo.

Visualizando descrições de fluxos

É possível visualizar o conteúdo de uma descrição de fluxo em um navegador da web clicando em uma opção na caixa de diálogo de propriedades do fluxo. O conteúdo da descrição depende das opções especificadas na guia Implementação da caixa de diálogo. Consulte o tópico [“Opções de Implementação do Fluxo”](#) na página 217 para obter informações adicionais.

Para visualizar uma descrição do fluxo:

1. No menu principal do IBM SPSS Modelador, clique em:

Ferramentas > Propriedades do Fluxo > Implementação

2. Configure o tipo de implementação, o nó de escoragem designado e quaisquer parâmetros de escoragem.
3. Se o tipo de implementação for Atualização de Modelo, opcionalmente, você poderá selecionar:
 - Nó de modelagem e quaisquer parâmetros de construção de modelo
 - Nugget de modelo na ramificação de escoragem do fluxo
4. Clique no botão **Visualizar Descrição do Fluxo**.

Exportando descrições de fluxo

É possível exportar conteúdo da descrição do fluxo para um arquivo HTML.

Para exportar uma descrição de fluxo:

1. No menu principal, clique em:

Arquivo > Descrição do Fluxo de Exportação

2. Insira um nome para o arquivo HTML e clique em **Salvar**.

Fluxos em execução

Uma vez que você especifica as opções necessárias para fluxos e conecte os nós necessários, você pode executar o fluxo executando os dados através de nós no fluxo. Há várias maneiras de se executar um fluxo dentro do IBM SPSS Modelador. Será possível:

- Clique em **Executar** no menu Ferramentas.
- Clique em um dos **Run ...** botões na barra de ferramentas. Esses botões permitem executar o fluxo inteiro ou simplesmente o nó terminal selecionado. Veja o tópico [“IBM SPSS Modelador Barra de ferramentas”](#) na página 15 para obter mais informações.
- Execute um único fluxo de dados clicando com o botão direito em um nó terminal e clicando em **Executar** no menu pop-up.

- Execute parte de um fluxo de dados clicando com o botão direito em qualquer nó não terminal e clicando em **Executar Daqui** no menu pop-up. Isso faz com que somente aquelas operações após o nó selecionado sejam executadas.

Para interromper a rodagem de um fluxo em andamento, é possível clicar no botão vermelho **Parar** na barra de ferramentas ou selecionar **Stop Execution** no menu Ferramentas. Note que clicar no botão Parar uma vez diz ao Modeler para ter o Modeler Server parar a execução. Em alguns casos a execução irá parar imediatamente, mas em outros casos deve terminar a etapa atual antes que ela possa parar toda a execução. Então o tempo vai variar. Se você clicar duas vezes no botão, a conexão do servidor será eliminada e uma nova conexão será estabelecida. Na maioria dos casos, isso fechará todos os processos do servidor (e pode levar algum tempo). Mas em alguns casos, algum processo de servidor não será interrompido.

Se algum fluxo levar mais tempo do que três segundos para ser executado, a caixa de diálogo Feedback de Execução será exibida para indicar o progresso.

Alguns nós têm exibições adicionais fornecendo informações adicionais sobre a execução do fluxo. Selecione a linha correspondente na caixa de diálogo para ver essas exibições. A primeira linha é selecionada automaticamente.

Trabalhando com modelos

Se um fluxo incluir um nó de modelagem (ou seja, da guia Modelagem ou Modelagem da Base de Dados da paleta de nós), um **nugget de modelo** será criado quando o fluxo for executado. Um nugget do modelo é um contêiner para um **modelo**, ou seja, o conjunto de regras, fórmulas ou equações que permite gerar previsões com relação a seus dados de origem e que está no centro da análise preditiva.



Figura 12. Nugget do modelo

Quando você executa um nó de modelagem com sucesso, um nugget do modelo correspondente é colocado na tela de fluxo, onde é representado por um ícone em forma de diamante dourado (por isso o nome "nugget"). É possível abrir o nugget e navegar por seu conteúdo para visualizar detalhes sobre o modelo. Para visualizar as previsões, você anexa e executa um ou mais nós terminais, cuja saída apresenta as previsões em formato legível.

Um fluxo de modelagem típico consiste em duas ramificações. A **ramificação de modelagem** contém o nó de modelagem, com os nós de processamento e de origem que o precedem. A **ramificação de escoragem** é criada quando você executa o nó de modelagem e contém o nugget do modelo e o nó ou nós terminais que podem ser usados para visualizar as previsões.

Para obter mais informações, consulte o guia *Nós de Modelagem do IBM SPSS Modelador*.

Incluindo comentários e anotações em nós e fluxos

Você pode precisar descrever um fluxo para outras pessoas em sua organização. Para ajudar com isso, é possível anexar comentários explicativos aos fluxos, nós e nuggets de modelo.

Outras pessoas poderão então visualizar esses comentários na tela ou é possível imprimir uma imagem do fluxo que inclui os comentários.

É possível listar todos os comentários para um fluxo ou SuperNode, alterar a ordem dos comentários na lista, editar o texto do comentário e alterar as cores do primeiro plano e do segundo plano de um comentário. Consulte o tópico ["Listando comentários do fluxo"](#) na página 55 para obter informações adicionais.

Também é possível incluir notas em forma de anotações de texto em fluxos, nós e nuggets por meio da guia Anotações de uma caixa de diálogo de propriedades do fluxo, caixa de diálogo do nó ou janela do nugget de modelo. Essas notas ficam visíveis somente quando a guia Anotações está aberta, exceto

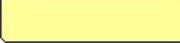
que as anotações de fluxo também podem ser mostradas como comentários na tela. Consulte o tópico “Anotações” na página 56 para obter informações adicionais.

Comentários

Comentários têm forma de caixas de texto nas quais é possível inserir qualquer quantidade de texto, e é possível incluir quantos comentários você quiser. Um comentário pode ser independente (não anexado a nenhum objeto de fluxo) ou ser conectado a um ou mais nós ou nuggets de modelo no fluxo. Os comentários independentes normalmente são usados para descrever o propósito geral do fluxo; os comentários conectados descrevem o nó ou nugget ao qual estão conectados. Nós e nuggets podem ter mais de um comentário conectado, e o fluxo pode ter qualquer número de comentários independentes.

Nota: Também é possível mostrar anotações de fluxo como comentários na tela, embora elas não possam ser anexadas a nós ou nuggets. Veja o tópico [“Convertendo anotações em comentários”](#) na página 55 para obter mais informações.

A aparência da caixa de texto muda para indicar o atual modo do comentário (ou anotação mostrada como um comentário), como mostra a tabela a seguir.

<i>Tabela 3. Modos de caixa de texto de comentário e anotação</i>				
Caixa de texto do comentário	Caixa de texto da anotação	Modo	Indica	Obtido ao...
		Editar	Comentário é aberto para edição.	Criar um novo comentário ou anotação ou dar um clique duplo em um existente.
		Última seleção	Comentário pode ser movido, redimensionado ou excluído.	Clicar no segundo plano do fluxo após a edição ou clicar em um comentário ou anotação existente.
		Visualizar	Edição está concluída.	Clicar em outro nó, comentário ou anotação após a edição.

Quando você cria um novo comentário independente, ele é exibido inicialmente no canto superior esquerdo da tela de fluxo.

Se você estiver anexando um comentário a um nó ou nugget, o comentário será exibido inicialmente acima do objeto de fluxo ao qual está anexado.

A caixa de texto tem a cor branca para mostrar que um texto pode ser inserido. Quando você tiver inserido o texto, clique fora da caixa de texto. O segundo plano do comentário muda para amarelo para mostrar que a entrada de texto foi concluída. O comentário permanece selecionado, o que permite que ele possa ser movido, redimensionado ou excluído.

Quando você clicar novamente, as bordas mudarão para linhas sólidas para mostrar que a edição foi concluída.

Um clique duplo em um comentário muda a caixa de texto para o modo de edição--o segundo plano muda para branco e o texto do comentário pode ser editado.

Também é possível anexar comentários a SuperNodes.

Operações envolvendo comentários

É possível executar inúmeras operações nos comentários. Será possível:

- Incluir um comentário independente
- Anexar um comentário a um nó ou nugget

- Editar um comentário
- Redimensionar um comentário
- Mover um comentário
- Desconectar um comentário
- Excluir um comentário
- Mostrar ou ocultar todos os comentários para um fluxo

Para incluir um comentário independente

1. Assegure que nada esteja selecionado no fluxo.

2. Faça o seguinte:

- No menu principal, clique em:

Inserir > Novo comentário

- Clique com o botão direito no segundo plano do fluxo e clique em **Novo Comentário** no menu pop-up.
 - Clique no botão **Novo Comentário** na barra de ferramentas.
3. Insira o texto do comentário (ou cole-o da área de transferência).
4. Clique em um nó no fluxo para salvar o comentário.

Para anexar um comentário a um nó ou nugget

1. Selecione um ou mais nós ou nuggets na tela de fluxo.

2. Faça o seguinte:

- No menu principal, clique em:

Inserir > Novo comentário

- Clique com o botão direito no segundo plano do fluxo e clique em **Novo Comentário** no menu pop-up.
 - Clique no botão **Novo Comentário** na barra de ferramentas.
3. Insira um texto de comentário.
4. Clique em outro nó no fluxo para salvar o comentário.

Como alternativa, é possível:

5. Inserir um comentário independente (consulte a seção anterior).

6. Faça o seguinte:

- Selecione o comentário, pressione F2 e, em seguida, selecione o nó ou nugget.
- Selecione o nó ou nugget, pressione F2 e, em seguida, selecione o comentário.
- (Somente mouses com três botões) Mova o ponteiro do mouse sobre o comentário, mantenha o botão do meio pressionado, arraste o ponteiro do mouse sobre o nó ou nugget e solte o botão.

Para anexar um comentário a um nó ou nugget adicional

Se um comentário já estiver anexado a um nó ou nugget, ou se ele estiver atualmente no nível do fluxo e você desejar anexá-lo a um nó ou nugget adicional, execute o seguinte:

- Selecione o comentário, pressione F2 e, em seguida, selecione o nó ou nugget.
- Selecione o nó ou nugget, pressione F2 e, em seguida, selecione o comentário.
- (Somente mouses com três botões) Mova o ponteiro do mouse sobre o comentário, mantenha o botão do meio pressionado, arraste o ponteiro do mouse sobre o nó ou nugget e solte o botão.

Para editar um comentário existente

1. Faça o seguinte:

- Dê um clique duplo na caixa de texto do comentário.

- Selecione a caixa de texto e pressione Enter.
 - Clique com o botão direito na caixa de texto para exibir seu menu e clique em Editar.
2. Edite o texto do comentário. É possível usar teclas de atalho do Windows padrão durante a edição, por exemplo, Ctrl+C, para copiar texto. Outras opções durante a edição são listadas no menu pop-up para o comentário.
3. Clique fora da caixa de texto uma vez para exibir os controles de redimensionamento e depois novamente para concluir o comentário.

Para redimensionar uma caixa de texto de comentário

1. Selecione o comentário para exibir os controles de redimensionamento.
2. Clique e arraste um controle para redimensionar a caixa.
3. Clique fora da caixa de texto para salvar a mudança.

Para mover um comentário existente

Se desejar mover um comentário, mas não seus objetos anexados (se houver algum), execute o seguinte:

- Mova o ponteiro do mouse sobre o comentário, mantenha pressionado o botão esquerdo do mouse e arraste o comentário para a nova posição.
- Selecione o comentário, mantenha pressionada a tecla Alt e move o comentário usando as teclas de seta.

Se desejar mover um comentário com quaisquer nós ou nuggets aos quais o comentário está anexado:

1. Selecione todos os objetos que deseja mover.
2. Faça o seguinte:
 - Mova o ponteiro do mouse sobre um dos objetos, mantenha pressionado o botão esquerdo do mouse e arraste os objetos para a nova posição.
 - Selecione um dos objetos, mantenha pressionada a tecla Alt e move os objetos usando as teclas de seta.

Para desconectar um comentário de um nó ou nugget

1. Selecione um ou mais comentários para serem desconectados.
2. Faça o seguinte:
 - Pressione F3.
 - Clique com o botão direito em um comentário selecionado e clique em Desconectar em seu menu.

Para excluir um comentário

1. Selecione um ou mais comentários para serem excluídos.
2. Faça o seguinte:
 - Pressione a tecla Delete.
 - Clique com o botão direito em um comentário selecionado e clique em Excluir em seu menu.

Se o comentário estava anexado a um nó ou nugget, a linha de conexão será excluída também.

Se o comentário era originalmente um fluxo ou uma anotação SuperNode que foi convertida para um comentário independente, o comentário será excluído da tela, mas seu texto ficará retido na guia Anotações para o fluxo ou SuperNode.

Para mostrar ou ocultar comentários para um fluxo

1. Faça o seguinte:
 - No menu principal, clique em:
Visualizar > Comments
 - Clique no botão **Mostrar/Ocultar Comentários** na barra de ferramentas.

Listando comentários do fluxo

É possível visualizar uma lista de todos os comentários que foram feitos para um determinado fluxo ou SuperNode.

Nessa lista, é possível

- Alterar a ordem dos comentários
- Editar o texto do comentário
- Alterar as cores do primeiro plano e do plano de fundo de um comentário

Listando Comentários

Para listar os comentários feitos para um fluxo, execute uma das opções a seguir:

- No menu principal, clique em:

Ferramentas > Propriedades de Fluxo > Comments

- Clique com o botão direito em um fluxo na área de janela de gerenciadores e clique em **Propriedades do Fluxo**, depois em **Comentários**.
- Clique com o botão direito no segundo plano de um fluxo na tela e clique em **Propriedades do Fluxo**, depois em **Comentários**.

Texto. O texto do comentário. Dê um clique duplo no texto para alterar o campo para uma caixa de texto editável.

Links. O nome do nó ao qual o comentário é anexado. Se esse campo estiver vazio, o comentário se aplicará ao fluxo.

Botões de posicionamento. Movem um comentário selecionado para cima ou para baixo na lista.

Cores do comentário. Para alterar a cor do plano de fundo ou do primeiro plano de um comentário, selecione o comentário, marque a caixa de seleção **Cores Customizadas** e, em seguida, selecione uma cor da lista **Segundo Plano** ou **Primeiro Plano** (ou ambas). Clique em **Aplicar** e, em seguida, clique no segundo plano do fluxo para ver o efeito da em mudança. Clique em **OK** para salvar a mudança.

Convertendo anotações em comentários

Anotações feitas em fluxos ou SuperNodes podem ser convertidas em comentários.

No caso dos fluxos, a anotação é convertida em um comentário independente (ou seja, não é anexado a nenhum outro nó) na tela de fluxo.

Quando uma anotação SuperNode é convertida em um comentário, o comentário não é anexado ao SuperNode na tela de fluxo, mas fica visível quando você aumenta o zoom no SuperNode.

Para converter uma anotação de fluxo em um comentário

1. Clique em **Propriedades do Fluxo** no menu Ferramentas. (Alternativamente, é possível clicar com o botão direito em um fluxo na área de janela dos gerenciadores e clicar em **Propriedades do Fluxo**.)
2. Clique na guia **Anotações**.
3. Marque a caixa de seleção **Mostrar Anotação como Comentário**.
4. Clique em **OK**.

Para converter uma anotação SuperNode em um comentário

1. Dê um clique duplo no ícone SuperNode na tela.
2. Clique na guia **Anotações**.
3. Marque a caixa de seleção **Mostrar Anotação como Comentário**.
4. Clique em **OK**.

Anotações

Nós, fluxos e modelos podem ser anotados de inúmeras formas. É possível incluir anotações descritivas e especificar um nome customizado. Essas opções são úteis principalmente durante a geração de relatórios para fluxos incluídos na área de janela do projeto. Para nós e nuggets de modelo, também é possível incluir um texto de dica de ferramenta para ajudar a distinguir entre nós semelhantes na tela de fluxo.

Incluindo Anotações

A edição de um nó ou nugget de modelo abre uma caixa de diálogo em aba contendo uma guia Anotações usada para configurar uma variedade de opções de anotação. Também é possível abrir a guia Anotações diretamente.

1. Para anotar um nó ou nugget, clique com o botão direito no nó ou nugget na tela de fluxo e clique em **Renomear e Anotar**. A caixa de diálogo de edição abre com a guia Anotações visível.
2. Para anotar um fluxo, clique em **Propriedades do Fluxo** no menu Ferramentas. (Alternativamente, é possível clicar com o botão direito em um fluxo na área de janela dos gerenciadores e clicar em **Propriedades do Fluxo**.) Clique na guia Anotações.

Nome. Selecione **Customizado** para ajustar o nome gerado automaticamente ou para criar um nome exclusivo para o nó conforme exibido na tela de fluxo.

Texto de dica de ferramenta. (Somente para nós e nuggets de modelo) Insira o texto usado como uma dica de ferramenta na tela de fluxo. Isso é particularmente útil quando se trabalha com um grande número de nós semelhantes.

palavras-chave. Especifique palavras-chave para serem usadas em relatórios de projeto e durante a procura por nós em um fluxo ou durante o rastreamento de objetos armazenados no repositório (consulte “Sobre o IBM SPSS Reppositório de serviços de colaboração e implantação” na página 205). Várias palavras-chave podem ser separadas por ponto e vírgula, por exemplo, `income; crop type; claim value`. Espaços em branco no início e no final de cada palavra-chave são removidos -- por exemplo, `income ; crop type` produzirá os mesmos resultados que `income;crop type`. (No entanto, espaços em branco dentro de palavras-chave não são aparados. Por exemplo, `crop type` com um espaço e `crop type` com dois espaços não são iguais.)

A área de texto principal pode ser usada para inserir anotações longas referentes às operações do nó ou decisões tomadas no nó. Por exemplo, quando você está compartilhando e reutilizando fluxos, é útil fazer anotações sobre decisões, como descartar um campo com inúmeros espaços em branco usando um nó de Filtro. A anotação do nó armazena essas informações com o nó. Também é possível escolher incluir essas anotações em um relatório do projeto criado a partir da área de janela do projeto. Consulte o tópico “Introdução aos Projetos” na página 227 para obter informações adicionais.

Mostrar anotação como comentário. (Somente para anotações de fluxo e SuperNode) Marque esta caixa para converter a anotação em um comentário independente que ficará visível na tela de fluxo. Veja o tópico “Incluindo comentários e anotações em nós e fluxos” na página 51 para obter mais informações.

ID. Exibe um ID exclusivo que pode ser usado para referenciar o nó para o propósito de script ou automação. Esse valor é gerado automaticamente quando o nó é criado e não mudará. Observe também que para evitar confusão com a letra “O,” zeros não são usados em IDs de nós. Use o botão de cópia à direita para copiar e colar o ID nos scripts ou em onde for necessário.

Salvando fluxos de dados

Depois de criar um fluxo, você pode salvá-lo para reutilização futura.

Para salvar um fluxo

1. No menu Arquivo, clique em **Salvar Fluxo** ou em **Salvar Fluxo Como**.
2. Na caixa de diálogo Salvar, navegue para a pasta em que deseja salvar o arquivo de fluxo.
3. Insira um nome para o fluxo na caixa de texto Nome do Arquivo.

4. Selecione **Incluir no Projeto** se quiser incluir o fluxo salvo no projeto atual.

Um clique em **Salvar** armazena o fluxo com a extensão `*.str` no diretório especificado.

Arquivos de backup automático. Cada vez que um fluxo é salvo, a versão salva anteriormente do arquivo é automaticamente preservada como um backup, com um hífen anexado ao nome do arquivo (por exemplo, `mystream.str-`). Para restaurar a versão submetida a backup, simplesmente exclua o hífen e reabra o arquivo.

Salvando estados

Além de fluxos, é possível salvar **estados**, que incluem o diagrama de fluxo atualmente exibido e quaisquer nuggets de modelo que você criou (listados na guia Modelos na área de janela de gerenciadores).

Para salvar um estado

1. No menu Arquivo, clique em:

Estado > Salvar Estado ou Salvar Estado Como

2. Na caixa de diálogo Salvar, navegue para a pasta em que deseja salvar o arquivo de estado.

Um clique em **Salvar** armazena o estado com a extensão `*.cst` no diretório especificado.

Salvando nós

Também é possível salvar um nó individual clicando nele com o botão direito na tela de fluxo e clicando em **Salvar Nó** no menu pop-up. Use a extensão do arquivo `*.nod`.

Salvando vários objetos de fluxo

Ao sair do IBM SPSS Modelador com vários objetos não salvos, como fluxos, projetos ou nuggets de modelo, será solicitado que você salve antes de fechar completamente o software. Se escolher salvar itens, uma caixa de diálogo exibirá opções para salvar cada objeto.

1. Basta marcar as caixas de seleção para os objetos que deseja salvar.

2. Clique em **OK** para salvar cada objeto no local necessário.

Será aberto um prompt com a caixa de diálogo Salvar padrão para cada objeto. Depois de terminar de salvar, o aplicativo fechará.

Salvando saída

Tabelas, gráficos e relatórios gerados a partir dos nós de saída do IBM SPSS Modelador podem ser salvos em formato de objeto de saída (`*.cou`).

1. Durante a visualização da saída que você deseja salvar, nos menus da janela de saída, clique em:

Arquivo > Salvar

2. Especifique um nome e uma localização para o arquivo de saída.

3. Opcionalmente, selecione **Incluir Arquivo no Projeto** na caixa de diálogo Salvar para incluir o arquivo no projeto atual. Veja o tópico ["Introdução aos Projetos"](#) na página 227 para obter mais informações.

Alternativamente, é possível clicar com o botão direito em qualquer objeto de saída listado na área de gerenciadores e selecionar **Salvar** no menu pop-up.

Criptografando e decriptografando informações

Quando você salva um fluxo, nó, projeto, arquivo de saída ou nugget do modelo, é possível criptografá-los para evitar seu uso desautorizado. Para isso, você seleciona uma opção extra ao salvar e inclui uma senha no item sendo salvo. Essa criptografia pode ser configurada para qualquer um dos itens salvos e inclui segurança extra; isso não será o mesmo que a criptografia SSL usada se você estiver passando arquivos entre o IBM SPSS Modelador e o ServidorIBM SPSS Modeler.

Quando você tenta abrir um item criptografado, é solicitado que você insira uma senha. Após a inserção da senha correta, o item é decriptografado automaticamente e aberto como usual.

Para criptografar um item

1. Na caixa de diálogo Salvar, para o item ser criptografado, clique em **Opções**. A caixa de diálogo Opções de Criptografia é aberta.
2. Selecione **Criptografar este Arquivo**.
3. Opcionalmente, para maior segurança, selecione **Máscara de Senha**. Isso exibe qualquer coisa que você inserir como uma série de pontos.
4. Insira a senha. *Aviso:* Se você esqueceu a senha, o arquivo ou modelo não poderá ser aberto.
5. Se você selecionou **Máscara de Senha**, insira a senha novamente para confirmar que ela foi inserida corretamente.
6. Clique em **OK** para retornar à caixa de diálogo Salvar.

Nota: Se você salvar uma cópia de qualquer item protegido por criptografia, o novo item será salvo automaticamente em um formato criptografado usando a senha original, a menos que você altere as configurações na caixa de diálogo Opções de Criptografia.

Carregando arquivos

É possível recarregar inúmeros objetos salvos no IBM SPSS Modelador:

- Fluxos (.str)
- Estados (.cst)
- Modelos (.gm)
- Paleta de modelos (.gen)
- Nós (.nod)
- Saída (.cou)
- Projetos (.cpj)

Abrindo novos arquivos

Fluxos podem ser carregados diretamente no menu Arquivo.

- No menu Arquivo, clique em **Abrir Fluxo**.

Todos os outros tipos de arquivo podem ser abertos usando os itens de submenu disponíveis no menu Arquivo. Por exemplo, para carregar um modelo, no menu Arquivo, clique em:

Modelos > Paleta de Open Model ou Load Models

Abrindo arquivos usados recentemente

Para o carregamento rápido de arquivos usados recentemente, é possível usar as opções para parte inferior do menu Arquivo.

Selecione **Fluxos Recentes**, **Projetos Recentes** ou **Estados Recentes** para expandir uma lista de arquivos usados recentemente.

Mapeando fluxos de dados

Usando a ferramenta de mapeamento, é possível conectar uma nova origem de dados a um fluxo pré-existente. A ferramenta de mapeamento não só irá configurar a conexão, mas também ajudará a especificar como os campos na nova origem substituirão aqueles no fluxo existente. Em vez de recriar um fluxo de dados inteiro para uma nova origem de dados, é possível simplesmente se conectar a um fluxo existente.

A ferramenta de mapeamento de dados permite unir dois fragmentos de fluxo e se certificar de que todos os nomes de campos (essenciais) sejam correspondentes. Essencialmente, os dados de mapeamento resultam simplesmente na criação de um novo nó de Filtro, o que faz a correspondência dos campos apropriados por meio da renomeação.

Há duas maneiras equivalentes de se mapear dados:

Selecionar o nó de substituição. Esse método começa com o nó a ser substituído. Primeiro, você clica com o botão direito do mouse em substituição; em seguida, usando a opção **Mapeamento de Dados > Selecionar Nó de Substituição** no menu pop-up, selecione o nó com o qual irá substituí-lo.

Mapear para. Esse método começa com o nó a ser introduzido no fluxo. Primeiro, clique com o botão direito do mouse para introduzir; em seguida, usando a opção **Mapeamento de Dados > Mapear Para** no menu pop-up, selecione o nó ao qual ele deve aderir. Esse método é útil principalmente para o mapeamento para um nó terminal. *Nota:* Não é possível mapear para os nós de Mesclagem ou Anexação. Em vez disso, é necessário simplesmente conectar o fluxo ao nó de Mesclagem de maneira normal.

O mapeamento de dados é fortemente integrado à construção do fluxo. Se tentar se conectar a um nó que já tem uma conexão, você terá a opção de substituir a conexão ou mapeamento para esse nó.

Mapeando dados para um modelo

Para substituir a origem de dados para um fluxo de modelo por um novo nó de origem trazendo seus próprios dados para o IBM SPSS Modelador, é necessário usar a opção **Selecionar Nó de Substituição** no menu pop-up Mapeamento de Dados. Essa opção está disponível para todos os nós, exceto nós de Mesclagem, Agregado e todos os nós terminais. O uso da ferramenta de mapeamento de dados para executar essa ação ajuda a assegurar que os campos sejam adequadamente correspondentes entre as operações de fluxo existentes e a nova origem de dados. Os passos a seguir fornecem uma visão geral do processo de mapeamento de dados.

Passo 1: Especificar campos essenciais no nó de origem original. Para que operações de fluxo sejam executadas adequadamente, os campos essenciais devem ser especificados. Veja o tópico [“Especificando campos essenciais”](#) na página 60 para obter mais informações.

Passo 2: Incluir uma nova origem de dados na tela de fluxo. Usando um dos nós de origem, traga novos dados de substituição.

Passo 3: Substituir o nó de origem do modelo. Usando a opção Mapeamento de Dados no menu pop-up para o nó de origem do modelo, clique em **Selecionar Nó de Substituição** e, em seguida, selecione o nó de origem para os dados de substituição.

Passo 4: Verificar campos mapeados. Na caixa de diálogo que é aberta, verifique se o software está mapeando os campos de maneira adequada da origem de dados de substituição para o fluxo. Quaisquer campos essenciais não mapeados serão exibidos em vermelho. Esses campos são usados em operações de fluxo e devem ser substituídos por um campo semelhante na nova origem de dados para que operações de recebimento de dados funcionem adequadamente. Veja o tópico [“Examinando campos mapeados”](#) na página 60 para obter mais informações.

Após o uso da caixa de diálogo para assegurar que todos os campos essenciais sejam mapeados adequadamente, a origem de dados antiga é desconectada e a nova origem de dados é conectada ao fluxo usando um nó de Filtro chamado *Mapear*. Esse nó de Filtro direciona o mapeamento real dos campos no fluxo. Um nó de Filtro *Remover Mapeamento* também é incluído na tela de fluxo. O nó de filtro *Remover Mapeamento* pode ser usado para inverter o mapeamento de nome de campo incluindo-o no fluxo. Isso desfará os campos mapeados, mas observe que você terá que editar quaisquer nós terminais de recebimento de dados para selecionar novamente os campos e as sobreposições.

Mapeamento entre fluxos

Assim como a conexão de nós, esse método de mapeamento de dados não requer a configuração de campos essenciais antecipadamente. Com esse método, você simplesmente se conecta de um fluxo a outro usando **Mapear para** no menu pop-up Mapeamento de Dados. Esse tipo de mapeamento de dados

é útil para mapear para nós terminais e para copiar e colar entre fluxos. *Nota:* Usando a opção **Mapear para**, não é possível mapear para os nós de Mesclagem, Anexação e todos os tipos de nós de origem.

Para mapear dados entre fluxos

1. Clique com o botão direito no nó que deseja usar para se conectar ao novo fluxo.
2. No menu, clique em:

Mapeamento de Dados > Mapa para

3. Use o cursor para selecionar um nó de destino no fluxo de destino.
4. Na caixa de diálogo que é aberta, assegure que os campos estejam correspondidos de forma adequada e clique em **OK**.

Especificando campos essenciais

Durante o mapeamento para um fluxo existente, normalmente campos essenciais são especificados pelo autor do fluxo. Esses campos essenciais indicam se um determinado campo é usado nas operações de recebimento de dados. Por exemplo, o fluxo existente pode construir um modelo que usa um campo chamado *Perda de Cliente*. Nesse fluxo, *Perda de Cliente* é um campo essencial porque não seria possível construir o modelo sem ele. Da mesma forma, os campos usados em nós de manipulação, como nó de Derivação, são necessários para derivar o novo campo. A configuração desse tipo de campo explicitamente ajuda a assegurar que os campos adequados no nó de origem sejam mapeados para eles. Se campos obrigatórios não forem mapeados, você receberá uma mensagem de erro. Se você decidir que certas manipulações ou nós de saída são desnecessários, é possível excluir os nós do fluxo e remover os campos apropriados da lista Campos Essenciais.

Para configurar campos essenciais

1. Clique com o botão direito no nó de origem no fluxo de modelo que será substituído.
2. No menu, clique em:

Mapeamento De Dados > Especificar Campos Essenciais

3. Usando o Seletor de Campo, é possível incluir ou remover campos da lista. Para abrir o Seletor de Campo, clique no ícone à direita da lista de campos.

Examinando campos mapeados

Após ter selecionado o ponto em que uma origem de dados ou um fluxo de dados será mapeado para outro, uma caixa de diálogo será exibida para você selecionar campos para mapeamento ou para assegurar que o mapeamento padrão do sistema esteja correto. Se os campos essenciais tiverem sido configurados para o fluxo ou origem de dados e não forem correspondentes, eles serão exibidos em vermelho. Quaisquer campos não mapeados a partir da origem de dados passarão pelo nó de Filtro não alterado, mas observe que é possível mapear um campo não essencial também.

Original. Lista todos os campos no modelo ou fluxo existente - todos os campos que estão presentes além do recebimento de dados. Os campos da nova origem de dados serão mapeados para esses campos.

Mapeado. Lista os campos selecionados para mapeamento para campos modelo. Esses são os campos cujos nomes podem precisar de mudança para corresponderem aos campos originais usados nas operações de fluxo. Clique na célula de tabela para um campo para ativar uma lista de campos disponíveis.

Se você não tiver certeza de quais campos mapear, talvez seja útil examinar os dados de origem mais de perto antes do mapeamento. Por exemplo, é possível usar a guia **Tipos** no nó de origem para revisar um resumo dos dados de origem.

Dicas e atalhos

Trabalhe de maneira rápida e fácil se familiarizando com os atalhos e dicas a seguir:

- **Construir fluxos rapidamente dando um clique duplo.** Basta dar um clique duplo em um nó na paleta para incluí-lo e conectá-lo ao fluxo atual.
- **Usar combinações de teclas para selecionar nós de recebimento de dados.** Pressione Ctrl+Q e Ctrl+W para alternar a seleção de todos os nós de recebimento de dados.
- **Usar teclas de atalho para conectar e desconectar nós.** Quando um nó for selecionado na tela, pressione F2 para iniciar uma conexão, pressione Tab para se mover para o nó necessário e pressione Shift+Barra de Espaço para concluir a conexão. Pressione F3 para desconectar todas as entradas e saídas do nó selecionado.
- **Customizar a guia Paleta de Nós com seus nós favoritos.** No menu Ferramentas, clique em **Gerenciar Paletas** para abrir uma caixa de diálogo para incluir, remover ou mover os nós mostrados na Paleta de Nós.
- **Renomear nós e incluir Dicas de ferramenta.** Cada caixa de diálogo inclui uma guia Anotações na qual é possível especificar um nome customizado para nós na tela, bem como incluir Dicas de ferramenta para ajudar a organizar seu fluxo. Também é possível incluir anotações detalhadas para controlar o progresso, salvar detalhes do processo e denotar quaisquer decisões de negócios necessárias ou arquivadas.
- **Inserir valores automaticamente em uma expressão do CLEM.** Usando o construtor de expressões, acessível em várias caixas de diálogo (como aquelas para os nós de Derivação e Filtro), é possível inserir valores de campo automaticamente em uma expressão do CLEM. Clique no botão de valores no construtor de expressões para escolher entre os valores de campo existentes.



Figura 13. Botão de valores

- **Navegar para arquivos rapidamente.** Durante a navegação em arquivos em uma caixa de diálogo Abrir, use a lista Arquivos (clique no botão de diamante amarelo na parte superior da caixa de diálogo e, em seguida, no campo Verificar) para acessar os diretórios usados anteriormente, bem como diretórios padrão do IBM SPSS Modelador. Use os botões Avançar e Voltar para rolar pelos diretórios acessados.
- **Minimizar desordem na janela de saída.** É possível fechar e excluir uma saída rapidamente usando o botão X vermelho no canto superior direito de todas as janelas de saída. Isso permite que você mantenha somente os resultados mais promissores e interessantes na guia Saída da área de janelas de gerenciadores.

Um intervalo completo de atalhos de teclado está disponível para o software. Consulte o tópico “[Teclado de acessibilidade](#)” na página 254 para obter informações adicionais.

Você sabia que é possível...

- Arrastar e selecionar um grupo de nós na tela de fluxo usando seu mouse.
- Copiar e colar nós de um fluxo em outro.
- Acessar Ajuda de cada caixa de diálogo e janela de saída.
- Obter Ajuda sobre CRISP-DM o processo padrão de vários segmentos de mercados para mineração de dados. (No menu Ajuda, clique em **Ajuda do CRISP-DM**.)

Capítulo 6. Gráficos de construção

serviço necessário

Estúdio Watson

Formato de dados

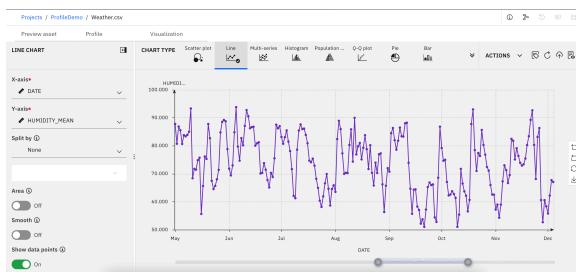
Tabulares: Avro, CSV, JSON, Parquet, TSV, SAV, arquivos .xls e .xlsx Microsoft Excel, SAS, arquivos de texto delimitados e dados conectados.

Para obter mais informações sobre as origens de dados suportadas, consulte [Conectores](#)

Tamanho dos dados

Sem limites

É possível criar gráficos semelhantes ao exemplo a seguir, que mostra como os valores de umidade variam com o tempo.



Iniciando o criador de gráficos

- Clique com o botão direito do mouse no nó de dados com o qual deseja trabalhar e selecione **Exibir dados**.

Catálogo de gráficos

O catálogo do Gráfico fornece opções para encontrar gráficos adequados, navegando nas categorias de gráfico, usando busca por palavra-chave ou filtragem por características do gráfico. O catálogo do Gráfico pode ser acessado a partir de qualquer guia de dataset ativo.

A partir de qualquer guia de dataset ativo clique no controle **builder Gráfico** que está disponível na barra de ferramentas. Alternadamente, acesse o catálogo através do **Visualize > builder gráfico ...**. Existem vários métodos que você pode usar para pesquisar tipos de gráficos adequados.

- O campo **Find chart type** pode ser usado para procurar gráficos específicos e palavras-chave. À medida que você insere um nome de gráfico ou palavras-chave, o número de filtros de azulejos de gráficos disponíveis para fornecer resultados para apenas os nomes de gráficos especificados ou palavras-chave.

Nota: Você pode usar a visualização de azulejos e listar ícones de visualização ao lado do campo **Find chart type** para controlar como os tipos de gráficos são apresentados no catálogo Gráfico.

- Os azulejos do catálogo do Gráfico são agrupados de acordo com as suas categorias correspondentes. Você pode rolar pelos azulejos disponíveis até localizar o tipo de gráfico desejado.
- A pane **Filtro** fornece opções para filtragem de gráficos de acordo com suas categorias correspondentes, propósito, variáveis dependentes e estatísticas.

Notes:

- O número próximo a cada título de pane **Filtro** identifica o número de filtros que são selecionados atualmente para cada seção.
- Clique em **Reset filter** para limpar as configurações da pane **Filtrar**.

Layout e termos

Tela

O canvas é a área do diálogo Gráfico Builder onde você constrói o gráfico.

Tipo de gráfico

Lista os tipos de gráficos disponíveis. Os elementos gráficos são os itens do gráfico que representam dados (barras, pontos, linhas, entre outros).

Área de janela Detalhes

A pane Detalhes fornece os blocos de construção do gráfico básico.

Configurações de gráfico

Fornece opções para selecionar quais variáveis são usadas para construir o gráfico, método de distribuição, campos de título e subtítulo e assim por diante. Dependendo do tipo de gráfico selecionado, as opções de painel Detalhes podem variar. Para obter mais informações, consulte “[Tipos de gráficos](#)” na página 64.

Ações

Fornece opções para download de arquivos de configuração do gráfico, baixando gráficos como arquivos de imagem, redefinindo gráficos e definindo as preferências do gráfico global.

Construindo um gráfico a partir da galeria de tipos de gráficos

Use galeria de tipos de gráficos para construção de gráficos. Siga os passos gerais para construir um gráfico a partir da galeria.

1. Na seção **Tipo Gráfico**, selecione uma categoria de gráfico. Uma versão pré-visualização do tipo de gráfico selecionado é mostrada na tela do gráfico.
2. Se o canvas já exibe um gráfico, o novo gráfico substitui o conjunto de eixos e elementos gráficos do gráfico.
 - a. Dependendo do tipo de gráfico selecionado, as variáveis disponíveis são apresentadas sob uma série de diferentes títulos na área de janela Detalhes (por exemplo, **Category** para gráficos de barras, **eixo X** e **eixo Y** para gráficos de linha). Selecione as variáveis apropriadas para o tipo de gráfico selecionado.

Tipos de gráficos

A galeria contém uma coleção dos gráficos mais comumente usados.

A guia **Visualização** contém uma coleção dos gráficos mais usados.

Gráficos 3D

Os gráficos 3D são comumente usados para representar funções múltiplas-variáveis e incluem uma variável de eixo z que é uma função das variáveis x e y-eixo.

Criando um gráfico 3D simples

1. Na seção **Tipo Gráfico**, clique no ícone **3D**.



O canvas atualiza para exibir um modelo de gráfico 3D.

2. Selecione o gráfico **Tipo** a partir da lista suspensa.
3. Selecione uma variável **X-axis** da lista suspensa.
4. Selecione uma variável **eixo Y** a partir da lista suspensa.
5. Selecione uma variável **Z-axis** da lista suspensa.

Opções

Tipo

Lista os tipos de gráfico que estão disponíveis para representar os dados.

eixo X

Lista variáveis que estão disponíveis para o eixo X do gráfico.

eixo Y

Lista variáveis que estão disponíveis para o eixo Y do gráfico.

Eixo Z

Lista variáveis que estão disponíveis para o eixo Z-eixo do gráfico.

Informações da dica de ferramenta

Lista as variáveis que podem ser usadas as informações da dica de ferramentas de geração quando o cursor pairar sobre um ponto de dados.

Mapa de cores

Lista variáveis de mapa de cores disponíveis. Essas variáveis usam progressão de cor, com base na faixa de valores na coluna especificada, para representarem a si mesmos nos pontos da trama. Os mapas de cores também são conhecidos como mapas de choropleth.

Dimensionar mapa

Lista variáveis de mapa de tamanho disponíveis. Essas variáveis usam tamanhos diferenciados para representar a si mesmos nos pontos da trama.

Razão Z

Configura a escala dos valores de dados do eixo Z, relativos aos eixos X e Y.

Girar

Ativa e desativa a rotação do gráfico.

Dicas de ferramentas de pontos de dados

Controles onde o ponto de dados tooltips display (direito de pontos de dados, superior direito de gráfico ou ocultar).

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Gráficos de barras

Os gráficos de barras são úteis para resumir variáveis categóricas. Por exemplo, você pode usar um gráfico de barras para mostrar o número de homens e o número de mulheres que participaram de uma pesquisa. Você também pode usar um gráfico de barras para mostrar o salário médio dos homens e o salário médio das mulheres.

Criação de um gráfico de barras simples

1. Na seção **Chart Type (Tipo de gráfico)**, clique no ícone **Bar (Barra)**.



A tela é atualizada para exibir um modelo de gráfico de barras.

2. Selecione uma variável categórica (nominal ou ordinal) como a variável **de categoria**. Você pode usar uma variável de escala, mas os resultados são úteis apenas em alguns casos especiais. Um gráfico de barras fica melhor com um número limitado de valores distintos. Se você criar um gráfico de barras com um eixo de **Categoria de escala**, as barras serão finas porque cada barra é desenhada em um valor exato e não pode sobrepor outros valores contínuos.
3. Selecione uma estatística na lista **Resumo**. O resultado de qualquer estatística determina a altura das barras. Se a estatística desejada não aparecer na lista **Resumo**, ela pode exigir uma variável. Selecione

uma variável na lista **Value (Valor)** e verifique se a estatística está disponível agora. Podem existir outras limitações de tipo de gráfico. Por exemplo, os gráficos de barras de erro podem ser calculados somente para estatísticas específicas.

Opções

Categoria

Lista as variáveis que estão disponíveis para o eixo X do gráfico.

Ordenar com base em

Selecione uma opção de classificação para as categorias dentro da variável.

Nome da categoria

Use os rótulos de categoria para classificar as categorias da variável. Os rótulos aparecem no gráfico, geralmente como rótulos de tique ou de legenda.

Valor categórico

Use o valor que está armazenado no conjunto de dados para classificar as categorias da variável. O valor da categoria é o que identifica a categoria no conjunto de dados. Muitas vezes, é diferente de seu rótulo e não é necessariamente descritivo. Por exemplo, o valor pode ser um número (por exemplo, 1), enquanto o rótulo é uma descrição de texto da categoria (por exemplo, *Feminino*).

Ordem da categoria

Selecione a ordem em que as categorias de variáveis são classificadas.

Como lido

As categorias de variáveis são apresentadas conforme aparecem no conjunto de dados.

Crescente

Classifique as categorias de variáveis em ordem crescente.

Decrescente

Classifique as categorias de variáveis em ordem decrescente.

Resumo

Selecione uma função de resumo estatístico para o elemento gráfico. O resultado da estatística determina a posição dos elementos gráficos no eixo Y. Em um gráfico 2-D, a estatística é calculada para cada valor no eixo X. Em um gráfico 3D, ele é calculado para a interseção dos valores no eixo X e no eixo Z.

Há dois tipos de funções de resumo estatístico disponíveis. A distinção é importante porque determina se você precisa especificar uma variável **Value**.

- **Funções que não requerem uma variável de valor.** Funções que não requerem uma variável. Todas as estatísticas de contagem e porcentagem estão nessa categoria. Essas estatísticas estão disponíveis quando a variável **Value** não está definida.
- **Funções que de fato exigem uma variável de valor.** Funções que de fato exigem uma variável **Value**. Por exemplo, a função " **Mean** " requer uma variável na qual a média é calculada. Essas estatísticas estão disponíveis quando a variável **Value** não está definida.

Valor

Esse campo é exibido quando uma função **de resumo** que requer uma variável de valor é selecionada. Selecione uma variável para servir como valor.

Dividir por

Selecione uma variável categórica que crie uma tabela de gráficos, com uma célula para cada categoria na variável Split by. Assim como o agrupamento, a divisão por variáveis essencialmente adiciona mais dimensões ao seu gráfico, exibindo informações para cada categoria de variável.

Tipo de divisão

Quando uma variável **Split by** é selecionada, você pode optar por exibir as barras de categoria resultantes como empilhadas ou agrupadas. O agrupamento e o empilhamento adicionam dimensionalidade ao gráfico. O agrupamento divide uma barra em várias barras, e o empilhamento cria segmentos em cada barra. Tenha cuidado ao escolher a estatística correta para o empilhamento.

Quando os valores são adicionados (empilhados), o resultado deve fazer sentido. Por exemplo, a adição e o empilhamento de valores médios (médias) geralmente não são significativos.

Tipo de barra

Selecione o tipo de gráfico de barras entre as opções fornecidas.

- eixo X
- eixo Y
- Eixo X inverso
- Eixo Y inverso
- Eixo de ângulo polar
- Eixo de raio polar
- Arco-íris polar

Posição do rótulo

Selecione a posição do rótulo do gráfico no menu suspenso.

- Nenhum
- parte superior
- esquerda
- direita
- parte inferior
- por dentro
- insideLeft
- insideRight
- insideTop
- insideBottom
- insideTopLeft
- insideBottomLeft
- insideTopRight
- insideBottomRight

Mostrar linha de referência

O controle de alternância ativa e desativa a exibição de linhas de referência no gráfico. As opções disponíveis são **Min**, **Max** e **Average**, que exibem linhas de referência nos valores mínimo, máximo e médio do gráfico.

Inserir um valor de linha de referência

Quando a opção **Mostrar linha de referência** está ativada, essa configuração oferece a opção de especificar um valor de linha de referência. Clique em **Adicionar outra coluna** para especificar mais valores de linha de referência.

Transpor

Quando ativado, os eixos X e Y do gráfico são transpostos.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Rótulo do Eixo x

O rótulo do eixo x.

Rótulo do Eixo y

O rótulo do eixo y.

box plots

Um gráfico de plot de caixa mostra as cinco estatísticas (mínimo, primeiro quartil, mediano, terceiro quartil e máximo). Ele é útil para exibir a distribuição de uma variável de escala e pinapontando outliers.

Criando uma trama de caixa simples

1. Na seção **Tipo Gráfico**, clique no ícone **Caixa de caixa**.



As atualizações do canvas para exibir um modelo de gráfico de plot de caixa.

2. Selecione uma ou mais variáveis de escala como a variável **Columns**.

Nota: A estatística para um enredo de ponto é Box plot. Não é possível alterar esta configuração.

Opções

Colunas

Lista variáveis que estão disponíveis para o eixo X do gráfico.

Clique em **Adicionar outra coluna** para adicionar mais colunas.

Dividir por

Selecionar uma variável categórica que crie uma tabela de gráficos, com uma célula para cada categoria na variável Split by. Assim como o agrupamento, a divisão por variáveis essencialmente adiciona mais dimensões ao seu gráfico, exibindo informações para cada categoria de variável.

Ordem da categoria

Selecionar a ordem em que as categorias de variáveis são classificadas.

Como lido

As categorias de variáveis são apresentadas conforme aparecem no conjunto de dados.

Crescente

Classifique as categorias de variáveis em ordem crescente.

Decrescente

Classifique as categorias de variáveis em ordem decrescente.

Portador de força de IQR

O titular da força da gama inter-quartil ($N \times IQR$). O valor padrão N é 1.5.

Normalizar dados

Quando ativado, esta configuração transforma dados em uma distribuição normal comparando dados de vários conjuntos de dados ou múltiplas colunas. Esta configuração cria 100% de empilhamento para contagens e converte estatísticas em percents.

Transpor

Quando ativado, os eixos X e Y do gráfico são transpostos.

Mostrar o valor discrepante

Quando ativado, outliers são exibidos no gráfico.

Mostrar o estilo de valor discrepante extremo

Quando ativado, estilo outlier extremo exibe no gráfico.

Mostrar rótulo

Quando ativado, etiquetas de coluna são exibidas no gráfico. Apenas dados de séries de dispersão são suportados.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Rótulo do Eixo x

O rótulo do eixo x.

Rótulo do Eixo y

O rótulo do eixo y.

gráficos de bolhas

Gráficos de bolhas exibem categorias em seus grupos como círculos lotados não hierárquicos. O tamanho de cada círculo (bolha) é proporcional ao seu valor. Gráficos de bolhas são úteis para comparar relacionamentos em seus dados.

Criando um gráfico de bolhas simples

1. Na seção **Tipo Gráfico**, clique no ícone **Bubble**.

Bubble



O canvas atualiza para exibir um modelo de gráfico de bolhas.

2. Selecione uma variável **Colunas** a partir da lista suspensa.

Nota: Clique em **Adicionar outra coluna** para incluir mais variáveis de coluna.

Opções

Colunas

Lista variáveis que estão disponíveis para o gráfico.

Clique em **Adicionar outra coluna** para adicionar mais colunas.

Cor do grupo

Ligue ou desative os agrupamentos de cores.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Gráficos Candlestick (candelabro)

Os gráficos de candlestick são um estilo de gráficos financeiros que são usados para descrever movimentos de preços de uma segurança, derivativos ou moeda. Cada elemento candlestick tipicamente mostra um dia. Um gráfico de um mês pode mostrar os 20 dias de negociação como 20 elementos de candlesticks. Os gráficos de candlestick são mais frequentemente utilizados na análise de padrões de equidade e de preço da moeda e são semelhantes a tramas de caixa.

O conjunto de dados que é usado para criar um gráfico de candlestick deve conter valores abertos, altos, baixos e próximos para cada período de tempo que você deseja exibir.

Criando um gráfico de Candlestick simples

1. Na seção **Tipo Gráfico**, clique no ícone **Candlestick**.



As atualizações do canvas para exibir um modelo de gráfico Candlestick.

2. Selecione uma variável como a variável **X-axis** .
3. Selecione uma variável como a variável **Alta** .
4. Selecione uma variável como a variável **Low** .

Opções

eixo X

Lista variáveis que estão disponíveis para o eixo X do gráfico.

Resumo

Quando ativado, cálculos sumários para disponíveis para as seguintes opções.

Alto

Lista variáveis que estão disponíveis para o alto valor de preço do gráfico.

Resumo de campo alto

Selecionar uma função de resumo estatístico para a alta variável selecionada.

Baixo

Lista variáveis que estão disponíveis para o baixo valor de preço do gráfico.

Abrir

Lista variáveis que estão disponíveis para o valor do preço de abertura do gráfico.

Fechar

Lista variáveis que estão disponíveis para o valor do preço de fechamento do gráfico.

da Linha de Base

Lista variáveis que estão disponíveis para as barras de volume do gráfico.

Ordem da categoria

Selecionar a ordem em que as categorias de variáveis são classificadas.

Como lido

As categorias de variáveis são apresentadas conforme aparecem no conjunto de dados.

Crescente

Classifique as categorias de variáveis em ordem crescente.

Decrescente

Classifique as categorias de variáveis em ordem decrescente.

Candlestick

Alternar os dados do gráfico para exibir como o candlestick ou linha.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Rótulo do Eixo x

O rótulo do eixo x.

Rótulo do Eixo y

O rótulo do eixo y.

Gráficos de embalagem de círculo

Os gráficos de embalagem de círculo exibem dados hierárquicos como um conjunto de áreas aninhadas para visualizar uma grande quantidade de dados estruturados hierarquicamente. É semelhante a um treemap, mas usa círculos em vez de retângulos. Os gráficos de embalagem de círculo utilizam a contenção (nesting) para exibir dados de hierarquia.

Como criar um gráfico de embalagem de círculo simples

1. Na seção **Tipo Gráfico**, clique no ícone **plot Time**.

Circle packing



As canvas atualizam para exibir um modelo de gráfico de embalagem de círculo.

2. Selecione uma variável **Colunas** a partir da lista suspensa.

Nota: Clique em **Adicionar outra coluna** para incluir mais variáveis de coluna.

Opções

Colunas

Lista variáveis que estão disponíveis para o gráfico.

Clique em **Adicionar outra coluna** para adicionar mais colunas.

Cor do grupo

Ligue ou desative os agrupamentos de cores.

Resumo

Selecione uma função de resumo estatístico (o método que é usado para resumir cada categoria).

Há dois tipos de funções de resumo estatístico disponíveis. A distinção é importante porque determina se você precisa especificar uma variável **Value**.

- **Funções que não requerem uma variável de valor.** Funções que não requerem uma variável. Todas as estatísticas de contagem e porcentagem estão nessa categoria. Essas estatísticas estão disponíveis quando a variável **Value** não está definida.
- **Funções que de fato exigem uma variável de valor.** Funções que de fato exigem uma variável **Value**. Por exemplo, a função " **Mean** " requer uma variável na qual a média é calculada. Essas estatísticas estão disponíveis quando a variável **Value** não está definida.

Valor

Esse campo é exibido quando uma função **de resumo** que requer uma variável de valor é selecionada. Selecione uma variável para servir como valor.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Gráficos Personalizados

A opção de gráficos personalizados oferece opções para colar ou editar o código JSON para criar o gráfico desejado.

Criando um gráfico personalizado

1. Na seção **Chart Type (Tipo de gráfico)**, clique no ícone **Customized (Personalizado)** Ícone 

2. Cole o código JSON que contém as especificações do gráfico no campo **script JSON** fornecido no painel **DetailsCUSTOMIZEDCHART**.

Gráficos de Dendrogram

Os gráficos do Dendrogram são semelhantes a gráficos de árvores e são tipicamente utilizados para ilustrar uma estrutura de rede (por exemplo, uma estrutura hierárquica). Os gráficos do Dendrogram consistem em um nó raiz que é conectado a nós subordinados através de bordas ou filiais. Os últimos nós da hierarquia são chamados de folhas.

Criando um gráfico de dendrograma simples

1. Na seção **Tipo Gráfico**, clique no ícone **Dendrograma**.



O canvas atualiza para exibir um modelo de gráfico Dendrogram.

2. Selecione pelo menos duas variáveis a partir da lista **Nome do campo**. Você pode selecionar **SELECT ALL** para selecionar todas as variáveis disponíveis. Recomenda-se uma quantidade máxima de número 300 variáveis.

Opções

Nome do campo

A lista fornece todas as variáveis disponíveis que representam os nós de licença. É necessário selecionar pelo menos duas variáveis.

Mostrar eixo

O controle de comutação possibilita ou desativa a exibição do eixo de valor da distância.

Ligação

O critério de ligação determina a distância entre conjuntos de observações como uma função das distâncias pairwise entre as observações. Selecione **Distância média**, **distância Minou A distância Max.**

Layout de árvore

Esquerda para a direita

O nó raiz exibe na esquerda e os nós de folhas são exibidos à direita.

Direita para a esquerda

O nó raiz exibe na direita e os nós de folhas são exibidos à esquerda.

Superior a inferior

O nó raiz exibe na parte superior e os nós de folhas são exibidos na parte inferior.

De baixo para cima

O nó raiz exibe na parte inferior e os nós de folhas são exibidos na parte superior.

Radial

O nó raiz exibe no meio e os nós de folhas irradiam da raiz.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Gráficos de eixos duplos Y

Um gráfico de Y-axes duplo resume ou trama duas variáveis de Y-axes que possuem domínios diferentes. Por exemplo, você pode traçar o número de casos em um eixo e o salário médio em outro. Este gráfico também pode ser uma mistura de diferentes elementos gráficos para que o gráfico dual Y-axes englobe vários dos diferentes tipos de gráficos. Os gráficos dual Y-axes podem exibir as contagens como uma linha e a média de cada categoria como um bar.

Criando um gráfico simples de Dual Y-axes

1. Na seção **Tipo Gráfico**, clique no ícone **Dual Y-axes**.



O canvas atualiza para exibir um modelo de gráfico Dual Y-axes.

2. Selecione uma variável como a variável **X-axis**.
3. Selecione variável para a primeira variável **Eixo Y** e, em seguida, selecione um tipo de gráfico para representar a variável (**Bar**, **Line** ou **Scatter plot**).
4. Selecione variável para a segunda variável **Eixo Y** e, em seguida, selecione um tipo de gráfico para representar a variável (**Bar**, **Line** ou **Scatter plot**).

Nota: Você pode usar os controles de seta para cima e para baixo para alterar a ordem Y-axes.

Opções

eixo X

Lista variáveis que estão disponíveis para o eixo X do gráfico.

eixo Y

Lista variáveis que estão disponíveis para o dual Y-axes do gráfico.

Resumo

Quando ativados, são exibidas opções para seleção do método que é utilizado para resumir cada categoria.

Resumo do eixo Y esquerdo

Configura o método sumário para o eixo Y que exibe no lado esquerdo do gráfico. As opções incluem **Sum**, **Mean**, **Máximo** e **mínimo**.

Resumo do eixo Y direito

Configura o método sumário para o eixo Y que exibe no lado direito do gráfico. As opções incluem **Sum**, **Mean**, **Máximo** e **mínimo**.

Normalizar dados

Quando ativados, os dados são transformados em uma distribuição normal, o que permite que dados de vários conjuntos de dados ou colunas sejam facilmente comparados.

Linhas do segundo eixo

Quando ativado, a segunda linha do eixo do gráfico é mostrada.

Ordenar novamente

Quando ativado, os dados do gráfico são reordenados com base nos valores do eixo X e Y.

Orientação da legenda

Configura a orientação da legenda do gráfico. As opções disponíveis são **Horizontal**, **Vertical** e **Vertical bottom**.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Rótulo do Eixo x

O rótulo do eixo x.

Rótulo do Eixo y

O rótulo do eixo y.

Gráficos de barras de erro

Os gráficos de barras de erro representam a variabilidade dos dados e indicam o erro (ou incerteza) em uma medição relatada. As barras de erro ajudam a determinar se as diferenças são estatisticamente significativas. As barras de erro também podem sugerir a bondade de ajuste para uma função específica.

Como criar um gráfico de barras de erro simples

1. Na seção **Tipo Gráfico**, clique no ícone **Barra de erro**.



O canvas atualiza para exibir um modelo de gráfico de barras de erro.

2. Selecione uma variável de escala como a variável **Category** (a variável cujos dados são representados no eixo X).
3. Selecione uma variável como a variável **Y-axis** (a variável cujos dados são representados no eixo Y).

Opções

Categoría

Lista as variáveis que estão disponíveis para o eixo X do gráfico.

eixo Y

Lista variáveis que estão disponíveis para o eixo Y do gráfico.

Ordem da categoría

Selecionar a ordem em que as categorias de variáveis são classificadas.

Como lido

As categorias de variáveis são apresentadas conforme aparecem no conjunto de dados.

Crescente

Classifique as categorias de variáveis em ordem crescente.

Decrescente

Classifique as categorias de variáveis em ordem decrescente.

Dividir por

Selecionar uma variável categórica que crie uma tabela de gráficos, com uma célula para cada categoria na variável Split by. Assim como o agrupamento, a divisão por variáveis essencialmente adiciona mais dimensões ao seu gráfico, exibindo informações para cada categoria de variável.

Linha de referência

Quando ativado, exibe uma linha de referência no gráfico. A linha de referência correlaciona-se com o selecionado **Método Estatístico**.

Barras de erros

Quando ativados, as linhas que representam a faixa de erro são exibidas no gráfico.

Medida

Selecionar o tipo de medida que é representado pelas barras de erro:

Intervalos de confiança

Configura os intervalos de confiança para as variáveis selecionadas. O valor padrão é 0.95 (95%), como refletido no campo **Valor Represente**.

Erro padrão

Mede o erro padrão das variáveis selecionadas.

Desvio padrão

Mede os desvios-padrão das variáveis selecionadas.

Nível de confiança

Este valor representa os intervalos de confiança para a **Medida** selecionada. O valor padrão é 0.95 (95%).

Método estatístico

Selecione o método para descrever a tendência central:

Média

O resultado de somar os rácios e dividir o resultado pelo número total de proporções.

Mediano

O valor tal esse número de proporções menor que este valor e o número de proporções maior que este valor são os mesmos.

Modo de exibição

Selecione como os displays de seleção **Método Estatístico** (barra, linha ou círculo).

Orientação da legenda

Configura a orientação da legenda do gráfico. As opções disponíveis são **Horizontal**, **Vertical** e **Vertical bottom**.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Rótulo do Eixo x

O rótulo do eixo x.

Rótulo do Eixo y

O rótulo do eixo y.

Gráficos de avaliação

Os gráficos de avaliação são semelhantes a histogramas ou gráficos de coleta. Os gráficos de avaliação mostram como os modelos precisos estão em prever resultados particulares. Eles trabalham sorteando registros com base no valor previsto e na confiança da predição, dividindo os registros em grupos de igual tamanho (quantiles) e, em seguida, traçando o valor do critério para cada quantil, do mais alto para o mais baixo. Diversos modelos são mostrados como linhas separadas no gráfico.

Os resultados são tratados definindo um valor ou faixa de valores específicos como um "hit". As ocorrências geralmente indicam sucesso de algum tipo (como uma venda para um cliente) ou um evento de interesse (como um diagnóstico médico específico).

Sinalização

Os campos de saída são diretos; os acertos correspondem a valores true .

Nominal

Para campos de saída nominais, o primeiro valor no conjunto define um acerto.

Contínuo

Para campos de saída contínua, atinge valores iguais maiores do que o ponto médio do intervalo do campo.

Os gráficos de avaliação também podem ser cumulativos para que cada ponto seja igual ao valor para o quantil correspondente mais todos os quantiles superiores. Os gráficos acumulativos geralmente conduzem melhor o desempenho geral dos modelos, ao passo que os gráficos não acumulativos geralmente se destacam ao indicar áreas de problema específicas para os modelos.

Criando um gráfico de Avaliação simples

1. Na seção **Tipo Gráfico** , clique no ícone **Avaliação** .

Evaluation



O canvas atualiza para exibir um modelo de gráfico de Avaliação.

2. Configure o **Campo de Destino**, **Campo de Prever** e as variáveis **Campo de Confiança**. O campo de destino pode ser qualquer sinalização ou campo nominal instanciado com dois ou mais valores. O campo prever define a variável que é usada como o valor previsto. O campo de confiança define a variável que é usada para estabelecer a confiança da predição.

Nota: O Tipo de variável **prever campo** deve combinar com o tipo de variável que é selecionado para o **Campo de destino**.

3. Especifique uma condição customizada usada para indicar o **Usuário definido pelo usuário**. Esta opção é útil para definir o resultado de interesse, ao invés de deduzi-lo a partir do tipo de campo de destino e da ordem de valores.

Você deve especificar uma expressão CLEM para uma condição de acerto. Por exemplo, @TARGET = "YES" é uma condição válida que indica um valor de Yes para o campo de destino é contado como um acerto na avaliação. A condição especificada é usada para todos os campos de destino.

Opções

Campo de resposta

Lista sinalização instanciada ou variáveis de campo nominais com dois ou mais valores.

Ocorrência definida pelo usuário

Especificar um valor de acerto. Os acertos indicam eventos de interesse (por exemplo, um diagnóstico médico específico).

Campo de previsão

Lista variáveis que podem ser usadas como o valor previsto.

Campo de confiança

Lista variáveis que podem estabelecer a confiança da predição.

Gráfico acumulado

Crie um gráfico acumulativo quando ativado. Os valores nos gráficos acumulativos são representados para cada quantil, mais todos os quantis maiores.

Modo de exibição

O controle de configurações que gráficos exibem no modo de visualização preview e na saída.

Modo único

Quando selecionado, o gráfico Ajuste de Classificação Modelo está no único gráfico que exibe no modo de visualização preview e na saída.

Modo clássico

Quando selecionados, o Ajuste de Classificação Modelo, Cutoff, Matriz Bar, ROC, Gains, ROI e gráficos do Profit exibem no modo de visualização e na saída.

Modo completo

Quando selecionados, os gráficos de Classificação de Classificação Modelo, Corte, Barra Matriz, ROC, Ganhos, ROI, Lucro, GINI, Lift e Resposta gráficos são exibidos no modo de visualização preview e na saída.

Gráficos de avaliação

Corte

O gráfico de corte mostra os valores previstos versus reais para variáveis selecionadas para um valor de corte especificado.

Barra de matriz

Os gráficos da Matriz Bar são uma boa maneira de determinar se as correlações lineares existem entre múltiplas variáveis.

ROC

ROC (Receiver Operating Característica) avalia o desempenho de esquemas de classificação em que os assuntos são classificados para uma variável com duas categorias.

Ganhos

Os ganhos são definidos como a proporção do total de ocorrências que ocorre em cada quantil. Os ganhos são computados como (number of hits in quantile / total number of hits) $\times 100\%$.

ROI

O ROI (retorno sobre investimento) é similar ao lucro no que concerne à definição de receitas e despesas. O ROI compara os lucros com os custos do quantil. O ROI é computado como (profits for quantile / costs for quantile) $\times 100\%$.

Lucro

O lucro é igual a receita para cada registro meno o custo do registro. Os lucros para um quantil são a soma dos lucros para todos os registros no quantil. Assume-se que as receitas se aplicam apenas a ocorrências, mas os custos se aplicam a todos os registros. Os lucros e os custos podem ser corrigidos ou podem ser definidos por campos nos dados. Os lucros são calculados como (soma de receitas dos registros no quantil - soma dos custos dos registros no quantil).

Kolmogorov-Smirnov

Compara a função de distribuição cumulativa observada para uma variável com distribuição teórica especificada, que pode ser normal, uniforme, exponencial ou Poisson.

GINI

A GINI mede a dispersão estatística e tem como objetivo representar a distribuição de renda ou riqueza. É a medição mais comumente usada da desigualdade.

Elevar

A elevação compara a porcentagem de registros em cada quantil que forem ocorrências com a porcentagem geral de ocorrências nos dados de treinamento. Ele é computado como (hits in quantile / records in quantile) / (total hits / total records).

Resposta

Resposta é a porcentagem de registros no quantil que são acertos. Resposta é computada como (hits in quantile / records in quantile) $\times 100\%$.

Configurações do gráfico de avaliação

As configurações a seguir aplicam-se apenas a gráficos de lucro e ROI.

Custos

Especificar o custo fixo associado a cada registro.

Receita

Especificar a receita fixa associada a cada registro que representa um acerto.

Ponderação

Se os registros em seus dados representarem mais de uma unidade, será possível usar ponderações de frequência para ajustar os resultados. Especificar o peso fixo associado a cada registro.

Gráficos de mapas de utilização

Gráficos de mapas de calor apresentam dados em que os valores individuais que estão contidos em uma matriz são representados como cores.

Criando um gráfico de mapa de calor simples

1. Na seção **Tipo Gráfico**, clique no ícone **Heat map**.



O canvas atualiza para exibir um modelo de gráfico de mapas de calor.

2. Selecione uma variável como a variável **Coluna**. Cada categoria de variável é representada como uma coluna de gráfico individual.
3. Selecione uma variável como a variável **Row**. Cada categoria de variável é representada como uma linha de gráfico individual.

Opções

Coluna

Lista variáveis que estão disponíveis para as colunas do gráfico. Cada categoria de variável é representada como uma coluna de gráfico individual.

Linha

Lista variáveis que estão disponíveis para as linhas do gráfico. Cada categoria de variável é representada como uma linha de gráfico individual.

Ordem da categoria

Selecione a ordem em que as categorias de variáveis são classificadas.

Como lido

As categorias de variáveis são apresentadas conforme aparecem no conjunto de dados.

Crescente

Classifique as categorias de variáveis em ordem crescente.

Decrescente

Classifique as categorias de variáveis em ordem decrescente.

Resumo

Selecione uma função de resumo estatístico para o elemento gráfico. O resultado da estatística determina a posição dos elementos gráficos no eixo Y. Em um gráfico 2-D, a estatística é calculada para cada valor no eixo X. Em um gráfico 3D, ele é calculado para a interseção dos valores no eixo X e no eixo Z.

Há dois tipos de funções de resumo estatístico disponíveis. A distinção é importante porque determina se você precisa especificar uma variável **Value**.

- **Funções que não requerem uma variável de valor.** Funções que não requerem uma variável. Todas as estatísticas de contagem e porcentagem estão nessa categoria. Essas estatísticas estão disponíveis quando a variável **Value** não está definida.
- **Funções que de fato exigem uma variável de valor.** Funções que de fato exigem uma variável **Value**. Por exemplo, a função " **Mean** " requer uma variável na qual a média é calculada. Essas estatísticas estão disponíveis quando a variável **Value** não está definida.

Valor

Esse campo é exibido quando uma função **de resumo** que requer uma variável de valor é selecionada. Selecione uma variável para servir como valor.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Rótulo do Eixo x

O rótulo do eixo x.

Rótulo do Eixo y

O rótulo do eixo y.

Gráficos de histograma

Um histograma é semelhante na aparência a um gráfico de barras, mas em vez de comparar categorias ou procurar tendências ao longo do tempo, cada barra representa como os dados são distribuídos em uma única categoria. Cada barra representa uma gama contínua de dados ou o número de frequências para um ponto de dados específico.

Os histogramas são úteis para mostrar a distribuição de uma variável de escala única. Os dados são vinculados e resumidos usando uma estatística de contagem ou porcentagem. Uma variação de um

histograma é um polígono de frequência, que é como um histograma típico, exceto que o elemento gráfico de área é usado em vez do elemento gráfico de barras.

Outra variação do histograma é a pirâmide populacional. Seu nome é derivado de seu uso mais comum: resumir dados populacionais. Quando usado com dados populacionais, ele é dividido por gênero para fornecer dois histogramas de costas horizontais de dados de idade. Em países com uma população jovem, a forma do gráfico resultante se assemelha a uma pirâmide.

Criando um gráfico de histograma

1. Na seção **Tipo Gráfico**, clique no ícone **Histograma**.



O canvas atualiza para exibir um modelo de gráfico de histograma.

2. Selecione uma variável de escala como a variável **X-axis**.

Nota: A estatística para um histograma é Histograma ou Histograma Percent. Essas estatísticas bin os dados e calculam uma contagem para cada bin.

Opções

eixo X

Lista variáveis que estão disponíveis para o eixo X do gráfico.

Dividir por

Seleciona uma variável categórica que crie uma tabela de gráficos, com uma célula para cada categoria na variável Split by. Assim como o agrupamento, a divisão por variáveis essencialmente adiciona mais dimensões ao seu gráfico, exibindo informações para cada categoria de variável.

Método de categorizar

Especifique um método bin que é usado para criar as barras de gráfico. A opção disponível inclui **Auto bin**, **By bin width** e **By bin num**.

Mostrar curva kde

Quando ativado, a curva de estimativa de densidade do kernel é mostrada no gráfico.

Mostrar curva de distribuição

Quando ativada, a curva de ajuste de distribuição é mostrada no gráfico.

Distribuição

A lista suspensa fornece para as seguintes opções de distribuição.

Distribuição de ajuste automático

Automaticamente se encaixa na distribuição (a configuração padrão).

Beta

Retorna o valor a partir de uma distribuição Beta com parâmetros de forma especificados.

Exponencial

Retorna o valor a partir de uma distribuição exponencial.

Gama

Retorna o valor a partir da distribuição Gamma, com os parâmetros de forma e escala especificados.

Log-normal

Retorna o valor a partir de um log-distribuição normal com parâmetros especificados.

Normal

Retorna o valor a partir de uma distribuição normal com média e desvio padrão especificado.

Triangular

Retorna o valor a partir de uma distribuição triangular com parâmetros especificados.

Uniforme

Retorna o valor a partir da distribuição uniforme entre o mínimo e o máximo.

Weibull

Retorna o valor a partir de uma distribuição Weibull com parâmetros especificados.

Largura da categoria

O slider controla o tamanho do intervalo que é usado para dividir os dados em grupos.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Rótulo do Eixo x

O rótulo do eixo x.

Rótulo do Eixo y

O rótulo do eixo y.

Gráficos de linhas

Um gráfico de linha trama uma série de pontos de dados em um gráfico e os conecta com linhas. Um gráfico de linha é útil para mostrar linhas de tendência com diferenças sutis, ou com linhas de dados que cruzam uma outra. Você pode usar um gráfico de linha para resumir variáveis categóricas, nesse caso ele é semelhante a um gráfico de barras (ver “[Gráficos de barras](#)” na página 65). Os gráficos de linha também são úteis para dados de séries temporais.

Como criar um gráfico de linha de tempo simples

1. Na seção **Tipo Gráfico**, clique no ícone **Linha**.



O canvas atualiza para exibir um modelo de gráfico de linha.

2. Selecione uma variável de data como a variável **X-axis**.
3. Selecione uma variável de escala como a variável **Y-axis** (a variável cujos valores foram registrados ao longo do tempo).

Opções

eixo X

Lista variáveis que estão disponíveis para o eixo X do gráfico.

eixo Y

Lista variáveis que estão disponíveis para o eixo Y do gráfico.

Dividir por

Selecione uma variável categórica que crie uma tabela de gráficos, com uma célula para cada categoria na variável Split by. Assim como o agrupamento, a divisão por variáveis essencialmente adiciona mais dimensões ao seu gráfico, exibindo informações para cada categoria de variável.

Área

Quando ativada, a área abaixo da linha é mostrada em uma cor diferente.

Suave

Quando ativado, o gráfico mostra uma curva suave.

Mostrar pontos de dados

Quando ativado, o ponto de dados é mostrado no gráfico.

Ordenar novamente

O toggle controle reordena dados com base em valores X e Y-axis.

Linha de ajuste

Em uma linha de encaixe, os pontos de dados são encaixados em uma linha que geralmente não passa por todos os pontos de dados. A linha fit representa a tendência dos dados. Algumas linhas de encaixe são baseadas em regressão. Outros são baseados em quadrados iterativos menos ponderados. Selecione uma opção de linha fit na lista suspensa.

Mostrar linha de referência

Quando ativada, a opção mostra uma linha de referência no gráfico que é baseado nos valores especificados **xAxis** e **yAxis**.

Insira um valor de linha de referência no eixo x

Quando a **Mostrar linha de referência** está ativada, esta configuração fornece a opção de especificar um valor de linha de referência específico para o eixo X. Clique em **Adicionar outra coluna** para especificar mais valores de linha de referência.

Insira um valor de linha de referência no eixo y

Quando a **Mostrar linha de referência** está ativada, esta configuração fornece a opção de especificar um valor de linha de referência específico para o eixo Y. Clique em **Adicionar outra coluna** para especificar mais valores de linha de referência.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Rótulo do Eixo x

O rótulo do eixo x.

Rótulo do Eixo y

O rótulo do eixo y.

Gráficos de mapas

Os gráficos de mapas são comumente usados para comparar valores e mostrar categorias através de regiões geográficas. Os gráficos de mapas são mais benéficos quando os dados contêm informações geográficas (países, regiões, estados, condados, códigos postais, entre outros).

Criando um gráfico de mapas simples

1. Na seção **Tipo Gráfico**, clique no ícone **Mapa**.



O canvas atualiza para exibir um modelo de gráfico de mapas.

2. Selecione um serviço a ser usado para servir imagens do mapa a partir da lista suspensa **Mapa do serviço**. A lista fornece opções que abrangem regiões globais específicas.
3. Selecione um gráfico de mapas **Tipo** a partir do menu suspenso. As opções a seguir estão disponíveis, dependentes do tipo de gráfico selecionado:

Longitude

Selecione uma variável para servir como valor longitudinal a partir da lista suspensa.

Latitude

Selecione uma variável para servir como valor latitudinal a partir da lista suspensa.

Grupo

Selecione uma variável que agrupe os locais de ponto de dados do menu suspenso.

Categoria

Selecione uma variável de coluna que você deseja visualizar.

Resumo

Selecione uma função de resumo estatístico (o método que é usado para resumir cada categoria).

Há dois tipos de funções de resumo estatístico disponíveis. A distinção é importante porque determina se você precisa especificar uma variável **Value**.

- **Funções que não requerem uma variável de valor.** Funções que não requerem uma variável. Todas as estatísticas de contagem e porcentagem estão nessa categoria. Essas estatísticas estão disponíveis quando a variável **Value** não está definida.
- **Funções que de fato exigem uma variável de valor.** Funções que de fato exigem uma variável **Value**. Por exemplo, a função " **Mean** requer uma variável na qual a média é calculada. Essas estatísticas estão disponíveis quando a variável **Value** não está definida.

Valor

Esse campo é exibido quando uma função **de resumo** que requer uma variável de valor é selecionada. Selecione uma variável para servir como valor.

Opções

Serviço de mapa

Lista os serviços que estão disponíveis para fornecer imagens de mapa. Veja "["Opções de serviço de mapas"](#) na página 83.

Tipo

Lista os tipos de gráfico que estão disponíveis para representar os dados.

Longitude

Lista as variáveis que estão disponíveis para servir como valor longitudinal.

Latitude

Lista as variáveis que estão disponíveis para servir como valor latitudinal.

Grupo

Lista as variáveis que podem ser usadas para agrupar os locais do ponto de dados.

Categoria

Lista variáveis de coluna.

Resumo

Selecione uma função de resumo estatístico (o método que é usado para resumir cada categoria).

Há dois tipos de funções de resumo estatístico disponíveis. A distinção é importante porque determina se você precisa especificar uma variável **Value**.

- **Funções que não requerem uma variável de valor.** Funções que não requerem uma variável. Todas as estatísticas de contagem e porcentagem estão nessa categoria. Essas estatísticas estão disponíveis quando a variável **Value** não está definida.
- **Funções que de fato exigem uma variável de valor.** Funções que de fato exigem uma variável **Value**. Por exemplo, a função " **Mean** requer uma variável na qual a média é calculada. Essas estatísticas estão disponíveis quando a variável **Value** não está definida.

Valor

Esse campo é exibido quando uma função **de resumo** que requer uma variável de valor é selecionada. Selecione uma variável para servir como valor.

Informações da dica de ferramenta

Lista as variáveis que podem ser usadas as informações da dica de ferramentas de geração quando o cursor pairar sobre um ponto de dados.

Dimensionar mapa

Lista variáveis de mapa de tamanho disponíveis. Essas variáveis usam tamanhos diferenciados para representar a si mesmos nos pontos da trama.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Opções de serviço de mapas

Na área de janela Detalhes para gráficos de mapas, é possível selecionar o serviço de mapas a ser usado para fornecer imagens do mapa.

Configurando opções de serviço de mapas

1. Pare o ServidorIBM SPSS Modeler.
2. Se você for aplicar arquivos de mapas formatados do geojson local, coloque-os no diretório <Modeler_installation_directory>/dataview/conf/public/mapfiles. Se o diretório não existir, crie-o.
3. Abra o arquivo <Modeler_installation_directory>/dataview/conf/application.conf em um editor de texto e inclua as opções a seguir na seção map.resources. Neste exemplo, são adicionados dois novos mapas denominados World - Local Map e World - Online Map .

```
map.resources {
  "mapservices" :
  [
    {
      "type"      : "geojson",
      "name"      : "World (Bulit-in)",
      "location"  : "built_in_world",
      "id"        : "built_in_world"
    },
    {
      "type"      : "geojson",
      "name"      : "Blank (Built-in)",
      "location"  : "built_in_blank",
      "id"        : "built_in_blank"
    },
    {
      "type"      : "geojson",
      "name"      : "World - Local Map",
      "location"  : "/mapfiles/world.json",
      "useProxy"  : "false",
      "id"        : "world_local"
    },
    {
      "type"      : "geojson",
      "name"      : "World - Online Map",
      "location"  : "http://map.example.com/map/world.json",
      "useProxy"  : "true",
      "id"        : "world_online"
    }
  ]
}
```

Os seguintes parâmetros estão disponíveis para as opções do mapa:

Tabela 4. Parâmetros de serviço de mapas

Opção	Valor	Tipo	Obrigatório?	Descrição
type	geojson	Sequência	True	Atualmente, apenas o valor geojson é suportado para esta opção.

Tabela 4. Parâmetros de serviço de mapas (continuação)

Opção	Valor	Tipo	Obrigatório?	Descrição
name	Nome do mapa	Sequência	True	Insira um nome para o mapa. Este nome é mostrado no suspenso Mapa do serviço na área de janela Detalhes para gráficos de mapas.
location	Localização do mapa	Sequência	True	Digite o endereço URL ou o local do arquivo local do mapa. Para URL, ele deve começar com <code>http://</code> ou <code>https://</code> . Para arquivo local, ele deve começar com o caminho <code>/mapfiles/</code> .
useProxy	true ou false	Sequência	True	Use <code>true</code> se o local do mapa for um URL. Use <code>false</code> se o local do mapa for um arquivo local.
id	Mapear ID	Sequência	True	Use um ID exclusivo para cada mapa.

Nota:

- Não remova as opções padrão **World (Built-in)** e **Blank (Built-in)**.
- Lembre-se de usar uma vírgula entre os blocos de conteúdo como mostrado no exemplo, e tenha certeza de que sua nova seção `map.resources` é um conteúdo formatado JSON válido.
- O ServidorIBM SPSS Modeler (ou o standalone ModeladorSPSS Client sem Servidor) deve ser capaz de acessar o local do mapa.

4. Inicie o ServidorIBM SPSS Modeler.

Agora enquanto trabalha com gráficos de mapas, na pane **Detalhes**, você verá o novo serviço de mapas adicionado.

Gráficos de curva de matemática

Um gráfico de curva de matemática trama curvas de equação matemática que são baseadas em expressões inseridas pelo usuário.

Criando um gráfico de curva de matemática simples

1. Na seção **Tipo Gráfico**, clique no ícone **Curva de matemática**.

Math curve



O canvas atualiza para exibir um modelo de gráfico de curva de matemática.

2. Insira um valor inicial para o eixo X no campo **X valor inicia a partir de**.
3. Insira um valor final para o eixo X no campo **X value ends com**.
4. Entre em uma equação que plota a curva de gráfico no campo **equações**. Clique em **Adicionar outra coluna** para incluir mais equações.

Cada equação trata x como uma variável independente. As equações a seguir são permitidas.

- +, -, *, /, %, and ^
- abs(x)
- ceil(x)
- floor(x)
- log(x)
- max(a,b,c...)
- min(a,b,c...)
- random()
- round(x)
- sqrt(x)
- sin
- cos
- exp
- tan
- atan
- atan2
- asin
- acos

Opções

O valor X começa em

O valor inicial do eixo X.

O valor X termina em

O valor final do eixo X.

equações

Equações inseridas pelo usuário que traçam a curva do gráfico.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Rótulo do Eixo x

O rótulo do eixo x.

Rótulo do Eixo y

O rótulo do eixo y.

Gráficos de multigráficos

Os gráficos de multigráficos fornecem opções para criação de vários gráficos. Os gráficos podem ser dos mesmos ou diferentes tipos, e podem incluir diferentes variáveis a partir do mesmo conjunto de dados.

Criando um gráfico de múltiplos gráficos simples

1. Na seção **Tipo Gráfico**, clique no ícone **Multi-gráfico**.

Multi-chart



O canvas atualiza para exibir um modelo de gráfico de vários gráficos.

2. Clique em **Adicionar outro sub-gráfico** para adicionar um subgráfico.
3. Selecione um tipo de gráfico a partir da lista suspensa **Tipo**.
4. Dependendo do tipo de gráfico selecionado, as opções a seguir estão disponíveis:

Gráficos de Barra e Pie

Selecione uma variável **Category** a partir da lista suspensa.

Gráficos de gráficos de linha e Scatter

Selecione uma variável **X-axis** da lista suspensa.

Selecione uma variável **eixo Y** a partir da lista suspensa.

5. Clique em **Adicionar outro sub-gráfico** para incluir mais tipos de gráficos.

Opções

Título

O título do subgráfico.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Vários gráficos de séries

Vários gráficos de séries são semelhantes a gráficos de linha, com exceção de que é possível traçar múltiplas variáveis no eixo Y.

Como criar um gráfico de múltiplas séries simples

1. Na seção **Tipo Gráfico**, clique no ícone **Multi-series**.

Multi-series



O canvas atualiza para exibir um modelo de gráfico de várias séries.

2. Selecione uma variável como a variável **X-axis**.
3. Selecione pelo menos duas variáveis de escala como as variáveis **Eixo Y**.

Opções

eixo X

Lista variáveis que estão disponíveis para o eixo X do gráfico.

eixo Y

Lista variáveis que estão disponíveis para os Y-axes do gráfico.

Selecione o tipo de gráfico (barra, linha ou dispersão) da lista suspensa.

Clique em **Adicionar outra coluna** para incluir mais colunas no gráfico.

Estilo da série

Fornece opções para definir a orientação de Y-axes. As opções a seguir estão disponíveis.

- Padrão
- Separar eixos Y
- Mostrar um eixo Y secundário
- Eixos Y duplos

Eixo Y secundário

Lista variáveis que estão disponíveis para os Y-axes secundários do gráfico.

Selecione o tipo de gráfico (barra, linha ou dispersão) da lista suspensa.

Clique em **Adicionar outra coluna** para incluir mais colunas no gráfico.

Normalizar dados

Quando ativado, esta configuração transforma dados em uma distribuição normal comparando dados de vários conjuntos de dados ou múltiplas colunas. Esta configuração cria 100% de empilhamento para contagens e converte estatísticas em percents.

Ordenar novamente

Quando ativado, os dados do gráfico são reordenados com base nos valores do eixo X e Y.

Mostrar rótulo

Quando ativado, etiquetas de coluna são exibidas no gráfico. Apenas dados de séries de dispersão são suportados.

Campo de rótulo

O menu de campo fornece variáveis para exibir como etiquetas de gráfico.

Orientação da legenda

Configura a orientação da legenda do gráfico. As opções disponíveis são **Horizontal**, **Vertical** e **Vertical bottom**.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Rótulo do Eixo x

O rótulo do eixo x.

Rótulo do Eixo y

O rótulo do eixo y.

Gráficos paralelos

Os gráficos paralelos são úteis para visualização da geometria dimensional alta e para análise de dados multivariados. Os gráficos paralelos se assemelham a gráficos de linha para dados de séries temporais, mas os eixos não correspondem a pontos no tempo (uma ordem natural não está presente).

Criando um gráfico paralelo simples

1. Na seção **Tipo Gráfico**, clique no ícone **Parallel**.



O canvas atualiza para exibir um modelo de gráfico paralelo.

2. Selecione pelo menos duas variáveis como as variáveis **Columns**. Cada coluna representa um eixo vertical, paralelo, no gráfico.

Nota: A ordem de coluna é importante para encontrar recursos. Em uma análise de dados típica, você pode precisar reordenar as colunas inúmeras vezes.

Opções

Colunas

Lista variáveis que estão disponíveis para os Y-axes do gráfico.

Clique em **Adicionar outra coluna** para adicionar mais colunas.

Mapa de cores

Lista variáveis de mapa de cores disponíveis. Essas variáveis usam progressão de cor, com base na faixa de valores na coluna especificada, para representarem a si mesmos nos pontos da trama. Os mapas de cores também são conhecidos como mapas de choropleth.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Gráficos de Pareto

Os gráficos de pareto contêm ambos os barras e um gráfico de linha. As barras representam categorias variáveis individuais e o gráfico de linhas representa o total acumulado.

Criando um gráfico de Pareto simples

1. Na seção **Tipo Gráfico**, clique no ícone **Pareto**.



O canvas atualiza para exibir um modelo de gráfico de Pareto.

2. Selecione uma variável como a variável **Category**. As categorias de variáveis selecionadas são desenhadas no eixo X do gráfico.

Opções

Categoria

Lista variáveis que estão disponíveis para o eixo X do gráfico.

Alinhamento do eixo YA

O controle de comutação possibilita e desativa o alinhamento dos dois eixos Y no gráfico.

Barra de gradiente

O controle de comutação ativa e desativa a exibição de gradientes de cor nas barras de gráfico.

Destacar poucos vitais

O controle de comutação possibilita e desativa o destaque de categorias variáveis consideradas vitais.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Gráficos de setores circulares

Um gráfico de pizza é útil para comparar proporções. Por exemplo, você pode usar um gráfico de pizza para demonstrar que uma proporção maior de europeus é cadastrada em uma determinada classe.

Criando um gráfico de pizza simples

1. Na seção **Tipo Gráfico**, clique no ícone **Pie**.



O canvas atualiza para exibir um modelo de gráfico de pizza.

2. Selecione uma variável categórica (nominal ou ordinal) na lista **Categoria**. As categorias nesta variável determinam o número de fatias no gráfico de pizza.
3. Selecione uma função de resumo estatístico para o elemento gráfico. Para gráficos de tortas, você tipicamente quer uma estatística baseada em contagem ou uma soma. O resultado da estatística determina o tamanho de cada fatia.

Opções

Categoria

Selecione uma variável categórica (nominal ou ordinal) que determine o número de fatias no gráfico de pizza.

Resumo

Selecione uma função de resumo estatístico para o elemento gráfico. Para gráficos de tortas, você tipicamente quer uma estatística baseada em contagem ou uma soma. O resultado da estatística determina o tamanho de cada fatia.

Há dois tipos de funções de resumo estatístico disponíveis. A distinção é importante porque determina se você precisa especificar uma variável **Value**.

- **Funções que não requerem uma variável de valor.** Funções que não requerem uma variável. Todas as estatísticas de contagem e porcentagem estão nessa categoria. Essas estatísticas estão disponíveis quando a variável **Value** não está definida.
- **Funções que de fato exigem uma variável de valor.** Funções que de fato exigem uma variável **Value**. Por exemplo, a função "Mean" requer uma variável na qual a média é calculada. Essas estatísticas estão disponíveis quando a variável **Value** não está definida.

Valor

Este campo exibe quando uma função **Sumário** que requer uma variável de escala, é selecionada. Selecione uma variável para servir como variável de escala.

Tipo de gráfico de pizza

Os estilos a seguir estão disponíveis.

Normal

Os segmentos de tortas são exibidos como fatias normais.

Anel

Os segmentos de tortas são exibidos como um anel. Este estilo também é conhecido como um gráfico de rosca.

Rosa

Ao contrário do gráfico de pizza normal, que usa um raio comum, os tamanhos do segmento de tortas variam de acordo com o seu valor.

Área rosa

Ao contrário do gráfico de pizza normal, que usa um raio comum, os tamanhos do segmento de tortas variam dependendo de sua área.

Anel rosa

Ao contrário do gráfico de pizza normal, que usa um raio comum, os tamanhos do segmento de tortas variam dependendo de seu valor e os segmentos são exibidos como um anel.

Metade rosa

Mesma que a Rose, exceto o gráfico é representado como meia de uma torta.

Mostrar valor

Quando ativado, os valores de fatia de pizza são exibidos na legenda.

Campo externo

A lista fornece variáveis que podem ser usadas como valor de contagem do anel externo.

Ângulo de etiqueta externo

O valor espécie a localização do valor variável de campo externo no anel externo.

Orientação da legenda

Configura a orientação da legenda do gráfico. As opções disponíveis são **Horizontal**, **Vertical** e **Vertical bottom**.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Gráficos de pirâmide populacional

Os gráficos de pirâmide populacional (também conhecidos como "pirâmides de idade-sexo") são comumente usados para apresentar e analisar informações populacionais baseadas na idade e no gênero.

Criando um gráfico de pirâmide de População simples

1. Na seção **Tipo Gráfico**, clique no ícone **pirâmide de População**.

Population ...



As atualizações do canvas atualizam um modelo de gráfico de pirâmide de População.

2. Selecione uma variável **eixo Y** a partir da lista suspensa.

3. Selecione uma variável **Split by** a partir da lista suspensa.

Opções

eixo Y

Lista variáveis que estão disponíveis para o eixo Y do gráfico.

Dividir por

Seleciona uma variável categórica que crie uma tabela de gráficos, com uma célula para cada categoria na variável Split by. Assim como o agrupamento, a divisão por variáveis essencialmente adiciona mais dimensões ao seu gráfico, exibindo informações para cada categoria de variável.

Largura da categoria

O slider controla o tamanho do intervalo que é usado para dividir os dados em grupos.

Mostrar curva de distribuição

Quando ativada, a curva de ajuste de distribuição é mostrada no gráfico.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Gráficos Q-Q

As parcelas de Q-Q (quantile-quantile) comparam duas distribuições de probabilidade, trafegando seus quantiles uns contra os outros. Uma trama Q-Q é usada para comparar as formas de distribuições, fornecendo uma visão gráfica de como propriedades como localização, escala e defasagem são semelhantes ou diferentes nas duas distribuições.

Criando um gráfico de plot Q-Q simples

1. Na seção **Tipo Gráfico**, clique no ícone **Q-Q plot**.



O canvas atualiza para exibir um modelo de gráfico de plot Q-Q.

2. Selecione uma variável como a variável **X-axis**.

Opções

eixo X

Lista variáveis que estão disponíveis para o eixo X do gráfico.

Distribuição

A lista suspensa fornece todos os métodos de distribuição disponíveis.

Distribuição de ajuste automático

Automaticamente se encaixa na distribuição (a configuração padrão).

Beta

Retorna o valor a partir de uma distribuição Beta com parâmetros de forma especificados.

Exponencial

Retorna o valor a partir de uma distribuição exponencial.

Gama

Retorna o valor a partir da distribuição Gamma, com os parâmetros de forma e escala especificados.

Log-normal

Retorna o valor a partir de um log-distribuição normal com parâmetros especificados.

Normal

Retorna o valor a partir de uma distribuição normal com média e desvio padrão especificado.

Uniforme

Retorna o valor a partir da distribuição uniforme entre o mínimo e o máximo.

T de Student

Retorna o valor da distribuição t do Student, com graus de liberdade especificados.

Tipo de gráfico

Selecione uma trama Q-Q (quantile-quantile) ou uma trama P-P (percent-percentual).

Ajuste automático

Quando ativados, os parâmetros de dados para os **Distribuição** selecionados são automaticamente estimados. Quando não selecionados, os valores de distribuição **Shape** e **Scale** são exibidos.

Shape1

Configura o valor shape1 para a distribuição **Beta**. Esta configuração está disponível apenas quando **Auto fit** não é ativado e **Beta** é selecionado como a **Distribuição**.

Shape2

Configura o valor shape2 para a distribuição **Beta**. Esta configuração está disponível apenas quando **Auto fit** não é ativado e **Beta** é selecionado como a **Distribuição**.

Forma

Configura o valor de forma para a distribuição selecionada. Esta configuração está disponível apenas quando **Auto fit** não estiver habilitado e **Exponential**, **Gamma**, ou **Log-normal** é selecionado como a **Distribuição**.

Escala

Configura o valor de escala para a distribuição selecionada. Esta configuração está disponível apenas quando **Auto fit** não estiver habilitado e **Exponential**, **Gamma**, ou **Log-normal** é selecionado como a **Distribuição**.

Média

Configura o valor de significa para a distribuição **Normal**. Esta configuração está disponível apenas quando **Auto fit** não é ativado e **Normal** é selecionado como a **Distribuição**.

Erro dev.

Configura o valor de desvio padrão para a distribuição **Normal**. Esta configuração está disponível apenas quando **Normal** é selecionado como a **Distribuição**.

Mín.

Configura o valor mínimo para a distribuição **Uniforme**. Esta configuração está disponível apenas quando **Auto fit** não é ativado e **Uniform** é selecionado como a **Distribuição**.

Máx.

Configura o valor máximo para a distribuição **Uniforme**. Esta configuração está disponível apenas quando **Auto fit** não é ativado e **Uniform** é selecionado como a **Distribuição**.

Graus de liberdade (df)

Defina os graus de valor da liberdade para a distribuição **Student t**. Esta configuração está disponível apenas quando **Auto fit** não é ativado e **Student t** é selecionado como a **Distribuição**.

Não seguir a tendência

Quando ativado, exibição de tramas detrendidas no gráfico.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Rótulo do Eixo x

O rótulo do eixo x.

Rótulo do Eixo y

O rótulo do eixo y.

Gráfico de radar

Os gráficos de radar comparam múltiplas variáveis quantitativas e são úteis para visualizar quais variáveis possuem valores semelhantes, ou se outliers existem entre as variáveis. Os gráficos de radar consistem em uma sequência de spokes, com cada um falou representando uma única variável. Os Gráficos de Radar também são úteis para determinar quais variáveis estão pontuando alto ou baixo dentro de um conjunto de dados.

Criando um gráfico de radar simples

1. Na seção **Tipo Gráfico**, clique no ícone **Radar**.

Radar



As atualizações de tela atualizam para exibir um modelo de gráfico de radar.

2. Selecione uma variável **Colunas** a partir da lista suspensa.

Nota: Clique em **Adicionar outra coluna** para incluir mais colunas. Pelo menos três variáveis de colunas devem ser definidas.

Opções

Categoria

Selecione uma variável categórica (nominal ou ordinal). Se você selecionar **Nenhum** como os valores da categoria, todos os valores serão mostrados separadamente, e nenhum método sumário será aplicado.

Resumo

Quando uma variável categórica é selecionada, as estatísticas de resumo a seguir estão disponíveis.

Contagem

Número total de casos.

Soma

Soma dos valores.

Média

Média aritmética; a soma dividida pelo número de casos.

Máximo

Maior (mais alto) valor.

Mínimo

Menor (mais baixo) valor.

Layout do radar

Determina o layout de imagem de fundo para o gráfico de radar:

Círculo

Quando selecionado, o gráfico de radar é desenhado sobre um layout circular.

Polígono

Quando selecionado, o gráfico de radar é desenhado sobre um layout poligonal.

Dividir por

Seleciona uma variável categórica que crie uma tabela de gráficos, com uma célula para cada categoria na variável Split by. Assim como o agrupamento, a divisão por variáveis essencialmente adiciona mais dimensões ao seu gráfico, exibindo informações para cada categoria de variável.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Gráficos de relacionamento

Um gráfico de relacionamento é útil para determinar como as variáveis se relacionam umas com as outras.

Criando um gráfico de relacionamento simples

1. Na seção **Tipo Gráfico**, clique no ícone **Relacionamento**.



O canvas atualiza para exibir um modelo de gráfico relacional.

2. Selecione pelo menos duas variáveis como variáveis **Columns**.

Nota: Clique em **Adicionar outra coluna** para incluir mais variáveis de coluna.

Opções

Colunas

Lista as variáveis disponíveis.

Clique em **Adicionar outra coluna** para adicionar mais colunas.

Estilo da linha

Controla o estilo de linha entre os pontos de dados relacionados.

Curva

Quando selecionadas, linhas curvas são desenhadas entre os pontos de dados relacionados.

Linear

Quando selecionados, linhas retas são desenhadas entre os pontos de dados relacionados.

Limite de rótulo

Exibe etiquetas para pontos de dados cujos valores excedem o valor definido.

Orientação da legenda

Configura a orientação da legenda do gráfico. As opções disponíveis são **Horizontal**, **Vertical** e **Vertical bottom**.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Tramas de dispersão e tramas de pontos

Várias categorias amplas de gráficos são criadas com o elemento gráfico de ponto.

gráficos de dispersão

Os gráficos de dispersão são úteis para plotar dados multivariados. Eles podem ajudá-lo a determinar relacionamentos potenciais entre as variáveis de escala. Uma trama de dispersão simples usa um sistema de coordenadas 2-D para traçar duas variáveis. Uma trama de dispersão 3-D usa um sistema de coordenadas 3-D para traçar três variáveis. Quando você precisa traçar mais variáveis, você pode tentar overlay scatter plots e scatter plot matrices (SPLOMs). Uma trama de dispersão de sobreposição exibe pares sobrepostos de variáveis X-Y, com cada par distinguido por cor ou forma. Um SPLOM cria uma matriz de tramas de dispersão de 2-D, com cada variável plotada contra cada outra variável no SPLOM.

Parcelas de ponto

Como os histogramas, as parcelas de ponto são úteis para mostrar a distribuição de uma variável de escala única. Os dados são binarizados, mas, em vez de um valor para cada bin (como uma contagem), todos os pontos de cada bin são exibidos e empilhados. Esses gráficos às vezes são chamados de parcelas de densidade.

Parcelas do ponto de resumo

As parcelas do ponto sumário são semelhantes a gráficos de barras, exceto que os pontos são desenhados no lugar da parte superior das barras. Para obter mais informações, consulte “[Gráficos de barras](#)” na página 65.

Gráficos de linhas de queda

Os gráficos de linhas de largada são um tipo especial de enredo de ponto de resumo. Os pontos são agrupados e uma linha é desenhada através dos pontos em cada categoria. O gráfico de queda de linha é útil para comparar uma estatística através de variáveis categóricas.

Criando uma trama de dispersão simples

1. Na seção **Tipo Gráfico**, clique no ícone **Scatter plot**.



O canvas atualiza para exibir um modelo de gráfico de trama de dispersão.

2. Selecione uma variável de escala como a variável **X-axis**.
3. Selecione uma variável de escala como a variável **Y-axis**. Você não precisa especificar uma estatística porque as parcelas de dispersão geralmente mostram valores brutos.

Opções

eixo X

Lista variáveis que estão disponíveis para o eixo X do gráfico.

eixo Y

Lista variáveis que estão disponíveis para o eixo Y do gráfico.

Mapa de cores

Lista variáveis de mapa de cores disponíveis. Essas variáveis usam progressão de cor, com base na faixa de valores na coluna especificada, para representarem a si mesmos nos pontos da trama. Os mapas de cores também são conhecidos como mapas de choropleth.

Dimensionar mapa

Lista variáveis de mapa de tamanho disponíveis. Essas variáveis usam tamanhos diferenciados para representar a si mesmos nos pontos da trama.

Formato do mapa

Lista variáveis de mapa de forma disponíveis. Essas variáveis usam formas diferenciadas para representar a si mesmas nos pontos da trama.

Linha de ajuste

Em uma linha de encaixe, os pontos de dados são encaixados em uma linha que geralmente não passa por todos os pontos de dados. A linha fit representa a tendência dos dados. Algumas linhas de encaixe são baseadas em regressão. Outros são baseados em quadrados iterativos menos ponderados. Selecione uma opção de linha fit na lista suspensa.

Bolha Gradiente

O controle de comutação ativa e desativa a exibição de gradientes de cor e efeitos 3D nas bolhas de gráfico. A configuração não está disponível quando uma variável **Color map** é selecionada.

Tamanho mínimo da bolha

Configura o tamanho mínimo da bolha. Digite um valor o intervalo 5-20.

Tamanho máximo da bolha

Configura o tamanho máximo da bolha. Digite um valor na faixa de 20-80.

Mostrar linha de referência

Quando ativada, a opção mostra uma linha de referência no gráfico que é baseado nos valores especificados **xAxis** e **yAxis**.

Insira um valor de linha de referência no eixo x

Quando a **Mostrar linha de referência** está ativada, esta configuração fornece a opção de especificar um valor de linha de referência específico para o eixo X. Clique em **Adicionar outra coluna** para especificar mais valores de linha de referência.

Insira um valor de linha de referência no eixo y

Quando a **Mostrar linha de referência** está ativada, esta configuração fornece a opção de especificar um valor de linha de referência específico para o eixo Y. Clique em **Adicionar outra coluna** para especificar mais valores de linha de referência.

Mostrar rótulo

Quando ativado, etiquetas de coluna são exibidas no gráfico. Apenas dados de séries de dispersão são suportados.

Campo de rótulo

O menu de campo fornece variáveis para exibir como etiquetas de gráfico.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Rótulo do Eixo x

O rótulo do eixo x.

Rótulo do Eixo y

O rótulo do eixo y.

Gráficos de matriz de dispersão

As matrizes de enredo de dispersão são uma boa maneira de determinar se as correlações lineares existem entre múltiplas variáveis.

Criando um gráfico de matriz de dispersão

1. Na seção **Tipo Gráfico**, clique no ícone **Matriz de Scatter**.



O canvas atualiza para exibir um modelo de gráfico de matriz de dispersão.

2. Selecione pelo menos duas variáveis de escala **Columns**.

Nota: Clique em **Adicionar outra coluna** para incluir mais variáveis de coluna.

Cada variável selecionada é plotada em relação a cada outra variável para criar uma matriz de tramas individuais de dispersão.

Opções

Colunas

Selecione pelo menos duas variáveis de matriz. As variáveis devem ser numéricas (mas não formatar formato).

Clique em **Adicionar outra coluna** para adicionar mais colunas.

Mapa de cores

Lista variáveis de mapa de cores disponíveis. Essas variáveis usam progressão de cor, com base na faixa de valores na coluna especificada, para representarem a si mesmos nos pontos da trama. Os mapas de cores também são conhecidos como mapas de choropleth.

Correlação

Quando ativada, as informações de correlação linear (Strong, Medium, Fracas) são mostradas para as variáveis selecionadas.

Mostrar curva kde

Quando ativado, a curva de estimativa de densidade do kernel é mostrada no gráfico.

Mostrar o histograma

Quando ativado, os gráficos de histograma são exibidos para as variáveis de coluna selecionadas.

Bolha Gradiente

O controle de comutação ativa e desativa a exibição de gradientes de cor e efeitos 3D nas bolhas de gráfico. A configuração não está disponível quando uma variável **Color map** é selecionada.

Tamanho mínimo da bolha

Configura o tamanho mínimo da bolha. Digite um valor o intervalo 5-20.

Tamanho máximo da bolha

Configura o tamanho máximo da bolha. Digite um valor na faixa de 20-80.

Mostrar rótulo

Quando ativado, etiquetas de coluna são exibidas no gráfico. Apenas dados de séries de dispersão são suportados.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Gráficos de matriz de série

Os gráficos de matriz da série incluem sub gráficos individuais e exibem o eixo Y para todos os sub gráficos na legenda.

Criando um gráfico de array de série

1. Na seção **Tipo Gráfico**, clique no ícone **Série Array**.

Series array



O canvas atualiza para exibir um modelo de gráfico de matriz de série.

2. Clique em **Adicionar outro sub-gráfico** para definir o sub gráfico da matriz da primeira série.
3. Selecione uma variável como a variável **X-axis**.
4. Selecione uma variável como a variável **Y-axis**.
5. Clique em **Adicionar outro sub-gráfico** para definir mais sub gráficos de matriz de série.

Opções

eixo X

Lista variáveis que estão disponíveis para o eixo x do sub gráfico.

eixo Y

Lista variáveis que estão disponíveis para o eixo y do sub gráfico.

Dividir por

Selecionar uma variável categórica que crie uma tabela de gráficos, com uma célula para cada categoria na variável Split by. Assim como o agrupamento, a divisão por variáveis essencialmente adiciona mais dimensões ao seu gráfico, exibindo informações para cada categoria de variável.

Rótulo de legenda

Especificar um título de legenda sub-gráfico.

Título

Especificar um título de sub-gráfico.

Tipo

Selecionar um tipo de sub-gráfico. As opções disponíveis são **line**, **bare** e **scatter**.

Suave

Quando ativado, o gráfico mostra uma curva suave.

Mostrar pontos de dados

Quando ativado, o ponto de dados é mostrado no gráfico.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Rótulo do Eixo x

O rótulo do eixo x.

Rótulo do Eixo y

O rótulo do eixo y.

Gráficos de explosão solar

Um gráfico de sunburst é útil para visualização de estruturas de dados hierárquicos. Um gráfico de sunburst consiste em um círculo interno que é rodeado por anéis de níveis de hierarquia mais profundos.

O ângulo de cada segmento proporcional a um valor ou dividido igualmente sob o seu segmento interno. Os segmentos de gráficos são coloridos com base na categoria ou nível hierárquico ao qual pertencem.

Criando um gráfico de sunburst simples

1. Na seção **Tipo Gráfico**, clique no ícone **Sunburst**.

Sunburst



O canvas atualiza para exibir um modelo de gráfico de sunburst.

2. Selecione uma variável categórica (nominal ou ordinal) a partir da lista **Colunas**. As categorias nesta variável determinam o número de segmentos no gráfico.
3. Clique em **Adicionar outra coluna** e selecione outra variável categórica (nominal ou ordinal) a partir da lista **Colunas**. As categorias nesta variável determinam o número de segmentos no segundo anel do gráfico e representam um nível hierárquico.
4. Selecione uma função de resumo estatístico para o elemento gráfico (Estatística baseada em contagem ou uma soma). O resultado da estatística determina o tamanho de cada segmento. Quando **Sum** for selecionado, escolha uma variável de escala a partir da lista **Valor** para representar o valor no conjunto de dados para resumir.
5. Selecione uma opção **Layout Sunburst** (seja **Tradicional** ou **Divergente**).

Opções

Colunas

Selecione uma variável categórica (nominal ou ordinal) que determine o número de segmentos no gráfico.

Resumo

Selecione uma função de resumo estatístico para o elemento gráfico (Estatística baseada em contagem ou uma soma). O resultado da estatística determina o tamanho de cada fatia.

Dois tipos de funções de resumo estatístico estão disponíveis. A distinção é importante porque determina se você precisa especificar uma variável **Valor**.

Funções que não requerem uma variável de valor

Todas as estatísticas de contagem e porcentagem estão nesta categoria. Essas estatísticas estão disponíveis quando não há uma variável **Value** definida.

Funções que requerem uma variável de valor

Por exemplo, a função **Sum** requer uma variável sobre a qual o resumo é calculado.

Valor

Este campo exibe quando uma função **Sumário** que requer uma variável de escala, é selecionada. Selecione uma variável para servir como variável de escala.

Layout de explosão solar

As opções disponíveis são **Tradicional** e **Divergente**.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

gráficos t-SNE

T-distribuído Stochastic Neighbor Embedding (t-SNE) é um algoritmo de aprendizado de máquina para visualização. Gráficos t-SNE modelam cada objeto de alta dimensional por um ponto dimensional de dois

ou três dimensões de tal forma que objetos semelhantes são modelados por pontos próximos e objetos desparecidos são modelados por pontos distantes com alta probabilidade.

Criando um gráfico de t-SNE simples

1. Na seção **Tipo Gráfico**, clique no ícone **t-SNE**.



As atualizações do canvas para exibir um modelo de gráfico t-SNE.

2. Configure os valores **Perplexidade**, **Taxa de aprendizado** e **Iterações Máximas**.
3. Opcional: Selecione uma variável **Color map**.

Opções

Perplexidade

Configura um número que estabelece um palpite educado quanto ao número de vizinhos próximos para cada ponto de dados. O propósito é equilibrar os aspectos locais e globais para os seus dados.

Taxa de aprendizado

Esse valor afeta a velocidade do aprendizado especificando as alterações de tamanho do peso em cada iteração.

Máximo de iterações

O número máximo de iterações a serem executadas.

Mapa de cores

Lista variáveis de mapa de cores disponíveis. Essas variáveis usam progressão de cor, com base na faixa de valores na coluna especificada, para representarem a si mesmos nos pontos da trama. Os mapas de cores também são conhecidos como mapas de choropleth.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Tramas de tempo

As tramas do tempo ilustram os pontos de dados em intervalos sucessivos de tempo. A série temporal que você trama deve conter valores numéricos e que são assumidos para ocorrer ao longo de um intervalo de tempo em que os períodos são uniformes. As tramas do tempo fornecem uma análise preliminar das características dos dados da série temporal sobre estatísticas básicas e teste e, assim, geram insights úteis sobre seus dados antes da modelagem. As parcelas do tempo incluem métodos de análise como decomposição, teste de Dickey-Fuller ampliado (ADF), correlações (ACF/PACF) e análise espectral.

Criando um complô de tempo simples

1. Na seção **Tipo Gráfico**, clique no ícone **plot Time**.

Time plot



O canvas atualiza para exibir um modelo de gráfico de tempo de plot.

2. A partir do drop-down **Valores**, selecione um valor para o eixo y.

Opções

Valores

Lista variáveis que estão disponíveis para os valores de plot de tempo.

Data

Seleciona uma data a partir do drop-down, se aplicável. Cada observação é separada pelo mesmo intervalo de tempo. Se você selecionar uma variável de data, uma opção resample será mostrada. Você pode usar essa opção para agrregar campos de valor para combinar com o intervalo especificado.

Algoritmo de gráfico de tempo

O algoritmo de plot de tempo para usar para analisar os dados da série temporal.

Decomposição

Decompõe uma série temporal como três componentes (trend-cycle, sazonal, e irregular). A decomposição é executada de forma aditiva.

Teste de ADF

O Augmented Dickey-Fuller (ADF) testa a hipótese nula de que uma raiz unitária existe na série e a série não é estacionária. Se o resultado do teste rejeitar a hipótese nula, a série é estacionária, ou pode ser representada para ser estacionária com um modelo de diferença.

ACF/PACF

Correlações de séries.

Análise espectral

Uma ferramenta analítica no domínio de frequência. A frequência mais alta é marcada como diamante.

Trocar posição do gráfico

Reverte as posições dos gráficos sobre o sistema de coordenadas retangulares planimétricas e o sistema de coordenadas polares.

Mostrar o ponto de ajuste

Mostra ou esconde os pontos de virada nas paradas. A série é explorada para determinar se ela tem uma tendência geral ou tem alguns pontos de viragem que mudam a direção do padrão de tendência.

Mostrar o valor discrepante

Mostra ou esconde qualquer outliers. Os outliers da série temporal são analisados a partir do componente irregular da decomposição da série temporal.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Rótulo do Eixo x

O rótulo do eixo x.

Rótulo do Eixo y

O rótulo do eixo y.

Gráficos do Rio Tema

Um rio temático é um gráfico de fluxo especializado que mostra mudanças ao longo do tempo.

Como criar um gráfico de rio de tema simples

1. Na seção **Tipo Gráfico**, clique no ícone **Rio Tema**.

Theme River



O canvas atualiza para exibir um modelo de gráfico do rio tema.

2. Selecione uma variável **X-axis**.

3. Selecione uma variável **Category**.

Restrição: Se um campo de categoria possui mais de 50 categorias distintas, apenas as primeiras 50 categorias máximas são usadas como eventos no gráfico.

Opções

eixo X

Lista variáveis que estão disponíveis para o eixo X.

Ordenar com base em

Dependendo do valor do eixo X que você selecionar, você pode ser capaz de especificar se a ordem de categoria é baseada em nome da categoria ou valor da categoria.

Ordem da categoria

Dependendo do valor do eixo X que você selecionar, você pode ser capaz de especificar se a ordem de categoria é crescente, decrescente ou como lida a partir do conjunto de dados.

Categoria

Lista variáveis que estão disponíveis como categorias.

Resumo

Selecione uma função de resumo estatístico (o método que é usado para resumir cada categoria).

Há dois tipos de funções de resumo estatístico disponíveis. A distinção é importante porque determina se você precisa especificar uma variável **Value**.

- **Funções que não requerem uma variável de valor.** Funções que não requerem uma variável. Todas as estatísticas de contagem e porcentagem estão nessa categoria. Essas estatísticas estão disponíveis quando a variável **Value** não está definida.
- **Funções que de fato exigem uma variável de valor.** Funções que de fato exigem uma variável **Value**. Por exemplo, a função " **Mean** " requer uma variável na qual a média é calculada. Essas estatísticas estão disponíveis quando a variável **Value** não está definida.

Valor

Esse campo é exibido quando uma função **de resumo** que requer uma variável de valor é selecionada. Selecione uma variável para servir como valor.

Orientação da legenda

Configura a orientação da legenda do gráfico. As opções disponíveis são **Horizontal**, **Vertical** e **Vertical bottom**.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Gráficos de árvore

Os gráficos de árvores representam hierarquia em uma estrutura semelhante a árvore. A estrutura de um gráfico de árvore consiste em um nó raiz (não possui nenhum nó pai), conexões de linha (ramificações

denominadas) e nós de folhas (não possuem nós filhos). As conexões de linha representam as relações e as conexões entre os membros.

Criando um gráfico de árvore simples

1. Na seção **Tipo Gráfico**, clique no ícone **Tree**.

Tree



As canvas atualizam para exibir um modelo de gráfico de árvore.

2. Selecione uma variável **Colunas** a partir da lista suspensa.

Nota: Clique em **Adicionar outra coluna** para incluir mais variáveis de coluna.

Opções

Colunas

Lista variáveis que estão disponíveis para representar colunas de gráfico.

Resumo

Selecione uma função de resumo estatístico para o elemento gráfico. O resultado da estatística determina a posição dos elementos gráficos.

Dois tipos de funções de resumo estatístico estão disponíveis. A distinção é importante porque determina se você precisa especificar uma variável **Valor**.

Funções que não requerem uma variável de valor

Todas as estatísticas de contagem e porcentagem estão nesta categoria. Essas estatísticas estão disponíveis quando não há uma variável **Value** definida.

Funções que requerem uma variável de valor

Por exemplo, a função **Mean** requer uma variável sobre a qual a média é calculada. Essas estatísticas estão disponíveis quando há uma variável **Value** definida.

Valor

Este campo exibe quando uma função **Sumário** que requer uma variável de valor, é selecionada. Selecione uma variável para servir como base para o valor do resumo.

Layout de árvore

Esquerda para a direita

O nó raiz exibe na esquerda e os nós de folhas são exibidos à direita.

Direita para a esquerda

O nó raiz exibe na direita e os nós de folhas são exibidos à esquerda.

Superior a inferior

O nó raiz exibe na parte superior e os nós de folhas são exibidos na parte inferior.

De baixo para cima

O nó raiz exibe na parte inferior e os nós de folhas são exibidos na parte superior.

Radial

O nó raiz exibe no meio e os nós de folhas irradiam da raiz.

Espessura da folha

Configura o valor de nível de drill down para os nós de folhas.

Mostrar etiqueta de folhas

Quando ativado, etiquetas são exibidas para cada nó da folha.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Gráficos de treemap

Os gráficos do Treemap são um método alternativo para visualizar a estrutura hierárquica de diagramas de árvores enquanto também exibia quantidades para cada categoria. Os gráficos do Treemap são úteis para identificar padrões em dados. Os galhos de árvore são representados por retângulos, com cada sub ramo representado por retângulos menores.

Criando um gráfico de Treemap simples

1. Na seção **Tipo Gráfico**, clique no ícone **Treemap**.



O canvas atualiza para exibir um modelo de gráfico Treemap.

2. Selecione uma variável **Colunas** a partir da lista suspensa.

Nota: Clique em **Adicionar outra coluna** para incluir mais variáveis de coluna.

Opções

Colunas

Lista variáveis que estão disponíveis para representar colunas de gráfico.

Resumo

Selecione uma função de resumo estatístico para o elemento gráfico. O resultado da estatística determina a posição dos elementos gráficos.

Dois tipos de funções de resumo estatístico estão disponíveis. A distinção é importante porque determina se você precisa especificar uma variável **Valor**.

Funções que não requerem uma variável de valor

Todas as estatísticas de contagem e porcentagem estão nesta categoria. Essas estatísticas estão disponíveis quando não há uma variável **Value** definida.

Funções que requerem uma variável de valor

Por exemplo, a função **Mean** requer uma variável sobre a qual a média é calculada. Essas estatísticas estão disponíveis quando há uma variável **Value** definida.

Valor

Este campo exibe quando uma função **Sumário** que requer uma variável de valor é selecionada. Selecione uma variável para servir como base para o valor do resumo.

Espessura da folha

Configura o valor de nível de drill down para os nós de folhas.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Gráficos de nuvem de palavras

Os gráficos de nuvem de palavras apresentam dados como palavras, em que o tamanho e o posicionamento de qualquer palavra individual é determinado por como ele é ponderado.

Criando um gráfico de nuvem de palavras simples

1. Na seção **Tipo Gráfico**, clique no ícone **nuvem Word**.

O canvas atualiza para exibir um modelo de gráfico de nuvem de palavras.

2. Selecione uma variável como a variável **Fonte**. Cada categoria de variável é representada no gráfico com base em seu valor de peso.
3. Selecione um valor **Shape** para o gráfico. Os dados do gráfico são apresentados na forma selecionada.

Opções

Origem

Lista variáveis que estão disponíveis como fonte do gráfico. Cada categoria de variável é representada no gráfico com base em seu valor de peso.

Forma

Lista as formas de gráfico disponíveis. Os dados do gráfico são apresentados na forma selecionada.

Título principal

O título do gráfico.

Subtítulo

O subtítulo do gráfico.

Nota de rodapé

A nota de rodapé do gráfico.

Preferências de visualização global

É possível substituir as configurações padrão para títulos, régua de controle de intervalo, linhas de grade e rastreamento do mouse para todas as visualizações de um ativo de dados.. Você também pode especificar um modelo de esquema de cores diferente. Os modelos de tema de cores são úteis para dados com agrupamentos ou variáveis de empilhamento

Com as opções **Ferramentas**, é possível usar uma **Range Slider** para controlar quantos dados você deseja que sejam exibidos em um gráfico ou para incluir opções como zoom, salvar como imagem, restaurar, selecionar dados e limpar seleção. A opção **Aria** quando ativada torna o conteúdo da web e os aplicativos da web mais acessíveis.

Para mudar as configurações padrão de todas as visualizações para um ativo de dados:

1. Na guia **Visualização**, Após clicar com o botão direito em um nó de dados e selecionar **Visualizar Dados**, clique no controle **Preferências de Visualização Global** na seção **Ações**.



Figura 14. Controle de preferências de visualização global

- a. O diálogo **Preferências de visualização global** fornece as configurações a seguir: Essas configurações são aplicadas a todas as visualizações para o ativo atual

Títulos

Fornece configurações de título do gráfico global.

Títulos globais

Ativa ou desativa os títulos globais para todos os gráficos.

Título principal global

Ativa ou desativa a exibição de títulos de gráficos globais e primários. Quando ativado, o título de gráfico de nível superior que você insere aqui é aplicado a todos os gráficos, efetivamente anulando a configuração individual do **Título primário** de cada gráfico.

Subtítulo global

Ativa ou desativa a exibição de legendas de gráfico global. Quando ativado, o subtítulo do gráfico que você insere aqui é aplicado a todos os gráficos, efetivamente anulando a configuração individual do **Subtítulo** de cada gráfico.

Títulos padrão

Ativa ou desabilita os títulos padrão para todos os gráficos.

Alinhamento do título

Fornece as opções de alinhamento de título **Esquerda**, **Center** (a configuração padrão) e **Direita**.

Ferramentas

Fornece opções que controlam o comportamento do gráfico.

Régua de controle do intervalo

Ativa ou desabilita o slider de alcance para cada gráfico. Quando ativado, é possível controlar a quantidade de dados do gráfico que é exibido com uma régua de corte de gama que é fornecida para cada gráfico.

Linhas de grade

Controla a exibição de linhas do eixo X (vertical) e do eixo Y (horizontal).

Rastreador do mouse

Quando ativado, o local do cursor do mouse, em relação aos dados do gráfico, é rastreado e exibido quando colocado em qualquer lugar sobre o gráfico.

Caixa de Ferramentas

Ativa ou desabilita a caixa de ferramentas para cada gráfico. Dependendo do tipo de gráfico, a caixa de ferramentas à direita da tela fornece ferramentas como zoom, salvar como imagem, restaurar, selecionar dados e limpar a seleção.

ARIA

Quando ativados, o conteúdo web e os aplicativos web são mais acessíveis para os usuários com deficiência.

Filtrar nulo

Ativa ou desativa a filtragem de dados de gráficos nulos.

Eixo X em zero

Quando ativado, o eixo X encontra-se na posição de origem do outro. Quando não ativado, o eixo X sempre começa em 0.

Eixo Y em zero

Quando ativado, o eixo Y encontra-se na posição de origem do outro. Quando não ativado, o eixo Y sempre começa em 0.

Mostrar rótulo de Eixo x

Ativa ou desativa o rótulo do xAis

Mostrar etiqueta do eixo y

Ativa ou desativa o rótulo do yAis

Mostrar linha de Eixo x

Ativa ou desativa a linha do xAis

Mostrar linha de Eixo y

Ativa ou desativa a linha do yAis

Mostrar o nome do eixo X

Ativa ou desativa o nome do xAis

Mostrar o nome do eixo Y

Ativa ou desativa o nome do yAis

Local do nome do Eixo y

A lista suspensa fornece opções para especificar o local do nome do yAis As opções incluem **Start**, **Middle** e **End**.

Comprimento do truncamento

O valor especificado configura o comprimento da cadeia. Strings que são mais compridas do que o comprimento especificado são truncados. O valor padrão é 10. Quando 0 é especificado, o truncamento é desligado.

Decimal do rótulo de marcação de XAxis

Configura o valor decimal do rótulo de marcação para xAxis. O valor padrão é 3.

Decimal do rótulo de marcação de YAxis

Configura o valor decimal do rótulo de marcação para yAxis. O valor padrão é 3.

Girar rótulo de marcação do eixo X

Configura o valor de rotação do rótulo de marcação xAxis O valor padrão é 0 (sem rotação). Você pode especificar valor na faixa de -90 90 graus.

Tema

Selecione um modelo para alterar as cores que são usadas em gráficos que possuem uma variável de agrupamento ou empilhamento. Quaisquer atributos de elemento definidos no arquivo de template selecionado substituem as configurações de template padrão para esses atributos do elemento.

2. Escolha as configurações necessárias para. As configurações selecionadas são aplicadas a todos os gráficos.
3. Clique em **Aplicar** para salvar suas configurações ou em **Cancelar** para desconsiderar as mudanças

As **Preferências de visualização global** são aplicadas a todas as visualizações para o ativo de dados atual

Para obter uma lista de tipos de gráfico, consulte “[Tipos de gráficos](#)” na página 64

Capítulo 7. Painel

Você pode criar um layout de painel de gráficos para visualização de vários gráficos de cada vez. Salve layouts como templates e arraste e baixe seus gráficos salvos para as posições em seu layout.

1. Após clicar com o botão direito do mouse em um nó de dados e selecionar **Visualizar Dados**, clique no controle do Painel na seção **Ações**.



Figura 15. Controle do painel

O Dashboard exibe e fornece as configurações a seguir.

Modelo

Nesta seção, você pode criar novos modelos de layout ou escolher a partir de modelos predefinidos ou de seus modelos salvos.

Escolher um modelo de layout

Escolha a partir de modelos de layout disponíveis ou comece com um novo modelo.

Ações

Clique no ícone **Editar o layout** para editar o modelo de layout selecionado. Você também pode importar um arquivo de painel. Quando finalizado, salve seu template.

Clique no ícone **Deixe o modo de edição do layout** quando você terminar.

Conteúdo

Use esta seção para arrastar e soltar itens para o layout do seu painel.

Escolher um gráfico salvo

Esta seção fornece uma lista de seus gráficos salvos. Gráficos de arrastar e soltar para a posição desejada no layout do seu painel.

Escolher um objeto

Você também pode arrastar texto HTML ou imagens para o layout do seu painel.

Capítulo 8. Trabalhando com saída

Ao executar alguns fluxos, os resultados estão disponíveis no Viewer através da guia **Modelo** ou na guia **Avançado** de nós de nugget do modelo. No Visualizador, é possível navegar facilmente para a saída que deseja ver. Também é possível manipular a saída e criar um documento que contenha precisamente a saída desejada. Alguma saída de gráfico também usa o Viewer.

O Viewer é usado para a seguinte saída em IBM SPSS Modelador:

- Nuggets do modelo TCM
- Nuggets do modelo STP
- Nuggets do modelo de Cluster TwoStep-AS
- Nuggets do modelo GSAR
- Nό de gráfico de Visualização de Mapa

Viewer

Resultados são exibidos no Visualizador. É possível utilizar o Visualizador para:

- Procurar resultados
- Mostrar ou ocultar tabelas e gráficos selecionados
- Alterar a ordem de exibição dos resultados ao mover os itens selecionados
- Mover itens entre o Visualizador e outros aplicativos

O Visualizador é dividido em duas áreas de janela

- A área de janela esquerda contém a visualização da estrutura de tópicos do conteúdo.
- A área de janela direita contém tabelas de estatísticas, gráficos e saída texto.

É possível clicar em um item na estrutura de tópicos para acessar diretamente a tabela ou gráfico correspondente. É possível clicar e arrastar a borda direita da área de janela da estrutura de tópicos para alterar a largura da área de janela da estrutura de tópicos.

Mostrando e ocultando resultados

No Viewer, você pode mostrar seletivamente e ocultar tabelas individuais ou resultados de um procedimento inteiro. Esse processo é útil quando desejar encurtar a quantia de saída visível na área de janela de conteúdo.

Para ocultar tabelas e gráficos

1. Clique duas vezes no ícone de livro do item na área de janela da estrutura de tópicos do Visualizador.
ou
2. Clique no item para selecioná-lo.
3. No menu, escolha:

Visualizar > Ocultar

ou

4. Clique no ícone de livro fechado (Ocultar) na barra de ferramentas da Estrutura de Tópicos.

O ícone de livro aberto (Mostrar) se torna o ícone ativo, indicando que o item está agora oculto.

Para ocultar os resultados do procedimento

1. Clique na caixa à esquerda do nome do procedimento na área de janela da estrutura de tópicos.
- Isso oculta todos os resultados do procedimento e reduz a visualização da estrutura de tópicos.

Mover, excluir e copiar saída

É possível reorganizar os resultados ao copiar, mover ou excluir um item ou um grupo de itens.

Movimentação de movimentação no Viewer

1. Selecione os itens na área de janela de estrutura de tópicos ou de conteúdo.
 2. Arraste e solte os itens selecionados em uma localização diferente.
- O item que você corta na etapa 1, é movido para o local selecionado.

Excluindo saída no Viewer

1. Selecione os itens na área de janela de estrutura de tópicos ou de conteúdo.
 2. Pressione a tecla **Excluir**.
- ou
3. No menu, escolha:

Editar > Excluir

Alterando o alinhamento inicial

Por padrão, todos os resultados são inicialmente alinhados à esquerda. Para alterar o alinhamento inicial dos novos itens de saída

1. No menu, escolha:
- Editar > Opções**
2. Clique na guia **Visualizador**.
 3. No grupo Estado de Saída Inicial, selecione o tipo de item (por exemplo, tabela dinâmica, gráfico e saída de texto).
 4. Selecione a opção de alinhamento desejada.

Alterando o alinhamento dos itens de saída

1. Na área de janela de estrutura de tópicos ou de conteúdo, selecione os itens que deseja alinhar.
 2. No menu, escolha:
- Formatar > Alinhar à Esquerda**
- ou
- Formato > Centro**
- ou
- Formato > Alinhar à direita**

Viewer esboço

A pane de esboço fornece uma tabela de conteúdo do documento do espectador. A área de janela da estrutura de tópicos pode ser usada para navegar pelos seus resultados e controlar a exibição. A maioria das ações na área de janela da estrutura de tópicos possui um efeito correspondente na área de janela de conteúdo.

- Selecionar um item na área de janela da estrutura de tópicos exibe o item correspondente na área de janela de conteúdo.
- Mover um item na área de janela da estrutura de tópicos move o item correspondente na área de janela de conteúdo.
- Reduzir a visualização da estrutura de tópicos oculta os resultados de todos os itens nos níveis reduzidos.

Controlando a exibição da estrutura de tópicos

Para controlar a exibição da estrutura de tópicos, é possível:

- Expandir e reduzir a visualização da estrutura de tópicos
- Alterar o nível da estrutura de tópicos para itens selecionados
- Alterar o tamanho dos itens na exibição da estrutura de tópicos
- Alterar a fonte que é usada na exibição da estrutura de tópicos

Reducindo e expandindo a visualização da estrutura de tópicos

1. Clique na caixa à esquerda do item de estrutura de tópicos que deseja reduzir ou expandir.

ou

2. Clique no item na estrutura de tópicos.

3. No menu, escolha:

Visualizar > Reduzir

ou

Visualizar > Expandir

Mudando o nível da estrutura de tópicos

1. Clique no item na área de janela da estrutura de tópicos.

2. No menu, escolha:

Editar > Estrutura de Tópicos > Promover

ou

Editar > Estrutura de tópicos > Rebaixar

Para alterar o tamanho dos itens da estrutura de tópicos

1. No menu, escolha:

Visualizar > Tamanho da estrutura de tópicos

2. Selecione o tamanho da estrutura de tópicos (**Pequeno, Médio ou Grande**).

Para alterar a fonte na estrutura de tópicos

1. No menu, escolha:

Visualizar > Fonte delineada...

2. Selecione uma fonte.

Editando e adicionando itens para o Viewer

No Viewer, é possível editar e adicionar itens como títulos, novo texto, gráficos ou material de outras aplicações.

Incluindo um título ou texto

Os itens de texto que não estão conectados a uma tabela ou gráfico podem ser adicionados ao Viewer.

1. Clique na tabela, no gráfico ou em outro objeto que precederá o título ou o texto.
2. No menu, escolha:

Inserir > Novo Título

ou

Inserir > Novo Texto

3. Clique duas vezes no novo objeto.
4. Insira o texto.

Para incluir um arquivo de texto

1. Na área de janela da estrutura de tópicos ou na área de janela de conteúdo do Visualizador, clique na tabela, em um gráfico ou em outro objeto que precederá o texto.
2. No menu, escolha:

Inserir > Arquivo de texto...

3. Selecione um arquivo de texto.

Para editar o texto, clique duas vezes nele.

Colando objetos no Viewer

Objetos de outros aplicativos podem ser colados no Viewer. Você pode usar o **Colar Depois** ou **Colar Especial**. Qualquer um desses tipos de colagem coloca o novo objeto após o objeto selecionado atualmente no Visualizador. Use **Colar especial** quando quiser escolher o formato do objeto colado.

Como localizar e substituir informações no Viewer

1. Para localizar ou substituir informações no Visualizador, nos menus escolha:

Editar > Localizar

ou

Editar > Substituir

É possível utilizar Localizar e Substituir para:

- Procurar o documento inteiro ou somente os itens selecionados.
- Procurar para cima ou para baixo na localização atual.
- Procurar áreas de janela ou restringir a procura para a área de janela da estrutura de tópicos ou de conteúdo.
- Procurar itens ocultos. Esses incluem quaisquer itens ocultos na área de janela de conteúdo (por exemplo, tabelas de Notas, que ficam ocultas por padrão) e linhas e colunas ocultas nas tabelas dinâmicas.
- Restringir o critério de procura para correspondências com distinção entre maiúsculas e minúsculas.
- Restringir o critério de procura em tabelas dinâmicas para correspondências do conteúdo da célula inteira.
- Restringir o critério de procura em tabelas dinâmicas somente para marcadores de nota de rodapé. Essa opção não estará disponível se a seleção no Visualizador incluir qualquer item diferente de tabelas dinâmicas.

Itens Ocultos e Camadas de Tabela Dinâmica

- As estratos sob a estrato atualmente visível de uma tabela dinâmica multidimensional não são consideradas ocultas, e serão incluídas na área de procura mesmo quando itens ocultos não forem incluídos na procura.
- Itens ocultos incluem itens ocultos na área de janela de conteúdo (itens com ícones de um livro fechado na área de janela da estrutura de tópicos ou incluídos em blocos reduzidos de página de estrutura de tópico), e linhas e colunas em tabelas dinâmicas ocultas por padrão (por exemplo, linhas e colunas vazias ficam ocultas por padrão) ou ocultas manualmente ao editar a tabela e ocultar seletivamente linhas ou colunas específicas. Os itens ocultos serão incluídos na procura somente se você selecionar explicitamente **Incluir itens ocultos**.
- Em ambos os casos, o elemento oculto ou não visível que contiver o texto ou o valor da procura é exibido quando localizado, mas o item retorna para o seu estado original posteriormente.

Localizar um intervalo de linhas nas tabelas dinâmicas

Para localizar valores que caem em um intervalo de valores especificado nas tabelas dinâmicas:

1. Ative uma tabela dinâmica ou selecione uma ou mais tabelas dinâmicas no Visualizador. Assegure-se de que somente tabelas dinâmicas sejam selecionadas. Se quaisquer outros objetos forem selecionados, a opção Intervalo não estará disponível.

2. No menu, escolha:

Editar > Localizar

3. Clique na guia **Intervalo**.

4. Selecione o tipo de intervalo: Entre, Maior ou igual a ou Menor ou igual a.

5. Selecione o valor ou os valores que definem o intervalo.

- Se algum valor contiver caracteres não numéricos, ambos os valores serão tratados como sequências de caracteres.
- Se os dois valores forem números, somente valores numéricos serão procurados.
- Não é possível usar a guia Intervalo para substituir valores.

Esse recurso não está disponível para tabelas anteriores. Consulte o tópico [“Tabelas anteriores” na página 138](#) para obter informações adicionais.

Copiando a saída em outros aplicativos

Objetos de saída podem ser copiados e colados em outros aplicativos, como um programa de processamento de texto ou uma planilha. É possível colar saída em uma variedade de formatos.

Dependendo do aplicativo de destino e dos objetos de saída selecionados, alguns ou todos os formatos a seguir podem estar disponíveis:

Metarquivo. Formato de metarquivo WMF e EMF. Esses formatos estão disponíveis apenas em sistemas operacionais Windows.

RTF (formato Rich Text). Vários objetos, saídas de texto e tabelas dinâmicas selecionados podem ser copiados e colados no formato RTF. Para tabelas dinâmicas, na maioria dos aplicativos, isto significa que as tabelas são coladas como tabelas, que podem, então, ser editadas no outro aplicativo. Tabelas dinâmicas que são muito grandes para a largura do documento serão quebradas, reduzidas para caber na largura do documento, ou serão deixados inalterados, dependendo das configurações de opções de tabela dinâmica. Veja o tópico [“Opções da tabela dinâmica” na página 139](#) para obter mais informações.

Nota: O Microsoft Word pode não exibir tabelas extremamente largas adequadamente.

Imagem. Formatos de imagem JPG e PNG.

BIFF. As tabelas dinâmicas e as saídas de texto podem ser coladas em uma planilha no formato BIFF. Os números nas tabelas dinâmicas retêm precisão numérica. Esse formato está disponível somente em sistemas operacionais Windows.

Texto. As tabelas dinâmicas e a saídas de texto podem ser copiadas e coladas como texto. Este processo pode ser útil para aplicativos como e-mail, em que o aplicativo pode aceitar ou transmitir apenas texto.

Objeto Gráfico Microsoft Office. Os gráficos que suportam esse formato podem ser copiados para aplicativos Microsoft Office e editados nesses aplicativos como gráficos nativos do Microsoft Office. Por causa das diferenças entre os gráficos SPSS Estatísticas/ModeladorSPSS e os gráficos do Microsoft Office, alguns recursos de gráficos SPSS Estatísticas/ModeladorSPSS não são retidos na versão copiada. A cópia de múltiplos gráficos selecionados no formato Microsoft Office Graphic Object não é suportada.

Se o aplicativo de destino suporta vários formatos disponíveis, ele pode ter um item de menu Colar especial que permite selecionar o formato, ou ele pode exibir automaticamente uma lista de formatos disponíveis.

Nota: O Microsoft Office versão 16 (ou superior) é necessário ao copiar e colar Boxplots e Histogramas.

Copiando e colando múltiplos objetos de saída

As limitações a seguir são aplicadas ao colar vários objetos de saída em outros aplicativos:

- **Formato RTF.** Na maioria dos aplicativos, as tabelas dinâmicas são coladas como tabelas que podem ser editadas neste aplicativo. Gráficos, árvores e visualizações do modelo são colados como imagens.
- **Formatos de metarquivo e imagem.** Todos os objetos de saída selecionados são colados como um objeto único em outro aplicativo.
- **Formato BIFF.** Gráficos, árvores e visualizações de modelo são excluídos.

Copiar especial

Ao copiar e colar grandes quantias de saída, particularmente tabelas dinâmicas muito grandes, é possível melhorar a velocidade da operação usando **Editar > Copiar especial** para limitar o número de formatos copiados para a área de transferência.

Também é possível salvar os formatos selecionados como o conjunto padrão de formatos para copiar para a área de transferência. Esta configuração persiste entre as sessões.

Copiar como

É possível clicar com o botão direito em um objeto selecionado no Visualizador de saída e selecionar **Editar > Copiar como** para copiar para os formatos mais populares (por exemplo, **Todos**, **Imagem** ou **Objeto gráfico do Microsoft Office**). A seleção de **Editar > Copiar** copia **Todos**. Observe que, se **Copiar como** estiver esmaecido ou não estiver presente para um objeto selecionado, esse formato de cópia não está disponível para esse objeto específico.

Resultado interativo

Os objetos de resultado interativo contêm diversos objetos de saída relacionados. A seleção em um objeto pode alterar o que é exibido ou destacado no outro objeto. Por exemplo, selecionar uma linha em uma tabela pode destacar uma área em um mapa ou exibir um gráfico para uma categoria diferente.

Objetos de resultado interativo não suportam a edição de recursos, como mudança de texto, cores, fontes e bordas de tabela. Objetos individuais podem ser copiados do objeto interativo para o Visualizador. As tabelas copiadas do resultado interativo podem ser editadas no editor de tabela dinâmica.

Copiando objetos do resultado interativo

Arquivo > Copiar para o Visualizador copia objetos de saída individuais para a janela do **Visualizador**.

- As opções disponíveis dependem do conteúdo do resultado interativo.
- **Gráfico e Mapa** cria objetos de gráfico.
- **Tabela** cria uma tabela dinâmica que pode ser editada no editor de tabela dinâmica.

- **Captura instantânea** cria uma imagem da visualização atual.
 - **Modelo** cria uma cópia do objeto de resultado interativo atual.
- Editar>Copiar objeto** copia objetos de saída individuais para a área de transferência.
- Colar o objeto copiado no Visualizador equivale a **Arquivo>Copiar para o Visualizador**.
 - Colar o objeto em outro aplicativo cola o objeto como uma imagem.

Zoom e Panorama

Para mapas, é possível usar **Visualizar>Zoom** para alterar o zoom da visualização do mapa. Em uma visualização de mapa com zoom, é possível usar **Visualizar>Panoramizar** para mover a visualização.

Configurações de impressão

Arquivo>Configurações de impressão controla como os objetos interativos são impressos.

- **Imprimir somente visualização visível.** Imprime somente a visualização que estiver atualmente exibida. Essa opção é a configuração padrão.
- **Imprimir todas as visualizações.** Imprime todas as visualizações contidas no resultado interativo.
- A opção selecionada também determina a ação padrão para exportar o objeto de saída.

Exportar saída

Saída de exportação salva Visualizador saída em HTML, texto, Word / RTF, Excel, PowerPoint (requer PowerPoint 97 ou posterior) e formatos PDF. Os gráficos também podem ser exportados em um número de diferentes formatos gráficos.

Nota: Exportar para PowerPoint está disponível somente em sistemas operacionais Windows.

Exportar Saída

1. Faça o Viewer o diálogo ativo (clique em qualquer lugar no diálogo).
2. Clique no botão **Exportar** na barra de ferramentas ou clique com o botão direito na janela de saída e selecione **Exportar**.
3. Insira um nome do arquivo (ou um prefixo para gráficos) e selecione um formato de exportação.

Objetos para Exportação. É possível exportar todos os objetos no Visualizador, todos os objetos visíveis ou somente os objetos selecionados.

Tipo de Documento. As opções disponíveis são:

- **Word/RTF (*.doc).** Tabelas dinâmicas são exportadas como tabelas do Word com todos os atributos de formatação intactos (por exemplo, bordas de célula, estilos de fonte e cores do plano de fundo). A saída de texto é exportada como RTF formatado. Gráficos, diagramas de árvore e visualizações do modelo são incluídos no formato PNG. Observe que o Microsoft Word poderá não exibir corretamente tabelas extremamente largas.
- **Excel 97-2004 (*.xls) /Excel 2007 e Superior (*.xlsx).** Linhas, colunas e células de tabela dinâmica são exportadas como linhas, colunas e células do Excel com todos os atributos de formatação intactos (por exemplo, bordas de célula, estilos de fonte e cores do plano de fundo). A saída de texto é exportada como todos os atributos de fonte intactos. Cada linha na saída de texto é uma linha no arquivo Excel, com o conteúdo inteiro da linha em uma única célula. Gráficos, diagramas de árvore e visualizações do modelo são incluídos no formato PNG. A saída pode ser exportada como *Excel 97-2004* ou *Excel 2007 e superior*.
- **HTML (*.htm).** Tabelas dinâmicas são exportadas como tabelas HTML. A saída de texto é exportada como HTML pré-formatado. Gráficos, diagramas de árvore e visualizações do modelo são integrados no documento no formato de gráfico selecionado. Um navegador compatível com o HTML 5 é necessário para visualizar a saída que é exportada no formato HTML.

- **Portable Document Format (*.pdf).** Toda a saída é exportada como se ela aparecesse na Visualização de Impressão, com todos os atributos de formatação intactos.
- **Texto- Plain/UTF8/UTF16 (*.txt).** Os formatos de saída de texto incluem texto simples, UTF-8 e UTF-16. Tabelas dinâmicas podem ser exportadas em formato separado por tabulação ou separado por espaço. Toda a saída de texto é exportada em formato separado por espaço. Para gráficos, diagramas de árvore e visualizações do modelo, uma linha é inserida no arquivo de texto para cada gráfico, indicando o nome do arquivo de imagem.
- **Nenhum (Somente Gráficos).** Os formatos de exportação disponíveis incluem: EPS, JPEG, TIFF, PNG e BMP. Em sistemas operacionais Windows, o formato EMF (metarquivo aprimorado) também está disponível.

Abrir a pasta de conteúdo. Abre a pasta que contém os arquivos que são criados pela exportação.

Opções de HTML

A exportação HTML requer um navegador que seja compatível com HTML 5.

As opções a seguir estão disponíveis para exportar a saída em formato do HTML:

Camadas em tabelas dinâmicas. Por padrão, a inclusão ou a exclusão de camadas de tabela dinâmica é controlada pelas propriedades da tabela de cada tabela dinâmica. É possível substituir essa configuração e incluir todas as camadas ou excluir todas, exceto a camada visível atualmente. Veja o tópico [“Propriedades de tabela: impressão” na página 133](#) para obter mais informações.

Exportar tabelas em camadas como interativas. Tabelas em camadas são exibidas conforme elas aparecem no Visualizador, e é possível alterar interativamente a camada exibida no navegador. Se essa opção não estiver selecionada, cada camada de tabela será exibida como uma tabela separada.

Tabelas como HTML. Isso controla as informações de estilo incluídas para tabelas dinâmicas exportadas.

- **Exportar com estilos e largura de coluna fixa.** Todas as informações de estilo de tabela dinâmica (estilos de fonte, cores do plano de fundo, etc.) e larguras de coluna são preservadas.
- **Exportar sem estilos.** As tabelas dinâmicas são convertidas em tabelas HTML padrão. Nenhum atributo de estilo é preservado. A largura da coluna é determinada automaticamente.

Incluir notas de rodapé e legendas. Controla a inclusão ou a exclusão de todas as notas de rodapé e legendas de tabela dinâmica.

Visualizações de Modelos. Por padrão, a inclusão ou a exclusão de visualizações do modelo é controlada pelas propriedades de modelo de cada modelo. É possível substituir essa configuração e incluir todas as visualizações ou excluir todas, exceto a visualização aberta atualmente. (Nota: todas as visualizações do modelo, incluindo tabelas, são exportadas como gráficos).

Nota: Para HTML, também é possível controlar o formato de arquivo de imagem para gráficos exportados. Veja o tópico [“Opções de formato de gráficos” na página 121](#) para obter mais informações.

Para configurar as opções de exportação de HTML

1. Selecione **HTML** como o formato de exportação.
2. Clique em **Alterar Opções**.

Opções de relatório da web

Um relatório da web é um documento interativo compatível com a maioria dos navegadores. Muitos dos recursos interativos de tabelas dinâmicas disponíveis no Visualizador também estão disponíveis nos relatórios da web.

Título do Relatório. O título que é exibido no cabeçalho do relatório. Por padrão, o nome do arquivo é utilizado. É possível especificar um título customizado a ser usado ao invés do nome do arquivo.

Formato. Há duas opções para o formato do relatório:

- **Relatório da Web SPSS (HTML 5).** Esse formato requer um navegador que seja compatível com HTML 5.
- **Cognos Active Report (mht).** Esse formato requer um navegador que suporte os arquivos de formato MHT ou o aplicativo Cognos Active Report.

Excluir Objetos. É possível excluir os tipos de objeto selecionados do relatório:

- **Texto** Objetos de texto que não são logs. Essa opção inclui objetos de texto que contêm informações sobre o conjunto de dados ativo.
- **Logs.** Objetos de texto que contêm uma lista da sintaxe de comando que foi executada. Os itens de log também incluem mensagens de aviso e de erro que forem encontradas pelos comandos que não produzem nenhuma saída do Visualizador.
- **Tabelas de Notas.** A saída dos procedimentos estatísticos e de gráfico inclui uma tabela de Notas. Essa tabela contém informações sobre o conjunto de dados que foi usado, valores omissos e a sintaxe de comando que foi utilizada para executar o procedimento.
- **Mensagens de Aviso e de Erro.** Mensagens de aviso e de erro a partir de procedimentos estatísticos e de gráfico.

Redefinir estilo de tabelas e de gráficos para corresponder ao Web Report. Essa opção aplica o estilo de relatório da web padrão a todas as tabelas e gráficos. Essa configuração substitui quaisquer fontes, cores ou outros estilos na saída, conforme exibidos no Visualizador. Não é possível modificar o estilo do relatório da web padrão.

Conexão do Servidor da Web. É possível incluir a localização da URL de um ou mais servidores de aplicativos que estiverem executando o IBM SPSS Estatísticas Web Report Application Server. O servidor de aplicativos da web fornece recursos para tabelas dinâmicas, editar gráficos e salvar relatórios da web modificados.

- Selecione **Usar** para cada servidor de aplicativos que desejar incluir no relatório da web.
- Se um relatório da web contiver uma especificação de URL, o relatório da web se conectará a esse servidor de aplicativos para fornecer os recursos de edição adicionais.
- Se diversas URLs forem especificadas, o relatório da web tentará conectar-se a cada servidor na ordem em que eles estiverem especificados.

O IBM SPSS Estatísticas Web Report Application Server pode ser transferido por download a partir de <http://www.ibm.com/developerworks/spssdevcentral>.

Opções de Palavra

As opções a seguir estão disponíveis para exportar a saída em formato do Word:

Preservar pontos de interrupção. Se tiver definido pontos de interrupção, essas configurações serão preservadas nas tabelas do Word.

Incluir notas de rodapé e legendas. Controla a inclusão ou a exclusão de todas as notas de rodapé e legendas de tabela dinâmica.

Visualizações de Modelos. Por padrão, a inclusão ou a exclusão de visualizações do modelo é controlada pelas propriedades de modelo de cada modelo. É possível substituir essa configuração e incluir todas as visualizações ou excluir todas, exceto a visualização aberta atualmente. (Nota: todas as visualizações do modelo, incluindo tabelas, são exportadas como gráficos).

Configuração da Página para Exportação. Isso abre um diálogo no qual é possível definir o tamanho e as margens da página para o documento exportado. A largura do documento usada para determinar o comportamento de quebra ou de redução é a largura da página menos as margens esquerda e direita.

Para configurar as opções de exportação do Word

1. Selecione **Word/RTF** como o formato de exportação.
2. Clique em **Alterar Opções**.

Opções do Excel

As opções a seguir estão disponíveis para exportar a saída em formato do Excel:

Criar uma planilha ou modificar uma planilha existente. Por padrão, uma nova planilha é criada. Se um arquivo com o nome especificado já existir, ele será sobreescrito. Se selecionar a opção para criar uma planilha, se uma planilha com o nome especificado já existir no arquivo especificado, ela será sobreescrita. Se selecionar a opção para modificar uma planilha existente, deve-se também especificar o nome da planilha. (Isso é opcional para criar uma planilha). Os nomes de planilha não podem exceder 31 caracteres e não podem conter barras ou barras invertidas, colchetes, pontos de interrogação ou asteriscos.

Ao exportar no Excel 97-2004, se modificar uma planilha existente, os gráficos, visualizações do modelo e diagramas de árvore não serão incluídos na saída exportada.

Localização na planilha. Controla a localização dentro da planilha para a saída exportada. Por padrão, a saída exportada será incluída após a última coluna que tiver qualquer conteúdo, iniciando na primeira linha, sem modificar nenhum conteúdo existente. Essa é uma boa opção para incluir novas colunas em uma planilha existente. Incluir saída exportada após a última linha é uma boa opção para incluir novas linhas em uma planilha existente. Incluir saída exportada iniciando em uma localização de célula específica sobreescrivendo qualquer conteúdo existente na área em que a saída exportada for incluída.

Camadas em tabelas dinâmicas. Por padrão, a inclusão ou a exclusão de camadas de tabela dinâmica é controlada pelas propriedades da tabela de cada tabela dinâmica. É possível substituir essa configuração e incluir todas as camadas ou excluir todas, exceto a camada visível atualmente. Veja o tópico [“Propriedades de tabela: impressão” na página 133](#) para obter mais informações.

Incluir notas de rodapé e legendas. Controla a inclusão ou a exclusão de todas as notas de rodapé e legendas de tabela dinâmica.

Visualizações de Modelos. Por padrão, a inclusão ou a exclusão de visualizações do modelo é controlada pelas propriedades de modelo de cada modelo. É possível substituir essa configuração e incluir todas as visualizações ou excluir todas, exceto a visualização aberta atualmente. (Nota: todas as visualizações do modelo, incluindo tabelas, são exportadas como gráficos).

Para configurar as opções de exportação do Excel

1. Selecione **Excel** como o formato de exportação.
2. Clique em **Alterar Opções**.

Opções do PowerPoint

As opções a seguir estão disponíveis para o PowerPoint:

Camadas em tabelas dinâmicas. Por padrão, a inclusão ou a exclusão de camadas de tabela dinâmica é controlada pelas propriedades da tabela de cada tabela dinâmica. É possível substituir essa configuração e incluir todas as camadas ou excluir todas, exceto a camada visível atualmente. Veja o tópico [“Propriedades de tabela: impressão” na página 133](#) para obter mais informações.

Tabelas Dinâmicas Largas. Controla o tratamento das tabelas que forem muito largas para a largura do documento definida. Por padrão, a tabela é quebrada para ajustar-se. A tabela é dividida em seções, e rótulos de linha são repetidos para cada seção da tabela. Como alternativa, é possível reduzir tabelas largas ou não fazer nenhuma mudança nessas tabelas e permitir que elas se estendam para além da largura do documento definida.

Incluir notas de rodapé e legendas. Controla a inclusão ou a exclusão de todas as notas de rodapé e legendas de tabela dinâmica.

Usar entradas de estrutura de tópicos do Visualizador como títulos de slide. Inclui um título em cada slide que for criado pela exportação. Cada slide contém um item único que é exportado do Visualizador. O título é formado a partir da entrada de estrutura de tópicos para o item na área de janela da estrutura de tópicos do Visualizador.

Visualizações de Modelos. Por padrão, a inclusão ou a exclusão de visualizações do modelo é controlada pelas propriedades de modelo de cada modelo. É possível substituir essa configuração e incluir todas as visualizações ou excluir todas, exceto a visualização aberta atualmente. (Nota: todas as visualizações do modelo, incluindo tabelas, são exportadas como gráficos).

Configuração da Página para Exportação. Isso abre um diálogo no qual é possível definir o tamanho e as margens da página para o documento exportado. A largura do documento usada para determinar o comportamento de quebra ou de redução é a largura da página menos as margens esquerda e direita.

Para configurar as opções de exportação do PowerPoint

1. Selecione **PowerPoint** como o formato de exportação.
2. Clique em **Alterar Opções**.

Nota: Exportar para PowerPoint está disponível somente em sistemas operacionais Windows.

opções de PDF

As opções a seguir estão disponíveis para PDF:

Favoritos Integrados Essa opção inclui favoritos no documento PDF que correspondem às entradas de estrutura de tópicos do Visualizador. Assim como a área de janela da estrutura de tópicos do Visualizador, os favoritos podem facilitar muito navegar pelos documentos com um número grande de objetos de saída.

Fontes Integradas. As fontes integradas asseguram que o documento PDF tenha a mesma aparência em todos os computadores. Caso contrário, se algumas fontes usadas no documento não estiverem disponíveis no computador que estiver sendo usado para visualizar (ou imprimir) o documento PDF, a substituição de fonte poderá gerar resultados insatisfatórios.

Camadas em tabelas dinâmicas. Por padrão, a inclusão ou a exclusão de camadas de tabela dinâmica é controlada pelas propriedades da tabela de cada tabela dinâmica. É possível substituir essa configuração e incluir todas as camadas ou excluir todas, exceto a camada visível atualmente. Veja o tópico [“Propriedades de tabela: impressão”](#) na página 133 para obter mais informações.

Visualizações de Modelos. Por padrão, a inclusão ou a exclusão de visualizações do modelo é controlada pelas propriedades de modelo de cada modelo. É possível substituir essa configuração e incluir todas as visualizações ou excluir todas, exceto a visualização aberta atualmente. (Nota: todas as visualizações do modelo, incluindo tabelas, são exportadas como gráficos).

Para configurar as opções de exportação em PDF

1. Selecione **Portable Document Format** como o formato de exportação.
2. Clique em **Alterar Opções**.

Outras Configurações Que Afetam a Saída PDF

Configuração e Atributos de Página. O tamanho da página, orientação, margens, conteúdo e exibição dos cabeçalhos e notas de rodapé da página e o tamanho do gráfico impresso em documentos PDF são controlados pelas opções de configuração e de atributo da página.

Propriedades da Tabela/TableLooks. O ajuste de escala de tabelas largas e/ou longas e a impressão de camadas de tabela são controlados pelas propriedades de cada tabela. Essas propriedades também podem ser salvas nos TableLooks.

Impressora Padrão/Atual. A resolução (DPI) do documento PDF é a configuração de resolução atual para a impressora padrão ou atualmente selecionada (que pode ser alterada usando Configuração de Página). A resolução máxima é 1200 DPI. Se a configuração da impressora for superior, a resolução do documento PDF será 1200 DPI.

Nota: Documentos de alta resolução podem gerar baixos resultados quando impressos em impressoras de resolução inferior.

Opções de texto

Configurando opções de saída de texto

As opções a seguir estão disponíveis para exportação de texto:

Formato de Tabela Dinâmica. Tabelas dinâmicas podem ser exportadas em formato separado por tabulação ou separado por espaço. Para o formato separado por espaço, também é possível controlar:

- **Largura da Coluna. Autoajuste** não quebra nenhum conteúdo da coluna, e cada coluna é tão larga quanto o rótulo ou valor mais largo nessa coluna. **Customizado** configura uma largura da coluna máxima que é aplicada a todas as colunas na tabela, e os valores que excederem essa largura são quebrados na próxima linha nessa coluna.
- **Caractere de Borda Linha/Coluna.** Controla os caracteres usados para criar bordas de linha e de coluna. Para suprimir a exibição de bordas de linha e de coluna, insira espaços em branco para os valores.

Camadas em tabelas dinâmicas. Por padrão, a inclusão ou a exclusão de camadas de tabela dinâmica é controlada pelas propriedades da tabela de cada tabela dinâmica. É possível substituir essa configuração e incluir todas as camadas ou excluir todas, exceto a camada visível atualmente. Veja o tópico “[Propriedades de tabela: impressão](#)” na [página 133](#) para obter mais informações.

Incluir notas de rodapé e legendas. Controla a inclusão ou a exclusão de todas as notas de rodapé e legendas de tabela dinâmica.

Visualizações de Modelos. Por padrão, a inclusão ou a exclusão de visualizações do modelo é controlada pelas propriedades de modelo de cada modelo. É possível substituir essa configuração e incluir todas as visualizações ou excluir todas, exceto a visualização aberta atualmente. (Nota: todas as visualizações do modelo, incluindo tabelas, são exportadas como gráficos).

Para configurar as opções de exportação de texto

1. Selecione **Texto** como o formato de exportação.
2. Clique em **Alterar Opções**.

Imagens apenas opções

As opções a seguir estão disponíveis para exportação de imagens apenas:

1. Faça a guia Workbook a guia ativa (clique em qualquer lugar da guia).
2. Selecione os itens que deseja exportar, clique no controle de elipse vertical e, em seguida, selecione **Exportar** no menu. Você também pode clicar no botão **Exportar** na barra de ferramentas. O diálogo **Exportar workbook** exibe.

Configuração do documento

Tipo

Selecione **Imagen apenas** como o documento **Tipo**.

Formato da imagem

Selecione um formato de imagem para imagens exportadas.

PNG

Quando selecionadas, as imagens são exportadas no formato Portable Network Graphics (*.png). A configuração fornece opções para seleção da imagem **Escala (%)** e **Color profundidade**.

JPG

Quando selecionados, as imagens são exportadas em formato Joint Photographic Experts Group (*.jpg). A configuração fornece opções para selecionar a imagem **Scale (%)** e **Convert to grayscale**.

TIFF

Quando selecionadas, as imagens são exportadas no formato de Tag Image File Format (*.tiff). A configuração fornece opções para selecionar a imagem **Escala (%)**.

BMP

Quando selecionadas, as imagens são exportadas em formato bitmap (*.bmp). A configuração fornece opções para selecionar a imagem **Scale (%)** e **Compress image para reduzir o tamanho do arquivo**.

Abra a pasta que contém o arquivo

Opcionalmente, selecione a configuração para abrir a pasta de localização do arquivo salvar exportação após a exportação.

Objetos a serem exportados

Objetos a serem incluídos

Selecionar para exportar **Todos** ou **Selecionados** (a configuração padrão) objetos de saída.

3. Clique em **Exportar** após especificar a imagem apenas configurações de exportação.

Opções de formato de gráficos

Para documentos HTML e de texto e somente para exportação de gráficos, é possível selecionar o formato gráfico e, para cada formato gráfico, é possível controlar várias configurações opcionais.

Para selecionar o formato gráfico e as saídas para os gráficos exportados:

1. Selecione **HTML**, **Texto**, ou **Nenhum (somente Gráficos)** como o tipo de documento.
2. Selecione o formato de arquivo gráfico na lista suspensa.
3. Clique em **Alterar Opções** para alterar as opções para o formato de arquivo gráfico selecionado.

Opções de exportação de gráfico JPEG

Tamanho da imagem (%)

Porcentagem do tamanho do gráfico original, até 200%.

Converter para escala de cinzas

Converte cores em tons de cinza.

Opções de exportação de gráfico BMP

Tamanho da imagem (%)

Porcentagem do tamanho do gráfico original, até 200%.

Imagem compacta para reduzir o tamanho do arquivo

Uma técnica de compactação sem perdas que cria arquivos menores sem afetar a qualidade da imagem.

Opções de exportação de gráfico PNG

Tamanho da imagem (%)

Porcentagem do tamanho do gráfico original, até 200%.

Profundidade de Cor

Determina o número de cores no gráfico exportado. Um gráfico que é salvo com qualquer intensidade terá um número mínimo de cores que são realmente utilizadas e um número máximo de cores que são permitidas na intensidade. Por exemplo, se o gráfico contém três cores (vermelho, branco e preto) e você o salva como 16 cores, o gráfico permanecerá como três cores.

Nota: Se o número de cores no gráfico exceder o número de cores para essa intensidade, as cores serão misturadas para replicar as cores no gráfico.

Opções de exportação de gráfico EMF e TIFF

Tamanho da imagem. Porcentagem do tamanho do gráfico original, até 200%.

Note: O formato EMF (metarquivo aprimorado) está disponível somente em sistemas operacionais Windows.

Opções de exportação de gráfico EPS

Tamanho da imagem. É possível especificar o tamanho como uma porcentagem do tamanho da imagem original (até 200%), ou especificar uma largura da imagem em pixels (com a altura determinada pelo valor de largura e a proporção de aspecto). A imagem exportada é sempre proporcional ao original.

Incluir imagem de visualização TIFF. Salva uma visualização com a imagem EPS em formato TIFF para exibição em aplicativos que não podem exibir imagens EPS na tela.

Fontes. Controla o tratamento das fontes em imagens EPS.

- **Usar referências de fonte.** Se as fontes que forem usadas no gráfico estiverem disponíveis no dispositivo de saída, as fontes serão utilizadas. Caso contrário, o dispositivo de saída usará fontes alternativas.
- **Substituir fontes por curvas.** Transforma fontes em dados de curva do PostScript. O texto em si não é mais editável como texto em aplicativos que podem editar gráficos EPS. Essa opção será útil se as fontes que forem usadas no gráfico não estiverem disponíveis no dispositivo de saída.

Impressão do visualizador

Há duas opções para impressão de conteúdos da janela Visualizador:

Todas as saídas visíveis. Imprime apenas itens que são atualmente exibidos na área de janela de conteúdo. Os itens ocultos (itens com um ícone de livro fechado na área de janela da estrutura de tópicos ou ocultos nas estratos de estrutura de tópicos reduzidas) não são impressos.

Seleção. Imprime apenas itens que estão atualmente selecionados na estrutura de tópicos e/ou nas áreas de janela de conteúdo.

Para imprimir saída e gráficos

1. Torne o Visualizador a janela ativa (clique em qualquer lugar na janela).
2. No menu, escolha:

Arquivo > Imprimir...

3. Selecione as configurações de impressão desejadas.
4. Clique em **OK** para imprimir.

Visualização da impressão

Visualização de Impressão mostra o que será impresso em cada página dos documentos do Visualizador. É recomendado verificar a Visualização de Impressão antes de realmente imprimir um documento do Visualizador, já que a Visualização de Impressão mostra itens que podem não estar visíveis ao observar a área de janela de conteúdo do Visualizador, incluindo:

- Quebras de página
- Camadas ocultas de tabelas dinâmicas
- Quebras em tabelas largas
- Cabeçalhos e rodapés que são impressos em cada página.

Se qualquer saída estiver selecionada atualmente no Visualizador, a visualização exibirá somente a saída selecionada. Para obter uma visualização de toda a saída, assegure-se de que nada esteja selecionado no Visualizador.

Atributos de Página: Cabeçalhos e Rodapés

Cabeçalhos e rodapés são as informações que são impressas na parte superior e inferior de cada página. É possível inserir qualquer texto que desejar utilizar como cabeçalhos e rodapés. Também é possível usar a barra de ferramentas no meio da caixa de diálogo para inserir:

- Data e hora
- Números de página
- Nome do arquivo do visualizador
- Rótulos de título de estrutura de tópicos
- Títulos e subtítulos de página
- **Tornar Padrão** usa as configurações especificadas como as configurações padrão para novos documentos de Visualizador. (Nota: isso torna as configurações atuais na guia Cabeçalho/Rodapé e na guia Opções as configurações padrão).
- Rótulos de cabeçalho de estrutura de tópico indicam cabeçalho de estrutura de tópicos de primeiro, segundo e terceiro nível para o primeiro item em cada página.
- Títulos e subtítulos de página imprimem os títulos e subtítulos da página atual. Esses podem ser criados com o Novo Título da Página no menu Inserção de Visualizador ou com os comandos TITLE e SUBTITLE. Se você não tiver especificado nenhum título ou subtítulo de página, essa configuração será ignorada.

Nota: Características de fonte para novos títulos e subtítulos de página são controladas na guia Visualizador da caixa de diálogo Opções (acessada ao escolher Opções no menu Editar). Características de fonte para títulos e subtítulos de página existente podem ser alteradas ao editar os títulos no Visualizador.

Para ver a aparência dos cabeçalhos e rodapés na página impressa, escolha Visualização de Impressão no menu Arquivo.

Para inserir cabeçalhos e rodapés de página

1. Torne o Visualizador a janela ativa (clique em qualquer lugar na janela).
2. Nos menus, escolha:
Arquivo > Cabeçalho e Rodapé...
3. Insira o cabeçalho e/ou o rodapé que deseja que apareça em cada página.

Atributos da página: opções

Essa caixa de diálogo controla o tamanho do gráfico impresso, o espaço entre os itens de saída impressos e a numeração de página.

- **Tamanho do Gráfico Impresso.** Controla o tamanho do gráfico impresso relativo ao tamanho da página definido. A proporção de aspecto do gráfico (razão entre largura e altura) não é afetada pelo tamanho do gráfico impresso. O tamanho impresso geral de um gráfico é limitado pela sua altura e largura. Quando as bordas externas de um gráfico atingem as bordas esquerda e direita da página, o tamanho do gráfico não pode aumentar ainda mais para preencher mais altura da página.
- **Espaço entre os itens.** Controla o espaço entre os itens impressos. Cada tabela dinâmica, gráfico e objeto de texto é um item separado. Essa configuração não afeta a exibição dos itens no Visualizador.
- **Número de páginas iniciando com.** Numera as páginas sequencialmente, iniciando com o número especificado.
- **Tornar o Padrão.** Essa opção usa as configurações que são especificadas como as configurações padrão para os novos documentos do Visualizador. (Observe que essa configuração torna padrão as configurações atuais de Cabeçalho/Nota de rodapé e as configurações de Opções)

Para alterar o tamanho do gráfico impresso, a numeração de página e o espaço entre os itens impressos

1. Torne o Visualizador a janela ativa (clique em qualquer lugar na janela).
2. No menu, escolha:

Arquivo > Atributos da página...

3. Clique na guia **Opções**.
4. Altere as configurações e clique em **OK**.

Salvar saída

O conteúdo do Visualizador pode ser salvo.

- **Objeto de saída (*.cou)**. Esse formato salva o contêiner de saída inteiro, incluindo o gráfico, guias, anotações, e assim por diante. Esse formato pode ser aberto e visualizado no IBM SPSS Modelador, incluído em projetos e publicado e controlado usando o IBM SPSS Repositório de serviços de colaboração e implantação. Esse formato não é compatível com o IBM SPSS Estatísticas .
- **Arquivos do espectador (*.spv)**. O formato que é usado para exibir arquivos na janela **Viewer**. Ao salvar nesse formato em um nugget do modelo do IBM SPSS Modelador, somente o conteúdo do Visualizador na guia **Modelo** é salvo.

Para controlar as opções para salvar relatórios da web ou salvar resultados em outros formatos (por exemplo, texto, Word, Excel), use **Exportar** no menu **Arquivo**.

Salvando um Viewer documento

1. Nos menus da janela do Visualizador escolha:

Arquivo > Salvar

2. Insira o nome do documento e, em seguida, clique em **Salvar**.

Opcionalmente, é possível fazer o seguinte:

Bloqueie arquivos para evitar edição no IBM SPSS Leitor inteligente

Se um documento Visualizador estiver bloqueado, será possível manipular tabelas dinâmicas (trocar linhas e colunas, alterar a estrato de exibição, etc.), mas não será possível editar nenhuma saída ou salvar alguma mudança no documento Visualizador IBM SPSS Leitor inteligente (um produto separado para trabalhar com os documentos do Visualizador). Essa configuração não tem efeitos em documentos do Visualizador abertos no IBM SPSS Estatísticas ou no IBM SPSS Modelador.

Criptografar arquivos com uma senha

É possível proteger informações confidenciais armazenadas em um documento do Visualizador ao criptografar o documento com uma senha. Uma vez criptografado, o documento poderá ser aberto somente mediante fornecimento de senha. os usuários do IBM SPSS Leitor inteligente também poderão precisar fornecer a senha para abrir o arquivo.

Para criptografar um documento do Visualizador:

- a. Selecione **Criptografar arquivo com senha** na caixa de diálogo Salvar Saída Como.
- b. Clique em **Salvar**.
- c. Na caixa de diálogo Criptografar Arquivo, forneça uma senha e a insira novamente na caixa de texto Confirmar senha. As senhas são limitadas a 10 caracteres e fazem distinção entre maiúsculas e minúsculas.

Aviso: As senhas não poderão ser recuperadas se elas forem perdidas. Se a senha for perdida, o arquivo não poderá ser aberto.

Criando senhas fortes

- Use oito ou mais caracteres.
- Inclua números, símbolos e até mesmo pontuação em sua senha.
- Evite sequências de números ou caracteres, como "123" e "abc", e evite repetição, como "111aaa".
- Não crie senhas que usem informações pessoais, como aniversários e apelidos.
- Altere a senha periodicamente.

Nota: Armazenar arquivos criptografados em um IBM SPSS Repositório de serviços de colaboração e implantação não é suportado.

Modificando arquivos criptografados

- Se você abrir um arquivo criptografado, faça modificações nele e escolha Arquivo > Salvar, o arquivo modificado será salvo com a mesma senha.
- É possível alterar a senha em um arquivo criptografado ao abrir o arquivo, repetir os passos para criptografá-lo e especificar uma senha diferente na caixa de diálogo Criptografar Arquivo.
- É possível salvar uma versão não criptografada de um arquivo criptografado ao abrir o arquivo, escolhendo Arquivo > Salvar Como e cancelar a seleção de **Criptografar arquivo com senha** na caixa de diálogo Salvar Saída Como.

Nota: Arquivos de dados criptografados e documentos de saída não podem ser abertos em versões do IBM SPSS Estatísticas anteriores a 21. Arquivos de sintaxe criptografados não podem ser abertos em versões anteriores a 22.

Armazenar informações do modelo requerido no documento de saída

Essa opção se aplica somente quando houver itens do visualizador de modelo no documento de saída que requerem informações auxiliares que permitem alguns dos recursos interativos. Clique em **Mais Informações** para exibir uma lista desses itens de visualizador de modelo e os recursos interativos que requerem informações auxiliares. Armazenar essas informações com o documento de saída pode aumentar substancialmente o tamanho do documento. Se optar por não armazenar essas informações, ainda será possível abrir esses itens de saída, mas as variáveis interativas especificadas não estarão disponíveis.

Tabelas dinâmicas

Tabelas dinâmicas

Muitos resultados são apresentados nas tabelas que podem ser articuladas interativamente. Ou seja, é possível reorganizar as linhas, colunas e estratos.

Manipulando uma tabela dinâmica

As opções para manipular uma tabela dinâmica incluem:

- Transpondo linhas e colunas
- Movendo linhas e colunas
- Criando camadas multidimensionais
- Agrupar e desagrupar linhas e colunas
- Mostrar e ocultar linhas, colunas e outras informações
- Girar rótulos de linha e de coluna
- Localizar definições de termos

Ativando uma tabela pivot

Antes de poder manipular ou modificar uma tabela dinâmica, é necessário **ativar** a tabela. Para ativar uma tabela:

1. Dê um clique duplo na tabela.
ou
2. Clique com o botão direito na tabela e no menu pop-up escolha **Editar Conteúdo**.
3. No submenu, escolha **No Visualizador** ou **Na janela separada**.

Definindo uma tabela como dinâmica

Uma tabela possui três dimensões: linhas, colunas e estratos. Uma dimensão pode conter diversos elementos (ou nenhum elemento). É possível alterar a organização da tabela ao mover os elementos entre ou dentro das dimensões. Para mover um elemento, apenas arraste-o e solte-o no lugar desejado.

Alterando a ordem de exibição de elementos dentro de uma dimensão

Para alterar a ordem de exibição de elementos dentro de uma dimensão de tabela (linha, coluna ou estrato):

1. Se as bandejas de dinamização ainda não estiverem ativas, no menu de Tabela Dinâmica escolha: **Dinâmica > Bandejas de dinamização**
2. Arraste e solte os elementos dentro da dimensão na bandeja de dinamização.

Movendo linhas e colunas dentro de um elemento de dimensão

1. Na própria tabela (não em bandejas de dinamização), clique no rótulo para a linha ou coluna que deseja mover.
2. Arraste o rótulo para a nova posição.

Transpondo linhas e colunas

Se desejar apenas inverter as linhas e colunas, há uma alternativa simples para uso das bandejas de dinamização:

1. No menu, escolha:

Tabela Dinâmica > Transpor Linhas e Colunas

Isso possui o mesmo efeito que arrastar todos os elementos da linha para a dimensão da coluna e arrastar todos os elementos da coluna para a nova dimensão.

Agrupando linhas ou colunas

1. Selecione os rótulos das linhas e colunas que deseja agrupar (clique e arraste ou dê um Shift+clique para selecionar diversos rótulos).
2. No menu, escolha:

Editar > Grupo

Um rótulo de grupo é inserido automaticamente. Clique duas vezes no rótulo de grupo para editar o texto do rótulo.

Nota: Para incluir linhas ou colunas em um grupo existente, primeiro deve-se desagrupar os itens que estiverem atualmente no grupo. Em seguida, é possível criar um novo grupo que inclua os itens adicionais.

Desagrupando linhas ou colunas

Desagrupar automaticamente exclui o rótulo de grupo.

Girando rótulos de linha e de coluna

É possível girar rótulos entre a exibição horizontal e vertical para os rótulos de coluna mais internos e para os rótulos de linha mais externos em uma tabela.

1. No menu, escolha:

Formatar > Girar rótulos da coluna interna

ou

Formatar > Girar rótulos da coluna externa

Somente os rótulos de coluna mais internos e os rótulos de linha mais externos podem ser girados.

classificando linhas

Para ordenar as linhas de uma tabela dinâmica:

1. Ative a tabela.
2. Selecione qualquer célula na coluna que deseja usar para ordenar. Para ordenar somente um grupo de linhas selecionado, selecione duas ou mais células contínuas na coluna que deseja usar para ordenar.
3. Nos menus, escolha:

Editar > Classificar linhas

4. Selecione **Crescente** ou **Decrescente** no submenu.
 - Se a dimensão da linha contiver grupos, a ordenação afetará somente o grupo que contiver a seleção.
 - Não é possível ordenar entre os limites de grupo.
 - Não é possível ordenar tabelas com mais de um item na dimensão da linha.

Inserindo linhas e colunas

Para inserir uma linha ou coluna em uma tabela dinâmica:

1. Ative a tabela.
2. Selecione qualquer célula na tabela.
3. Nos menus, escolha:

Inserir antes

ou

Inserir depois

No submenu, escolha:

Linha

ou

Coluna

- Um sinal de mais (+) é inserido em cada célula da nova linha ou coluna para evitar que a nova linha ou coluna seja ocultada automaticamente por ela estar vazia.
- Em uma tabela com dimensões aninhadas ou em estratos, uma coluna ou linha é inserida em cada nível de dimensão correspondente.

Controlando a exibição de rótulos de variáveis e de valor

Se variáveis contiverem rótulos de variável ou de valor descritivos, será possível controlar a exibição de nomes e rótulos de variável, de valores de dados e de rótulos de valor em tabelas dinâmicas.

1. Ative a tabela dinâmica.

2. Nos menus, escolha:

Visualização > Rótulos variáveis

ou

Visualização > Rótulos de valor

3. Selecione uma das opções a seguir no submenu:

- **Nome** ou **Valor**. Somente nomes de variável (ou valores) são exibidos. Rótulos descritivos não são exibidos.
- **Rótulo**. Somente rótulos descritivos são exibidos. Nomes de variável (ou valores) não são exibidos.
- **Ambos**. Nomes (ou valores) e rótulos descritivos são exibidos.

Alterando o idioma de saída

Para alterar o idioma de saída em uma tabela dinâmica:

1. Ative a tabela

2. Nos menus, escolha:

Visualizar > Idioma

3. Selecione um dos idiomas disponível.

Alterar o idioma afeta somente o texto que é gerado pelo aplicativo, como títulos de tabela, rótulos de linha e de coluna e o texto da nota de rodapé. Os nomes de variável e os rótulos de variável e de valor descritivos não são afetados.

Navegando em tabelas grandes

Para usar a janela de navegação para navegar em tabelas grandes:

1. Ative a tabela.

2. No menu, escolha:

Visualizar > Navegação

Desfazendo mudanças

É possível desfazer a maioria da mudança recente ou todas as mudanças feitas em uma tabela dinâmica ativada. Ambas as ações se aplicam somente às mudanças feitas desde a ativação mais recente da tabela.

Para desfazer a maioria da mudança recente:

1. Nos menus, escolha:

Editar > Desfazer

Para desfazer todas as mudanças:

2. Nos menus, escolha:

Editar > Restaurar

Trabalhando com estratos

É possível exibir uma tabela bidimensional separada para cada categoria ou combinação de categorias. Imagine a tabela como sendo empilhada em estratos, com apenas a estrato superior visível.

Criando e exibindo estratos

Para criar estratos:

1. Se as bandejas de dinamização ainda não estiverem ativas, no menu **Tabela Dinâmica** escolha:

Tabela dinâmica > Bandejas de dinamização

2. Arraste um elemento da dimensão da linha ou de coluna para a dimensão da estrato.

Mover os elementos para a dimensão da estrato cria uma tabela multidimensional, mas somente uma única "fatia" bidimensional é exibida. A tabela visível é a tabela para a estrato superior. Por exemplo, se uma variável categórica de sim/não estiver na dimensão da estrato, então a tabela multidimensional terá duas estratos: uma para a categoria "sim" e outra para a categoria "não".

Alterando a estrato exibida

1. Escolha uma categoria na lista suspensa de estratos (na própria tabela dinâmica, não na bandeja de dinamização).

Acesse a categoria de camada

Acessar Categoria de Camada permite alterar camadas em uma tabela dinâmica. Essa caixa de diálogo é útil principalmente quando houver muitas camadas ou quando a camada selecionada possuir muitas categorias.

Mostrando e ocultando itens

Muitos tipos de células podem ser ocultos, incluindo:

- Rótulos de dimensão
- Categorias, incluindo a célula de rótulo e células de dados em uma linha ou coluna
- Rótulos de categoria (sem ocultar as células de dados)
- Notas de rodapé, títulos e legendas

Ocultando linhas e colunas em uma tabela

Mostrando linhas e colunas ocultas em uma tabela

1. No menu, escolha:

Visualizar > Mostrar Todas as Categorias

Isso exibe todas as linhas e colunas ocultas na tabela. (Se **Ocultar linhas e colunas vazias** for selecionada em Propriedades da Tabela para essa tabela, uma linha ou coluna totalmente vazia permanecerá oculta).

Ocultando e mostrando rótulos de dimensão

1. Selecione o rótulo de dimensão ou qualquer rótulo de categoria dentro da dimensão.

2. No menu Visualizar ou no menu pop-up, escolha **Ocultar Rótulo de Dimensão** ou **Mostrar Rótulo de Dimensão**.

Ocultando e mostrando títulos de tabela

Para ocultar um título:

1. Ative a tabela dinâmica.
2. Selecione o título.
3. No menu Visualizar, selecione **Ocultar**.

Para mostrar títulos ocultos:

4. No menu Visualizar, selecione **Mostrar Tudo**.

TableLooks

Um TableLook é um conjunto de propriedades que definem a aparência de uma tabela. É possível selecionar um TableLook definido anteriormente ou criar seu próprio TableLook.

- Antes ou após um TableLook ser aplicado, é possível alterar a formatação de célula para células individuais ou de grupos de células ao usar as propriedades de célula. Os formatos de célula editados permanecem intactos, mesmo quando aplicar um novo TableLook.
- Opcionalmente, é possível reconfigurar todas as células para os formatos de célula que estiverem definidos no TableLook atual. Esta opção reconfigura quaisquer células que tiverem sido editadas. Se **Conforme Exibido** for selecionada na lista Arquivos do TableLook, quaisquer células editadas serão reconfiguradas para as propriedades da tabela atuais.
- Apenas as propriedades da tabela que são definidas no diálogo Propriedades da Tabela são salvas em TableLooks. TableLooks não incluem modificações de célula individual.

Para aplicar uma TableLook

1. Ative uma tabela dinâmica.
2. Nos menus, escolha:

Formato > TableLooks...

3. Selecione um TableLook na lista de arquivos. Para selecionar um arquivo a partir de outro diretório, clique em **Navegar**.
4. Clique em **OK** para aplicar o TableLook à tabela dinâmica selecionada.

Para editar ou criar um TableLook

1. Na caixa de diálogo TableLooks, selecione um TableLook na lista de arquivos.
 2. Clique em **Editar verificação**.
 3. Ajuste as propriedades da tabela para os atributos que você deseja e, em seguida, clique em **OK**.
 4. Clique em **Salvar verificação** para salvar o TableLook editado ou clique em **Salvar como** para salvá-lo como um novo TableLook.
- Editar um TableLook afeta apenas a tabela dinâmica selecionada. Um TableLook editado não é aplicado em qualquer outra tabela que usa este TableLook, a menos que essas tabelas sejam selecionadas e o TableLook seja reaplicado.
 - Apenas as propriedades da tabela que são definidas no diálogo Propriedades da Tabela são salvas em TableLooks. TableLooks não incluem modificações de célula individual.

Propriedades da tabela

Propriedades da Tabela permite configurar propriedades gerais de uma tabela e configurar estilos de célula para várias partes de uma tabela. Será possível:

- Controlar propriedades gerais, como ocultar linhas ou colunas vazias e ajustar propriedades de impressão.
- Controlar o formato e a posição de marcadores de nota de rodapé.
- Determinar formatos específicos para células na área de dados, para rótulos de linha e de coluna, e para outras áreas da tabela.
- Controlar a largura e a cor das linhas que formam as bordas de cada área da tabela.

Para alterar as propriedades da tabela dinâmica:

1. No menu, escolha:

Formato > Propriedades da tabela...

2. Selecione uma guia (**Geral, Notas de Rodapé, Formatos de Célula, Bordas ou Impressão**).
3. Selecione as opções desejadas.
4. Clique em **OK** ou em **Aplicar**.

As novas propriedades são aplicadas à tabela dinâmica selecionada.

Propriedades da tabela: geral

Várias propriedades se aplicam à tabela como um todo. Será possível:

- Mostrar ou ocultar linhas e colunas vazias. (Uma linha ou coluna vazia não possui nada em nenhuma das células de dados).
- Controlar o posicionamento dos rótulos de linha, que podem estar no canto superior esquerdo ou aninhados.
- Controlar as larguras de coluna máxima e mínima (expressas em pontos).

Para alterar propriedades gerais da tabela:

1. Clique na guia **Geral**.
2. Selecione as opções desejadas.
3. Clique em **OK** ou em **Aplicar**.

Configurar linhas para exibição

Nota: Esse recurso se aplica somente a tabelas anteriores.

Por padrão, tabelas com muitas linhas são exibidas em seções de 100 linhas. Para controlar o número de linhas exibidas em uma tabela:

1. Selecione **Exibir tabelas por linha**.
 2. Clique em **Configurar Linhas para Exibição**.
- ou
3. No menu Visualizar de uma tabela dinâmica ativada, escolha **Exibir tabela por linhas** e **Configurar Linhas para Exibição**.

Linhas para exibição. Controla o número máximo de linhas para exibir de uma vez. Os controles de navegação permitem mover para seções diferentes da tabela. O valor mínimo é 10. O padrão é 100.

Tolerância a viúva/órfão. Controla o número máximo de linhas da dimensão da linha mais interna da tabela para dividir entre as visualizações exibidas da tabela. Por exemplo, se houver seis categorias em cada grupo de dimensão da linha mais interna, especificar um valor de seis evita que qualquer grupo seja dividido entre as visualizações exibidas. Essa configuração pode fazer com que o número total de linhas em uma visualização exibida exceda o número máximo especificado de linhas para exibição.

Propriedades da tabela: notas

A guia Notas do diálogo Propriedades de Tabela controla a formatação da nota de rodapé e o texto de comentário da tabela.

Notas de Rodapé. As propriedades dos marcadores de nota de rodapé incluem o estilo e a posição com relação ao texto.

- O estilo dos marcadores de notas de rodapé é o de números (1, 2, 3, ...) ou letras (a, b, c, ...).
- Os marcadores de nota de rodapé podem ser anexados ao texto como sobrescritos ou subscritos.

Texto de Comentário. É possível incluir um texto de comentário em cada tabela.

- O texto de comentário é exibido em uma dica de ferramenta ao passar o mouse sobre uma tabela no Visualizador.
- Os leitores de tela leem o texto de comentário quando a tabela possuir foco.

- A dica de ferramentas no Visualizador exibe apenas os 200 primeiros caracteres do comentário, mas os leitores de tela leem o texto inteiro.
- Ao exportar a saída em HTML, o texto de comentário é usado como um texto alternativo.

Propriedades de tabela: formatos de célula

Para formatação, uma tabela é dividida em áreas: título, camadas, rótulos de canto, rótulos de linha, rótulos de coluna, dados, legenda e notas de rodapé. Para cada área de uma tabela, é possível modificar os formatos de célula associados. Os formatos de célula incluem características de texto (como fonte, tamanho, cor e estilo), alinhamento horizontal e vertical, cores do plano de fundo e margens de célula internas.

Os formatos de célula são aplicados em áreas (categorias de informações). Eles não representam características de células individuais. Essa distinção é uma consideração importante ao definir uma tabela como dinâmica.

Por exemplo,

- Se você especificar uma fonte em negrito como um formato de célula de rótulos de coluna, os rótulos de coluna aparecerão em negrito, independentemente das informações exibidas atualmente na dimensão da coluna. Se você mover um item da dimensão da coluna para outra dimensão, ele não reterá a característica de negrito dos rótulos de coluna.
- Se você colocar os rótulos de coluna em negrito simplesmente ao destacar as células em uma tabela dinâmica ativada e clicar no botão Negrito na barra de ferramentas, o conteúdo dessas células permanecerá em negrito, independentemente da dimensão que for movida, e os rótulos de coluna não reterão a característica de negrito de outros itens movidos para a dimensão de coluna.

Para alterar os formatos de célula:

1. Selecione a guia **Formatos da Célula**.
2. Selecione uma Área na lista suspensa ou clique em uma área da amostra.
3. Selecione características para a área. Suas seleções são refletidas na amostra.
4. Clique em **OK** ou em **Aplicar**.

Alternando cores de linha

Para aplicar uma cor de plano de fundo ou de texto diferente em linhas alternativas na área de Dados da tabela:

1. Selecione **Dados** na lista suspensa Área.
2. Selecione (marque) **Cor de linha alternativa** no grupo Cor do Plano de Fundo.
3. Selecione as cores a serem usadas para plano de fundo e texto da linha alternativa.

As cores de linha alternativa afetam apenas a área de Dados da tabela. Elas não afetam as áreas de rótulo de linha e de coluna.

Propriedades da tabela: bordas

Para cada localização de borda em uma tabela, é possível selecionar um estilo da linha e uma cor. Se selecionar **Nenhum** como o estilo, não haverá nenhuma linha na localização selecionada.

Para alterar bordas da tabela:

1. Clique na guia **Bordas**.
2. Selecione a localização da borda clicando no seu nome na lista ou clicando em uma linha na área Amostra.
3. Selecione um estilo da linha e selecione **Nenhum**.
4. Selecione uma cor.
5. Clique em **OK** ou em **Aplicar**.

Propriedades de tabela: impressão

É possível controlar as propriedades a seguir para tabelas dinâmicas impressas:

- Imprimir todas as camadas ou somente a camada superior da tabela e imprimir cada camada em uma página separada.
- Reduzir uma tabela horizontalmente ou verticalmente para ajustar a página para impressão.
- Controlar linhas viúvas/órfãs ao controlar o número mínimo de linhas e colunas que estarão contidas em qualquer sessão impressa de uma tabela se ela for muito larga e/ou muito longa para o tamanho da página definido.

Nota: Se uma tabela for muito longa para ajustar na página atual porque há outra saída acima dela, mas se ajustando dentro da largura de página definida, a tabela é impressa automaticamente em uma nova página, independentemente da configuração de viúva/órfã.

- Incluir texto de continuação para tabelas que não se ajustarem em uma única página. É possível exibir o texto de continuação na parte inferior e na parte superior de cada página. Se nenhuma opção for selecionada, o texto da continuação não será exibido.

Para controlar as propriedades de impressão de tabela dinâmica:

1. Clique na guia **Impressão**.
2. Selecione as opções de impressão desejadas.
3. Clique em **OK** ou em **Aplicar**.

Propriedades de célula

As propriedades da célula são aplicadas a uma célula selecionada. É possível alterar a fonte, o formato do valor, o alinhamento, margens e cores. As propriedades da célula substituem as propriedades da tabela, portanto, se alterar as propriedade da tabela, não altere nenhuma propriedade da célula aplicada individualmente.

Para alterar as propriedades da célula:

1. Selecione as células na tabela.
2. No menu Formato ou no menu pop-up, escolha **Propriedades da Célula**.

Fonte e segundo plano

A guia Fonte e Segundo Plano controla o estilo e a cor da fonte e a cor do plano de fundo para as células selecionadas na tabela.

Valor de formato

A guia Valor de Formato controla os formatos de valor para as células selecionadas. É possível selecionar formatos para números, datas, hora ou moedas, e também ajustar o número de dígitos decimais que são exibidos.

Alinhamento e margens

A guia Alinhamento e Margens controla o alinhamento horizontal e vertical dos valores e as margens superior, inferior, esquerda e direita para as células selecionadas. O alinhamento horizontal **Combinado** alinha o conteúdo de cada célula de acordo com seu tipo. Por exemplo, as datas são alinhadas à direita e os valores de texto são alinhados à esquerda.

Notas de rodapé e legendas

É possível incluir notas de rodapé e legendas em uma tabela. Também é possível ocultar notas de rodapé ou legendas, alterar os marcadores de nota de rodapé e numerar novamente notas de rodapé.

Incluindo notas de rodapé e legendas

Para incluir uma legenda em uma tabela:

1. No menu Inserir, escolha **Legendas**.

Uma nota de rodapé pode ser anexada a qualquer item e uma tabela. Para incluir uma nota de rodapé:

1. Clique em um título, célula ou legenda dentro de uma tabela dinâmica ativada.
2. No menu Inserir, escolha **Notas de Rodapé**.
3. Insira o texto da nota de rodapé na área fornecida.

Para ocultar e mostrar uma legenda

Para ocultar uma legenda:

1. Selecione a legenda.
2. No menu Visualizar, selecione **Ocultar**.

Para mostrar legendas ocultas:

1. No menu Visualizar, selecione **Mostrar Tudo**.

Para ocultar e mostrar uma nota de rodapé em uma tabela

Para ocultar uma nota de rodapé:

1. Clique com o botão direito na célula que contém a referência de nota de rodapé e selecione **Ocultar Notas de Rodapé** no menu pop-up
ou
2. Selecione a nota de rodapé na área de nota de rodapé da tabela e selecione **Ocultar** no menu pop-up.

Nota: Para tabelas anteriores, selecione área de nota de rodapé da tabela, selecione **Editar Nota de Rodapé** no menu pop-up e, em seguida, desmarque (limpe) a propriedade Visível para quaisquer notas de rodapé que desejar ocultar.

Se uma célula contiver diversas notas de rodapé, use o último método para ocultar seletivamente notas de rodapé.

Para ocultar todas as notas de rodapé em uma tabela

1. Selecione todas as notas de rodapé na área de nota de rodapé da tabela (use clicar e arrastar ou Shift+clique para selecionar as notas de rodapé) e selecione **Ocultar** no menu Visualizar.

Para mostrar notas de rodapé ocultas:

1. Selecione **Mostrar Todas as Notas de Rodapé** no menu Visualizar.

Marcador de nota de rodapé

Marcador de Nota de Rodapé altera as características que podem ser usadas para marcar uma nota de rodapé. Por padrão, os marcadores de nota de rodapé padrão são letras ou números sequenciais, dependendo das configurações das propriedades de tabela. Também é possível designar um marcador especial. Marcadores especiais não são afetados ao renumerar notas de rodapé ou alternar entre números e letras para marcadores padrão. A exibição de números ou letras para marcadores padrão e a posição do subscrito ou sobrescrito dos marcadores de nota de rodapé são controladas pela guia Notas de Rodapé no diálogo Propriedades de Tabela.

Para alterar marcadores de nota de rodapé:

1. Selecione uma nota de rodapé.
2. No menu **Formato**, escolha **Marcador de Nota de Rodapé**.

Marcadores especiais são limitados a 2 caracteres. Notas de rodapé com marcadores especiais precedem notas de rodapé com letras ou números sequenciais na área de nota de rodapé da tabela, de modo que alterar para um marcador especial poderá reordenar a lista de notas de rodapé.

Numerando novamente notas de rodapé

Ao ter definido uma tabela como dinâmica ao alternar linhas, colunas e estratos, as notas de rodapé podem estar fora de ordem. Para numerar novamente as notas de rodapé:

1. No menu Formatar, escolha **Numerar Novamente Notas de Rodapé**.

Editando notas de rodapé em tabelas anteriores

Para tabelas anteriores, é possível usar o diálogo Editar Notas de Rodapé para inserir e modificar as configurações de texto e de fonte da nota de rodapé, alterar marcadores da nota de rodapé e ocultar ou excluir seletivamente notas de rodapé.

Ao inserir uma nova nota de rodapé em uma tabela anterior, o diálogo Editar Notas de Rodapé é aberto automaticamente. Para usar o diálogo Editar Notas de Rodapé para editar notas de rodapé existentes (sem criar uma nova nota de rodapé):

Marcador. Por padrão, os marcadores de nota de rodapé padrão são letras ou números sequenciais, dependendo das configurações das propriedades de tabela. Para designar um marcador especial, simplesmente insira o novo valor de marcador na coluna Marcador. Marcadores especiais não são afetados ao renumerar notas de rodapé ou alternar entre números e letras para marcadores padrão. A exibição de números ou letras para marcadores padrão e a posição do subscrito ou sobrescrito dos marcadores de nota de rodapé são controladas pela guia Notas de Rodapé no diálogo Propriedades de Tabela. Consulte o tópico [“Propriedades da tabela: notas”](#) na página 131 para obter mais informações.

Para alterar um marcador especial de volta para o marcador padrão, clique com o botão direito no marcador no diálogo Editar Notas de Rodapé, selecione **Marcador de Nota de Rodapé** no menu pop-up e selecione o marcador Padrão na caixa de diálogo Marcador de Nota de Rodapé.

Nota de Rodapé. O conteúdo da nota de rodapé. A exibição reflete as configurações atuais de fonte e de segundo plano. As configurações de fonte podem ser alteradas para notas de rodapé individuais usando o subdiálogo Formatar. Consulte o tópico [“Configurações de fonte e de cor de nota de rodapé”](#) na página 135 para obter mais informações. Uma cor do plano de fundo é aplicada a todas as notas de rodapé e pode ser alterada na guia Fonte e Segundo Plano do diálogo Propriedades da Célula. Consulte o tópico [“Fonte e segundo plano”](#) na página 133 para obter mais informações.

Visível. Todas as notas de rodapé estão visíveis por padrão. Limpe (desmarque) a caixa de seleção Visível para ocultar a nota de rodapé.

Configurações de fonte e de cor de nota de rodapé

Para tabelas anteriores, é possível usar o diálogo Formato para alterar a família da fonte, o estilo, o tamanho e a cor de uma ou mais notas de rodapé selecionadas:

1. No diálogo Editar Notas de Rodapé, selecione (clique) em uma ou mais notas de rodapé na grade Notas de Rodapé.
2. Clique no botão **Formatar**.

A família da fonte, estilo, tamanho e cores selecionados são aplicados a todas as notas de rodapé selecionadas.

A cor do plano de fundo, alinhamento e margens podem ser configurados no diálogo Propriedades de Célula e aplicados a todas as notas de rodapé. Não é possível alterar essas configurações para notas de rodapé individuais. Consulte o tópico [“Fonte e segundo plano”](#) na página 133 para obter informações adicionais.

Larguras da célula de dados

Configurar Largura da Célula de Dados é usado para configurar todas as células de dados para a mesma largura.

Para configurar a largura para todas as células de dados:

1. No menu, escolha:

Formato > Configurar larguras de células de dados...

2. Insira um valor para a largura da célula.

Alterando a largura da coluna

1. Clique em arraste a borda da coluna

Exibindo bordas ocultas em uma tabela dinâmica

Para tabelas sem muitas bordas visíveis, é possível exibir as bordas ocultas. Isso pode simplificar tarefas como alterar larguras da coluna.

1. No menu Visualizar, selecione **Linhas de Grade**.

Selecionando linhas, colunas e células em uma tabela dinâmica

É possível selecionar uma linha ou coluna inteira ou um conjunto especificado de células de dados e de rótulo.

Para selecionar diversas células:

Selecionar > Células de Dados e de Rótulo

Imprimindo tabelas dinâmicas

Vários fatores podem afetar o modo de aparência das tabelas dinâmicas impressas, fatores esses que podem ser controlados ao alterar os atributos de tabela dinâmica.

- Para tabelas dinâmicas multidimensionais (tabelas com estratos), é possível imprimir todas as estratos ou somente a estrato superior (visível). Consulte o tópico [“Propriedades de tabela: impressão” na página 133](#) para obter informações adicionais.
- Para tabelas dinâmicas longas e largas, é possível redimensionar automaticamente a tabela para ajustar à página ou controlar a localização das quebras de tabela e das quebras de página. Consulte o tópico [“Propriedades de tabela: impressão” na página 133](#) para obter informações adicionais.

Use Visualização de Impressão no menu Arquivo para ver como as tabelas impressas se parecerão.

Controlando quebras de tabela para tabelas largas e longas

As tabelas dinâmicas que forem muito largas ou muito longas para imprimir dentro do tamanho da página definido são divididas e impressas automaticamente em diversas seções. Será possível:

- Controlar as localizações de linha e de coluna nas quais tabelas grandes são divididas.
- Especificar linhas e colunas que devem ser mantidas juntas quando as tabelas forem divididas.
- Ajustar novamente a escala de tabelas grandes para ajustar o tamanho de página definido.

Para especificar as quebras de linha e de coluna para tabelas dinâmicas impressas:

1. Ative a tabela dinâmica.
2. Clique em uma célula na coluna à esquerda de onde deseja inserir a quebra ou clique em qualquer célula na linha antes da linha em que deseja inserir a quebra.
3. Nos menus, escolha:

Formatar > Pontos de interrupção > Ponto de interrupção vertical

ou

Formato > Pontos de interrupção > Ponto de interrupção horizontal

1. Ative a tabela dinâmica.
2. Clique em uma célula na coluna à esquerda de onde deseja inserir a quebra ou clique em qualquer célula na linha antes da linha em que deseja inserir a quebra.
3. Nos menus, escolha:

Formatar > Pontos de interrupção > Ponto de interrupção vertical

ou

Formato > Pontos de interrupção > Ponto de interrupção horizontal

Para especificar linhas ou colunas para manter juntas:

1. Selecione os rótulos das linhas e colunas que deseja manter juntas. Clique e arraste ou dê um Shift+clique para selecionar diversos rótulos de linha ou de coluna.
2. Nos menus, escolha:

Formatar > Pontos de interrupção > Manter juntos

Para visualizar pontos de interrupção e manter grupos juntos:

1. Nos menus, escolha:

Formatar > Pontos de interrupção > Exibir pontos de interrupção

Os pontos de interrupção são mostrados como linhas verticais ou horizontais. Manter grupos juntos aparece como regiões retangulares esmaecidas delimitadas por uma borda mais escura.

Nota: Exibir pontos de interrupção e manter grupos juntos não são suportados para tabelas anteriores.

Para limpar pontos de interrupção e manter grupos juntos:

Para limpar um ponto de interrupção:

1. Clique em uma célula na coluna à esquerda de um ponto de interrupção vertical ou clique em qualquer célula na linha acima de um ponto de interrupção horizontal.
2. Nos menus, escolha:

Formatar > Pontos de interrupção > Limpar Ponto de interrupção ou Grupo

Para limpar um manter grupo junto:

3. Selecione os rótulos de coluna ou de linha que especificam o grupo.
4. Nos menus, escolha:

Formatar > Pontos de interrupção > Limpar Ponto de interrupção ou Grupo

Todos os pontos de interrupção e manter grupos juntos são limpos automaticamente quando definir como tabela dinâmica ou reordenar qualquer linha ou coluna. Esse comportamento não se aplica a tabelas anteriores.

Criando um gráfico a partir de uma tabela dinâmica

1. Clique duas vezes na tabela dinâmica para ativá-la.
2. Selecione as linhas, colunas ou células que deseja exibir no gráfico.
3. Clique com o botão direito em qualquer lugar na área selecionada.
4. Escolha **Criar Gráfico** no menu pop-up e selecione um tipo de gráfico.

Tabelas anteriores

É possível optar por renderizar tabelas como tabelas anteriores (referidas como tabelas de recursos completos na liberação 19) que são totalmente compatíveis com liberações do IBM SPSS Estatísticas anteriores a 20. Tabelas anteriores podem renderizar lentamente e serão recomendadas somente se a compatibilidade com liberações anteriores a 20 for necessária. Para obter informações sobre como criar tabelas anteriores, consulte “[Opções da tabela dinâmica](#)” na página 139.

Opções

Opções

As opções controlam várias configurações.

Para alterar configurações de opções

1. Nos menus, escolha:

Editar > Opções...

2. Clique nas guias para as configurações que deseja alterar.

3. Altere as configurações.

4. Clique em **OK** ou em **Aplicar**.

Opções gerais

Número Máximo de Encadeamentos

O número de encadeamentos que os procedimentos multiencadeados utilizam ao calcular resultados. A configuração **Automático** baseia-se no número de núcleos de processamento disponíveis. Especifique um valor menor se desejar tornar os recursos de processamento disponíveis para outros aplicativos enquanto procedimentos multiencaminhados estiverem em execução. Esta opção é desativada no modo de análise distribuído.

Saída

Exibir um zero à esquerda para valores decimais. Exibe zeros à esquerda para valores numéricos que consistirem somente em uma parte decimal. Por exemplo, quando zeros à esquerda são exibidos, o valor .123 é exibido como 0,123. Essa configuração não se aplica a valores numéricos que possuírem um formato de moeda ou de porcentagem. Exceto para arquivos ASCII (*.dat) fixos, zeros à esquerda não são incluídos quando os dados são salvos em um arquivo externo.

Sistema de Medição. O sistema de medição usado (pontos, polegadas ou centímetros) para especificar atributos, como margens de célula de tabela dinâmica, larguras de célula e espaço entre tabelas para impressão.

Opções do Visualizador

As opções de exibição de saída do Visualizador afetam somente a nova saída que é produzida após alterar as configurações. A saída que já é exibida no Visualizador não é afetada pelas mudanças nessas configurações.

Estado de saída inicial

Controla quais itens são exibidos ou ocultados automaticamente sempre que executar um procedimento e como os itens são alinhados inicialmente. É possível controlar a exibição dos itens a seguir: log, avisos, notas, títulos, tabelas dinâmicas, gráficos, diagramas de árvore e saída de texto. Também é possível ativar ou desativar a exibição de comandos no log. É possível copiar a sintaxe de comando a partir do log e salvá-la em um arquivo de sintaxe.

Nota: Todos os itens de saída são exibidos alinhados à esquerda no Visualizador. Somente o alinhamento da saída impressa é afetado pelas configurações de justificação. Itens centralizados e alinhados à direita são identificados por um símbolo pequeno.

Título

Controla o estilo, o tamanho e a cor da fonte para novos títulos de saída. A lista de **Tamanho** da fonte fornece um conjunto de tamanhos predefinidos, mas é possível inserir manualmente outros valores de tamanho suportados.

Título da página

Controla o estilo, o tamanho e a cor da fonte para novos títulos de página e títulos de página gerados pela sintaxe de comando 'TITLE e 'SUBTITLE ou criados por **Novo título de página** no menu

Inserir. A lista de **Tamanho** da fonte fornece um conjunto de tamanhos predefinidos, mas é possível inserir manualmente outros valores de tamanho suportados.

Saída de Texto

A fonte usada para a saída de texto. A saída de texto é projetada para uso com uma fonte monoespaçada (pitch fixo). Se selecionar uma fonte proporcional, a saída tabular não estará alinhada corretamente. A lista de **Tamanho** da fonte fornece um conjunto de tamanhos predefinidos, mas é possível inserir manualmente outros valores de tamanho suportados.

Configuração Padrão da Página

Controla as opções padrão para orientação e margens para impressão.

Opções da tabela dinâmica

As opções de Tabela Dinâmica configuram várias opções para a exibição de tabelas dinâmicas.

TableLook

Selecione um TableLook na lista de arquivos e clique em **OK** ou em **Aplicar**. É possível usar uma das TableLooks fornecidas com o IBM SPSS Estatísticas ou criar sua própria no Editor de Tabela Dinâmica (escolha **TableLooks** no menu Formato).

- **Procurar.** Permite selecionar uma TableLook a partir de outro diretório.
- **Configurar Diretório do TableLook.** Permite alterar o diretório do TableLook padrão. Use **Procurar** para navegar para o diretório que deseja utilizar, selecione um TableLook nesse diretório e, em seguida, selecione **Configurar Diretório do TableLook**.

Nota: os TableLooks criados em versões anteriores ao IBM SPSS Estatísticas não podem ser usados na versão 16.0 ou posterior.

Larguras da coluna

Essas opções controlam o ajustamento automático das larguras da coluna nas tabelas dinâmicas.

- **Ajuste para rótulos somente.** Ajusta a largura da coluna para a largura do rótulo da coluna. Isso produz tabelas mais compactas, mas valores de dados mais largos do que o rótulo podem ser truncados.
- **Ajuste para rótulos e dados para todas as tabelas.** Ajusta a largura da coluna para o que for mais largo: o rótulo da coluna ou o valor dos dados maior. Isso produz tabelas mais largas, mas assegura que todos os valores sejam exibidos.

Modo de Edição Padrão

Essa opção controla a ativação das tabelas dinâmicas na janela do Visualizador ou em uma janela separada. Por padrão, clicar duas vezes em uma tabela dinâmica ativa todas as tabelas, exceto as muito grandes, na janela do Visualizador. É possível optar por ativar tabelas dinâmicas em uma janela separada ou selecionar uma configuração de tamanho que abrirá tabelas dinâmicas menores na janela do Visualizador e tabelas dinâmicas maiores em uma janela separada.

Copiando tabelas largas para a área de transferência no formato rich text

Quando tabelas dinâmicas são coladas em formato Word/RTF, as tabelas que forem muito grandes para a largura do documento serão quebradas, escaladas para baixo para ajustar a largura do documento ou deixadas inalteradas.

Opções de saída

As opções de saída controlam a configuração padrão para um número de opções de saída.

Acessibilidade do Leitor de Tela. Controla como os rótulos de linha e de coluna de tabela dinâmica são lidos pelos leitores de tela. É possível ler rótulos de linha e de coluna integrais para cada célula de dados ou ler somente os rótulos que se alteram conforme você percorre as células de dados na tabela.

Capítulo 9. Tratamento de valores ausentes

Visão geral de valores omissos

Durante a fase Preparação de Dados da mineração de dados, geralmente você deseja substituir valores omissos nos dados. **Valores ausentes** são valores no conjunto de dados que são desconhecidos, não coletados ou inseridos incorretamente. Geralmente esses valores são inválidos para seus campos. Por exemplo, o campo *Sexo* deve conter os valores *M* e *F*. Se você descobrir os valores *Y* ou *Z* no campo, pode-se presumir com segurança que tais valores são inválidos e devem, portanto, ser interpretados como espaços em branco. Da mesma forma, um valor negativo para o campo *Idade* não faz sentido e também deve ser interpretado como um espaço em branco. Muitas vezes, esses valores obviamente incorretos são inseridos, ou campos deixados em branco, propositalmente durante um questionário para indicar a ausência de uma resposta. Às vezes, você pode querer examinar esses campos em branco mais de perto para determinar se uma ausência de resposta, como uma recusa em fornecer a idade de alguém, é um fator na predição de um resultado específico.

Algumas técnicas de modelagem manipulam dados omissos melhor que outras. Por exemplo, C5.0 e Apriori lidam bem com valores declarados explicitamente como "omissos" em um nó Tipo. Outras técnicas de modelagem têm problemas para lidar com valores omissos e enfrentam tempos de treinamento mais longos, resultando em modelos menos precisos.

Existem vários tipos de valores ausentes reconhecidos por IBM SPSS Modelador:

- **Valores nulos ou ausentes do sistema.** Esses são valores de não sequência de caracteres que foram deixados em branco no arquivo de origem ou banco de dados e não foram definidos especificamente como "omissos" em um nó Tipo ou de origem. Os valores omissos do sistema são exibidos como `$null$`. Observe que sequências de caracteres vazias não são consideradas nulas no IBM SPSS Modelador, embora possam ser tratadas como nulas por certos bancos de dados.
- **sequências de caracteres vazias e espaço branco.** Valores de sequência de caracteres vazia e espaços em branco (sequências de caracteres sem caracteres visíveis) são tratados como sendo diferentes de valores nulos. Sequências de caracteres vazias são tratadas como sendo equivalentes a espaço em branco para a maioria dos propósitos. Por exemplo, se você selecionar a opção para tratar um espaço em branco como brancos em um nó Tipo ou de origem, essa configuração se aplicará também às sequências de caracteres vazias.
- **Valores ausentes em branco ou definidos pelo usuário.** Estes são valores como *unknown*, *99* ou *-1* que são explicitamente definidos em um nó de origem ou Tipo de nó como ausente. Opcionalmente, também é possível escolher tratar nulos e espaços em branco como brancos, o que permite que eles sejam sinalizados para tratamento especial e sejam excluídos da maioria dos cálculos. Por exemplo, é possível usar a função `@BLANK` para tratar esses valores, com outros tipos de valores omissos, como brancos.

Leitura de dados mistos. Observe que quando você está lendo campos com armazenamento numérico (número inteiro, real, hora, registro de data e hora ou data), quaisquer valores não numéricos são configurados para *nulo* ou *omissos do sistema*. Isso acontece porque, ao contrário de alguns aplicativos, não é permitido combinar tipos de armazenamento dentro de um campo. Para evitar isso, quaisquer campos com dados combinados devem ser lidos como sequências de caracteres alterando o tipo de armazenamento no nó de origem ou aplicativo externo, conforme necessário.

Leitura de sequências de caracteres vazias do Oracle. Ao ler ou gravar em um banco de dados Oracle, esteja ciente de que, ao contrário do IBM SPSS Modelador e da maioria dos outros bancos de dados, o Oracle trata e armazena valores de sequências de caracteres vazias como equivalentes a valores nulos. Isso significa que os mesmos dados extraídos de um banco de dados Oracle podem se comportar diferentemente quando extraídos de um arquivo ou outro banco de dados, e os dados podem retornar resultados diferentes.

Manipulando valores omissos

Você deve decidir como tratar valores omissos considerando seus conhecimentos de negócios ou domínio. Para facilitar o tempo de treinamento e aumentar a precisão, você pode querer remover espaços em branco do conjunto de dados. Por outro lado, a presença de valores em branco pode levar a novas oportunidades de negócios ou insights adicionais. Ao escolher a melhor técnica, é necessário considerar os seguintes aspectos de seus dados:

- Tamanho do conjunto de dados
- Número de campos contendo espaços em branco
- Quantidade de informações omissas

Em termos gerais, há duas abordagens que podem ser seguidas:

- É possível excluir campos ou registros com valores omissos
- É possível imputar, substituir ou forçar valores omissos usando uma variedade de métodos

Ambas as abordagens podem ser amplamente automatizadas usando o nó Auditoria de Dados. Por exemplo, é possível gerar um nó de Filtro que exclua campos com muitos valores omissos para ser útil na modelagem e gerar um Supernode que impute valores omissos para qualquer ou para todos os campos restantes. É aí que entra o real poder da auditoria, permitindo que você não apenas avalie o estado atual dos seus dados, mas tome providas com base na avaliação.

Manipulando registros com valores omissos

Se a maioria dos valores omissos estiver concentrada em um pequeno número de registros, é possível simplesmente excluir esses registros. Por exemplo, um banco geralmente mantém registros completos e detalhados sobre seus clientes de empréstimo. No entanto, se o banco for menos restritivo para aprovar empréstimos para os membros de sua própria equipe, os dados reunidos para empréstimos da equipe provavelmente terão vários campos em branco. Nesse caso, há duas opções para manipular esses valores omissos:

- É possível usar um nó de Seleção para remover registros da equipe.
- Se os dados configurados forem grandes, é possível descartar todos os registros com valores em branco.

Manipulando campos com valores omissos

Se a maioria dos valores omissos estiver concentrada em um pequeno número de campos, é possível abordá-la em nível de campo em vez de em nível de registro. Essa abordagem também permite experimentar a importância relativa de determinados campos antes de você decidir sobre uma abordagem para manipular valores omissos. Se um campo for irrelevante na modelagem, provavelmente não valerá a pena mantê-lo, independentemente de quantos valores omissos ele tem.

Por exemplo, uma empresa de pesquisa de mercado pode coletar dados de um questionário geral com 50 questões. Duas das questões abordam idade e opinião política, informações que muitas pessoas relutam em fornecer. Nesse caso, *Age* e *Political_persuasion* têm muitos valores ausentes.

Nível de Medição de Campo

Ao determinar qual método usar, também é necessário considerar o nível de medição dos campos com valores omissos.

Campos numéricos. Para tipos de campo numéricos, como *Contínuo*, é sempre necessário eliminar quaisquer valores não numéricos antes da construção de um modelo, pois muitos modelos não funcionarão se espaços em branco forem incluídos em campos numéricos.

Campos categóricos. Para campos categóricos, como *Nominal* e *Flag*, a alteração de valores omissos não é necessária, mas aumenta a precisão do modelo. Por exemplo, um modelo que usa o campo *Sexo* ainda funcionará com valores sem sentido, como *Ye Z.*, mas remover todos os valores diferentes de *M* e *F* aumentará a precisão do modelo.

Rastreando ou Removendo Campos

Para rastrear campos com muitos valores omissos, você tem várias opções:

- É possível usar um nó Auditoria de Dados para filtrar campos com base na qualidade.
- É possível usar um nó Seleção de Variável para rastrear campos com mais de uma porcentagem especificada de valores omissos e ranquear campos com base na importância relativa a uma resposta especificada.
- Em vez de remover os campos, é possível usar um nó Tipo para configurar o papel do campo para **Nenhum**. Isso manterá os campos no conjunto de dados, mas os excluirá dos processos de modelagem.

Tratamento de registros com valores omissos do sistema

O que são valores omissos do sistema?

Valores omissos do sistema representam valores de dados que não são conhecidos ou não são aplicáveis. Em bancos de dados, esses valores costumam ser chamados de valores *NULL*.

Valores omissos do sistema são diferentes de valores em branco. Valores em branco normalmente são definidos no nó Tipo como valores específicos, ou intervalos de valores, que podem ser considerados omissos definidos pelo usuário. Valores em branco são manipulados diferentemente no contexto de modelagem.

Construindo valores omissos do sistema

Valores omissos do sistema podem estar presentes nos dados que são lidos a partir de uma origem de dados (por exemplo, tabelas de banco de dados podem conter valores *NULL*).

Os valores ausentes do sistema podem ser construídos usando o valor **undef** em expressões. Por exemplo, a expressão do CLEM a seguir retorna Idade, se for menor que ou igual a 30, ou um valor omissso, se for maior que 30:

```
if Age > 30 then undef else Age endif
```

Valores omissos também podem ser criados quando uma junção externa é executada, quando um número é dividido por zero, quando a raiz quadrada de um número negativo é calculada e em outras situações.

Exibindo valores omissos do sistema

Valores omissos do sistema são exibidos em tabelas e outras saídas como `$null$`.

Testando valores omissos do sistema

Use a função especial **@NULL** para retornar `true` se o valor de argumento for um valor omissso do sistema, por exemplo:

```
if @NULL(MyFieldName) then 'It is null' else 'It is not null' endif
```

Valores omissos do sistema passados para funções

Valores omissos do sistema que são passados para funções geralmente propagam valores omissos para a saída. Por exemplo, se o valor do campo `f1` for um valor omissso do sistema em uma determinada linha, a expressão `log(f1)` também será avaliada para um valor omissso do sistema para essa linha. Uma exceção é a função **@NULL**.

Valores omissos do sistema em expressões que envolvem operadores aritméticos

A aplicação de operadores aritméticos em valores que incluem um valor omissos do sistema resulta em um valor omissos do sistema. Por exemplo, se o valor de campo f1 for um valor ausente do sistema em uma determinada linha, então a expressão $f1 + 10$ também será avaliada a um valor ausente do sistema para essa linha.

Valores omissos do sistema em expressões que envolvem operadores lógicos

Quando você trabalha com valores omissos do sistema em expressões que envolvem operadores lógicos, as regras de três lógicas de valor (*verdadeiro*, *falso* e *omissos*) se aplicam e podem ser descritas em tabelas de verdade. As tabelas de verdade para os operadores lógicos comuns de *not*, *and* e *or* são mostradas nas seguintes tabelas.

<i>Tabela 5. Tabela de verdade para NOT</i>	
Operando	Operando NOT
verdadeiro	Não
Não	verdadeiro
ausente	ausente

<i>Tabela 6. Tabela de verdade para AND</i>		
Operand1	Operand2	Operand1 AND Operand2
verdadeiro	verdadeiro	verdadeiro
verdadeiro	Não	Não
verdadeiro	ausente	ausente
Não	verdadeiro	Não
Não	Não	Não
Não	ausente	Não
ausente	verdadeiro	ausente
ausente	Não	Não
ausente	ausente	ausente

<i>Tabela 7. Tabela de verdade para OR</i>		
Operand1	Operand2	Operand1 OR Operand2
verdadeiro	verdadeiro	verdadeiro
verdadeiro	Não	verdadeiro
verdadeiro	ausente	verdadeiro
Não	verdadeiro	verdadeiro
Não	Não	Não
Não	ausente	ausente
ausente	verdadeiro	verdadeiro
ausente	Não	ausente
ausente	ausente	ausente

Valores omissos do sistema em expressões que envolvem operadores de comparação

Quando você compara um valor omissos do sistema e um valor não omissos do sistema, o resultado é avaliado para um valor omissos do sistema em vez de para um resultado verdadeiro ou falso. Valores omissos do sistema podem ser comparados uns com os outros; dois valores omissos do sistema são considerados iguais.

Valores omissos do sistema em expressões if/then/else/endif

Quando você usa uma expressão condicional, e a expressão condicional retorna um valor omissos do sistema, o valor da cláusula else é retornado da expressão condicional.

Valores omissos do sistema no nó de Seleção

Quando, para um determinado registro, a expressão de seleção é avaliada para um valor omissos, o registro não é emitido do nó de Seleção (essa ação se aplica aos modos Inclusão e Descarte).

Valores omissos do sistema no nó de Mesclagem

Quando você mescla usando uma chave, nenhum registro que tem valores omissos do sistema em um campo-chave é mesclado.

Valores omissos do sistema em agregação

Quando você agrupa dados em colunas, valores omissos não são incluídos no cálculo. Por exemplo, em uma coluna com três valores { 1, 2 e undef }, a soma dos valores na coluna é calculada como 3; o valor médio é calculado como 1.5.

Imputando ou preenchendo valores omissos

Em casos em que há somente alguns valores omissos, talvez seja útil inserir valores para substituir espaços em branco. Isso pode ser feito a partir do relatório Auditoria de Dados, que permite especificar opções para campos específicos conforme apropriado e gerar um SuperNode que impõe valores usando inúmeros métodos. Esse é o método mais flexível, além de permitir especificar a manipulação de grandes números de campos em um único nó.

Os métodos a seguir estão disponíveis para imputar valores omissos:

Fixo. Substitui um valor fixo (a média do campo, o ponto médio do intervalo ou a constante especificada).

Aleatório. Substitui um valor aleatório com base em uma distribuição uniforme ou normal.

Expressão. Permite especificar uma expressão customizada. Por exemplo, você poderia substituir valores por uma variável global criada pelo nó Configurar Globais.

Algoritmo. Substitui um valor predito por um modelo baseado no algoritmo C&RT. Para cada campo imputado usando esse método, haverá um modelo C&RT separado, com um nó de Preenchimento que substitui brancos e nulos pelo valor predito pelo modelo. Um nó de Filtro é usado para remover os campos de predição gerados pelo modelo.

Alternativamente, para forçar valores para campos específicos, é possível usar um nó Tipo para assegurar que os tipos de campo cubram somente valores legais e configurar a coluna Verificação para **Forçar** para os campos cujos valores em branco precisam de substituição.

Funções do CLEM para valores omissos

Há várias funções usadas para manipular valores omissos. As funções a seguir geralmente são usadas nos nós de Seleção e Preenchimento para descartar ou preencher valores omissos:

- `count_nulls(LIST)`
- `@BLANK(FIELD)`
- `@NULL(FIELD)`
- `undef`

As funções @ podem ser usadas com a função @FIELD para identificar a presença de valores em branco ou nulos em um ou mais campos. Os campos podem simplesmente ser sinalizados quando houver valores em branco ou nulos presentes, ou podem ser preenchidos com valores de substituição ou usados em várias outras operações.

É possível contar os nulos em uma lista de campos, como a seguir:

```
count_nulls(['cardtenure' 'card2tenure' 'card3tenure'])
```

Durante o uso de qualquer uma das funções que aceitam uma lista de campos como entrada, as funções especiais @FIELDS_BETWEEN e @FIELDS_MATCHING podem ser usadas, conforme mostrado no exemplo a seguir:

```
count_nulls(@FIELDS_MATCHING('card*'))
```

Você pode usar a função `undef` para preencher campos com o sistema-valor ausente, exibido como **\$null\$**. Por exemplo, para substituir qualquer valor numérico, você poderia usar uma instrução condicional, tais como:

```
if not(Age > 17) or not(Age < 66) then undef else Age endif
```

Isso substitui qualquer coisa que não esteja no intervalo com um valor ausente do sistema, exibido como **\$null\$**. Ao usar a função `not()`, você pode pegar todos os outros valores numéricos, incluindo quaisquer negativos. Consulte o tópico “[Funções manipulando valores em branco e nulos](#)” na página 201 para obter mais informações.

Nota sobre Descarte de Registros

Durante o uso de um nó de Seleção para descartar registros, observe que a sintaxe usa lógica de três valores e inclui automaticamente valores nulos em instruções de seleção. Para excluir valores nulos (omissos do sistema) em uma expressão de seleção, deve-se especificar isso explicitamente usando `and not` na expressão. Por exemplo, para selecionar e incluir todos os registros nos quais a droga do tipo de prescrição é Drog C, você usaria a seguinte instrução de seleção:

```
Drug = 'drugC' and not(@NULL(Drug))
```

Versões anteriores de valores nulos excluídos nessas situações.

Capítulo 10. Construindo expressões CLEM

Sobre CLEM

Control Language for Expression Manipulation (CLEM) é uma poderosa linguagem para analisar e manipular os dados que fluem ao longo dos fluxos do IBM SPSS Modelador. Os mineradores de dados usam CLEM extensivamente em operações de fluxo para executar tarefas tão simples quanto derivar lucros dos dados de custo e renda ou tão complexas quanto transformar dados de log da web em um conjunto de campos e registros com informações utilizáveis.

CLEM é usado dentro do IBM SPSS Modelador para:

- Comparar e avaliar condições em campos de registro.
- Derivar valores para novos campos.
- Derivar novos valores para campos existentes.
- Motivar a sequência de registros.
- Inserir dados de registros em relatórios.

Expressões do CLEM são indispensáveis para preparação de dados no IBM SPSS Modelador e podem ser usadas em uma grande variedade de nós-de operações de registro e campo (Seleção, Balanceamento, Preenchimento) a plots e saída (Análise, Relatório, Tabela). Por exemplo, é possível usar CLEM em um nó de Derivação para criar um novo campo em uma fórmula, como razão.

Expressões do CLEM também podem ser usadas para operações de procura e substituição globais. Por exemplo, a expressão @NULL (@FIELD) pode ser usada em um nó de Preenchimento para substituir **valores omissos do sistema** com o valor de número inteiro 0. (Para substituir **valores omissos do sistema**, também chamados em branco, use a função @BLANK.)

Expressões do CLEM mais complexas também podem ser criadas. Por exemplo, você pode derivar novos campos com base em um conjunto condicional de regras, como uma nova categoria de valor criada usando-se as seguintes expressões: If: CardID = @OFFSET(CardID,1), Then: @OFFSET(ValueCategory,1), Else: 'exclude'.

Esse exemplo usa a função @OFFSET para dizer "se o valor do campo *CardID* para um determinado registro for igual para o registro anterior, retorne o valor do campo denominado *ValueCategory* para o registro anterior. Caso contrário, designe a sequência de caracteres "exclude." Em outras palavras, se *CardIDs* para registros adjacentes forem os mesmos, eles deverão ser designados à mesma categoria de valor. (Registros com a sequência de caracteres *exclude* podem ser selecionados posteriormente usando um nó de Seleção.)

Exemplos de CLEM

Para ilustrar a sintaxe correta, bem como os tipos de expressões possíveis com o CLEM, algumas expressões de exemplo são mostradas a seguir.

Expressões Simples

Fórmulas podem ser tão simples quanto estas, que derivam um novo campo com base nos valores dos campos *After* e *Before*:

```
(After - Before) / Before * 100.0
```

Observe que os nomes de campos não são citados entre aspas quando se referem aos valores do campo.

Da mesma forma, a expressão a seguir simplesmente retorna o log de cada valor para o campo *salary*.

```
log(salary)
```

Expressões Complexas

Expressões também podem ser longas e mais complexas. As expressões a seguir retornam *true* se o valor dos dois campos (\$KX-Kohonen e \$KY-Kohonen) estiver dentro dos intervalos especificados. Observe que aqui os nomes de campo são colocados entre aspas simples, pois os nomes de campo contêm caracteres especiais.

```
( '$KX-Kohonen' >= -0.2635771036148072 and '$KX-Kohonen' <= 0.3146203637123107  
and '$KY-Kohonen' >= -0.18975617885589602 and '$KY-Kohonen' <= 0.17674794197082522) -> T
```

Várias funções, como funções de sequência de caracteres, requerem a inserção de vários parâmetros usando a sintaxe correta. No exemplo a seguir, a função `subscr$` é usada para retornar o primeiro caractere de um campo *produce_ID*, indicando se um item é orgânico, geneticamente modificado ou convencional. Os resultados de uma expressão são descritos por `-> 'result'`.

```
subscr$(1,produce_ID) -> 'c'
```

Da mesma forma, a expressão a seguir é:

```
stripchar('3','123') -> '12'
```

É importante observar que os caracteres estão sempre contidos entre crases simples.

Combinando funções em uma expressão

Frequentemente, expressões do CLEM consistem em uma combinação de funções. A função a seguir combina `subscr$` e `lowertoupper$` para retornar o primeiro caractere de *produce_ID* e convertê-lo em maiúscula.

```
lowertoupper$(subscr$(1,produce_ID)) -> 'C'
```

Essa mesma expressão pode ser escrita com taquigrafia como:

```
lowertoupper$(produce_ID(1)) -> 'C'
```

Outra combinação de funções bastante usada é:

```
locchar_back('n', (length(web_page)), web_page)
```

Essa expressão localiza o caractere 'n' dentro dos valores do campo *web_page*, fazendo a leitura reversa a partir do último caractere do valor do campo. Incluindo a função `length` também, a expressão calcula dinamicamente o comprimento do valor atual em vez de usar um número estático, como 7, que será inválido para valores com menos de sete caracteres.

Funções especiais

Há inúmeras funções especiais (precedidas por um símbolo @) disponíveis. As funções mais usadas incluem:

```
@BLANK('referrer ID') -> T
```

Frequentemente, funções especiais são usadas em combinação, que é um método comumente usado de flag de espaços em branco em mais de um campo por vez.

```
@BLANK(@FIELD) -> T
```

Exemplos adicionais são discutidos em toda a documentação do CLEM. Veja o tópico “CLEM Visão Geral de Referência” na página 163 para obter mais informações.

Valores e tipos de dados

Expressões do CLEM são semelhantes às fórmulas construídas a partir de valores, nomes de campo, operadores e funções. A expressão do CLEM válida mais simples é um valor ou um nome de campo. Exemplos de valores válidos são:

```
3
1.79
'banana'
```

Exemplos de nomes de campo são:

```
Product_ID
'$P-NextField'
```

em que *Produto* é o nome de um campo de um conjunto de dados de cesta de compras, '\$P-NextField' é o nome de um parâmetro e o valor da expressão é o valor do campo nomeado. Geralmente, os nomes de campo começam com uma letra e podem também conter dígitos e sublinhados (_). Você pode usar nomes que não seguem essas regras se você colocar o nome dentro de aspas. Os valores do CLEM podem ser os seguintes:

- Sequências de caracteres--por exemplo, "c1", "Type 2", "a piece of free text"
- Números inteiros--por exemplo, 12, 0, -189
- Números reais--por exemplo, 12.34, 0.0, -0.0045
- Campos de data/hora--por exemplo, 05/12/2002, 12/05/2002, 12/05/02

Também é possível usar os elementos a seguir:

- Códigos de caracteres -- por exemplo, `a` or 3
- Listas de itens -- por exemplo, [1 2 3], ['Type 1' 'Type 2']

Códigos de caractere e listas geralmente não ocorrem como valores de campo. Geralmente eles são usados como argumentos de funções do CLEM.

Regras de Citação

Embora o software seja flexível durante a determinação de campos, valores, parâmetros e sequências de caracteres usados em uma expressão do CLEM, as regras gerais a seguir fornecem uma lista de "melhores práticas" para usar durante a criação de expressões:

- **Sequências de caracteres**—Sempre usam aspas duplas durante a gravação de sequências de caractere ("Type 2" ou "value"). Aspas simples podem ser usadas, mas há um risco de confusão com os campos citados.
- **Caracteres**—Sempre usam crase simples como esta ` . Por exemplo, observe o caractere d na função `stripchar(`d` , "drugA")`. A única exceção para isso é quando você está usando um número inteiro para se referir a um caractere específico em uma sequência de caracteres. Por exemplo, observe o caractere 5 na função `lowertoupper("drugA" (5))` → "A". *Nota:* Em um teclado padrão do Reino Unido ou dos EUA, a tecla para o caractere de acento crase (acento grave, Unicode 0060) pode estar localizada bem abaixo da tecla Esc.
- **Campos**—Os campos são tipicamente pouco citados quando usados em expressões CLEM (`subscr(2, arrayID)`) → CHAR). Você pode usar citações únicas quando necessário para aloniar espaços ou outros caracteres especiais ('Order Number'). Os campos que são citados mas indefinidos no conjunto de dados serão mal lidos como strings.
- **Parâmetros**—Sempre usam aspas simples ('\$P-threshold').

Expressões e condições

As expressões do CLEM podem retornar um resultado (usado na derivação de novos valores)--por exemplo:

```
Weight * 2.2
Age + 1
sqrt(Signal-Echo)
```

Ou podem ser avaliadas como *verdadeiras* ou *falsas* (usado durante a seleção de uma condição)--por exemplo:

```
Drug = "drugA"
Age < 16
not(PowerFlux) and Power > 2000
```

É possível combinar operadores e funções de forma arbitrária em expressões do CLEM--por exemplo:

```
sqrt(abs(Signal)) * max(T1, T2) + Baseline
```

Os colchetes e a precedência do operador determinam a ordem em que a expressão é avaliada. Nesse exemplo, a ordem de avaliação é:

- `abs(Signal)` é avaliada e `sqrt` é aplicada a seu resultado.
- `max(T1, T2)` é avaliada.
- Os dois resultados são multiplicados: `*` tem precedência maior que `+`.
- Por fim, `Baseline` é incluída no resultado.

A ordem decrescente de precedência (ou seja, operações que são executadas primeiro para operações executadas por último) é a seguinte:

- Argumentos da função
- Chamadas de função
- `xx`
- `x / mod div rem`
- `+` `-`
- `> < >= <= /== == = /=`

Se desejar substituir a precedência, ou se você estiver em dúvida quanto à ordem de avaliação, é possível usar parênteses para deixar isso explícito--por exemplo,

```
sqrt(abs(Signal)) * (max(T1, T2) + Baseline)
```

Parâmetros de Fluxo, Sessão e SuperNode

Parâmetros podem ser definidos para uso em expressões do CLEM e em scripts. Na verdade, eles são variáveis definidas pelo usuário que são salvas e persistem no fluxo, sessão ou SuperNode atuais e que podem ser acessadas a partir da interface com o usuário, bem como por meio de scripts. Se você salvar um fluxo, por exemplo, quaisquer parâmetros configurados para esse fluxo também serão salvos. (Isso os distingue das variáveis do script local, que podem ser usadas somente no script no qual elas são declaradas.) Geralmente os parâmetros são usados em scripts para controlar o comportamento do script, fornecendo informações sobre campos e valores que não precisam ser codificados permanentemente no script.

O escopo do parâmetro depende de onde ele está configurado:

- Parâmetros de fluxo podem ser configurados em um script de fluxo ou na caixa de diálogo de propriedades do fluxo e estão disponíveis para todos os nós no fluxo. Eles são exibidos na lista Parâmetros no construtor de expressões.
- Parâmetros de sessão podem ser configurados em um script independente ou na caixa de diálogo de parâmetros de sessão. Eles estão disponíveis para todos os fluxos usados na sessão atual (todos os fluxos listados na guia Fluxos na área de janela de gerenciadores).

Parâmetros também podem ser configurados para SuperNodes, em cujo caso ficam visíveis somente para os nós contidos nesse SuperNode.

Usando Parâmetros em CLEM Expressões

Parâmetros são representados em expressões do CLEM por `$P-pname`, em que `pname` é o nome do parâmetro. Quando usados em expressões do CLEM, os parâmetros devem ser colocados dentro de aspas simples—por exemplo, `'$P-scale'`.

Os parâmetros disponíveis são facilmente visualizados por meio do construtor de expressões. Para visualizar parâmetros atuais:

1. Em qualquer caixa de diálogo aceitando expressões do CLEM, clique no botão Construtor de Expressões.
2. Na lista Campos, selecione **Parâmetros**.

É possível selecionar parâmetros da lista para inserção na expressão do CLEM. Veja o tópico “[Selecionando campos, parâmetros e variáveis globais](#)” na página 159 para obter mais informações.

Trabalhando com sequências de caracteres

Há inúmeras operações disponíveis para sequências de caracteres, incluindo:

- Convertendo uma sequência em maiúsculas ou minúsculas - `uppertolower(CHAR)`.
- Removendo caracteres especificados, como `'ID_'` ou `'$'`, por meio de uma variável de sequência - `stripchar(CHAR, STRING)`.
- Determinando o comprimento (número de caracteres) para uma variável de sequência - `length(STRING)`.
- Verificando a ordenação alfabética de valores de sequência - `alphabefore(STRING1, STRING2)`.
- Remover espaço em branco inicial ou final dos valores—`trim(STRING)`, `trim_start(STRING)` ou `trimend(STRING)`.
- Extraia o primeiro ou último *n* caractere de uma sequência - `startstring(LENGTH, STRING)` ou `endstring(LENGTH, STRING)`. Por exemplo, suponhamos que você tenha um campo denominado *item* que combine um nome de produto com um código de ID de quatro dígitos (ACME CAMERA-D109). Para criar um novo campo que contenha apenas o código de quatro dígitos, especifique a fórmula a seguir em um nó de Derivação:

```
endstring(4, item)
```

- Correspondendo a um padrão específico - `STRING matches PATTERN`. Por exemplo, para selecionar pessoas com "mercado" em qualquer lugar em seu cargo de trabalho, seria possível especificar o seguinte em um nó de Seleção:

```
job_title matches "*market*"
```

- Substituindo todas as instâncias de uma substring dentro de uma `string-replace(SUBSTRING, NEWSUBSTRING, STRING)`. Por exemplo, para substituir todas as instâncias de um caractere não suportado, como um tubo vertical (|), com um ponto e vírgula anterior à mineração de texto, use a função `replace` em um nó Filler. Sob **Preencher Campos**: selecione todos os campos nos quais o

caractere pode ocorrer. Para a condição **Substituir**, selecione **Sempre** e especifique a condição a seguir sob **Substituir por**:

```
replace(' | ', ';' ,@FIELD)
```

- Derivar um campo de flag com base na presença de uma subsequência de caracteres específica. Por exemplo, você poderia usar uma função de sequência de caracteres em um nó de Derivação para gerar um campo de flag separado para cada resposta com uma expressão como:

```
hassubstring(museums , "museum_of_design")
```

Consulte o tópico “[Funções de sequência de caracteres](#)” na página 181 para obter informações adicionais.

Manipulando valores omissos e em branco

A substituição de valores omissos ou em branco é uma tarefa de preparação de dados comum para mineradores de dados. CLEM fornece inúmeras ferramentas para automatizar o tratamento de brancos. O nó de Preenchimento é o local mais comum para se trabalhar com espaços em branco; no entanto, as funções a seguir podem ser usadas em qualquer nó que aceita expressões do CLEM:

- @BLANK(FIELD) pode ser usado para determinar registros cujos valores estão em branco para um campo específico, como *Idade*.
- @NULL(FIELD) pode ser usado para determinar registros cujos valores estão ausentes do sistema para o(s) campo(s) especificado(s). No IBM SPSS Modelador, valores omissos do sistema são exibidos como valores **\$null\$**.

Consulte o tópico “[Funções manipulando valores em branco e nulos](#)” na página 201 para obter informações adicionais.

Trabalhando com números

Várias operações padrão em valores numéricos estão disponíveis no IBM SPSS Modelador, como:

- Calculando o seno do ângulo especificado - $\sin(\text{NUM})$
- Calculando o log natural de campos numéricos - $\log(\text{NUM})$
- Calculando a soma de dois números - $\text{NUM1} + \text{NUM2}$

Consulte o tópico “[Funções numéricas](#)” na página 176 para obter informações adicionais.

Trabalhando com datas e horas

Os formatos de data e hora podem variar, dependendo de sua origem de dados e do seu código de idioma. Os formatos de data e hora são específicos para cada fluxo e configurados na caixa de diálogo de propriedades do fluxo. Os exemplos a seguir são funções comumente usadas para se trabalhar com campos de data/hora.

Calculando o Tempo que Passou

É possível calcular facilmente o tempo que passou a partir de uma data de linha de base usando uma família de funções semelhantes à seguinte. Esta função retorna o tempo em meses a partir da data de referência até a data representada pela sequência de data DATE como um número real. Essa é uma figura aproximada com base em um mês de 30 dias.

```
date_in_months(Date)
```

Comparando Valores de Data/Hora

Valores de campos de data/hora podem ser comparados entre registros usando funções semelhantes à seguinte. Esta função retorna um valor de *true* se a sequência de data DATE1 representa uma data anterior à representada pela sequência de data DATE2. Caso contrário, essa função retornará um valor 0.

```
date_before(Date1, Date2)
```

Calculando Diferenças

Também é possível calcular a diferença entre dois tempos e duas datas usando funções, como:

```
date_weeks_difference(Date1, Date2)
```

Esta função retorna o tempo em semanas a partir da data representada pela sequência de data DATE1 até a data representada pela sequência de data DATE2 como um número real. Isso é baseado em uma semana de 7 dias. Se DATE2 for anterior a DATE1, essa função retornará um número negativo.

Data de hoje

A data atual pode ser incluída no conjunto de dados usando a função @TODAY. A data de hora é incluída como uma sequência de caracteres no campo especificado ou no novo campo usando o formato de data selecionado na caixa de diálogo de propriedades do fluxo. Veja o tópico [“Funções de Data e Hora” na página 187](#) para obter mais informações.

Resumindo diversos campos

A linguagem do CLEM inclui inúmeras funções que retornam estatísticas básicas de diversos campos. Essas funções podem ser úteis principalmente na análise de dados de pesquisas de opinião, em que várias respostas para uma questão podem ser armazenadas em diversos campos. Veja o tópico [“Trabalhando com dados de múltiplas respostas” na página 154](#) para obter mais informações.

Funções de comparação

É possível comparar valores em vários campos usando as funções `min_n` e `max_n`-por exemplo:

```
max_n(['card1fee' 'card2fee' 'card3fee' 'card4fee'])
```

Também é possível usar inúmeras funções de contagem para obter contagens de valores que atendem a critérios específicos, mesmo quando esses valores são armazenados em vários campos. Por exemplo, para contar o número de cartões que ficaram retidos por mais de cinco anos:

```
count_greater_than(5, ['cardtenure' 'card2tenure' 'card3tenure'])
```

Para contar valores nulos no mesmo conjunto de dados:

```
count_nulls(['cardtenure' 'card2tenure' 'card3tenure'])
```

Observe que esse exemplo conta o número de cartões sendo retidos, e não o número de pessoas retendo-os. Consulte o tópico [“Funções de comparação” na página 173](#) para obter informações adicionais.

Para contar o número de vezes que um valor especificado ocorre em vários campos, é possível usar a função `count_equal`. O exemplo a seguir conta o número de campos na lista que contêm o valor Y.

```
count_equal("Y", [Answer1, Answer2, Answer3])
```

Considerando os valores a seguir para os campos da lista, a função retorna os resultados para o valor Y como mostrado.

Tabela 8. Valores de função

Answer1	Answer2	Answer3	Contagem
Y	N	Y	2
Y	N	N	1

Funções numéricas

É possível obter estatísticas de vários campos usando funções `sum_n`, `mean_n` e `sdev_n`-por exemplo:

```
sum_n(['card1bal' 'card2bal''card3bal'])
```

```
mean_n(['card1bal' 'card2bal''card3bal'])
```

Consulte o tópico “[Funções numéricas](#)” na página 176 para obter informações adicionais.

Gerando listas de campos

Ao usar qualquer uma das funções que aceitam uma lista de campos como entrada, as funções especiais `@FIELDS_BETWEEN(start, end)` e `@FIELDS_MATCHING(pattern)` podem ser usadas como entrada. Por exemplo, supondo que a ordem de campos esteja como mostrado no exemplo `sum_n` anteriormente, o seguinte seria equivalente:

```
sum_n(@FIELDS_BETWEEN(card1bal, card3bal))
```

Alternativamente, para contar o número de valores nulos em todos os campos começando com “`card`”:

```
count_nulls(@FIELDS_MATCHING('card*'))
```

Consulte o tópico “[Campos Especiais](#)” na página 202 para obter informações adicionais.

Trabalhando com dados de múltiplas respostas

Inúmeras funções de comparação podem ser usadas para analisar dados de múltiplas respostas, incluindo:

- `value_at`
- `first_index / last_index`
- `first_non_null / last_non_null`
- `first_non_null_index / last_non_null_index`
- `min_index / max_index`

Por exemplo, imagine uma questão de múltiplas respostas feita pelo primeiro, segundo e terceiro motivos mais importantes para se decidir sobre uma determinada compra (por exemplo, preço, recomendação pessoal, revisão, fornecedor local, etc). Nesse caso, você pode determinar a importância do preço derivando o índice do campo no qual ele foi incluído primeiro:

```
first_index("price", [Reason1 Reason2 Reason3])
```

Da mesma forma, suponha que você tenha pedido aos clientes para ranquear três carros por ordem de probabilidade de compra e codificado as respostas em três campos separados, como a seguir:

Tabela 9. Exemplo de ranqueamento de carros

ID do cliente	car1	car2	car3
104	1	3	2
102	3	2	1

Tabela 9. Exemplo de ranqueamento de carros (continuação)

ID do cliente	car1	car2	car3
103	2	3	1

Neste caso, seria possível determinar o índice do campo para o carro que eles mais gostam (classificado #1 ou a classificação mais baixa) usando a função `min_index`:

```
min_index(['car1' 'car2' 'car3'])
```

Consulte o tópico “[Funções de comparação](#)” na página 173 para obter informações adicionais.

Referenciando conjuntos de múltiplas respostas

A função especial `@MULTI_RESPONSE_SET` pode ser usada para referenciar todos os campos em um conjunto de múltiplas respostas. Por exemplo, se os três campos `car` no exemplo anterior fossem incluídos em um conjunto de múltiplas respostas denominado `car_rankings`, o seguinte retornaria o mesmo resultado:

```
max_index(@MULTI_RESPONSE_SET("car_rankings"))
```

O construtor de expressões

É possível digitar expressões do CLEM manualmente ou usar o construtor de expressões, que exibe uma lista completa de funções e operadores do CLEM, bem como campos de dados do fluxo atual, permitindo construir expressões rapidamente sem memorizar os nomes exatos de campos e funções. Além disso, o construtor controla automaticamente a inclusão das aspas adequadas para campos e valores, facilitando a criação das expressões corretas sintaticamente.

Nota: O construtor de expressões não é suportado em configurações de parâmetro ou script.

Nota: Se desejar alterar sua origem de dados, antes de alterá-la, é necessário verificar se o construtor de expressões ainda pode suportar as funções que você selecionou. Como nem todos os bancos de dados suportam todas as funções, talvez você encontre um erro se executar com relação a uma nova origem de dados.

Acessando o construtor de expressões

O construtor de expressões está disponível em todos os nós nos quais expressões do CLEM são usadas, incluindo os nós de Seleção, Balanceamento, Derivação, Preenchimento, Análise, Relatório e Tabela. É possível abri-lo clicando no botão de calculadora à direita do campo de fórmula.

Criando expressões

O construtor de expressões fornece não apenas listas completas de campos, funções e operadores, mas também acesso aos valores de dados, caso seus dados sejam instanciados.

Para criar uma expressão usando o construtor de expressões

1. Digite no campo de expressão usando as listas de campo e função como referência.
or
2. Selecione os campos obrigatórios e as funções das listas de rolagem.
3. Dê um clique duplo ou clique no botão de seta amarelo para incluir o campo ou função no campo de expressão.
4. Use os botões de operando no centro da caixa de diálogo para inserir as operações na expressão.

Selecionando funções

A lista de funções exibe todas as funções e operadores do CLEM disponíveis. Role para selecionar uma função da lista ou, para facilitar a procura, use a lista suspensa para exibir um subconjunto de funções ou operadores. As funções disponíveis são agrupadas em categorias para facilitar a procura.

A maioria dessas categorias são descritas na seção Referência da linguagem de descrição do CLEM. Para obter mais informações, consulte “Referência de funções” na página 168.

As outras categorias são as seguintes.

- **Funções Gerais** contém uma seleção de algumas das funções usadas mais comumente.
- **Usado Recentemente** contém uma lista de funções do CLEM usadas dentro da sessão atual.
- **Funções @** contém uma lista de todas as funções especiais, que têm nomes precedidos por um sinal “@”.

Nota: As funções @DIFF1(FIELD1, FIELD2) e @DIFF2(FIELD1, FIELD2) requerem que os dois tipos de campo sejam iguais (por exemplo, ambos Inteiro ou ambos Longos ou ambos Reais).

- **Funções do banco de dados.** Se o fluxo incluir uma conexão de banco de dados (por meio de um nó de origem de banco de dados), essa seleção listará as funções disponíveis de dentro desse banco de dados, incluindo funções definidas pelo usuário (UDFs). Para obter mais informações, consulte “Funções de banco de dados” na página 156.
- **Agregados de banco de dados.** Se o fluxo inclui uma conexão ao banco de dados (por meio de um nó de origem da base de dados), esta seleção lista as opções de agregação disponíveis a partir desse banco de dados. Essas opções estão disponíveis no construtor de expressões do nó Agregado.
- **Agregados da janela do banco de dados.** Se o fluxo inclui uma conexão ao banco de dados (por meio de um nó de origem da base de dados), esta seleção lista a janela de opções de agregação que podem ser usadas nesse banco de dados. Essa opções estão disponíveis no Construtor de expressões dentro dos nós na paleta **Operações de campo** apenas.

Nota: Como o ModeladorSPSS obtém as **Funções Agregadas de Janela** a partir da Visualização do sistema de banco de dados, as opções disponíveis são dependentes do comportamento do banco de dados.

Embora chamadas “agregadas”, essas opções não são projetadas para serem usadas no nó Agregar, elas são mais aplicáveis aos nós como Derivar ou Selecionar. Isso ocorre porque a sua produção é escalar em vez de um agregado verdadeiro, ou seja, elas não reduzem a quantidade de dados mostrados na saída da mesma maneira que o nó Agregar não o faz. Por exemplo, você poderia usar este tipo de agregação para fornecer uma média móvel para baixo através de linhas de dados, como “média da linha atual e todas as linhas anteriores”.

- **Agregados integrados.** Contém uma lista dos modos possíveis de agregação que podem ser usados.
- **Operadores** lista todos os operadores que podem ser usados ao construir expressões. Os operadores também estão disponíveis nos botões no centro da caixa de diálogo.
- **Todos as funções** contém uma lista completa de funções disponíveis do CLEM .

Depois de ter selecionado um grupo de funções, dê um clique duplo para inserir as funções no campo expressão no ponto indicado pela posição do cursor.

Funções de banco de dados

Funções de banco de dados podem ser listadas em vários locais diferentes; a tabela a seguir mostra os locais onde o ModeladorSPSS procura detalhes da função. Esta tabela pode ser usada pelos administradores de banco de dados para assegurar que os usuários tenham privilégios de acesso às áreas necessárias para poderem usar as diferentes funções.

Além disso, a tabela lista as condições que são usadas para filtrar quando uma função está disponível para uso com base no banco de dados e no tipo de função.

Nota: Se estiver usando funções de banco de dados do Amazon Redshift, seu administrador de banco de dados pode precisar conceder a você permissões para os seis objetos de banco de dados a seguir. Os primeiros quatro são tabelas de catálogo do sistema e os dois últimos são esquemas.

- pg_type
- pg_proc
- pg_namespace
- pg_aggregate
- information_schema
- pg_catalog

<i>Tabela 10. Funções de banco de dados no Construtor de Expressões</i>			
Banco de dados	Tipo de função	Onde localizar funções	Condições usadas para filtrar funções
Db2 LUW	UDF	SYSCAT.ROUTINES SYSCAT.ROUTINEPARMS	ROUTINETYPE é F e FUNCTIONTYPE é S
Db2 LUW	UDA	SYSCAT.ROUTINES SYSCAT.ROUTINEPARMS	ROUTINETYPE é F e FUNCTIONTYPE é C
Db2 iSeries	UDF	QSYS2.SYSROUTINES QSYS2.SYSPARMS	ROUTINE_TYPE é F e FUNCTION_TYPE é S
Db2 iSeries	UDA	QSYS2.SYSROUTINES QSYS2.SYSPARMS	ROUTINE_TYPE é F e FUNCTION_TYPE é C
Db2 z/OS	UDF	SYSIBM.SYSROUTINES SYSIBM.SYSPARMS	ROUTINETYPE é F e FUNCTIONTYPE é S
Db2 z/OS	UDA	SYSIBM.SYSROUTINES SYSIBM.SYSPARMS	ROUTINETYPE é F e FUNCTIONTYPE é C
SQL Server	UDF	SYS.ALL_OBJECTS SYS.ALL_PARAMETERS SYS.TYPES	TYPE é FN ou FS
SQL Server	UDA	SYS.ALL_OBJECTS SYS.ALL_PARAMETERS SYS.TYPES	TYPE é AF
Oracle	UDF	ALL_ARGUMENTS ALL_PROCEDURES	Todas as condições a seguir são atendidas: <ul style="list-style-type: none"> • OBJECT_TYPE é FUNCTION • AGGREGATE é NO • PLSQL_TYPE não é NULL
Oracle	UDA	ALL_ARGUMENTS ALL_PROCEDURES	Todas as condições a seguir são atendidas: <ul style="list-style-type: none"> • ARGUMENT_NAME é NULL • AGGREGATE é YES • PLSQL_TYPE não é NULL

Tabela 10. Funções de banco de dados no Construtor de Expressões (continuação)

Banco de dados	Tipo de função	Onde localizar funções	Condições usadas para filtrar funções
Teradata	UDF	DBC.FUNCTIONS DBC.ALLRIGHTS	Todas as condições a seguir são atendidas: <ul style="list-style-type: none"> • FUNCTIONTYPE é F • COLUMNNNAME é RETURN0 • SPPARAMETERTYPE é 0 • ACCESSRIGHT é EF
Teradata	UDA	DBC.FUNCTIONS DBC.ALLRIGHTS	Todas as condições a seguir são atendidas: <ul style="list-style-type: none"> • FUNCTIONTYPE é A • COLUMNNNAME é RETURN0 • SPPARAMETERTYPE é 0 • ACCESSRIGHT é EF
Netezza	UDF	#H#H#.._V_FUNCTION NZA.._V_FUNCTION INZA.._V_FUNCTION	Para #H#H#.._V_FUNCTION, as condições a seguir se aplicam: <ul style="list-style-type: none"> • RESULT não contém uma sequência de caracteres com valores como: TABLE% • FUNCTION não contém uma sequência de caracteres com valores como: '/_%' escape '/' • VARARGS é FALSE Para NZA.._V_FUNCTION e INZA.._V_FUNCTION, as condições a seguir se aplicam: <ul style="list-style-type: none"> • RESULT não contém uma sequência de caracteres com valores como: TABLE% • FUNCTION não contém uma sequência de caracteres com valores como: '/_%' escape '/' • BUILTIN é f • VARARGS é FALSE
Netezza	UDA	#H#H#.._V_AGGREGATE NZA.._V_FUNCTION INZA.._V_FUNCTION	Ambas as condições a seguir são atendidas: <ul style="list-style-type: none"> • AGGTYPE é ANY ou GROUPED • VARARGS é FALSE

Tabela 10. Funções de banco de dados no Construtor de Expressões (continuação)

Banco de dados	Tipo de função	Onde localizar funções	Condições usadas para filtrar funções
Netezza	WUDA	####. . . _V_AGGREGATE NZA. . . _V_FUNCTION INZA. . . _V_FUNCTION	<p>Para ####. . . _V_AGGREGATE, as condições a seguir se aplicam:</p> <ul style="list-style-type: none"> • AGGTYPE é ANY ou ANALYTIC • AGGREGATE não é MAX_LABEL • VARARGS é FALSE <p>Para NZA. . . _V_FUNCTION e INZA. . . _V_FUNCTION, as condições a seguir se aplicam:</p> <ul style="list-style-type: none"> • AGGTYPE é ANY ou ANALYTIC • BUILTIN é f • VARARGS é FALSE

Chave para termos usados na tabela

- UDF Função Definida pelo Usuário
- UDA Função Agregada Definida pelo Usuário
- WUDA Função Agregada da Janela Definida pelo Usuário
- ##### O banco de dados ao qual você está atualmente conectado.

Selecionando campos, parâmetros e variáveis globais

A lista de campo exibe todos os campos disponíveis neste ponto no fluxo de dados. Role para selecionar um campo na lista. Dê um clique duplo ou clique no botão de seta amarela para incluir um campo na expressão.

Consulte o tópico [“Parâmetros de Fluxo, Sessão e SuperNode” na página 150](#) para obter informações adicionais.

Além dos campos, também é possível escolher entre os seguintes itens:

Conjuntos de múltiplas respostas. Para obter mais informações, consulte o guia *IBM SPSS Modeler Source, Process e Output Nodes*.

Usado recentemente contém uma lista de campos, conjuntos de múltiplas respostas, parâmetros e valores globais usados dentro da sessão atual.

Parâmetros. Veja o tópico [“Parâmetros de Fluxo, Sessão e SuperNode” na página 150](#) para obter mais informações.

Valores globais. Para obter mais informações, consulte o guia *IBM SPSS Modeler Source, Process e Output Nodes*.

Visualizando ou selecionando valores

Valores de campo podem ser visualizados a partir de inúmeros locais no sistema, incluindo o construtor de expressões, relatórios de auditoria de dados e durante a edição de valores futuros em um nó de Intervalos de Tempo. Observe que os dados devem ser completamente instanciados em um nó de origem ou Tipo para usar essa variável, de modo que o armazenamento, os tipos e os valores sejam conhecidos.

Para visualizar valores para um campo a partir do construtor de expressões ou de um nó Intervalos de Tempo, selecione o campo obrigatório e clique no botão do selecionador de valor para abrir uma caixa

de diálogo listando valores para o campo selecionado. Então é possível selecionar um valor e clicar em **Inserir** para colar o valor na atual expressão ou lista.



Figura 16. Botão do selecionador de valor

Para campos flag e nominais, todos os valores definidos são listados. Para campos contínuos (intervalo numérico), os valores mínimo e máximo são exibidos.

Verificando expressões CLEM

Clique em **Verificar** no construtor de expressões (canto inferior direito) para validar a expressão. As expressões que não foram verificadas são exibidas em vermelho. Se algum erro for localizado, uma mensagem indicando a causa será exibida.

Os itens a seguir são verificados:

- Aspas corretas nos valores e nomes de campos
- Uso correto de parâmetros e variáveis globais
- Uso válido de operadores
- Existência de campos referenciados
- Existência e a definição de globais referenciados

Se você encontrar erros na sintaxe, tente criar a expressão usando as listas e os botões do operador em vez de digitar a expressão manualmente. Esse método inclui automaticamente as aspas apropriadas para campos e valores.

Observe as limitações a seguir ao construir expressões no IBM Analytical Decision Management. As expressões não podem conter nenhum dos itens a seguir:

- Uma referência a um parâmetro de fluxo IBM SPSS Modelador
- Uma referência a um fluxo de IBM SPSS Modelador global
- Uma referência a uma função de banco de dados
- Uma referência a uma das funções de campo especial ou valor de campo @ a seguir:
 - @TARGET
 - @PREDICTED
 - @FIELD
 - @PARTITION_FIELD
 - @TRAINING_PARTITION
 - @TESTING_PARTITION
 - @VALIDATION_PARTITION

Nota: Os nomes de campo que contêm separadores devem ser colocados entre aspas simples. Para incluir cotações automaticamente, é possível criar expressões usando as listas e os botões do operador, em vez de digitar as expressões manualmente. Os seguintes caracteres em nomes de campo podem causar erros: • ! "# \$% & ' () = ~ | -^ ¥ @" "+ *" "<>? . , / : ; → (marca de seta), □ △ (marca gráfica, etc.)

Localizar e substituir

A caixa de diálogo Localizar/Substituir está disponível em locais onde você edita script ou texto de expressão, incluindo o editor de script, construtor de expressões do CLEM ou quando você define um modelo no nó de Relatório. Durante a edição de texto em qualquer uma dessas áreas, pressione **Ctrl+F** para acessar a caixa de diálogo, assegurando que o foco do cursor esteja na área de texto. Se estiver

trabalhando em um nó de Preenchimento, por exemplo, é possível acessar a caixa de diálogo a partir de qualquer uma das áreas de texto na guia Configurações ou do campo de texto no construtor de expressões.

1. Com o cursor em uma área de texto, pressione Ctrl+F para acessar a caixa de diálogo Localizar/ Substituir.
2. Insira o texto que deseja procurar ou escolha na lista suspensa de itens procurados recentemente.
3. Insira o texto de substituição, se houver.
4. Clique em **Localizar Próximo** para iniciar a procura.
5. Clique em **Substituir** para substituir a seleção atual ou em **Substituir Todos** para atualizar todas as instâncias ou as selecionadas.
6. A caixa de diálogo fecha após cada operação. Pressione F3 em qualquer área de texto para repetir a última operação Localizar ou pressione Ctrl+F para acessar a caixa de diálogo novamente.

Opções de procura

Caso de partida. Especifica se a operação Localizar faz distinção entre maiúsculas e minúsculas; por exemplo, se *myvar* corresponde a *myVar*. O texto de substituição é sempre inserido exatamente como digitado, independentemente dessa configuração.

Somente palavras inteiras. Especifica se a operação Localizar corresponde ao texto integrado às palavras. Se for selecionada, por exemplo, uma procura por *spider* não corresponderá a *spiderman* ou *spider-man*.

Expressões regulares. Especifica se a sintaxe de expressão regular é usada (consulte a próxima seção). Quando selecionada, a opção **Somente palavras inteiras** é desativada e seu valor é ignorado.

Somente texto selecionado. Controla o escopo da procura durante o uso da opção **Substituir Todos**.

sintaxe de expressão regular

Expressões regulares permitem procurar caracteres especiais, como guias ou caracteres de nova linha, classes ou intervalos de caracteres, como de *a* a *d*, qualquer dígito ou não dígito e limites, como início ou fim de uma linha. Os tipos de expressões a seguir são suportados.

Tabela 11. Correspondências de caractere	
Caracteres	Correspondências
x	O caractere x
\\	O caractere barra invertida
\On	O caractere com valor octal On (0 <= n <= 7)
\Onn	O caractere com valor octal Onn (0 <= n <= 7)
\0mnn	O caractere com valor octal 0mnn (0 <= m <= 3, 0 <= n <= 7)
\xhh	O caractere com valor hexadecimal 0xhh
\uhhhh	O caractere com valor hexadecimal 0xhhhh
\t	O caractere de tabulação ('\u0009')
\n	O caractere de nova linha (feed de linha) ('\u000A')
\r	O caractere de retorno de linha ('\u000D')
\f	O caractere de alimentação de formulário ('\u000C')
\a	O caractere de alerta (sino) ('\u0007')
\e	O caractere de escape ('\u001B')
\cx	O caractere de controle correspondente a x

Tabela 12. Classes de caracteres correspondentes

Classes de caracteres	Correspondências
[abc]	a, b ou c (classe simples)
[^abc]	Qualquer caractere, exceto a, b ou c (subtração)
[a-zA-Z]	a a z ou A a Z, inclusivo (intervalo)
[a-d[m-p]]	a a d ou m a p (união). Alternativamente, poderia ser especificado como [a-dm-p]
[a-z&&[def]]	a a z e d, e ou f (intersecção)
[a-z&&[^bc]]	a a z, exceto para b e c (subtração). Alternativamente, poderia ser especificado como [ad-z]
[a-z&&[^m-p]]	a a z, e não m a p (subtração). Alternativamente, poderia ser especificado como [a-lq-z]

Tabela 13. Classes de caractere predefinidas

Classes de caractere predefinidas	Correspondências
.	Qualquer caractere (pode ou não corresponder aos terminadores de linha)
\d	Qualquer dígito: [0-9]
\D	Um não dígito: [^0-9]
\s	Um caractere de espaço em branco: [\t\n\x0B\f\r]
\S	Um caractere de espaço não em branco: [^\\s]
\w	Um caractere de palavra: [a-zA-Z_0-9]
\W	Um caractere não de palavra: [^\\w]

Tabela 14. Correspondências de limite

Correspondentes de limite	Correspondências
^	O início de uma linha
\$	O final de uma linha
\b	Um limite de palavra
\B	Um limite que não seja de palavra
\A	O início da entrada
\Z	O fim da entrada, mas para o terminador final, se houver
\z	O fim da entrada

Capítulo 11. CLEM referência de idioma

CLEM Visão Geral de Referência

Esta seção descreve a Linguagem de Controle para Manipulação de Expressão (CLEM), que é uma poderosa ferramenta usada para analisar e manipular dados usados em fluxos do IBM SPSS Modelador. É possível usar CLEM dentro dos nós para executar tarefas que vão da avaliação de condições ou derivação de valores à inserção de dados em relatórios.

As expressões do CLEM consistem em valores, nomes de campo, operadores e funções. Usando a sintaxe correta, é possível criar uma grande variedade de operações de dados poderosas.

CLEM Datatypes

Os tipos de dados do CLEM podem ser compostos do seguinte:

- Números Inteiros
- Reais
- Caracteres
- Sequências
- Listas
- Campos
- Data/hora

Regras para Citação

Embora o IBM SPSS Modelador seja flexível quando você está determinando os campos, os valores, os parâmetros e as sequências de caracteres usados em uma expressão do CLEM, as regras gerais a seguir fornecem uma lista de "boas práticas" para serem usadas na criação de expressões:

- Strings—Sempre use citações duplas ao escrever strings, como "Type 2". Aspas simples podem ser usadas, mas há um risco de confusão com os campos citados.
- Campos—Use citações únicas apenas onde necessário para aloniar espaços ou outros caracteres especiais, como 'Order Number'. Os campos que estiverem entre aspas, mas estiverem indefinidos no conjunto de dados, serão mal interpretados como sequências de caracteres.
- Parâmetros—Sempre usam aspas simples durante o uso de parâmetros, como '\$P-threshold'.
- Caracteres—Sempre usam crases simples (`), como `stripchar(`d` , "drugA")`.

Essas regras são cobertas com mais detalhes nos tópicos a seguir.

Números Inteiros

Números inteiros são representados como uma sequência de dígitos decimais. Opcionalmente, é possível colocar um sinal de menos (–) antes do número inteiro para denotar um número negativo—por exemplo, 1234, 999, –77.

A linguagem do CLEM manipula números inteiros de precisão arbitrária. O tamanho máximo do número inteiro depende da sua plataforma. Se os valores forem muito grandes para serem exibidos em um campo de número inteiro, mudar o tipo de campo para Real geralmente restaura o valor.

Reais

Real refere-se a um número de vírgula flutuante. Reais são representados por um ou mais dígitos, seguidos por um separador decimal, seguido por um ou mais dígitos. Os reais do CLEM são mantidos em precisão dupla.

Opcionalmente, você pode colocar um sinal de menos (-) antes do real para denotar um número negativo -- por exemplo, 1.234, 0.999, -77.001. Use o formulário *<number>* e *<expoente>* para expressar um número real em notação exponencial--por exemplo, 1234.0e5, 1.7e-2. Quando o aplicativo IBM SPSS Modelador lê strings de número de arquivos e os converte automaticamente em números, números sem nenhum dígito antes do ponto decimal ou com nenhum dígito após o ponto são aceitos -- por exemplo, 999. ou .11. No entanto, essas formas são ilegais em expressões do CLEM.

Nota: Quando você fizer referência a números reais em expressões do CLEM, um período deverá ser usado como um separador decimal, independentemente de quaisquer configurações para o fluxo ou código de idioma atual. Por exemplo, especifique

```
Na > 0.6
```

em vez de

```
Na > 0,6
```

Isso se aplica mesmo se uma vírgula for selecionada como o símbolo decimal na caixa de diálogo de propriedades do fluxo e é consistente com a diretriz geral de que a sintaxe de código deve ser independente de qualquer convenção ou código de idioma específico.

Caracteres

Caracteres (geralmente mostrados como CHAR) normalmente são usados dentro de uma expressão do CLEM para executar testes em sequências de caracteres. Por exemplo, é possível usar a função *isuppercode* para determinar se o primeiro caractere de uma sequência de caracteres é maiúsculo. A expressão do CLEM a seguir usa um caractere para indicar se o teste deve ser executado no primeiro caractere da sequência de caracteres:

```
isuppercode(subscrs(1, "MyString"))
```

Para expressar o código (em contraste com o local) de um determinado caractere em uma expressão CLEM, use backquotes únicos do formulário `<character>` -- por exemplo, `A`, `Z`.

Nota: Não há um tipo de armazenamento CHAR para um campo, portanto, um campo é derivado ou preenchido com uma expressão que resulta em um CHAR, então esse resultado será convertido em uma sequência de caracteres.

Sequências

Geralmente é necessário colocar sequências de caracteres entre aspas duplas. Exemplos de sequências de caracteres são "c35product2" e "referrerID". Para indicar caracteres especiais em uma sequência de caracteres, use uma barra invertida--por exemplo, "\\$65443". (Para indicar um caractere de barra invertida, use uma barra invertida dupla, \\.) Você pode usar citações únicas em torno de uma string, mas o resultado é indistinguível de um campo citado ('referrerID'). Veja o tópico ["Funções de sequência de caracteres"](#) na página 181 para obter mais informações.

Listas

Uma lista é uma sequência ordenada de elementos, que pode ser de tipo misto. As listas são encerradas em colchetes ([]). Exemplos de listas são [1 2 4 16] e ["abc" "def"]. Listas não são usadas como o valor dos campos do IBM SPSS Modelador. Elas são usadas para fornecer argumentos para funções, como *member* e *oneof*.

Nota: As listas podem ser compostas apenas de objetos estáticos (por exemplo, uma sequência de caracteres, número ou nome de campo) e não chamadas de funções.

Campos

Nomes em expressões do CLEM que não são nomes de funções são considerados nomes de campo. É possível gravá-los simplesmente como Power, val27, state_flag, e assim por diante, mas se o nome começar com um dígito ou incluir caracteres não alfabéticos, como espaços (com a exceção do sublinhado), coloque o nome entre aspas simples--por exemplo, 'Power Increase', '2nd answer', '#101', '\$P-NextField'.

Nota: Os campos que estiverem entre aspas, mas estiverem indefinidos no conjunto de dados, serão mal interpretados como sequências de caracteres.

Datas

Cálculos de datas são baseados em uma data de "linha de base", que é especificada na caixa de diálogo de propriedades do fluxo. A data de linha de base padrão é 1 de janeiro de 1900.

A linguagem do CLEM suporta os formatos de data a seguir.

<i>Tabela 15. CLEM formatos de data de idioma</i>	
Formato	Exemplos
DDMMYY	150163
MMDDYY	011563
YYMMDD	630115
YYYYMMDD	19630115
YYYYDDD	Um ano de quatro dígitos seguido de um número de três dígitos representando o dia do ano, por exemplo, 2000032 representa o 32º dia de 2000 ou 1º de fevereiro de 2000.
DAY	Dia da semana no código de idioma atual, por exemplo, Monday, Tuesday, ..., em inglês.
MONTH	Mês no código de idioma atual, por exemplo, January, February,
DD/MM/YY	15/01/63
DD/MM/YYYY	15/01/1963
MM/DD/YY	01/15/63
MM/DD/YYYY	01/15/1963
DD-MM-YY	15-01-63
DD-MM-YYYY	15-01-1963
MM-DD-YY	01-15-63
MM-DD-YYYY	01-15-1963
DD.MM.YY	15.01.63
DD.MM.YYYY	15.01.1963
MM.DD.YY	01.15.63
MM.DD.YYYY	01.15.1963
DD-MON-YY	15-JAN-63, 15-jan-63, 15-Jan-63
DD/MON/YY	15/JAN/63, 15/jan/63, 15/Jan/63

Tabela 15. CLEM formatos de data de idioma (continuação)

Formato	Exemplos
DD.MON.YY	15.JAN.63, 15.jan.63, 15.Jan.63
DD-MON-YYYY	15-JAN-1963, 15-jan-1963, 15-Jan-1963
DD/MON/YYYY	15/JAN/1963, 15/jan/1963, 15/Jan/1963
DD.MON.YYYY	15.JAN.1963, 15.jan.1963, 15.Jan.1963
MON YYYY	Jan 2004
q Q YYYY	Data representada como um dígito (1-4) que representa o trimestre seguido da letra Q e um ano de quatro dígitos, por exemplo, o dia 25 de dezembro de 2004 seria representado como 4 Q 2004.
ww WK YYYY	Número de dois dígitos representando a semana do ano seguida pelas letras WK e um ano de quatro dígitos. A semana do ano é calculada supondo que o primeiro dia da semana seja segunda e haja pelo menos um dia na primeira semana.

Hora

A linguagem do CLEM suporta os seguintes formatos de hora.

Tabela 16. CLEM formatos de tempo de idioma

Formato	Exemplos
HHMMSS	120112, 010101, 221212
HHMM	1223, 0745, 2207
MMSS	5558, 0100
HH:MM:SS	12:01:12, 01:01:01, 22:12:12
HH:MM	12:23, 07:45, 22:07
MM:SS	55:58, 01:00
(H)H:(M)M:(S)S	12:1:12, 1:1:1, 22:12:12
(H)H:(M)M	12:23, 7:45, 22:7
(M)M:(S)S	55:58, 1:0
HH.MM.SS	12.01.12, 01.01.01, 22.12.12
HH.MM	12.23, 07.45, 22.07
MM.SS	55.58, 01.00
(H)H.(M)M.(S)S	12.1.12, 1.1.1, 22.12.12
(H)H.(M)M	12.23, 7.45, 22.7
(M)M.(S)S	55.58, 1.0

CLEM operadores

Os operadores a seguir estão disponíveis.

Tabela 17. CLEM operadores de idiomas

Operação	Comentários	Precedência (consulte a próxima seção)
or	Usado entre duas expressões do CLEM. Retorna um valor de verdadeiro se uma delas for verdadeira ou se ambas forem verdadeiras.	22
and	Usado entre duas expressões do CLEM. Retorna um valor de verdadeiro se ambas forem verdadeiras.	9
=	Usado entre dois itens comparáveis quaisquer. Retorna verdadeiro se ITEM1 for igual a ITEM2.	7
==	Idêntico a =.	7
/=	Usado entre dois itens comparáveis quaisquer. Retorna verdadeiro se ITEM1 não for igual a ITEM2.	7
/==	Idêntico a /=.	7
>	Usado entre dois itens comparáveis quaisquer. Retorna verdadeiro se ITEM1 for exatamente maior que ITEM2.	6
>=	Usado entre dois itens comparáveis quaisquer. Retorna verdadeiro se ITEM1 for maior que ou igual a ITEM2.	6
<	Usado entre dois itens comparáveis quaisquer. Retorna verdadeiro se ITEM1 for exatamente menor que ITEM2	6
<=	Usado entre dois itens comparáveis quaisquer. Retorna verdadeiro se ITEM1 for menor que ou igual a ITEM2.	6
&&=_0	Usado entre dois números inteiros. Equivalente à expressão booleana INT1 && INT2 = 0.	6
&&/=_0	Usado entre dois números inteiros. Equivalente à expressão booleana INT1 && INT2 /= 0.	6
+	Inclui dois números: NUM1 + NUM2.	5
><	Concatena duas sequências de caracteres; por exemplo, STRING1 >< STRING2.	5
-	Subtrai um número de outro: NUM1 - NUM2. Também pode ser usado na frente de um número: - NUM.	5
*	Usado para multiplicar dois números: NUM1 * NUM2.	4

Tabela 17. CLEM operadores de idiomas (continuação)

Operação	Comentários	Precedência (consulte a próxima seção)
<code>&&</code>	Usado entre dois números inteiros. O resultado é o bit a bit 'and' dos números inteiros INT1 e INT2.	4
<code>&&~~</code>	Usado entre dois números inteiros. O resultado é o bit a bit 'and' de INT1 e o complemento bit a bit de INT2.	4
<code> </code>	Usado entre dois números inteiros. O resultado é o bit a bit 'inclusive or' de INT1 e INT2.	4
<code>~~</code>	Usado na frente de um número inteiro. Produz o complemento bit a bit de INT.	4
<code> /</code>	Usado entre dois números inteiros. O resultado é o bit a bit 'exclusive or' de INT1 e INT2.	4
<code>INT1 << N</code>	Usado entre dois números inteiros. Produz o padrão de bits de INT deslocado para a esquerda N posições.	4
<code>INT1 >> N</code>	Usado entre dois números inteiros. Produz o padrão de bits de INT deslocado para a direita N posições.	4
<code>/</code>	Usado para dividir um número por outro: NUM1 / NUM2.	4
<code>**</code>	Usado entre dois números: BASE ** POWER. Retorna BASE elevado à potência de POWER.	3
<code>rem</code>	Usado entre dois números inteiros: INT1 rem INT2. Retorna o restante, INT1 - (INT1 div INT2) * INT2.	2
<code>div</code>	Usado entre dois números inteiros: INT1 div INT2. Executa divisão inteira.	2

Precedência do Operador

A precedência determina a análise sintática de expressões complexas, principalmente expressões sem colchetes com mais de um operador infix. Por exemplo,

`3 + 4 * 5`

parses como $3 + (4 * 5)$ em vez de $(3 + 4) * 5$ porque as precedentes relativas ditam que `*` deve ser analisado antes `+`. Todo operador na linguagem CLEM tem um valor de precedência associado a ele; quanto menor esse valor, mais importante ele está na lista de análise, significando que será processado mais cedo do que outros operadores com valores de precedência mais altos.

Referência de funções

As funções do CLEM a seguir estão disponíveis para o trabalho com dados no IBM SPSS Modelador. É possível inserir essas funções como código em diversas caixas de diálogo, como os nós de Derivação e

Configuração para Flag, ou usar o construtor de expressões para criar expressões válidas do CLEM sem memorizar listas de funções ou nomes de campos.

Tabela 18. CLEM funções para usar com IBM SPSS Modelador data

Tipo de função	Descrição
<u>Informações</u>	Usado para obter um insight dos valores de campo. Por exemplo, a função <code>is_string</code> retorna true para todos os registros cujo tipo é uma sequência de caracteres.
<u>Conversão</u>	Usado para construir novos campos ou converter o tipo de armazenamento. Por exemplo, a função <code>to_timestamp</code> converte o campo selecionado em um registro de data e hora.
<u>Comparação</u>	Usado para comparar valores de campo entre si ou com uma sequência especificada. Por exemplo, <code><=</code> é usado para comparar se os valores de dois campos são menores ou iguais.
<u>Lógico</u>	Usado para executar operações lógicas, como <code>if, then, else</code> .
<u>Numérico</u>	Usado para executar cálculos numéricos, como o log natural de valores de campo.
<u>Trigonométrico</u>	Usado para executar cálculos trigonométricos, como o arco cosseno de um ângulo especificado.
<u>Probabilidade</u>	Retorna probabilidades baseadas em várias distribuições, como a probabilidade de um valor da distribuição t do Estudante ser menor que um valor específico.
<u>Espacial</u>	Usado para executar cálculos espaciais em dados geoespaciais.
<u>Bit a bit</u>	Usado para manipular números inteiros como padrões de bits.
<u>Aleatório</u>	Usado para selecionar itens ou gerar números aleatoriamente.
<u>Sequência</u>	Usado para executar várias operações em sequências de caracteres, como <code>stripchar</code> , o que permite remover um caractere especificado.
<u>SoundEx</u>	Usado para localizar sequências de caracteres quando a ortografia precisa não é conhecida; com base em suposições fonéticas sobre como determinadas letras são pronunciadas.
<u>Data e Hora</u>	Usado para executar várias operações em campos de data, hora e registro de data e hora.
<u>sequência</u>	Usado para obter um insight na sequência de registro de um conjunto de dados ou executar operações que são baseadas nessa sequência.
<u>Global</u>	Usado para acessar valores globais criados por um nó <code>Configurar globais</code> . Por exemplo, <code>@MEAN</code> é usado para se referir à média de todos os valores para um campo em todo o conjunto de dados.
<u>Brancos e nulos</u>	Usado para acessar, sinalizar e, frequentemente, preencher espaços em branco especificados pelo usuário ou valores omissos do sistema. Por exemplo, <code>@BLANK(FIELD)</code> é usado para levantar um flag true para registros nos quais há valores em branco presentes.
<u>Campos especiais</u>	Usado para denotar os campos específicos que estão sendo examinados. Por exemplo, <code>@FIELD</code> é usado na derivação de campos múltiplos.

Convenções em descrições de função

As convenções a seguir são usadas em todo este guia ao referir-se aos itens em uma função.

<i>Tabela 19. Convenções em descrições de função</i>	
Convenção	Descrição
<i>bool</i>	Um booleano ou flag como true ou false.
<i>NUM, NUM1, NUM2</i>	Qualquer número.
<i>REAL, REAL1, REAL2</i>	Qualquer número real, como 1.234 ou -77.01.
<i>INT, INT1, INT2</i>	Qualquer número inteiro, como 1 ou -77.
<i>CHAR</i>	Um código de caractere, como 'A'.
<i>STRING</i>	Uma sequência de caracteres, como "referrerID".
<i>Lista</i>	Uma lista de itens, como ["abc" "def"].
<i>ITEM</i>	Um campo, como <i>Customer</i> ou <i>extract_concept</i> .
<i>DATE</i>	Um campo de data, como <i>start_date</i> , em que os valores estão em um formato como DD-MON-YYYY.
<i>TIME</i>	Um campo de tempo, como <i>power_flux</i> , no qual os valores estão em um formato como HHMMSS.

As funções neste guia são listadas com a função em uma coluna, o tipo de resultado (número inteiro, sequência de caracteres, etc) em outra e uma descrição (quando disponível) em uma terceira coluna. Por exemplo, a seguir está a descrição da função *rem*.

<i>Tabela 20. Descrição da função rem</i>		
Função	Resultado	Descrição
<i>INT1 rem INT2</i>	Número	Retorna o restante de <i>INT1</i> dividido por <i>INT2</i> . Por exemplo, <i>INT1</i> – (<i>INT1</i> div <i>INT2</i>) * <i>INT2</i> .

Os detalhes sobre convenções de uso, por exemplo, como listar itens ou especificar caracteres em uma função, são descritos em outro lugar. Consulte o tópico “[CLEM Datatypes](#)” na página 163 para obter informações adicionais.

Funções de informações

Funções de informações são usadas para obter um insight dos valores de um determinado campo. Normalmente elas são usadas para derivar campos de flag. Por exemplo, é possível usar a função *@BLANK* para criar um campo de flag indicando registros cujos valores estão em branco para o campo selecionado. Da mesma forma, é possível verificar o tipo de armazenamento para um campo usando qualquer uma das funções de tipo de armazenamento, como *is_string*.

<i>Tabela 21. Funções de informações do CLEM</i>		
Função	Resultado	Descrição
<i>@BLANK(FIELD)</i>	Booleano	Retorna verdadeiro para todos os registros cujos valores estão em branco de acordo com o conjunto de regras de manipulação de brancos em um nó Tipo de envio de dados ou nó de origem (guia Tipos).

Tabela 21. Funções de informações do CLEM (continuação)

Função	Resultado	Descrição
@NULL(ITEM)	Booleano	Retorna true para todos os registros cujos valores estão indefinidos. Valores indefinidos são valores nulos do sistema, exibidos em IBM SPSS Modelador como \$null\$.
is_date(ITEM)	Booleano	Retorna true para todos os registros cujo tipo é uma data.
is_datetime(ITEM)	Booleano	Retorna true para todos os registros cujo tipo é uma data, hora ou registro de data e hora.
is_integer(ITEM)	Booleano	Retorna true para todos os registros cujo tipo é um número inteiro.
is_number(ITEM)	Booleano	Retorna true para todos os registros cujo tipo é um número.
is_real(ITEM)	Booleano	Retorna true para todos os registros cujo tipo é um real.
is_string(ITEM)	Booleano	Retorna true para todos os registros cujo tipo é uma sequência de caracteres.
is_time(ITEM)	Booleano	Retorna true para todos os registros cujo tipo é uma hora.
is_timestamp(ITEM)	Booleano	Retorna true para todos os registros cujo tipo é um registro de data e hora.

Funções de conversão

Funções de conversão permitem construir novos campos e converter o tipo de armazenamento dos arquivos existentes. Por exemplo, é possível formar novas sequências de caracteres associando ou separando sequências de caracteres. Para associar duas sequências de caracteres, use o operador ><. Por exemplo, se o campo Site tiver o valor "BRAMLEY", então "xx" >< Site retorna "xxBRAMLEY". O resultado de >< é sempre uma sequência de caracteres, mesmo se os argumentos não forem sequências de caracteres. Assim, se o campo V1 é 3 e o campo V2 é 5, então V1 >< V2 retorna "35" (uma string, não um número).

Funções de conversão (e quaisquer outras funções que requerem um tipo específico de entrada, como valor de data e hora) dependem dos atuais formatos especificados na caixa de diálogo Opções de Fluxo. Por exemplo, se desejar converter um campo de sequência de caracteres com valores Jan 2003, Fev 2003, e assim por diante, selecione o formato de data correspondente **MÊS AAAA** como o formato de data padrão para o fluxo.

Tabela 22. Funções de conversão do CLEM

Função	Resultado	Descrição
ITEM1 >< ITEM2	Sequência	Concatena valores para dois campos e retorna a sequência de caracteres resultante como ITEM1ITEM2.
to_integer(ITEM)	Número inteiro	Converte o armazenamento do campo especificado em um número inteiro.
to_real(ITEM)	Real	Converte o armazenamento do campo especificado em um real.
to_number(ITEM)	Número	Converte o armazenamento do campo especificado em um número.

Tabela 22. Funções de conversão do CLEM (continuação)

Função	Resultado	Descrição
<code>to_string(ITEM)</code>	<i>Sequência</i>	Converte o armazenamento do campo especificado em uma sequência de caracteres. Quando um real é convertido em uma sequência de caracteres usando essa função, ele retorna um valor com 6 dígitos após o ponto radix.
<code>to_time(ITEM)</code>	<i>Horário</i>	Converte o armazenamento do campo especificado em hora.
<code>to_date(ITEM)</code>	<i>Data</i>	Converte o armazenamento do campo especificado em uma data.
<code>to_timestamp(ITEM)</code>	<i>Registro de data e hora</i>	Converte o armazenamento do campo especificado em um registro de data e hora.
<code>to_datetime(ITEM)</code>	<i>Data/hora</i>	Converte o armazenamento do campo especificado em um valor de data, hora ou registro de data e hora.
<code>datetime_date(ITEM)</code>	<i>Data</i>	Retorna o valor de data para um <i>número</i> , uma <i>sequência de caracteres</i> ou um <i>registro de data e hora</i> . Observe que essa é a única função que permite converter um número (em segundos) de volta em uma data. Se ITEM for uma sequência de caracteres, cria uma data analisando uma sequência de caracteres no formato de data atual. O formato de data especificado na caixa de diálogo de propriedades de fluxo deve estar correto para que essa função seja bem-sucedida. Se ITEM for um número, ele será interpretado como um número de segundos desde a data (ou época) de base. Frações de um dia são truncadas. Se ITEM for um registro de data e hora, a parte de data do registro de data e hora será retornada. Se ITEM for uma data, ela será retornada inalterada.
<code>stb_centroid_latitud e(ITEM)</code>	<i>Número inteiro</i>	Retorna um valor de número inteiro para latitude correspondente ao centroide do argumento geohash.
<code>stb_centroid_longitude(ITEM)</code>	<i>Número inteiro</i>	Retorna um valor de número inteiro para longitude correspondente ao centroide do argumento geohash.

Tabela 22. Funções de conversão do CLEM (continuação)

Função	Resultado	Descrição
to_geohash(ITEM)	Sequência	<p>Retorna a sequência de caracteres geohashed correspondente à latitude e longitude usando o número de bits especificado para a densidade.</p> <p>Um geohash é um código usado para identificar um conjunto de coordenadas geográficas com base nos detalhes de latitude e longitude. Os parâmetros de árvore para <code>to_geohash</code> são:</p> <ul style="list-style-type: none"> • <i>latitude</i>: Intervalo (-180, 180) e unidades são graus no sistema de coordenadas WGS84 • <i>longitude</i>: Intervalo (-90, 90) e unidades são graus no sistema de coordenadas WGS84 • <i>bits</i>: O número de bits a serem usados para armazenar o hash. Intervalo [1,75]. Isso afeta tanto o comprimento da sequência retornada (1 caractere é usado para cada 5 bits) e a precisão do hash. Por exemplo, 5 bits (1 caractere) representa aproximadamente 2500 quilômetros ou 45 bits (9 caracteres), representa aproximadamente 2.3 metros.

Funções de comparação

Funções de comparação são usadas para comparar valores de campo uns com os outros ou com uma sequência de caracteres especificada. Por exemplo, é possível verificar a igualdade das sequências de caracteres usando `=`. Um exemplo de verificação da igualdade da sequência de caracteres é: `Class = "class 1"`.

Para propósitos de comparação numérica, *maior* significa mais próximo do infinito positivo e *menor* significa mais próximo do infinito negativo. Ou seja, todos os números negativos são menores que qualquer número positivo.

Tabela 23. Funções de comparação do CLEM

Função	Resultado	Descrição
count_equal(ITEM1, LIST)	Número inteiro	Retorna o número de valores de uma lista de campos que são iguais a <i>ITEM1</i> ou nulos se <i>ITEM1</i> for nulo.
count_greater_than(ITEM1, LIST)	Número inteiro	Retorna o número de valores de uma lista de campos que são maiores que <i>ITEM1</i> ou nulos se <i>ITEM1</i> for nulo.
count_less_than(ITEM1, LIST)	Número inteiro	Retorna o número de valores de uma lista de campos que são menores que <i>ITEM1</i> ou nulos se <i>ITEM1</i> for nulo.
count_not_equal(ITEM1, LIST)	Número inteiro	Retorna o número de valores de uma lista de campos que não são iguais a <i>ITEM1</i> ou nulos se <i>ITEM1</i> for nulo.
count_nulls(LIST)	Número inteiro	Retorna o número de valores nulos de uma lista de campos.
count_non_nulls(LIST)	Número inteiro	Retorna o número de valores não nulos de uma lista de campos.
date_before(DATE1, DATE2)	Booleano	Usado para verificar a ordem dos valores de data. Retorna um valor real se <i>DATE1</i> for anterior a <i>DATE2</i> .

Tabela 23. Funções de comparação do CLEM (continuação)

Função	Resultado	Descrição
<code>first_index(ITEM, LIST)</code>	<i>Número inteiro</i>	Retorna um índice do primeiro campo contendo um ITEM de LIST de campos ou 0 se o valor não for localizado. Suportado somente para os tipos sequência de caracteres, número inteiro e real.
<code>first_non_null(LIST)</code>	<i>qualquer um</i>	Retorna o primeiro valor não nulo na lista de campos fornecida. Todos os tipos de armazenamento são suportados.
<code>first_non_null_index(LIST)</code>	<i>Número inteiro</i>	Retorna o índice do primeiro campo no LIST especificado contendo um valor não nulo ou 0 se todos os valores forem nulos. Todos os tipos de armazenamento são suportados.
<code>ITEM1 = ITEM2</code>	<i>Booleano</i>	Retorna true para registros nos quais <i>ITEM1</i> é igual a <i>ITEM2</i> .
<code>ITEM1 /= ITEM2</code>	<i>Booleano</i>	Retorna true se duas sequências de caracteres não forem idênticas ou 0 se elas forem idênticas.
<code>ITEM1 < ITEM2</code>	<i>Booleano</i>	Retorna true para registros nos quais <i>ITEM1</i> é menor que <i>ITEM2</i> .
<code>ITEM1 <= ITEM2</code>	<i>Booleano</i>	Retorna true para registros nos quais <i>ITEM1</i> é menor que ou igual a <i>ITEM2</i> .
<code>ITEM1 > ITEM2</code>	<i>Booleano</i>	Retorna true para registros nos quais <i>ITEM1</i> é maior que <i>ITEM2</i> .
<code>ITEM1 >= ITEM2</code>	<i>Booleano</i>	Retorna true para registros nos quais <i>ITEM1</i> é maior que ou igual a <i>ITEM2</i> .
<code>last_index(ITEM, LIST)</code>	<i>Número inteiro</i>	Retorna um índice do último campo contendo um ITEM de LIST de campos ou 0 se o valor não for localizado. Suportado somente para os tipos sequência de caracteres, número inteiro e real.
<code>last_non_null(LIST)</code>	<i>qualquer um</i>	Retorna o último valor não nulo na lista de campos fornecida. Todos os tipos de armazenamento são suportados.
<code>last_non_null_index(LIST)</code>	<i>Número inteiro</i>	Retorna o índice do último campo no LIST especificado contendo um valor não nulo ou 0 se todos os valores forem nulos. Todos os tipos de armazenamento são suportados.
<code>max(ITEM1, ITEM2)</code>	<i>qualquer um</i>	Retorna o maior dos dois itens-- <i>ITEM1</i> ou <i>ITEM2</i> .
<code>max_index(LIST)</code>	<i>Número inteiro</i>	Retorna o índice do campo contendo o valor máximo de uma lista de campos numéricos ou 0 se todos os valores forem nulos. Por exemplo, se o terceiro campo listado contiver o máximo, o valor de índice 3 será retornado. Se vários campos contiverem o valor máximo, o listado primeiro (primeiro da esquerda) será retornado.
<code>max_n(LIST)</code>	<i>Número</i>	Retorna o valor máximo de uma lista de campos numéricos ou nulo se todos os valores de campo forem nulos.

Tabela 23. Funções de comparação do CLEM (continuação)

Função	Resultado	Descrição
member(ITEM, LIST)	Booleano	Retorna true se <i>ITEM</i> for membro de um <i>LIST</i> especificado. Caso contrário, retorna false. Uma lista de nomes de campo também pode ser especificada.
min(ITEM1, ITEM2)	qualquer um	Retorna o menor dos dois itens-- <i>ITEM1</i> ou <i>ITEM2</i> .
min_index(LIST)	Número inteiro	Retorna o índice do campo contendo o valor mínimo de uma lista de campos numéricos ou 0 se todos os valores forem nulos. Por exemplo, se o terceiro campo listado contiver o mínimo, o valor de índice 3 será retornado. Se vários campos contiverem o valor mínimo, o listado primeiro (primeiro da esquerda) será retornado.
min_n(LIST)	Número	Retorna o valor mínimo de uma lista de campos numéricos ou nulo se todos os valores de campo forem nulos.
time_before(TIME1, TIME2)	Booleano	Usado para verificar a ordem dos valores de hora. Retorna um valor real se <i>TIME1</i> for anterior a <i>TIME2</i> .
value_at(INT, LIST)		Retorna o valor de cada campo listado no offset <i>INT</i> ou NULL, caso o offset esteja fora do intervalo de valores válidos (ou seja, menor que 1 ou maior que o número de campos listados). Todos os tipos de armazenamento são suportados.

Funções Lógicas

Expressões do CLEM podem ser usadas para executar operações lógicas.

Tabela 24. CLEM funções lógicas

Função	Resultado	Descrição
COND1 and COND2	Booleano	Esta operação é uma conjunção lógica e retorna um valor verdadeiro se <i>COND1</i> e <i>COND2</i> forem verdadeiras. Se <i>COND1</i> for falsa, <i>COND2</i> não será avaliada; isso possibilita ter conjunções nas quais <i>COND1</i> testa primeiro se a operação em <i>COND2</i> é legal. Por exemplo, <i>length(Label) >=6</i> e <i>Label(6) = 'x'</i> .
COND1 or COND2	Booleano	Esta operação é uma disjunção lógica (inclusiva) e retorna um valor de verdadeiro se <i>COND1</i> ou <i>COND2</i> for verdadeira ou se ambas forem verdadeiras. Se <i>COND1</i> for verdadeira, <i>COND2</i> não será avaliada.
not(COND)	Booleano	Esta operação é uma negação lógica e retorna um valor de verdadeiro se <i>COND</i> for falsa. Caso contrário, essa operação retornará um valor de 0.
if COND then EXPR1 else EXPR2 endif	qualquer um	Esta operação é uma avaliação condicional. Se <i>COND</i> for verdadeira, essa operação retornará o resultado de <i>EXPR1</i> . Caso contrário, o resultado da avaliação de <i>EXPR2</i> será retornado.

Tabela 24. CLEM funções lógicas (continuação)

Função	Resultado	Descrição
<code>if COND1 then EXPR1 elseif COND2 then EXPR2 else EXPR_N endif</code>	<i>qualquer um</i>	Esta operação é uma avaliação condicional com várias ramificações. Se <i>COND1</i> for verdadeira, essa operação retornará o resultado de <i>EXPR1</i> . Caso contrário, se <i>COND2</i> for verdadeira, essa operação retornará o resultado da avaliação de <i>EXPR2</i> . Caso contrário, o resultado da avaliação de <i>EXPR_N</i> será retornado.

Funções numéricas

CLEM contém inúmeras funções numéricas comumente usadas.

Tabela 25. CLEM funções numéricas

Função	Resultado	Descrição
<code>-NUM</code>	<i>Número</i>	Usado para negar <i>NUM</i> . Retorna o número correspondente com o sinal oposto.
<code>NUM1 + NUM2</code>	<i>Número</i>	Retorna a soma de <i>NUM1</i> e <i>NUM2</i> .
<code>NUM1 - NUM2</code>	<i>Número</i>	Retorna o valor de <i>NUM2</i> subtraído de <i>NUM1</i> .
<code>NUM1 * NUM2</code>	<i>Número</i>	Retorna o valor de <i>NUM1</i> multiplicado por <i>NUM2</i> .
<code>NUM1 / NUM2</code>	<i>Número</i>	Retorna o valor de <i>NUM1</i> dividido por <i>NUM2</i> .
<code>INT1 div INT2</code>	<i>Número</i>	Usado para executar divisão de número inteiro. Retorna o valor de <i>INT1</i> dividido por <i>INT2</i> .
<code>INT1 rem INT2</code>	<i>Número</i>	Retorna o restante de <i>INT1</i> dividido por <i>INT2</i> . Por exemplo, <i>INT1</i> - (<i>INT1 div INT2</i>) * <i>INT2</i> .
<code>INT1 mod INT2</code>	<i>Número</i>	Essa função foi descontinuada. Use a função <i>rem</i> no lugar.
<code>BASE ** POWER</code>	<i>Número</i>	Retorna <i>BASE</i> elevado à <i>POWER</i> potência, que pode ser qualquer número (exceto que <i>BASE</i> não deve ser zero se <i>POWER</i> for zero de algum tipo além do número inteiro 0). Se <i>POWER</i> for um número inteiro, o cálculo será executado multiplicando-se as potências de <i>BASE</i> sucessivamente. Assim, se <i>BASE</i> for um número inteiro, o resultado será um número inteiro. Se <i>POWER</i> for o número inteiro 0, o resultado será sempre um 1 do mesmo tipo do <i>BASE</i> . Caso contrário, se <i>POWER</i> não for um número inteiro, o resultado será calculado como $\exp(\text{POWER} * \log(\text{BASE}))$.
<code>abs(NUM)</code>	<i>Número</i>	Retorna o valor absoluto de <i>NUM</i> , que é sempre um número do mesmo tipo.
<code>exp(NUM)</code>	<i>Real</i>	Retorna <i>e</i> elevado à <i>NUM</i> potência, em que <i>e</i> é a base dos logaritmos naturais.
<code>fracof(NUM)</code>	<i>Real</i>	Retorna a parte fracionária de <i>NUM</i> , definida como <i>NUM</i> - <i>intof(NUM)</i> .
<code>intof(NUM)</code>	<i>Número inteiro</i>	Trunca seu argumento para um número inteiro. Ele retorna o número inteiro com o mesmo sinal de <i>NUM</i> e com a maior magnitude tal que $\text{abs}(\text{INT}) \leq \text{abs}(\text{NUM})$.
<code>log(NUM)</code>	<i>Real</i>	Retorna o logaritmo natural (base <i>e</i>) de <i>NUM</i> , que não deve ser zero de nenhum tipo.

Tabela 25. CLEM funções numéricas (continuação)

Função	Resultado	Descrição
log10(NUM)	Real	Retorna o logaritmo 10 de base de <i>NUM</i> , que não deve ser zero de nenhum tipo. Esta função é definida como $\log(\text{NUM}) / \log(10)$.
negate(NUM)	Número	Usado para negar <i>NUM</i> . Retorna o número correspondente com o sinal oposto.
round(NUM)	Número inteiro	Usado para arredondar <i>NUM</i> para um inteiro tomando <i>intof</i> (<i>NUM</i> +0.5) se <i>NUM</i> for positivo ou <i>intof</i> (<i>NUM</i> -0.5) se <i>NUM</i> for negativo.
sign(NUM)	Número	Usado para determinar o sinal de <i>NUM</i> . Essa operação retorna -1, 0 ou 1 se <i>NUM</i> for um número inteiro. Se <i>NUM</i> for real, ela retornará -1.0, 0.0 ou 1.0, dependendo se <i>NUM</i> for negativo, zero ou positivo.
sqrt(NUM)	Real	Retorna a raiz quadrada de <i>NUM</i> . <i>NUM</i> deve ser positivo.
sum_n(LIST)	Número	Retorna a soma dos valores de uma lista de campos numéricos ou nulo se todos os valores de campo forem nulos.
mean_n(LIST)	Número	Retorna o valor médio de uma lista de campos numéricos ou nulo se todos os valores de campo forem nulos.
sdev_n(LIST)	Número	Retorna o desvio padrão de uma lista de campos numéricos ou nulo se todos os valores de campo forem nulos.

Funções trigonométricas

Todas as funções nessa seção usam um ângulo como argumento ou retornam um como resultado. Em ambos os casos, as unidades do ângulo (radianos ou graus) são controladas pela configuração da opção de fluxo relevante.

Tabela 26. Funções trigonométricas do CLEM

Função	Resultado	Descrição
arccos(NUM)	Real	Calcula o arco cosseno do ângulo especificado.
arccosh(NUM)	Real	Calcula o arco cosseno hiperbólico do ângulo especificado.
arcsin(NUM)	Real	Calcula o arco seno do ângulo especificado.
arcsinh(NUM)	Real	Calcula o arco seno hiperbólico do ângulo especificado.
arctan(NUM)	Real	Calcula o arco tangente do ângulo especificado.
arctan2(NUM_Y, NUM_X)	Real	Calcula o arco tangente do <i>NUM_Y</i> / <i>NUM_X</i> e usa os sinais dos dois números para derivar informações do quadrante. O resultado é um real na faixa - $\pi < \text{ANGLE} \leq \pi$ (radians) - $180 < \text{ANGLE} \leq 180$ (degrees)
arctanh(NUM)	Real	Calcula o arco tangente hiperbólico do ângulo especificado.
cos(NUM)	Real	Calcula o cosseno do ângulo especificado.
cosh(NUM)	Real	Calcula o cosseno hiperbólico do ângulo especificado.
pi	Real	Essa constante é a melhor aproximação real para pi.
sin(NUM)	Real	Calcula o seno do ângulo especificado.

Tabela 26. Funções trigonométricas do CLEM (continuação)

Função	Resultado	Descrição
<code>sinh(NUM)</code>	<i>Real</i>	Calcula o seno hiperbólico do ângulo especificado.
<code>tan(NUM)</code>	<i>Real</i>	Calcula a tangente do ângulo especificado.
<code>tanh(NUM)</code>	<i>Real</i>	Calcula a tangente hiperbólica do ângulo especificado.

Funções de probabilidade

Funções de probabilidade retornam probabilidades baseadas em várias distribuições, como a probabilidade de um valor da distribuição t do Estudante ser menor que um valor específico.

Tabela 27. Funções de probabilidade do CLEM

Função	Resultado	Descrição
<code>cdf_chisq(NUM, DF)</code>	<i>Real</i>	Retorna a probabilidade de um valor da distribuição qui-quadrado com os graus de liberdade especificados serem menores que o número especificado.
<code>cdf_f(NUM, DF1, DF2)</code>	<i>Real</i>	Retorna a probabilidade de um valor da distribuição F , com graus de liberdade $DF1$ e $DF2$, ser menor que o número especificado.
<code>cdf_normal(NUM, MEAN, STDDEV)</code>	<i>Real</i>	Retorna a probabilidade de um valor da distribuição normal com a média especificada e um desvio padrão ser menor que o número testamento.
<code>cdf_t(NUM, DF)</code>	<i>Real</i>	Retorna a probabilidade de um valor da distribuição t do Estudante com os graus de liberdade especificados serem menores que o número especificado.

Funções espaciais

Funções espaciais podem ser usadas com dados geoespaciais. Por exemplo, elas permitem calcular as distâncias entre dois pontos, a área de um polígono, entre outras coisas. Também pode haver situações que requerem uma mesclagem de diversos conjuntos de dados geoespaciais baseados em um predicado espacial (dentro, perto de, entre outros), o que pode ser feito por meio de uma condição de mesclagem.

Essas funções espaciais funcionam em conjunto com o sistema de coordenadas especificado em **Ferramentas > Propriedades do Fluxo > Opções > Geospatial**.

Nota: Essas funções espaciais não se aplicam a dados tridimensionais. Se os dados tridimensionais forem importados no fluxo, somente as duas primeiras dimensões serão usadas por essas funções. Os valores do eixo Z serão ignorados.

Tabela 28. CLEM funções espaciais

Função	Resultado	Descrição
<code>close_to(SHAPE, SHAPE, NUM)</code>	<i>Booleano</i>	Testa se 2 formas estão dentro de uma certa DISTANCE uma da outra. Se um sistema de coordenadas projetadas for usado, a DISTANCE será em metros. Se nenhum sistema de coordenadas for usado, essa será uma unidade arbitrária.
<code>crosses(SHAPE, SHAPE)</code>	<i>Booleano</i>	Testa se 2 formas cruzam uma à outra. Essa função é adequada para 2 formas de linestring ou 1 linestring e um 1 polígono.

Tabela 28. CLEM funções espaciais (continuação)

Função	Resultado	Descrição
overlap(SHAPE, SHAPE)	Booleano	Testa se há uma interseção entre 2 polígonos e se a interseção é interior para ambas as formas.
within(SHAPE, SHAPE)	Booleano	Testa se a totalidade de SHAPE1 está contida em um POLYGON.
area(SHAPE)	Real	Retorna a área do POLYGON especificado. Se um sistema planejado for usado, a função retornará metros quadrados. Se nenhum sistema de coordenadas for usado, essa será uma unidade arbitrária. A forma deve ser POLYGON ou a MULTIPOLYGON.
num_points(SHAPE, LIST)	Número inteiro	Retorna o número de pontos de um campo de ponto (MULTIPOINT) que estão contidos dentro dos limites de um POLYGON. SHAPE1 deve ser POLYGON ou MULTIPOLYGON.
distance(SHAPE, SHAPE)	Real	Retorna a distância entre SHAPE1 e SHAPE2. Se um sistema de coordenadas projetadas for usado, a função retornará metros. Se nenhum sistema de coordenadas for usado, essa será uma unidade arbitrária. SHAPE1 e SHAPE2 podem ser qualquer tipo de medição geo.

Operações de número inteiro bit a bit

Essas funções permitem que números inteiros sejam manipulados como padrões de bits representando valores de dois complementos, em que a posição de bit N tem um peso 2^{**N} . Bits são numerados de 0 para cima. Essas operações agem como se o bit de sinal de um número inteiro fosse estendido indefinidamente para a esquerda. Assim, em qualquer lugar acima de seu bit mais significativo, um número inteiro positivo tem 0 bit e um número inteiro negativo tem 1 bit.

Tabela 29. CLEM Operações de número inteiro bit a bit

Função	Resultado	Descrição
<code>~~ INT1</code>	Número inteiro	Produz o complemento bit a bit do número inteiro $INT1$. Ou seja, há um 1 no resultado para cada posição de bit para a qual $INT1$ tem 0. É sempre verdade que $~~ INT = -(INT + 1)$.
<code>INT1 INT2</code>	Número inteiro	O resultado dessa operação é o bit a bit "or inclusivo" de $INT1$ e $INT2$. Ou seja, há um 1 no resultado para cada posição de bit para a qual há um 1 em $INT1$ ou $INT2$ ou ambos.
<code>INT1 /& INT2</code>	Número inteiro	O resultado dessa operação é o bit a bit "or exclusivo" de $INT1$ e $INT2$. Ou seja, há um 1 no resultado para cada posição de bit para a qual há um 1 em $INT1$ ou $INT2$, mas não ambos.
<code>INT1 && INT2</code>	Número inteiro	Produz o bit a bit "and" dos números inteiros $INT1$ e $INT2$. Ou seja, há um 1 no resultado para cada posição de bit para a qual há um 1 em $INT1$ e $INT2$.

Tabela 29. CLEM Operações de número inteiro bit a bit (continuação)

Função	Resultado	Descrição
<code>INT1 &&~~ INT2</code>	<i>Número inteiro</i>	Produz o bit a bit "and" de <i>INT1</i> e o complemento bit a bit de <i>INT2</i> . Ou seja, há um 1 no resultado para cada posição de bit para a qual há um 1 em <i>INT1</i> e um 0 em <i>INT2</i> . É o mesmo que <code>INT1 && (~INT2)</code> e é útil para limpar bits de <i>INT1</i> configurados em <i>INT2</i> .
<code>INT << N</code>	<i>Número inteiro</i>	Produz o padrão de bits de <i>INT1</i> deslocado para a esquerda <i>N</i> posições. Um valor negativo para <i>N</i> produz um deslocamento para a direita.
<code>INT >> N</code>	<i>Número inteiro</i>	Produz o padrão de bits de <i>INT1</i> deslocado para a direita <i>N</i> posições. Um valor negativo para <i>N</i> produz um deslocamento para a esquerda.
<code>INT1 &&=_0 INT2</code>	<i>Booleano</i>	Equivalente à expressão booleana <code>INT1 && INT2 /== 0</code> mas é mais eficiente.
<code>INT1 &&/=_0 INT2</code>	<i>Booleano</i>	Equivalente à expressão booleana <code>INT1 && INT2 == 0</code> mas é mais eficiente.
<code>integer_bitcount(INT)</code>	<i>Número inteiro</i>	Conta o número de 1 ou 0 bits na representação de dois complementos de <i>INT</i> . Se <i>INT</i> for não negativo, <i>N</i> será o número de 1 bit. Se <i>INT</i> for negativo, será o número de 0 bit. Devido à extensão do sinal, há um número infinito de 0 bit em um número inteiro não negativo ou 1 bit em um número inteiro negativo. Esse é sempre o caso de <code>integer_bitcount(INT) = integer_bitcount(- (INT+1))</code> .
<code>integer_leastbit(INT)</code>	<i>Número inteiro</i>	Retorna a posição de bit <i>N</i> do bit menos significativo configurado no número inteiro <i>INT</i> . <i>N</i> é a potência mais alta de 2 pela qual <i>INT</i> é dividido exatamente.
<code>integer_length(INT)</code>	<i>Número inteiro</i>	Retorna o comprimento em bits de <i>INT</i> como um número inteiro de dois complementos. Ou seja, <i>N</i> é o menor número inteiro tal que <code>INT < (1 << N)</code> if <code>INT >= 0</code> <code>INT >= (-1 << N)</code> if <code>INT < 0</code> . Se <i>INT</i> for não negativo, a representação de <i>INT</i> como um número inteiro não assinado requer um campo de pelo menos <i>N</i> bits. Alternativamente, no mínimo <i>N+1</i> bits é necessário para representar <i>INT</i> como um número inteiro assinado, independentemente do sinal.
<code>testbit(INT, N)</code>	<i>Booleano</i>	Testa o bit na posição <i>N</i> no número inteiro <i>INT</i> e retorna o estado do bit <i>N</i> como um valor booleano, que é verdadeiro para 1 e falso para 0.

Funções aleatórias

As funções a seguir são usadas para selecionar itens ou gerar números aleatoriamente.

Tabela 30. CLEM Funções aleatórias

Função	Resultado	Descrição
oneof(LIST)	qualquer um	Retorna um elemento escolhido aleatoriamente de <i>LIST</i> . Os itens da lista devem ser inseridos como [ITEM1, ITEM2, ..., ITEM_N]. Observe que uma lista de nomes de campo também pode ser especificada.
random(NUM)	Número	Retorna um número aleatório distribuído uniformemente do mesmo tipo (<i>INT</i> ou <i>REAL</i>), começando de 1 a <i>NUM</i> . Se você usar número inteiro, somente números inteiros serão retornados. Se você usar um número real (decimal), números reais serão retornados (precisão decimal determinada pelas opções de fluxo). O maior número aleatório retornado pela função poderia ser igual a <i>NUM</i> .
random0(NUM)	Número	Tem as mesmas propriedades de aleatório <code>random(NUM)</code> , mas começando em 0. O maior número aleatório retornado pela função nunca será igual a <i>NUM</i> .

Funções de sequência de caracteres

No CLEM, é possível executar as operações a seguir com sequências de caracteres:

- Comparar sequências de caracteres
- Criar sequências de caracteres
- Acessar caracteres

Em CLEM, uma string é qualquer sequência de caracteres entre marcas de aspas duplas correspondentes ("string quotes"). Caracteres (CHAR) podem ser qualquer caractere alfanumérico único. Eles são declarados em CLEM expressões usando backquotes únicos na forma de `<caractere>`, como `z`, `A` ou `2`. Caracteres que estiverem fora dos limites ou forem índices negativos para uma sequência de caracteres resultarão em um comportamento indefinido.

Nota: As comparações entre sequências de caracteres que usam e não usam SQL pushback podem gerar diferentes resultados onde existem espaços à direita.

Tabela 31. CLEM funções de sequências de caracteres

Função	Resultado	Descrição
allbutfirst(N, STRING)	Sequência	Retorna uma sequência de caracteres, que é <i>STRING</i> com os primeiros <i>N</i> caracteres removidos.
allbutlast(N, STRING)	Sequência	Retorna uma sequência de caracteres, que é <i>STRING</i> com os últimos caracteres removidos.
alphabefore(STRING1, STRING2)	Booleano	Usado para verificar a ordenação alfabética das sequências de caracteres. Retorna true se <i>STRING1</i> preceder <i>STRING2</i> .
endstring(LENGTH, STRING)	Sequência	Extrai os últimos caracteres <i>N</i> da sequência de caracteres especificada. Se o comprimento da sequência de caracteres for menor que ou igual ao comprimento especificado, ele permanecerá inalterado.
hasendstring(STRING, SUBSTRING)	Número inteiro	Esta função é a mesma de <code>isendstring(SUBSTRING, STRING)</code> .

Tabela 31. CLEM funções de sequências de caracteres (continuação)

Função	Resultado	Descrição
hasmidstring(STRING, SUBSTRING)	Número inteiro	Esta função é a mesma de <code>ismidstring(SUBSTRING, STRING)</code> (substring incorporada).
hasstartstring(STRING, SUBSTRING)	Número inteiro	Essa função é a mesma de <code>isstartstring(SUBSTRING, STRING)</code> .
hassubstring(STRING, N, SUBSTRING)	Número inteiro	Esta função é a mesma de <code>issubstring(SUBSTRING, N, STRING)</code> , onde <i>N</i> padronizado para 1.
count_substring(STRING, SUBSTRING)	Número inteiro	Retorna o número de vezes que a subsequência de caracteres especificada ocorre dentro da sequência de caracteres. Por exemplo, <code>count_substring("foooo.txt", "oo")</code> retorna 3.
hassubstring(STRING, SUBSTRING)	Número inteiro	Esta função é a mesma de <code>issubstring(SUBSTRING, 1, STRING)</code> , onde <i>N</i> padronizado para 1.
isalphacode(CHAR)	Booleano	Retorna um valor de true se <i>CHAR</i> for um caractere na sequência de caracteres especificada (geralmente um nome de campo) cujo código de caractere é uma letra. Caso contrário, essa função retornará um valor 0. Por exemplo, <code>isalphacode(produce_num(1))</code> .
isendstring(SUBSTRING, STRING)	Número inteiro	Se a sequência de caracteres <i>STRING</i> terminar com a subsequência de caracteres <i>SUBSTRING</i> , então essa função retornará o número inteiro subscrito de <i>SUBSTRING</i> em <i>STRING</i> . Caso contrário, essa função retornará um valor 0.
islowercode(CHAR)	Booleano	Retorna um valor de true se <i>CHAR</i> for um caractere de letra minúscula para a sequência de caracteres especificada (geralmente um nome de campo). Caso contrário, essa função retornará um valor 0. Por exemplo, <code>islowercode(``)</code> e <code>islowercode(country_name(2))</code> são expressões válidas.
ismidstring(SUBSTRING, STRING)	Número inteiro	Se <i>SUBSTRING</i> for uma subsequência de caracteres de <i>STRING</i> , mas não começar no primeiro caractere de <i>STRING</i> ou terminar no último, essa função retornará o subscrito no qual a subsequência de caracteres começa. Caso contrário, essa função retornará um valor 0.

Tabela 31. CLEM funções de sequências de caracteres (continuação)

Função	Resultado	Descrição
<code>isnumbercode(CHAR)</code>	Booleano	Retorna um valor de true se <i>CHAR</i> para a sequência de caracteres especificada (geralmente um nome de campo) for um caractere cujo código de caractere é um dígito. Caso contrário, essa função retornará um valor 0. Por exemplo, <code>isnumbercode(product_id(2))</code> .
<code>isstartstring(SUBSTRING, STRING)</code>	Número inteiro	Se a sequência de caracteres <i>STRING</i> começar com a subsequência de caracteres <i>SUBSTRING</i> , esta função retornará o subscrito 1. Caso contrário, essa função retornará um valor 0.
<code>issubstring(SUBSTRING, N, STRING)</code>	Número inteiro	Procura a sequência de caracteres <i>STRING</i> , começando de seu <i>N</i> º caractere, para uma sequência de caracteres igual à <i>SUBSTRING</i> . Se for localizada, essa função retornará a inscrição de número inteiro onde começa a subsequência de caracteres correspondente. Caso contrário, essa função retornará um valor 0. Se <i>N</i> não for fornecido, essa função será padronizada para 1.
<code>issubstring(SUBSTRING, STRING)</code>	Número inteiro	Procura a sequência de caracteres <i>STRING</i> , começando de seu <i>N</i> º caractere, para uma sequência de caracteres igual à <i>SUBSTRING</i> . Se for localizada, essa função retornará a inscrição de número inteiro onde começa a subsequência de caracteres correspondente. Caso contrário, essa função retornará um valor 0. Se <i>N</i> não for fornecido, essa função será padronizada para 1.
<code>issubstring_count(SUBSTRING, N, STRING):</code>	Número inteiro	Retorna o índice da <i>N</i> ª ocorrência de <i>SUBSTRING</i> dentro da <i>STRING</i> especificada. Se houver menos de <i>N</i> ocorrências de <i>SUBSTRING</i> , um 0 será retornado.
<code>issubstring_lim(SUBSTRING, N, STARTLIM, ENDLIM, STRING)</code>	Número inteiro	Essa função é igual à <code>issubstring</code> , mas a correspondência é restrinida para iniciar até o subscrito <i>STARTLIM</i> e para terminar até o subscrito <i>ENDLIM</i> . As restrições <i>STARTLIM</i> ou <i>ENDLIM</i> podem ser desativadas, fornecendo um valor de falso para qualquer argumento-por exemplo, <code>issubstring_lim(SUBSTRING, N, false, false, STRING)</code> é o mesmo que <code>issubstring</code> .
<code>isuppercode(CHAR)</code>	Booleano	Retorna um valor de true se <i>CHAR</i> for um caractere de letra maiúscula. Caso contrário, essa função retornará um valor 0. Por exemplo, <code>isuppercode(``)</code> and <code>isuppercode(country_name(2))</code> são expressões válidas.

Tabela 31. CLEM funções de sequências de caracteres (continuação)

Função	Resultado	Descrição
last(CHAR)	Sequência	Retorna o último caractere <i>CHAR</i> de <i>STRING</i> (que deve ter pelo menos um caractere de comprimento).
length(STRING)	Número inteiro	Retorna o comprimento da sequência de caracteres <i>STRING</i> --ou seja, o número de caracteres nele.
locchar(CHAR, N, STRING)	Número inteiro	<p>Usado para identificar a localização dos caracteres em campos simbólicos. A função procura a sequência de caracteres <i>STRING</i> para o caractere <i>CHAR</i>, começando a procura no <i>N</i>º caractere de <i>STRING</i>. Essa função retorna um valor indicando a localização (começando em <i>N</i>) onde o caractere está. Se o caractere não for localizado, esta função retorna o valor 0. Se a função tiver um offset inválido (<i>N</i>) (por exemplo, um offset além do comprimento da sequência de caracteres), essa função retornará \$null\$.</p> <p>Por exemplo, locchar(`n`, 2, web_page) procura o campo chamado <i>web_page</i> para o caractere 'n' iniciado no segundo caractere no valor de campo.</p> <p><i>Nota:</i> Certifique-se de usar crases simples para conter o caractere especificado.</p>
locchar_back(CHAR, N, STRING)	Número inteiro	<p>Semelhante a <i>locchar</i>, exceto que a procura é executada backward, começando no <i>N</i>º caractere. Por exemplo, locchar_back(`n`, 9, web_page) procura o campo <i>web_page</i>, começando no nono caractere e movendo backward para o início da sequência de caracteres. Se a função tiver um offset inválido (por exemplo, um offset além do comprimento da sequência de caracteres), essa função retornará \$null\$. O ideal é usar <i>locchar_back</i> com a função <i>length(<field>)</i> para usar dinamicamente o comprimento do valor atual do campo.</p> <p>Por exemplo, locchar_back(`n`, (length(web_page)), web_page).</p>

Tabela 31. CLEM funções de sequências de caracteres (continuação)

Função	Resultado	Descrição
lowertoupper(CHAR) lowertoupper (STRING)	CHAR ou String	A entrada pode ser uma sequência de caracteres ou um caractere, que é usada nessa função para retornar um novo item do mesmo tipo, com quaisquer caracteres minúsculos convertidos em seus equivalentes maiúsculos. Por exemplo, <code>lowertoupper('a')</code> , <code>lowertoupper("My string")</code> e <code>lowertoupper(field_name(2))</code> são todas expressões válidas.
matches	Booleano	Retorna true se uma sequência de caracteres corresponder a um padrão especificado. O padrão deve ser uma sequência de caracteres literal; ele não pode ser um nome contendo um padrão. Um ponto de interrogação (?) pode ser incluído no padrão para combinar exatamente um caractere; um asterisco (*) corresponde a zero ou mais caracteres. Para corresponder a um ponto de interrogação literal ou asterisco (em vez de usá-los como curinga), uma barra invertida (\) pode ser usada como um caractere de escape.
replace(SUBSTRING, NEWSUBSTRING, STRING)	Sequência	Dentro da STRING especificada, substitua todas as instâncias de SUBSTRING por NEWSUBSTRING.
replicate(COUNT, STRING)	Sequência	Retorna uma sequência de caracteres que consiste na sequência de caracteres original copiada o número de vezes especificado.
stripchar(CHAR, STRING)	Sequência	Permite remover caracteres especificados de uma sequência de caracteres ou campo. É possível usar essa função, por exemplo, para remover símbolos adicionais, como notações de moeda, dos dados para atingir um nome ou número simples. Por exemplo, o uso da sintaxe <code>stripchar('\$', 'Cost')</code> retorna um novo campo com o sinal de dólar removido de todos os valores. <i>Nota:</i> Certifique-se de usar crases simples para conter o caractere especificado.

Tabela 31. CLEM funções de sequências de caracteres (continuação)

Função	Resultado	Descrição
skipchar(CHAR, N, STRING)	Número inteiro	<p>Procura na sequência de caracteres <i>STRING</i> qualquer caractere além de <i>CHAR</i>, começando no <i>N</i>º caractere. Essa função retorna uma subsequência de caracteres de número inteiro, indicando o ponto em que um deles é localizado ou 0 se cada caractere do <i>N</i>º em diante for <i>CHAR</i>. Se a função tiver um offset inválido (por exemplo, um offset além do comprimento da sequência de caracteres), essa função retornará \$null\$.</p> <p><i>locchar</i> geralmente é usado com funções <i>skipchar</i> para determinar o valor de <i>N</i> (o ponto em que começa a procura da sequência de caracteres). Por exemplo, <i>skipchar(`s`, (locchar(`s`, 1, "MyString")))</i>, "MyString").</p>
skipchar_back(CHAR, N, STRING)	Número inteiro	Semelhante a <i>skipchar</i> , exceto que a procura é executada backward , começando no <i>N</i> º caractere.
startstring(LENGTH, STRING)	Sequência	Extrai os primeiros <i>N</i> caracteres da sequência de caracteres especificada. Se o comprimento da sequência de caracteres for menor que ou igual ao comprimento especificado, ele permanecerá inalterado.
strmember(CHAR, STRING)	Número inteiro	Equivalente à <i>locchar(CHAR, 1, STRING)</i> . Ele retorna uma subsequência de caracteres inteira indicando o ponto em que <i>CHAR</i> ocorre pela primeira vez, ou 0. Se a função tiver um offset inválido (por exemplo, um offset além do comprimento da sequência de caracteres), essa função retornará \$null\$.
subscrs(N, STRING)	CHAR	Retorna o <i>N</i> º caractere <i>CHAR</i> da sequência de caracteres de entrada <i>STRING</i> . Essa função também pode ser gravada em forma de taquigrafia como <i>STRING(N)</i> . Por exemplo, <i>lowertoupper("name"(1))</i> é uma expressão válida.
substring(N, LEN, STRING)	Sequência	Retorna uma sequência de caracteres <i>SUBSTRING</i> , que consiste nos caracteres <i>LEN</i> da sequência de caracteres <i>STRING</i> , começando no caractere no subscrito <i>N</i> .
substring_between(N1, N2, STRING)	Sequência	Retorna a subsequência de caracteres de <i>STRING</i> , que começa no subscrito <i>N1</i> e termina no subscrito <i>N2</i> .
trim(STRING)	Sequência	Remove os caracteres de espaço em branco inicial e final da sequência de caracteres especificada.

Tabela 31. CLEM funções de sequências de caracteres (continuação)

Função	Resultado	Descrição
trim_start(STRING)	Sequência	Remove os caracteres de espaço em branco iniciais da sequência de caracteres especificada.
trimend(STRING)	Sequência	Remove os caracteres de espaço em branco finais da sequência de caracteres especificada.
unicode_char(NUM)	CHAR	Entrada deve ser valores decimais, e não hexadecimais. Retorna o caractere com valor Unicode NUM.
unicode_value(CHAR)	Número	Retorna o valor Unicode de CHAR
uppertolower(CHAR) uppertolower (STRING)	CHAR ou String	A entrada pode ser uma sequência de caracteres ou um caractere e é usada nessa função para retornar um novo item do mesmo tipo, com quaisquer caracteres maiúsculos convertidos em seus equivalentes minúsculos. <i>Nota:</i> Lembre-se de especificar sequências de caracteres com aspas duplas e caracteres com crases simples. Nomes de campos simples devem ser especificados sem aspas.

Funções SoundEx

SoundEx é um método usado para localizar sequências quando o som é conhecido, mas a ortografia precisa não. Desenvolvido em 1918, ele procura palavras com sons similares com base em suposições fonéticas sobre como certas letras são pronunciadas. Ele pode ser usado para procurar nomes em um banco de dados, por exemplo, no qual ortografias e pronúncias para nomes similares podem variar. O algoritmo SoundEx básico é documentado em várias origens e, apesar das limitações conhecidas (por exemplo, combinações de letra inicial como ph e f não corresponderão, mesmo que soarem iguais), é suportado de alguma forma pela maioria dos bancos de dados.

Tabela 32. CLEM Funções SoundEx

Função	Resultado	Descrição
soundex(STRING)	Número inteiro	Retorna o código SoundEx de quatro caracteres para a STRING especificada.
soundex_difference(STRING1, STRING2)	Número inteiro	Retorna um número inteiro entre 0 e 4 indicando o número de caracteres que são iguais na mesma codificação SoundEx para as duas sequências de caracteres, em que 0 indica nenhuma semelhança e 4 indica grande semelhança ou sequências de caracteres idênticas.

Funções de Data e Hora

CLEM inclui uma família de funções para o tratamento de campos com armazenamento de data/hora das variáveis de sequencia de caracteres representando datas e horas. Os formatos de data e hora usados

são específicos para cada fluxo e especificados na caixa de diálogo de propriedades do fluxo. As funções de data e hora analisam sequências de caracteres de data e hora de acordo com o formato atualmente selecionado.

Quando você especifica um ano em uma data que usa somente dois dígitos (ou seja, o século não é especificado), IBM SPSS Modelador usa o século padrão que é especificado na caixa de diálogo de propriedades do fluxo.

Nota: Se a função de dados for enviada de volta para SQL ou IBM SPSS Servidor analítico, em uma ramificação que segue uma origem de dados do Servidor analítico, quaisquer sequências de formatações de data (to_date) dentro dos dados deverão corresponder ao formato de data especificado no fluxo do ModeladorSPSS.

Tabela 33. Funções de data e hora do CLEM		
Função	Resultado	Descrição
@TODAY	Sequência	Se você selecionar Dias/Minutos de Rolagem na caixa de diálogo de propriedades do fluxo, essa função retornará a data atual como uma sequência de caracteres no formato de data atual. Se você usa um formato de data de dois dígitos e não seleciona Rollover days / mins , esta função retorna \$null\$ no servidor atual.
to_time(ITEM)	Horário	Converte o armazenamento do campo especificado em hora.
to_date(ITEM)	Data	Converte o armazenamento do campo especificado em uma data.
to_timestamp(ITEM)	Registro de data e hora	Converte o armazenamento do campo especificado em um registro de data e hora.
to_datetime(ITEM)	Data/hora	Converte o armazenamento do campo especificado em um valor de data, hora ou registro de data e hora.
datetime_date(ITEM)	Data	Retorna o valor de data para um <i>número</i> , uma <i>sequência de caracteres</i> ou um <i>registro de data e hora</i> . Observe que essa é a única função que permite converter um número (em segundos) de volta em uma data. Se ITEM for uma sequência de caracteres, cria uma data analisando uma sequência de caracteres no formato de data atual. O formato de data especificado na caixa de diálogo de propriedades de fluxo deve estar correto para que essa função seja bem-sucedida. Se ITEM for um número, ele será interpretado como um número de segundos desde a data (ou época) de base. Frações de um dia são truncadas. Se ITEM for registro de data e hora, a parte de data do registro de data e hora será retornada. Se ITEM for uma data, ela será retornada inalterada.
date_add_years(NUM, DATE)	Data	Retorna o novo DATE ou registro de data e hora após incluir o valor especificado para a parte de data especificada YEAR em um DATE ou registro de data e hora.
date_add_years(NUM, TIMES TAMP)	Registro de data e hora	Retorna o novo DATE ou registro de data e hora após incluir o valor especificado para a parte de data especificada YEAR em um DATE ou registro de data e hora.

Tabela 33. Funções de data e hora do CLEM (continuação)

Função	Resultado	Descrição
date_add_months(NUM, DATE)	Data	Retorna o novo DATE ou registro de data e hora após incluir o valor especificado para a parte de data especificada MONTH em um DATE ou registro de data e hora.
date_add_months(NUM, TIME STAMP)	Registro de data e hora	Retorna o novo DATE ou registro de data e hora após incluir o valor especificado para a parte de data especificada MONTH em um DATE ou registro de data e hora.
date_add_weeks(NUM, DATE)	Data	Retorna o novo DATE ou registro de data e hora após incluir o valor especificado para a parte de data especificada WEEK em um DATE ou registro de data e hora.
date_add_weeks(NUM, TIMES TAMP)	Registro de data e hora	Retorna o novo DATE ou registro de data e hora após incluir o valor especificado para a parte de data especificada WEEK em um DATE ou registro de data e hora.
date_add_days(NUM, DATE)	Data	Retorna o novo DATE ou registro de data e hora após incluir o valor especificado para a parte de data especificada DAY em um DATE ou registro de data e hora.
date_add_days(NUM, TIMESTAMP)	Registro de data e hora	Retorna o novo DATE ou registro de data e hora após incluir o valor especificado para a parte de data especificada DAY em um DATE ou registro de data e hora.
date_before(DATE1, DATE2)	Booleano	Retorna um valor de true se DATE1 representar uma data ou um registro de data e hora anteriores aos representados por DATE2. Caso contrário, essa função retornará um valor 0.
date_days_difference(DATE1, DATE2)	Número inteiro	Retorna o tempo, em dias, a partir da data ou do registro de data e hora representados por DATE1 até os representados por DATE2 como um número inteiro. Se DATE2 for anterior a DATE1, essa função retornará um número negativo.
date_in_days(DATE)	Número inteiro	Retorna o tempo, em dias, a partir da data de linha de base até a data ou o registro de data e hora representados por DATE como um número inteiro. Se DATE for anterior à data da linha de base, essa função retornará um número negativo. Deve-se incluir uma data válida para o cálculo funcionar de maneira adequada. Por exemplo, você não deve especificar 29 de fevereiro de 2001 como a data. Como 2001 não é um ano bissexto, essa data não existe.

Tabela 33. Funções de data e hora do CLEM (continuação)

Função	Resultado	Descrição
date_in_months(DATE)	Real	Retorna o tempo, em meses, a partir da data de linha de base até a data ou o registro de data e hora representados por <i>DATE</i> como um número real. Essa é uma figura aproximada com base em um mês de 30,4375 dias. Se <i>DATE</i> for anterior à data da linha de base, essa função retornará um número negativo. Deve-se incluir uma data válida para o cálculo funcionar de maneira adequada. Por exemplo, você não deve especificar 29 de fevereiro de 2001 como a data. Como 2001 não é um ano bissexto, essa data não existe.
date_in_weeks(DATE)	Real	Retorna o tempo, em semanas, a partir da data de linha de base até a data ou o registro de data e hora representados por <i>DATE</i> como um número real. Isso é baseado em uma semana de 7 dias. Se <i>DATE</i> for anterior à data da linha de base, essa função retornará um número negativo. Deve-se incluir uma data válida para o cálculo funcionar de maneira adequada. Por exemplo, você não deve especificar 29 de fevereiro de 2001 como a data. Como 2001 não é um ano bissexto, essa data não existe.
date_in_years(DATE)	Real	Retorna o tempo, em anos, a partir da data de linha de base até a data ou o registro de data e hora representados por <i>DATE</i> como um número real. Essa é uma figura aproximada com base em um ano de 365,25 dias. Se <i>DATE</i> for anterior à data da linha de base, essa função retornará um número negativo. Deve-se incluir uma data válida para o cálculo funcionar de maneira adequada. Por exemplo, você não deve especificar 29 de fevereiro de 2001 como a data. Como 2001 não é um ano bissexto, essa data não existe.
date_months_difference(DATE1, DATE2)	Real	Retorna o tempo, em meses, a partir da data ou do registro de data e hora representados por <i>DATE1</i> até os representados por <i>DATE2</i> como um número real. Essa é uma figura aproximada com base em um mês de 30,4375 dias. Se <i>DATE2</i> for anterior a <i>DATE1</i> , essa função retornará um número negativo.
datetime_date(YEAR, MONTH, DAY)	Data	Cria um valor de data para <i>YEAR</i> , <i>MONTH</i> e <i>DAY</i> fornecidos. Os argumentos devem ser números inteiros.
datetime_day(DATE)	Número inteiro	Retorna o dia do mês de um <i>DATE</i> ou registro de data e hora fornecidos. O resultado é um número inteiro no intervalo de 1 a 31.
datetime_day_name(DAY)	Sequência	Retorna o nome completo do <i>DAY</i> fornecido. O argumento deve ser um número inteiro no intervalo de 1 (domingo) a 7 (sábado).
datetime_hour(TIME)	Número inteiro	Retorna a hora de um <i>TIME</i> ou registro de data e hora. O resultado é um número inteiro no intervalo de 0 a 23.

Tabela 33. Funções de data e hora do CLEM (continuação)

Função	Resultado	Descrição
datetime_in_seconds(TIME)	Real	Retorna a parte de segundos armazenada em <i>TIME</i> .
datetime_in_seconds(DATE), datetime_in_seconds(DATE TIME)	Real	Retorna o número acumulado, convertido em segundos, da diferença entre <i>DATE</i> ou <i>DATETIME</i> atuais e a data de linha de base (1900-01-01).
datetime_minute(TIME)	Número inteiro	Retorna o minuto de um <i>TIME</i> ou registro de data e hora. O resultado é um número inteiro no intervalo de 0 a 59.
datetime_month(DATE)	Número inteiro	Retorna o mês de um <i>DATE</i> ou registro de data e hora. O resultado é um número inteiro no intervalo de 1 a 12.
datetime_month_name(MONTH)	Sequência	Retorna o nome completo do <i>MONTH</i> fornecido. O argumento deve ser um número inteiro no intervalo de 1 a 12.
datetime_now	Registro de data e hora	Retorna a hora atual como um registro de data e hora.
datetime_second(TIME)	Número inteiro	Retorna o segundo de um <i>TIME</i> ou registro de data e hora. O resultado é um número inteiro no intervalo de 0 a 59.
datetime_day_short_name(DAY)	Sequência	Retorna o nome abreviado do <i>DAY</i> fornecido. O argumento deve ser um número inteiro no intervalo de 1 (domingo) a 7 (sábado).
datetime_month_short_name(MONTH)	Sequência	Retorna o nome abreviado do <i>MONTH</i> fornecido. O argumento deve ser um número inteiro no intervalo de 1 a 12.
datetime_time(HOUR, MINUTE, SECOND)	Horário	Retorna o valor do <i>HOUR</i> , <i>MINUTE</i> e <i>SECOND</i> especificados. Os argumentos devem ser números inteiros.
datetime_time(ITEM)	Horário	Retorna o valor de hora do <i>ITEM</i> fornecido.
datetime_timestamp(YEAR, MONTH, DAY, HOUR, MINUTE, SECOND)	Registro de data e hora	Retorna o valor de registro de data e hora para o <i>YEAR</i> , <i>MONTH</i> , <i>DAY</i> , <i>HOUR</i> , <i>MINUTE</i> e <i>SECOND</i> fornecidos.
datetime_timestamp(DATE, TIME)	Registro de data e hora	Retorna o valor de registro de data e hora para o <i>DATE</i> e <i>TIME</i> fornecidos.
datetime_timestamp(NUMBER)	Registro de data e hora	Retorna o valor de registro de data e hora do número de segundos fornecido.
datetime_weekday(DATE)	Número inteiro	Retorna o dia da semana de <i>DATE</i> ou registro de data e hora fornecidos.
datetime_year(DATE)	Número inteiro	Retorna o ano de um <i>DATE</i> ou registro de data e hora fornecidos. O resultado é um número inteiro como 2002.

Tabela 33. Funções de data e hora do CLEM (continuação)

Função	Resultado	Descrição
date_weeks_difference (DATE1, DATE2)	Real	Retorna o tempo em semanas da data ou registro de data e hora representados por <i>DATE1</i> até os representados por <i>DATE2</i> como um número real. Isso é baseado em uma semana de 7 dias. Se <i>DATE2</i> for anterior a <i>DATE1</i> , essa função retornará um número negativo.
date_years_difference (DATE1, DATE2)	Real	Retorna o tempo, em anos, a partir da data ou do registro de data e hora representados por <i>DATE1</i> até os representados por <i>DATE2</i> como um número real. Essa é uma figura aproximada com base em um ano de 365,25 dias. Se <i>DATE2</i> for anterior a <i>DATE1</i> , essa função retornará um número negativo.
date_from_ywd(YEAR, WEEK, DAY)	Número inteiro	Converte o ano, a semana em ano e o dia em semana, para uma data usando o padrão ISO 8601.
date_iso_day(DATE)	Número inteiro	Retorna o dia da semana a partir da data usando o padrão ISO 8601.
date_iso_week(DATE)	Número inteiro	Retorna a semana do ano a partir da data usando o padrão ISO 8601.
date_iso_year(DATE)	Número inteiro	Retorna o ano da data usando o padrão ISO 8601.
time_add_hours(NUM, DATE)	Data	Retorna o novo TIME ou registro de data e hora após incluir o valor especificado para a parte de tempo especificada HOUR em um TIME ou registro de data e hora.
time_add_hours(NUM, TIMES TAMP)	Registro de data e hora	Retorna o novo TIME ou registro de data e hora após incluir o valor especificado para a parte de tempo especificada HOUR em um TIME ou registro de data e hora.
time_add_minutes(NUM, DATE)	Data	Retorna o novo TIME ou registro de data e hora após incluir o valor especificado para a parte de tempo especificada MINUTE em um TIME ou registro de data e hora.
time_add_minutes(NUM, TIM ESTAMP)	Registro de data e hora	Retorna o novo TIME ou registro de data e hora após incluir o valor especificado para a parte de tempo especificada MINUTE em um TIME ou registro de data e hora.
time_add_seconds(NUM, DATE)	Data	Retorna o novo TIME ou registro de data e hora após incluir o valor especificado para a parte de tempo especificada SECOND em um TIME ou registro de data e hora.
time_add_seconds(NUM, TIM ESTAMP)	Registro de data e hora	Retorna o novo TIME ou registro de data e hora após incluir o valor especificado para a parte de tempo especificada SECOND em um TIME ou registro de data e hora.

Tabela 33. Funções de data e hora do CLEM (continuação)

Função	Resultado	Descrição
time_before(TIME1, TIME2)	Booleano	Retorna um valor de true se <i>TIME1</i> representar uma hora ou registro de data e hora anteriores aos representados por <i>TIME2</i> . Caso contrário, essa função retornará um valor 0.
time_hours_difference(TIME1, TIME2)	Real	Retorna a diferença de tempo em horas entre as horas ou registros de data e hora representados por <i>TIME1</i> e <i>TIME2</i> como um número real. Se você selecionar Dias/Minutos de Rolagem na caixa de diálogo de propriedades do fluxo, um valor superior de <i>TIME1</i> será usado para fazer referência ao dia anterior. Se você não selecionar a opção de rolagem, um valor superior de <i>TIME1</i> fará com que o valor retornado seja negativo.
time_in_hours(TIME)	Real	Retorna o tempo em horas representado por <i>TIME</i> como um número real. Por exemplo, sob o formato de hora HHMM, a expressão <code>time_in_hours('0130')</code> é avaliada para 1,5. <i>TIME</i> pode representar uma hora ou registro de data e hora.
time_in_mins(TIME)	Real	Retorna o tempo em minutos representado por <i>TIME</i> como um número real. <i>TIME</i> pode representar uma hora ou registro de data e hora.
time_in_secs(TIME)	Número inteiro	Retorna o tempo em segundos representado por <i>TIME</i> como um número inteiro. <i>TIME</i> pode representar uma hora ou registro de data e hora.
time_mins_difference(TIME1, TIME2)	Real	Retorna a diferença de tempo em minutos entre as horas ou registros de data e hora representados por <i>TIME1</i> e <i>TIME2</i> como um número real. Se você selecionar Dias/Minutos de Rolagem na caixa de diálogo de propriedades do fluxo, um valor superior de <i>TIME1</i> será usado para fazer referência ao dia anterior (ou à hora anterior, caso apenas minutos e segundos estejam especificados no formato atual). Se você não selecionar a opção de rolagem, um valor superior de <i>TIME1</i> fará com que o valor retornado seja negativo.
time_secs_difference(TIME1, TIME2)	Número inteiro	Retorna a diferença de tempo em segundos entre as horas ou registros de data e hora representados por <i>TIME1</i> e <i>TIME2</i> como um número inteiro. Se você selecionar Dias/Minutos de Rolagem na caixa de diálogo de propriedades do fluxo, um valor superior de <i>TIME1</i> será usado para fazer referência ao dia anterior (ou à hora anterior, caso apenas minutos e segundos estejam especificados no formato atual). Se você não selecionar a opção de rolagem, um valor superior de <i>TIME1</i> fará com que o valor retornado seja negativo.

Convertendo valores de data e hora

Observe que as funções de conversão (e quaisquer outras funções que requerem um tipo específico de entrada, como um valor de data e hora) dependem dos atuais formatos especificados na caixa de diálogo Opções de Fluxo. Por exemplo, se você tivesse um campo denominado *DATE* que é armazenado como uma sequência de caracteres com valores *Jan 2003, Feb 2003*, e assim por diante, você poderia convertê-lo em um armazenamento de datas da seguinte forma:

```
to_date(DATE)
```

Para essa conversão funcionar, selecione o formato de data correspondente **MÊS AAAA** como formato de data padrão para o fluxo.

Para ver um exemplo que converta valores da sequência de caracteres usando um nó de Preenchimento, consulte o fluxo *broadband_create_models.str*, instalado na pasta *|Demos* sob a subpasta *streams*.

Datas armazenadas como números. Observe que *DATE* no exemplo anterior é o nome de um campo, enquanto *to_date* é uma função do CLEM. Se você tiver datas armazenadas como números, é possível convertê-las usando a função *datetime_date*, em que o número é interpretado como um número de segundos desde a data base (ou época).

```
datetime_date(DATE)
```

Ao converter uma data em um número de segundos (e voltar), é possível executar cálculos como data atual mais ou menos um número fixo de dias, por exemplo:

```
datetime_date((date_in_days(DATE) - 7) * 60 * 60 * 24)
```

Funções de sequência

Para algumas operações, a sequência de eventos é importante. O aplicativo permite trabalhar com as sequências de registro a seguir:

- Sequências e séries temporais
- Funções de sequência
- Indexação de registro
- Valores de média, soma e comparação
- Mudança--diferenciação de monitoramento
- @SINCE
- Valores de compensação
- Recursos de sequência adicionais

Para muitos aplicativos, cada registro passando por um fluxo pode ser considerado um caso individual, independente de todos os outros. Nessas situações, a ordem dos registros geralmente é irrelevante.

Para algumas classes de problemas, no entanto, a sequência de registro é muito importante.

Normalmente são situações de séries temporais nas quais a sequência de registros representa uma sequência ordenada de eventos ou ocorrências. Cada registro representa uma captura instantânea em um determinado instante no tempo; muitas das informações mais ricas, porém, podem estar contidas não nos valores instantâneos, mas na maneira como esses valores estão mudando e se comportando ao longo do tempo.

É claro, o parâmetro relevante pode ser qualquer coisa além de tempo. Por exemplo, os registros poderiam representar análises executadas a distâncias ao longo de uma linha, mas os mesmos princípios se aplicariam.

Funções especiais e de sequência são imediatamente reconhecíveis pelas seguintes características:

- Todas elas são prefixadas por @.

- Seus nomes estão em letras maiúsculas.

Funções de sequência podem se referir ao registro atualmente sendo processado por um nó, aos registros que já passaram por um nó e, até, em alguns casos, aos registros que ainda têm que passar por um nó. Funções de sequência podem ser combinadas livremente com outros componentes de expressões do CLEM, embora algumas tenham restrições quanto ao que pode ser usado como seus argumentos.

Exemplos

Talvez você ache útil saber quanto tempo passou desde que ocorreu um certo evento ou que uma condição era verdadeira. Use a função @SINCE para fazer isto—por exemplo:

```
@SINCE(Income > Outgoings)
```

Essa função retorna o offset do último registro no qual essa condição era verdadeira—ou seja, o número de registros antes desse no qual a condição era verdadeira. Se a condição nunca foi verdadeira, @SINCE retornará @INDEX + 1.

Às vezes você pode querer se referir a um valor do registro atual na expressão usada por @SINCE. Isso pode ser feito usando a função @THIS, que especifica que um nome de campo sempre se aplica ao registro atual. Para localizar o offset do último registro que tinha um valor de campo Concentration duas vezes maior que o do registro atual, você poderia usar:

```
@SINCE(Concentration > 2 * @THIS(Concentration))
```

Em alguns casos, é verdadeira a condição dada a @SINCE do registro atual por definição—por exemplo:

```
@SINCE(ID == @THIS(ID))
```

Por esse motivo, @SINCE não avalia sua condição para o registro atual. Use uma função semelhante, @SINCE0, se desejar avaliar a condição para o registro atual, bem como anteriores; se a condição for verdadeira no registro atual, @SINCE0 retornará 0.

Tabela 34. CLEM funções de sequência

Função	Resultado	Descrição
MEAN(FIELD)	Real	Retorna a média dos valores para <i>FIELD</i> ou <i>FIELDS</i> especificados.
@MEAN(FIELD, EXPR)	Real	Retorna a média de valores para <i>FIELD</i> sobre os últimos registros <i>EXPR</i> recebidos pelo nó atual, incluindo o registro atual. <i>FIELD</i> deve ser o nome de um campo numérico. <i>EXPR</i> pode ser qualquer expressão avaliada para um número inteiro maior que 0. Se <i>EXPR</i> for omitido ou se exceder o número de registros recebidos até o momento, a média sobre todos os registros recebidos até o momento é devolvida.

Tabela 34. CLEM funções de sequência (continuação)

Função	Resultado	Descrição
@MEAN(FIELD, EXPR, INT)	Real	Retorna a média de valores para <i>FIELD</i> sobre os últimos registros <i>EXPR</i> recebidos pelo nó atual, incluindo o registro atual. <i>FIELD</i> deve ser o nome de um campo numérico. <i>EXPR</i> pode ser qualquer expressão avaliada para um número inteiro maior que 0. Se <i>EXPR</i> for omitido ou se exceder o número de registros recebidos até o momento, a média de todos os registros recebidos até o momento é retornada. <i>INT</i> especifica o número máximo de valores a reconsiderar. Isso é muito mais eficiente que usar apenas dois argumentos.
@DIFF1(FIELD)	Real	Retorna o primeiro diferenciado de <i>FIELD</i> . O formulário de argumento único simplesmente retorna a diferença entre o valor atual e o valor anterior do campo. Retorna \$null\$ se registros relevantes anteriores não existirem.
@DIFF1(FIELD1, FIELD2)	Real	O formulário de dois argumentos fornece o primeiro diferenciado de <i>FIELD1</i> com relação ao <i>FIELD2</i> . Retorna \$null\$ se registros relevantes anteriores não existirem. É calculado como @DIFF1(FIELD1) / @DIFF1(FIELD2).
@DIFF2(FIELD)	Real	Retorna o segundo diferenciado de <i>FIELD</i> . O formulário de argumento único simplesmente retorna a diferença entre o valor atual e o valor anterior do campo. Retorna \$null\$ se registros relevantes anteriores não existirem. @DIFF2 é calculado como @DIFF(@DIFF(FIELD)).
@DIFF2(FIELD1, FIELD2)	Real	O formulário de dois argumentos fornece o segundo diferenciado de <i>FIELD1</i> com relação ao <i>FIELD2</i> . Retorna \$null\$ se registros relevantes anteriores não existirem. Esse é um cálculo complexo -- @DIFF1(FIELD1) / @DIFF1(FIELD2) - @OFFSET(@DIFF1(FIELD1), 1) / @OFFSET(@DIFF1(FIELD2)) / @DIFF1(FIELD2).
@INDEX	Número inteiro	Retorna o índice do registro atual. Os índices são alocados aos registros à medida que chegam ao nó atual. Ao primeiro registro é dado o índice 1, e o índice é incrementado em 1 para cada registro subsequente.
@LAST_NON_BLANK(FIELD)	qualquer um	Retorna o último valor para <i>FIELD</i> que não estava em branco, conforme definido na origem de envio de dados ou nó Tipo. Se não houver valores não em branco para <i>FIELD</i> nos registros lidos até o momento, \$null\$ será retornado. Observe que os valores não em branco, também chamados valores omissos de usuário, podem ser definidos separadamente para cada campo.
@MAX(FIELD)	Número	Retorna o valor máximo para o <i>FIELD</i> especificado.

Tabela 34. CLEM funções de sequência (continuação)

Função	Resultado	Descrição
@MAX(FIELD, EXPR)	Número	Retorna o valor máximo para <i>FIELD</i> sobre os últimos registros <i>EXPR</i> recebidos até o momento, incluindo o registro atual. <i>FIELD</i> deve ser o nome de um campo numérico. <i>EXPR</i> pode ser qualquer expressão avaliada para um número inteiro maior que 0.
@MAX(FIELD, EXPR, INT)	Número	Retorna o valor máximo para <i>FIELD</i> sobre os últimos registros <i>EXPR</i> recebidos até o momento, incluindo o registro atual. <i>FIELD</i> deve ser o nome de um campo numérico. <i>EXPR</i> pode ser qualquer expressão avaliada para um número inteiro maior que 0. Se <i>EXPR</i> for omitido, ou se exceder o número de registros recebidos até o momento, o valor máximo sobre todos os registros recebidos até agora é retornado. <i>INT</i> especifica o número máximo de valores a reconsiderar. Isso é muito mais eficiente que usar apenas dois argumentos.
@MIN(FIELD)	Número	Retorna o valor mínimo para o <i>FIELD</i> especificado.
@MIN(FIELD, EXPR)	Número	Retorna o valor mínimo para <i>FIELD</i> sobre os últimos registros <i>EXPR</i> recebidos até o momento, incluindo o registro atual. <i>FIELD</i> deve ser o nome de um campo numérico. <i>EXPR</i> pode ser qualquer expressão avaliada para um número inteiro maior que 0.
@MIN(FIELD, EXPR, INT)	Número	Retorna o valor mínimo para <i>FIELD</i> sobre os últimos registros <i>EXPR</i> recebidos até o momento, incluindo o registro atual. <i>FIELD</i> deve ser o nome de um campo numérico. <i>EXPR</i> pode ser qualquer expressão avaliada para um número inteiro maior que 0. Se <i>EXPR</i> for omitido, ou se exceder o número de registros recebidos até o momento, o valor mínimo sobre todos os registros recebidos até o momento é retornado. <i>INT</i> especifica o número máximo de valores a reconsiderar. Isso é muito mais eficiente que usar apenas dois argumentos.

Tabela 34. CLEM funções de sequência (continuação)

Função	Resultado	Descrição
@OFFSET(FIELD, EXPR)	qualquer um	<p>Retorna o valor de <i>FIELD</i> no offset de registro do registro atual pelo valor de <i>EXPR</i>. Um offset positivo se refere a um registro que já passou (um "lookback"), enquanto um negativo especifica um "lookahead" para um registro que ainda vai chegar. Por exemplo, @OFFSET(Status, 1) retorna o valor do campo Status no registro anterior, enquanto @OFFSET(Status, -4) representa um "lookahead" de quatro registros na sequência (ou seja, para registros que ainda não passaram por esse nó) para obter o valor. <i>Observe que um offset negativo (lookahead) deve ser especificado como uma constante.</i> Somente para offsets positivos, <i>EXPR</i> também pode ser uma expressão do CLEM arbitrária avaliada para o registro atual para fornecer o offset. Nesse caso, a versão de três argumentos dessa função deve melhorar o desempenho (consulte a próxima função). Se a expressão retornar qualquer coisa que não seja um número inteiro não negativo, isso causará um erro—ou seja, não é legal ter offsets lookahead calculados.</p> <p><i>Nota:</i> Uma função @OFFSET de autorreferência não pode usar lookahead literal. Por exemplo, em um nó de Preenchimento, não é possível substituir o valor de <i>field1</i> usando uma expressão como @OFFSET(<i>field1</i>, -2).</p> <p><i>Nota:</i> No nó de Preenchimento, durante o preenchimento de um campo, há dois valores efetivamente diferentes desse campo, a saber, o valor pré-preenchido e o valor pós-preenchido. Quando @OFFSET se refere a si mesmo, ele se refere ao valor pré-preenchido. Esse valor pós-preenchido existe somente para linhas antigas, portanto, o @OFFSET de autorreferência só pode se referir a linhas antigas. Como o @OFFSET de autorreferência não pode se referir ao futuro, ele executa as seguintes verificações de offset:</p> <ul style="list-style-type: none"> • Se o offset for literal, e no futuro, um erro será relatado antes do início da execução. • Se o offset for uma expressão e for avaliado para o futuro no tempo de execução, @OFFSET retornará \$null\$. <p><i>Nota:</i> O uso de "lookahead" e "lookback" dentro de um nó não é suportado.</p>

Tabela 34. CLEM funções de sequência (continuação)

Função	Resultado	Descrição
<code>@OFFSET(FIELD, EXPR, INT)</code>	<i>qualquer um</i>	<p>Executa a mesma operação que a função <code>@OFFSET</code> com a adição de um terceiro argumento, <i>INT</i>, o que especifica o número máximo de valor para lookback. Nos casos em que o offset é calculado a partir de uma expressão, o terceiro argumento deve melhorar o desempenho.</p> <p>Por exemplo, em uma expressão como <code>@OFFSET(Foo, Month, 12)</code>, o sistema sabe manter somente os últimos doze valores de <i>Foo</i>; caso contrário, ele terá que armazenar cada valor se for o caso. Nos casos em que o valor do offset for uma constante—incluindo offsets "lookahead" negativos, que devem ser constantes—o terceiro argumento será sem pontos e a versão de dois argumentos dessa função deverá ser usada. Consulte também a nota sobre funções de autorreferência na versão de dois argumentos descrita anteriormente.</p> <p><i>Nota:</i> O uso de "lookahead" e "lookback" dentro de um nó não é suportado.</p>
<code>@SDEV(FIELD)</code>	<i>Real</i>	Retorna o desvio padrão dos valores para <i>FIELD</i> ou <i>FIELDS</i> especificados.
<code>@SDEV(FIELD, EXPR)</code>	<i>Real</i>	Retorna o desvio padrão de valores para <i>FIELD</i> sobre os últimos registros <i>EXPR</i> recebidos pelo nó atual, incluindo o registro atual. <i>FIELD</i> deve ser o nome de um campo numérico. <i>EXPR</i> pode ser qualquer expressão avaliada para um número inteiro maior que 0. Se <i>EXPR</i> for omitido, ou se exceder o número de registros recebidos até o momento, o desvio padrão sobre todos os registros recebidos até o momento é retornado.
<code>@SDEV(FIELD, EXPR, INT)</code>	<i>Real</i>	Retorna o desvio padrão de valores para <i>FIELD</i> sobre os últimos registros <i>EXPR</i> recebidos pelo nó atual, incluindo o registro atual. <i>FIELD</i> deve ser o nome de um campo numérico. <i>EXPR</i> pode ser qualquer expressão avaliada para um número inteiro maior que 0. Se <i>EXPR</i> for omitido, ou se exceder o número de registros recebidos até o momento, o desvio padrão sobre todos os registros recebidos até o momento é retornado. <i>INT</i> especifica o número máximo de valores a reconsiderar. Isso é muito mais eficiente que usar apenas dois argumentos.
<code>@SINCE(EXPR)</code>	<i>qualquer um</i>	Retorna o número de registros que passaram desde que <i>EXPR</i> , uma expressão do CLEM arbitrária, era verdadeira.
<code>@SINCE(EXPR, INT)</code>	<i>qualquer um</i>	A inclusão do segundo argumento, <i>INT</i> , especifica o número máximo de registros para lookback. Se <i>EXPR</i> nunca foi verdadeira, <i>INT</i> será <code>@INDEX+1</code> .

Tabela 34. CLEM funções de sequência (continuação)

Função	Resultado	Descrição
@SINCE0(EXPR)	qualquer um	Considera o registro atual, enquanto @SINCE não; @SINCE0 retornará 0 se EXPR for verdadeira para o registro atual.
@SINCE0(EXPR, INT)	qualquer um	A inclusão do segundo argumento, INT, especifica o número máximo de registros para o lookback.
@SUM(FIELD)	Número	Retorna a soma de valores para o FIELD ou FIELDS especificados.
@SUM(FIELD, EXPR)	Número	Retorna a soma de valores para FIELD sobre os últimos registros EXPR recebidos pelo nó atual, incluindo o registro atual. FIELD deve ser o nome de um campo numérico. EXPR pode ser qualquer expressão avaliada para um número inteiro maior que 0. Se EXPR for omitido, ou se exceder o número de registros recebidos até o momento, será retornada a soma de todos os registros recebidos até o momento.
@SUM(FIELD, EXPR, INT)	Número	Retorna a soma de valores para FIELD sobre os últimos registros EXPR recebidos pelo nó atual, incluindo o registro atual. FIELD deve ser o nome de um campo numérico. EXPR pode ser qualquer expressão avaliada para um número inteiro maior que 0. Se EXPR for omitido, ou se exceder o número de registros recebidos até o momento, será retornada a soma de todos os registros recebidos até o momento. INT especifica o número máximo de valores a reconsiderar. Isso é muito mais eficiente que usar apenas dois argumentos.
@THIS(FIELD)	qualquer um	Retorna o valor do campo denominado FIELD no registro atual. Usado somente em expressões @SINCE.

Funções globais

As funções @MEAN, @SUM, @MIN, @MAX e @SDEV trabalham em, no máximo, todos os registros lidos até e incluindo o atual. Em alguns casos, no entanto, é útil poder trabalhar para descobrir como os valores no registro atual se comparam com os valores vistos no conjunto de dados inteiro. Usando um nó Configurar Globais para gerar valores no conjunto de dados inteiro, é possível acessar esses valores em uma expressão do CLEM usando as funções globais.

Por exemplo,

```
@GLOBAL_MAX(Age)
```

retorna o valor mais alto de Age no conjunto de dados, enquanto a expressão

```
(Value - @GLOBAL_MEAN(Value)) / @GLOBAL_SDEV(Value)
```

expressa a diferença entre o Value desse registro e a média global como um número de desvios padrão. É possível usar valores globais somente após eles terem sido calculados por um nó Configurar Globais. Todos os valores globais atuais podem ser cancelados clicando no botão **Limpar Valores Globais** na guia Globais na caixa de diálogo de propriedades do fluxo.

Tabela 35. CLEM funções globais

Função	Resultado	Descrição
@GLOBAL_MAX(FIELD)	Número	Retorna o valor máximo para <i>FIELD</i> sobre o conjunto de dados inteiro conforme gerado anteriormente por um nó Configurar Globais. <i>FIELD</i> deve ser o nome de um campo numérico, de data / hora / data/hora ou de sequência de caracteres. Se o valor global correspondente não tiver sido configurado, ocorrerá um erro.
@GLOBAL_MIN(FIELD)	Número	Retorna o valor mínimo para <i>FIELD</i> sobre o conjunto de dados inteiro conforme gerado anteriormente por um nó Configurar Globais. <i>FIELD</i> deve ser o nome de um campo numérico, de data / hora / data/hora ou de sequência de caracteres. Se o valor global correspondente não tiver sido configurado, ocorrerá um erro.
@GLOBAL_SDEV(FIELD)	Número	Retorna o desvio padrão de valores para <i>FIELD</i> sobre o conjunto de dados inteiro conforme gerado anteriormente por um nó Configurar Globais. <i>FIELD</i> deve ser o nome de um campo numérico. Se o valor global correspondente não tiver sido configurado, ocorrerá um erro.
@GLOBAL_MEAN(FIELD)	Número	Retorna a média de valores para <i>FIELD</i> sobre o conjunto de dados inteiro conforme gerado anteriormente por um nó Configurar Globais. <i>FIELD</i> deve ser o nome de um campo numérico. Se o valor global correspondente não tiver sido configurado, ocorrerá um erro.
@GLOBAL_SUM(FIELD)	Número	Retorna a soma de valores para <i>FIELD</i> sobre o conjunto de dados inteiro conforme gerado anteriormente por um nó Configurar Globais. <i>FIELD</i> deve ser o nome de um campo numérico. Se o valor global correspondente não tiver sido configurado, ocorrerá um erro.

Funções manipulando valores em branco e nulos

Usando CLEM, é possível especificar que certos valores em um campo devem ser considerados valores "em branco" ou omissos. As funções a seguir trabalham com os em branco.

Tabela 36. CLEM funções de valor em branco e nulo

Função	Resultado	Descrição
@BLANK(FIELD)	Booleano	Retorna verdadeiro para todos os registros cujos valores estão em branco de acordo com o conjunto de regras de manipulação de brancos em um nó Tipo de envio de dados ou nó de origem (guia Tipos).
@LAST_NON_BLANK(FIELD)	qualquer um	Retorna o último valor para <i>FIELD</i> que não estava em branco, conforme definido na origem de envio de dados ou nó Tipo. Se não houver valores não em branco para <i>FIELD</i> nos registros lidos até o momento, \$null\$ será retornado. Observe que os valores não em branco, também chamados valores omissos de usuário, podem ser definidos separadamente para cada campo.

Tabela 36. CLEM funções de valor em branco e nulo (continuação)

Função	Resultado	Descrição
@NULL(FIELD)	Booleano	Retorna true se o valor de FIELD for o system-missing \$null\$. Retorna falso para todos os outros valores, incluindo espaços em branco definidos pelo usuário. Se desejar verificar ambos, use @BLANK(FIELD) e @NULL(FIELD).
undef	qualquer um	Usado geralmente no CLEM para inserir um valor \$null\$ – por exemplo, para preencher valores em branco com nulos no nó de Preenchimento.

Campos em branco devem ser "preenchidos" com o nó de Preenchimento. Em ambos os nós Filler e Derivam (modo múltiplo apenas), a função CLEM especial @FIELD refere-se ao (s) campo (s) atual (s) sendo examinado.

Campos Especiais

Funções especiais são usadas para denotar os campos específicos sendo examinados ou para gerar uma lista de campos como entrada. Por exemplo, ao derivar vários campos de uma vez, é necessário usar @FIELD para denotar "executar esta ação de derivação nos campos selecionados". O uso da expressão log(@FIELD) deriva um novo campo de log para cada campo selecionado.

Tabela 37. CLEM campos especiais

Função	Resultado	Descrição
@FIELD	qualquer um	Executa uma ação em todos os campos especificados no contexto de expressão.
@TARGET	qualquer um	Quando uma expressão CLEM é usada em uma função de análise definida pelo usuário, @TARGET representa o campo de destino ou "valor correto" para o par destino / previsto que está sendo analisado. Geralmente essa função é usada em um nó Análise.
@PREDICTED	qualquer um	Quando uma expressão do CLEM é usada em uma função de análise definida pelo usuário, @PREDICTED representa o valor previsto para o par de destino/previsto sendo analisado. Geralmente essa função é usada em um nó Análise.
@PARTITION_FIELD	qualquer um	Substitui o nome do atual campo de partição.

Tabela 37. CLEM campos especiais (continuação)

Função	Resultado	Descrição
@TRAINING_PARTITION	<i>qualquer um</i>	Retorna o valor da atual partição de treinamento. Por exemplo, para selecionar registros de treinamento usando um nó Select, use a expressão CLEM : @PARTITION_FIELD = @TRAINING_PARTITION Isto assegura que o nó Select sempre funcionará independentemente de quais valores são usados para representar cada partição nos dados.
@TESTING_PARTITION	<i>qualquer um</i>	Retorna o valor da atual partição de teste.
@VALIDATION_PARTITION	<i>qualquer um</i>	Retorna o valor da atual partição de validação.
@FIELDS_BETWEEN(start, end)	<i>qualquer um</i>	Retorna a lista de nomes de campo entre os campos de início e encerramento especificados (inclusivos) com base na ordem natural (ou seja, inserção) dos campos nos dados.
@FIELDS_MATCHING(pattern)	<i>qualquer um</i>	Retorna uma lista de nomes de campos correspondentes a um padrão especificado. Um ponto de interrogação (?) pode ser incluído no padrão para combinar exatamente um caractere; um asterisco (*) corresponde a zero ou mais caracteres. Para corresponder a um ponto de interrogação literal ou asterisco (em vez de usá-los como curinga), uma barra invertida (\) pode ser usada como um caractere de escape. Nota: Isto requer uma sequência literal como um argumento; não pode usar uma expressão aninhada para gerar o argumento.
@MULTI_RESPONSE_SET	<i>qualquer um</i>	Retorna a lista de campos no conjunto de múltiplas respostas.

Capítulo 12. Usando IBM SPSS Modelador com um repositório

Sobre o IBM SPSS Reppositório de serviços de colaboração e implantação

O ModeladorSPSS pode ser usado em conjunção com um repositório do IBM SPSS Serviços de colaboração e implantação, permitindo que você gerencie o ciclo de vida de modelos de mineração de dados e objetos preditivos relacionados, e permitindo que esses objetos sejam usados por aplicativos corporativos, ferramentas e soluções. IBM SPSS Modelador objetos que podem ser compartilhados desta forma incluem fluxos de fluxos, nós, saídas de fluxo, projetos e modelos. Os objetos são armazenados no repositório central, de onde eles podem ser compartilhados com outros aplicativos e controlados usando versão estendida, metadados e recursos de procura.

Antes que possa usar o ModeladorSPSS com o repositório, é necessário instalar um adaptador no host do repositório. Sem esse adaptador, você poderá ver a seguinte mensagem ao tentar acessar os objetos do repositório a partir de determinados nós ou modelos do ModeladorSPSS:

The repository may need updating to support new node, model and output types.

Para obter instruções sobre a instalação do adaptador, veja o guia *ModeladorSPSS Instalação de Implantação*, disponível como arquivo PDF como parte do seu download do produto. Os detalhes de como acessar os objetos de repositório do IBM SPSS Modelador a partir do IBM SPSS Gerente de implantação são fornecidos no *Guia de Implementação do ModeladorSPSS*.

As seções a seguir fornecem informações sobre como acessar o repositório de dentro do ModeladorSPSS.

Versão Extensiva e Suporte de Procura

O repositório fornece a versão do objeto e os recursos de procura abrangentes. Por exemplo, suponha que você crie um fluxo e armazene-o no repositório no qual ele pode ser compartilhado com pesquisadores de outras divisões. Se você atualizar o fluxo posteriormente no ModeladorSPSS, poderá incluir a versão atualizada no repositório sem sobrescrever a versão anterior. Todas as versões permanecem acessíveis e podem ser procuradas pelo nome, rótulo, campos usados ou outros atributos. Você pode, por exemplo, procurar por todas as versões do modelo, que usam receita líquida como uma entrada ou todos os modelos criados por um autor específico. (Para fazer isso com um sistema de arquivos tradicional, você teria que salvar cada versão com um nome de arquivo diferente, e os relacionamentos entre as versões seriam desconhecidos para o software).

conexão única

A variável de conexão única permite que os usuários se conectem ao repositório sem precisar inserir detalhes do nome de usuário e senha todas as vezes. Os detalhes de login da rede local existente do usuário fornecem a autenticação necessária para IBM SPSS Serviços de colaboração e implantação. Essa variável depende do seguinte:

- O IBM SPSS Serviços de colaboração e implantação deve ser configurado para usar um provedor de conexão única.
- O usuário deve estar com login efetuado em um host que seja compatível com o provedor.

Para obter mais informações, consulte “[Conectando ao repositório](#)” na página 206.

Armazenando e implementando objetos do repositório

Fluxos criados no IBM SPSS Modelador podem ser **armazenados** no repositório exatamente como eles são, como arquivos com a extensão **.str**. Dessa maneira, um único fluxo pode ser acessado por diversos usuários na empresa. Consulte o tópico “[Armazenando objetos no repositório](#)” na página 207 para obter mais informações.

Também é possível implementar um fluxo no repositório. Um fluxo implementado é armazenado como um arquivo com metadados adicionais. Um fluxo implementado pode tirar total proveito das variáveis de nível corporativo do IBM SPSS Serviços de colaboração e implantação, como escoragem automatizada e atualização de modelo. Por exemplo, um modelo pode ser automaticamente atualizado em intervalos planejados regularmente conforme novos dados se tornam disponíveis. Alternativamente, um conjunto de fluxos pode ser implementado para análise do Champion Challenger, na qual os fluxos são comparados para determinar quais contêm o modelo preditivo mais efetivo.

Nota: Os nós de modelagem Regras de Associação, STP e TCM não suportam os passos Avaliação de Modelo ou Champion Challenger no IBM SPSS Serviços de colaboração e implantação.

Você pode implementar um fluxo (com a extensão **.str**). A implementação como um fluxo permite que o fluxo seja usado pelo aplicativo thin client VantagensIBM SPSS Modeler. Consulte o tópico “[Abrindo um Fluxo em VantagensIBM SPSS Modeler](#)” na página 223 para obter mais informações.

Para obter mais informações, consulte “[Opcões de Implementação do Fluxo](#)” na página 217.

Outras opções de implementação

Enquanto o IBM SPSS Serviços de colaboração e implantação oferece as variáveis mais extensivas para gerenciar o conteúdo corporativo, inúmeros outros mecanismos para implementar ou exportar fluxos também estão disponíveis, incluindo:

- Exportar o fluxo e o modelo para uso posterior com o Editor de soluçõesIBM SPSS Modeler Runtime.
- Exportar um ou mais modelos em PMML, um formato baseado em XML para codificar informações de modelo. Consulte o tópico “[Importando e exportando modelos como PMML](#)” na página 223 para obter informações adicionais.

Conectando ao repositório

1. Para se conectar a um repositório, no menu principal do IBM SPSS Modelador, clique em:

Ferramentas > Re却tório > Opcões ...

2. Na **RepositoryURL**, digite ou selecione o caminho do diretório ou o endereço URL da instalação do repositório que deseja acessar. Só é possível se conectar a um repositório por vez.

As configurações são específicas de cada site ou instalação. Para detalhes de login específicos, entre em contato com o administrador do sistema local.

Definir Credenciais. Deixe essa caixa desmarcada para ativar a variável *conexão única*, que tenta efetuar seu login usando detalhes de nome de usuário e senha do computador local. Se a conexão única não for possível, ou se você selecionar esta opção para desativar a conexão única (por exemplo, para fazer login em uma conta de administrador), uma tela será exibida para você inserir suas credenciais.

Inserindo credenciais para o repositório

Dependendo de suas configurações, os campos a seguir podem ser necessários na caixa de diálogo Re却tório: Credenciais:

ID do usuário e senha. Especifique um nome de usuário válido e uma senha para efetuar login. Se for necessário, entre em contato com o administrador local para obter mais informações.

Provedor. Escolha um provedor de segurança para autenticação. O repositório pode ser configurado para usar diferentes provedores de segurança; se for necessário, entre em contato com seu administrador local para obter mais informações.

Lembrar repositório e ID do usuário. Salva as atuais configurações como o padrão para que não seja preciso reinseri-las cada vez que você desejar se conectar.

Navegação pelas credenciais do repositório

Quando você se conecta a um repositório a partir de um nó de origem do Analytic Server, Cognos, ODBC ou TM1, é possível selecionar credenciais registradas anteriormente para se conectar a um repositório. Essas credenciais são listadas na caixa de diálogo **Selecionar Credencial do Re却itório**. Para marcar essa caixa de diálogo, clique em Navegar próximo do campo **Credencial**.

Na caixa de diálogo **Selecionar Credencial do Re却itório**, destaque as credenciais na lista fornecida e clique em OK. Se a lista for muito grande, use o campo **Filtro** para inserir um nome, ou parte dele, para localizar as credenciais necessárias.

Navegando no conteúdo do repositório

O repositório permite navegar no conteúdo armazenado de uma maneira semelhante ao Windows Explorer; também é possível navegar em versões de cada objeto armazenado.

1. Para abrir a janela do IBM SPSS Re却itório de serviços de colaboração e implantação, nos menus do ModeladorSPSS, clique em:

Ferramentas > Re却itório > Explore ...

1. Especifique configurações de conexão para o repositório, caso seja necessário. Consulte o tópico “Conectando ao repositório” na página 206 para obter informações adicionais. Para porta, senha e outros detalhes de conexão específicos, entre em contato com o administrador do sistema local.

A janela do Explorer exibe inicialmente uma visualização em árvore da hierarquia de pastas. Clique no nome da pasta para exibir seu conteúdo.

Objetos que correspondem aos atuais critérios de procura ou seleção são listados na área de janela direita, com informações detalhadas sobre a versão selecionada exibida na área de janela inferior direita. Os atributos exibidos se aplicam à versão mais recente.

Armazenando objetos no repositório

É possível armazenar fluxos, nós, modelos, paletas de modelos, projetos e objetos de saída no repositório, onde eles podem ser acessados por outros usuários e aplicativos.

Também é possível publicar saída de fluxo no repositório em um formato que permita que outros usuários a visualizem pela Internet usando IBM SPSS Portal de implantação de serviços de colaboração e implantação.

Configurando propriedades do objeto

Quando você armazena um objeto, a caixa de diálogo Re却itório: Armazenamento é exibida, permitindo configurar os valores de inúmeras propriedades para o objeto. Será possível:

- Escolher o nome e a pasta do repositório sob os quais o objeto deve ser armazenado
- Incluir informações sobre o objeto, como rótulo de versão e outras propriedades pesquisáveis
- Designar um ou mais tópicos de classificação ao objeto
- Configurar opções de segurança para o objeto

As seções a seguir descrevem as propriedades que podem ser configuradas.

Escolhendo a localização para armazenar objetos

Na caixa de diálogo Re却itório: Armazenamento, insira o seguinte.

Salvar em. Mostra a pasta atual--a localização onde o objeto será armazenado. Dê um clique duplo no nome da pasta na lista para configurar essa pasta como a atual. Use o botão Pasta para Cima para navegar para a pasta pai. Use o botão Nova Pasta para criar uma pasta no nível atual.

Nome do arquivo. O nome sob o qual o objeto será armazenado.

Armazenamento. Armazena o objeto na localização atual.

Incluindo informações sobre objetos armazenados

Todos os campos na guia Informações da caixa de diálogo Repositório: Armazenamento são opcionais.

Autor. O nome de usuário do usuário criando o objeto no repositório. Por padrão, isso mostra o nome de usuário usado para conexão do repositório, mas é possível alterar esse nome aqui.

Rótulo de Versão. Selecione um rótulo da lista para indicar a versão do objeto ou clique em **Incluir** para criar um novo rótulo. Evite usar o caractere "[" no rótulo. Assegure-se de que nenhuma caixa esteja marcada se não desejar designar um rótulo a essa versão do objeto. Consulte o tópico ["Visualizando e editando propriedades do objeto"](#) na página 215 para obter mais informações.

DESCRIPTION. Uma descrição do objeto. Os usuários podem procurar objetos por descrição (consulte a nota).

Palavras-chave. Uma ou mais palavras-chave relacionadas ao objeto e que podem ser usadas para propósitos de procura (consulte a nota).

Expiração. Uma data após a qual o objeto não estará mais visível para usuários gerais, embora ainda possa ser visto por seu proprietário e pelo administrador do repositório. Para configurar uma data de expiração, selecione a opção **Data** e insira a data ou escolha uma usando o botão de calendário.

Armazenamento. Armazena o objeto na localização atual.

Nota: Informações nos campos **Descrição** e **Palavras-chave** são tratadas como distintas de tudo que é inserido no ModeladorSPSS na guia Anotações do objeto. Uma procura no repositório por descrição ou palavra-chave não retorna informações da guia Anotações. Veja o tópico ["Procurando objetos no repositório"](#) na página 212 para obter mais informações.

Designando tópicos a um objeto armazenado

Tópicos são um sistema de classificação hierárquico para o conteúdo armazenado no repositório. É possível escolher entre os tópicos disponíveis durante o armazenamento de objetos, e os usuários também podem procurar objetos por tópico. A lista de tópicos disponíveis é configurada pelos usuários do repositório com os privilégios apropriados (para obter mais informações, consulte *Deployment Manager User's Guide*).

Para designar um tópico ao objeto, na guia Tópicos da caixa de diálogo Repositório: Armazenamento:

1. Clique no botão **Incluir**.
2. Clique em um nome de tópico da lista de tópicos disponíveis.
3. Clique em **OK**.

Para remover uma designação de tópico:

4. Selecione o tópico na lista de tópicos designados.
5. Clique em **Excluir**.

Configurando opções de segurança para objetos armazenados

É possível configurar ou alterar inúmeras opções de segurança para um objeto armazenado na guia Segurança da caixa de diálogo Repositório: Armazenamento. Para um ou mais **principais** (ou seja, usuários ou grupos de usuários), é possível:

- Designar direitos de acesso ao objeto
- Modificar direitos de acesso ao objeto

- Remover direitos de acesso ao objeto

Principal. O nome de usuário do repositório do usuário ou grupo de tem direitos de acesso nesse objeto.

Permissões. Os direitos de acesso que este usuário ou grupo tem para o objeto.

Incluir. Permite incluir um ou mais usuários ou grupos na lista daqueles que têm direitos de acesso neste objeto. Veja o tópico [“Incluindo um usuário na lista de permissões”](#) na página 209 para obter mais informações.

Modificar. Permite modificar os direitos de acesso do usuário ou grupo selecionado para este objeto. O acesso de leitura é concedido por padrão. Essa opção permite conceder direitos de acesso adicionais, ou seja, Permissões de Proprietário, Gravar, Excluir e Modificar.

Excluir. Exclui o usuário ou grupo selecionado da lista de permissões para este objeto.

Incluindo um usuário na lista de permissões

Os campos a seguir estão disponíveis quando você seleciona **Incluir** na guia Segurança da caixa de diálogo Repositório: Armazenamento:

Selecionar provedor. Escolha um provedor de segurança para autenticação. O repositório pode ser configurado para usar diferentes provedores de segurança; se for necessário, entre em contato com seu administrador local para obter mais informações.

Localizar. Insira o nome de usuário do repositório do usuário ou grupo que deseja incluir e clique em **Procurar** para exibir o nome na lista de usuários. Para incluir mais de um nome de usuário por vez, deixe esse campo em branco e clique em **Procurar** para exibir uma lista de todos os nomes de usuário do repositório.

Lista de usuários. Selecione um ou mais nomes de usuário da lista e clique em OK para incluí-los na lista de permissões.

Modificando direitos de acesso para um objeto

Os campos a seguir estão disponíveis quando você seleciona **Modificar** na guia Segurança da caixa de diálogo Repositório: Armazenamento:

Proprietário. Selecione esta opção para dar a esse usuário ou grupo direitos de acesso de proprietário ao objeto. O proprietário tem controle total sobre o objeto, incluindo os direitos de acesso Excluir e Modificar.

Ler. Por padrão, um usuário ou grupo que não é proprietário do objeto tem somente direitos de acesso de Leitura ao objeto. Marque as caixas de seleção apropriadas para incluir direitos de acesso de Gravação, Exclusão e Modificação de Permissões para esse usuário ou grupo.

Armazenando fluxos

É possível armazenar um fluxo como um arquivo .str no repositório, de onde ele pode ser acessado por outros usuários.

Nota: Para obter informações sobre como implementar um fluxo, para aproveitar as vantagens das variáveis de repositório adicionais, consulte [“Implementando fluxos”](#) na página 217.

Para armazenar o fluxo atual:

1. No menu principal, clique em:

Arquivo > Loja > Loja como Fluxo ...

2. Especifique configurações de conexão para o repositório, caso seja necessário. Consulte o tópico [“Conectando ao repositório”](#) na página 206 para obter informações adicionais. Para porta, senha e outros detalhes de conexão específicos, entre em contato com o administrador do sistema local.
3. Na caixa de diálogo Repositório: Armazenamento, escolha a pasta na qual deseja armazenar o objeto, especifique quaisquer outras informações que deseja registrar e clique no botão **Armazenamento**. Veja o tópico [“Configurando propriedades do objeto”](#) na página 207 para obter mais informações.

Armazenando projetos

É possível armazenar um projeto do IBM SPSS Modelador completo como um arquivo *.cpj* no repositório para que ele possa ser acessado por outros usuários.

Como um arquivo de projeto é um contêiner para outros objetos do IBM SPSS Modelador, você precisa dizer ao IBM SPSS Modelador para armazenar os objetos do projeto no repositório. Isso é feito usando uma configuração na caixa de diálogo Propriedades do Projeto. Consulte o tópico [“Configurando propriedades do projeto”](#) na página 229 para obter mais informações.

Após configurar um projeto para armazenar objetos no repositório, sempre que você incluir um novo objeto no projeto, o IBM SPSS Modelador automaticamente solicitará que você armazene o objeto.

Após a conclusão da sessão do IBM SPSS Modelador, deve-se armazenar uma nova versão do arquivo de projeto para que ele se lembre de suas adições. O arquivo de projeto contém (e recupera) automaticamente as versões mais recentes de seus objetos. Se você não incluiu nenhum objeto em um projeto durante uma sessão do IBM SPSS Modelador, não será necessário restaurar o arquivo de projeto. No entanto, deve-se armazenar novas versões para os objetos do projeto (fluxos, saída, e assim por diante) que você alterou.

Para armazenar um projeto

1. Selecione o projeto na guia CRISP-DM ou Classes na área de janela de gerenciadores no IBM SPSS Modelador e, no menu principal, clique em:

Arquivo > Projeto > Projeto de Loja ...

2. Especifique configurações de conexão para o repositório, caso seja necessário. Consulte o tópico [“Conectando ao repositório”](#) na página 206 para obter informações adicionais. Para porta, senha e outros detalhes de conexão específicos, entre em contato com o administrador do sistema local.
3. Na caixa de diálogo Repositório: Armazenamento, escolha a pasta na qual deseja armazenar o objeto, especifique quaisquer outras informações que deseja registrar e clique no botão **Armazenamento**. Veja o tópico [“Configurando propriedades do objeto”](#) na página 207 para obter mais informações.

Armazenando nós

É possível armazenar uma definição de nó individual a partir do fluxo atual como um arquivo *.nod* no repositório, de onde ele pode ser acessado por outros usuários.

Para armazenar um nó:

1. Clique com o botão direito na tela de fluxo e clique em **Armazenar Nó**.
2. Especifique configurações de conexão para o repositório, caso seja necessário. Consulte o tópico [“Conectando ao repositório”](#) na página 206 para obter informações adicionais. Para porta, senha e outros detalhes de conexão específicos, entre em contato com o administrador do sistema local.
3. Na caixa de diálogo Repositório: Armazenamento, escolha a pasta na qual deseja armazenar o objeto, especifique quaisquer outras informações que deseja registrar e clique no botão **Armazenamento**. Veja o tópico [“Configurando propriedades do objeto”](#) na página 207 para obter mais informações.

Armazenando objetos de saída

É possível armazenar um objeto de saída a partir do fluxo atual como um arquivo *.cou* no repositório, onde ele pode ser acessado por outros usuários.

Para armazenar um objeto de saída:

1. Clique no objeto na guia Saídas da área de janela de gerenciadores no ModeladorSPSS e, no menu principal, clique em:

Arquivo > Outputs > Saída de Loja ...

2. Alternativamente, clique com o botão direito em um objeto na guia Saídas e clique em **Armazenar**.

3. Especifique configurações de conexão para o repositório, caso seja necessário. Consulte o tópico ["Conectando ao repositório"](#) na página 206 para obter informações adicionais. Para porta, senha e outros detalhes de conexão específicos, entre em contato com o administrador do sistema local.
4. Na caixa de diálogo Repositório: Armazenamento, escolha a pasta na qual deseja armazenar o objeto, especifique quaisquer outras informações que deseja registrar e clique no botão **Armazenamento**. Veja o tópico ["Configurando propriedades do objeto"](#) na página 207 para obter mais informações.

Armazenando modelos e paletas de modelos

É possível armazenar um modelo individual como um arquivo *.gm* no repositório, de onde ele pode ser acessado por outros usuários. Também é possível armazenar o conteúdo completo da paleta Modelos como um arquivo *.gen* no repositório.

Armazenando um modelo:

1. Clique no objeto na paleta Modelos no ModeladorSPSS e, no menu principal, clique em: **Arquivo > Modelos > Modelo de Loja ...**
2. Alternativamente, clique com o botão direito em um objeto na paleta Modelos e clique em **Armazenar Modelo**.
3. Continue em "Concluindo o procedimento de armazenamento".

Armazenando uma paleta Modelos:

1. Clique com o botão direito no segundo plano da paleta Modelos.
2. No menu pop-up, clique em **Armazenar Paleta**.
3. Continue em "Concluindo o procedimento de armazenamento".

Concluindo o procedimento de armazenamento:

1. Especifique configurações de conexão para o repositório, caso seja necessário. Consulte o tópico ["Conectando ao repositório"](#) na página 206 para obter informações adicionais. Para porta, senha e outros detalhes de conexão específicos, entre em contato com o administrador do sistema local.
2. Na caixa de diálogo Repositório: Armazenamento, escolha a pasta na qual deseja armazenar o objeto, especifique quaisquer outras informações que deseja registrar e clique no botão **Armazenamento**. Veja o tópico ["Configurando propriedades do objeto"](#) na página 207 para obter mais informações.

Recuperando objetos do repositório

É possível recuperar fluxos, modelos, paletas de modelos, nós, projetos e objetos de saída que foram armazenados no repositório.

Nota: Além de usar as opções de menu conforme descrito aqui, também é possível recuperar fluxos, objetos de saída, modelos e paletas de modelo clicando com o botão direito na guia apropriada da área de janelas de gerenciadores na parte superior direita da janela do ModeladorSPSS.

1. Para recuperar um fluxo, no menu principal do IBM SPSS Modelador, clique em:

Arquivo > Recuperar Stream ...

2. Para recuperar um modelo, paleta de modelos, projeto ou objeto de saída, no menu principal do IBM SPSS Modelador, clique em:

File > Models > Recuperar Modelo ...

or

File > Models > Recuperar Modelos Palette ...

or

Arquivo > Projetos > Recuperar Projeto ...

or

Arquivo > Saídas > Recuperar Saída ...

3. Alternativamente, clique com o botão direito na área de janela do projeto ou gerenciadores e clique em **Recuperar** no menu pop-up.
4. Para recuperar um nó, no menu principal do IBM SPSS Modelador, clique em:
Insert > Nó (ou SuperNode) do Repositório ...
 - a. Especifique configurações de conexão para o repositório, caso seja necessário. Consulte o tópico ["Conectando ao repositório"](#) na página 206 para obter informações adicionais. Para porta, senha e outros detalhes de conexão específicos, entre em contato com o administrador do sistema local.
5. Na caixa de diálogo Repositório: Recuperar, navegue para o objeto, selecione-o e clique no botão **Recuperar**. Consulte o tópico para obter mais informações.

Escolhendo um objeto para recuperar

Os campos a seguir estão disponíveis na caixa de diálogo Repositório: Recuperar/Procurar:

Procurar em. Mostra a hierarquia de pastas para a pasta atual. Para navegar para uma pasta diferente, selecione uma na lista para navegar diretamente ou navegue usando a lista de objetos abaixo desse campo.

Botão Pasta para Cima. Navega um nível acima da pasta atual na hierarquia.

Botão Nova Pasta. Cria uma nova pasta no nível atual na hierarquia.

Nome do arquivo. O nome do arquivo de repositório do objeto selecionado. Para recuperar esse objeto, clique em **Recuperar**.

Arquivos de tipo. O tipo de objeto que você escolheu para recuperar. Somente objetos desse tipo, com pastas, são mostrados na lista de objetos. Para exibir objetos de um tipo diferente para recuperação, selecione o tipo de objeto da lista.

Abrir como bloqueado. Por padrão, quando um objeto é recuperado, ele é bloqueado no repositório para que outras pessoas não possam atualizá-lo. Se não desejar que o objeto seja bloqueado na recuperação, desmarque essa caixa.

Descrição, Palavras-chave. Se detalhes adicionais sobre o objeto foram definidos quando o objeto foi armazenado, esses detalhes serão exibidos aqui. Veja o tópico ["Incluindo informações sobre objetos armazenados"](#) na página 208 para obter mais informações.

Versão. Para recuperar uma versão do objeto além da mais recente, clique neste botão. Informações detalhadas para todas as versões são exibidas, permitindo que você escolha a versão desejada.

Selecionando versão de um objeto

Para selecionar uma versão específica de um objeto do repositório, na caixa de diálogo Repositório: Selecionar Versão:

1. (Opcional) Ordene a lista por versão, rótulo, tamanho, data de criação ou usuário criador dando um clique duplo no cabeçalho da coluna apropriada.
2. Selecione a versão do objeto com o qual deseja trabalhar.
3. Clique em Continuar.

Procurando objetos no repositório

É possível procurar objetos por nome, pasta, tipo, rótulo, data ou outros critérios.

Procurando por objetos por nome

1. No menu principal do IBM SPSS Modelador, clique em:

Ferramentas > Repositório > Explore ...

a. Especifique configurações de conexão para o repositório, caso seja necessário. Consulte o tópico [“Conectando ao repositório”](#) na página 206 para obter informações adicionais. Para porta, senha e outros detalhes de conexão específicos, entre em contato com o administrador do sistema local.

2. Clique na guia **Procurar**.

3. No campo **Procurar Objetos Nomeados**, especifique o nome do objeto que deseja localizar.

Ao procurar objetos por nome, um asterisco (*) pode ser usado como um caractere curinga para corresponder a qualquer cadeia de caracteres e um ponto de interrogação (?) corresponde a qualquer caractere único. Por exemplo, `*cluster*` corresponde a todos os objetos que incluem o `cluster` de sequência de caracteres em qualquer lugar no nome. A sequência de caracteres de procura `m0?_*` corresponde a `M01_cluster.str` e `M02_cluster.str`, mas não `M01a_cluster.str`. As procura não fazem distinção entre maiúsculas e minúsculas (`cluster` corresponde a `Cluster` e a `CLUSTER`).

Nota: Se o número de objetos for grande, as buscas podem levar alguns momentos.

Procurando por outros critérios

É possível executar uma procura com base em título, rótulos, datas, autor, palavras-chave, conteúdo indexado ou descrição. Somente em objetos que correspondem a *todos* os critérios de procura especificados serão localizados. Por exemplo, você poderia localizar todos os fluxos contendo um ou mais modelos de armazenamento em cluster que também tenham um rótulo específico aplicado e que foram modificados após uma data específica.

Tipos de Objetos. Você pode restringir a pesquisa a modelos, fluxos, saídas, nós, SuperNodes, projetos, paletas de modelos ou outros tipos de objetos.

• **Modelos.** É possível procurar modelos por categoria (classificação, aproximação, armazenamento em cluster, etc.) ou por um algoritmo de modelagem específico, como Kohonen.

Também é possível procurar por campos usados-por exemplo, todos os modelos que usam um campo denominado *receita* como um campo de entrada ou saída (resposta).

• **Fluxos.** Para fluxos, é possível restringir a procura por campos usados ou tipo de modelo (categoria ou algoritmo) contidos no fluxo.

• É possível procurar modelos associados a tópicos específicos de uma lista configurada por usuários do repositório com os privilégios apropriados (para obter mais informações, consulte *Guia do Usuário do Deployment Manager*). Para obter a lista, marque essa caixa, clique no botão Incluir Tópicos que é exibido, selecione um ou mais tópicos na lista e clique em OK.

Rótulo. Restringe a procura a rótulos de versão do objeto específicos.

Datas. É possível especificar uma data de criação ou modificação e procurar objetos antes, depois ou entre o intervalo de dados especificado.

Autor. Restringe a procura a objetos criados por um usuário específico.

palavras-chave. Procura por palavras-chave específicas. No IBM SPSS Modelador, palavras-chave são especificadas na guia Anotação para um fluxo, modelo ou objeto de saída.

DESCRIPTION. Procura por termos específicos no campo de descrição. No IBM SPSS Modelador, a descrição é especificada na guia Anotação para um modelo, fluxo ou objeto de saída. Várias frases de procura podem ser separadas por ponto e vírgula-por exemplo, `income; crop type; claim value` (Observe que dentro da frase de procura, os espaços são importantes. Por exemplo, `crop type` com um espaço e `crop type` com dois espaços não são iguais.)

Modificando objetos do repositório

É possível modificar objetos existentes no repositório diretamente do ModeladorSPSS. Será possível:

- Criar, renomear ou excluir pastas
- Bloquear ou desbloquear objetos
- Excluir objetos

Criando, renomeando e excluindo pastas

1. Para executar operações em pastas no repositório, no menu principal do ModeladorSPSS, clique em: **Ferramentas > Re却torio > Explore ...**
 - a. Especifique configurações de conexão para o repositório, caso seja necessário. Consulte o tópico “Conectando ao repositório” na página 206 para obter informações adicionais. Para porta, senha e outros detalhes de conexão específicos, entre em contato com o administrador do sistema local.
2. Assegure que a guia **Pastas** esteja ativa.
3. Para criar uma nova pasta, clique com o botão direito na pasta pai e clique em **Nova Pasta**.
4. Para renomear uma pasta, clique nela com o botão direito e clique em **Renomear Pasta**.
5. Para excluir uma pasta, clique nela com o botão direito e clique em **Excluir Pasta**.

Bloqueando e desbloqueando objetos do repositório

É possível bloquear um objeto para evitar que outros usuários atualizem qualquer uma de suas versões existentes ou criem novas versões. Um objeto bloqueado é indicado por um símbolo de cadeado sobre o ícone do objeto.



Figura 17. Objeto bloqueado

Para bloquear um objeto

1. Na janela do explorador de repositório, clique com o botão direito no objeto obrigatório.
2. Clique em **Bloquear**.

Para desbloquear um objeto

1. Na janela do explorador de repositório, clique com o botão direito no objeto obrigatório.
2. Clique em **Desbloquear**.

Excluindo objetos do repositório

Antes de excluir um objeto do repositório, deve-se decidir se você deseja excluir todas as versões do objeto ou somente uma versão específica.

Para Excluir Todas as Versões de um Objeto

1. Na janela do explorador de repositório, clique com o botão direito no objeto obrigatório.
2. Clique em **Excluir Objetos**.

Para Excluir a Versão Mais Recente de um Objeto

1. Na janela do explorador de repositório, clique com o botão direito no objeto obrigatório.
2. Clique em **Excluir**.

Para Excluir uma Versão Anterior de um Objeto

1. Na janela do explorador de repositório, clique com o botão direito no objeto obrigatório.
2. Clique em **Excluir Versões**.
3. Selecione a(s) versão(ões) para excluir e clique em **OK**.

Gerenciando propriedades dos objetos do repositório

É possível controlar várias propriedades de objeto a partir do ModeladorSPSS. Será possível:

- Visualizar as propriedades de uma pasta

- Visualizar e editar as propriedades de um objeto
- Criar, aplicar e excluir rótulos de versão para um objeto

Visualizando propriedades da pasta

Para visualizar propriedades para qualquer pasta na janela de repositório, clique com o botão direito na pasta necessária. Clique em **Propriedades da Pasta**.

Guia geral

Esta guia exibe o nome da pasta e as datas de criação e modificação.

Guia Permissões

Nesta guia, você especifica permissões de leitura e gravação para a pasta. Todos os usuários e grupos com acesso à pasta pai são listados. Permissões seguem uma hierarquia. Por exemplo, se você não tiver permissão de leitura, não será possível ter permissão de gravação. Se você não tiver permissão de gravação, não será possível ter permissão de exclusão.

Usuários e Grupos. Lista os usuários e grupos do repositório que têm pelo menos acesso de Leitura a essa pasta. Marque as caixas de seleção Gravar e Excluir para incluir esses direitos de acesso para essa pasta em um determinado usuário ou grupo. Clique no ícone **Incluir Usuários/Grupos** no lado direito da guia Permissões para designar acesso a usuários e grupos adicionais. A lista de usuários e grupos disponíveis é controlada pelo administrador.

Permissões em Cascata. Escolha uma opção para controlar como as mudanças feitas na pasta atual são aplicadas em suas pastas filhas, se houver.

- **Todas as Permissões em Cascata.** Coloca em cascata todas as configurações de permissão, da pasta atual a todas as pastas filhas e descendentes. Essa é uma maneira rápida de configurar permissões para várias pastas de uma vez. Configure permissões conforme necessário para a pasta pai e, em seguida, forme a cascata.
- **Cascata Somente de Mudanças.** Coloca em cascata somente as mudanças feitas desde que as últimas mudanças foram aplicadas. Por exemplo, se um novo grupo tiver sido incluído e você desejar dar a ele acesso a todas as pastas sob a ramificação Vendas, é possível dar ao grupo acesso à pasta raiz Vendas e formar uma cascata das mudanças para todas as subpastas. Todas as outras permissões para subpastas existentes permanecem como antes.
- **Não Formar Cascata.** Quaisquer mudanças feitas se aplicam apenas à pasta atual e não formam uma cascata para pastas filhas.

Visualizando e editando propriedades do objeto

Na caixa de diálogo Propriedades do Objeto, é possível visualizar e editar propriedades. Embora algumas propriedades não possam ser alteradas, sempre é possível atualizar um objeto incluindo uma nova versão.

1. Na janela do repositório, clique com o botão direito no objeto necessário.
2. Clique em **Propriedades do Objeto**.

Guia Geral

Nome. O nome do objeto como ele é visualizado no repositório.

Criado em. Data em que o objeto (e não a versão) foi criado.

Última modificação. Data em que a versão mais recente foi modificada.

Autor. O nome de login do usuário.

DESCRIPTION. Por padrão, contém a descrição especificada na guia Anotação do objeto no ModeladorSPSS.

Tópicos vinculados. O repositório permite que modelos e objetos relacionados sejam organizados por tópicos, se necessário. A lista de tópicos disponíveis é configurada pelos usuários do repositório com os privilégios apropriados (para obter mais informações, consulte *Deployment Manager User's Guide*).

Palavras-chave. Você especifica palavras-chave na guia Anotação para um fluxo, modelo ou objeto de saída. Palavras-chave múltiplas devem ser separadas por espaços, com até no máximo 255 caracteres. (Se as palavras-chave contiverem espaços, use aspas para separá-las.)

Guia Versões

Objetos armazenados no repositório podem ter versões múltiplas. A guia Versões exibe informações sobre cada versão.

As propriedades a seguir podem ser especificadas ou modificadas para versões específicas de um objeto armazenado:

Versão. Identificador exclusivo para a versão gerada com base no horário em que a versão foi armazenada.

Rótulo. O rótulo atual da versão, se houver. Ao contrário do identificador de versão, os rótulos podem ser movidos de uma versão do objeto para outra.

O tamanho do arquivo, a data de criação e o autor também são exibidos para cada versão.

Editar Rótulos. Clique no ícone **Editar Rótulos** na parte superior direita da guia Versões para definir, aplicar, remover rótulos para objetos armazenados. Consulte o tópico “[Gerenciando rótulos de versão do objeto](#)” na página 216 para obter mais informações.

Guia Permissões

Na guia Permissões, é possível configurar permissões de leitura e gravação para o objeto. Todos os usuários e grupos com acesso ao objeto atual são listados. Permissões seguem uma hierarquia. Por exemplo, se você não tiver permissão de leitura, não será possível ter permissão de gravação. Se você não tiver permissão de gravação, não será possível ter permissão de exclusão.

Usuários e Grupos. Lista os usuários e grupos do repositório que têm pelo menos acesso de Leitura a esse objeto. Marque as caixas de seleção Gravar e Excluir para incluir esses direitos de acesso para esse objeto em um determinado usuário ou grupo. Clique no ícone **Incluir Usuários/Grupos** no lado direito da guia Permissões para designar acesso a usuários e grupos adicionais. A lista de usuários e grupos disponíveis é controlada pelo administrador.

Gerenciando rótulos de versão do objeto

A caixa de diálogo Editar Rótulos de Versão permite:

- Aplicar rótulos no objeto selecionado
- Remover rótulos do objeto selecionado
- Definir um novo rótulo e aplicá-lo ao objeto

Para aplicar rótulos no objeto

1. Selecione um ou mais rótulos na lista **Rótulos Disponíveis**.
2. Clique com o botão Seta à Direita para mover os rótulos selecionados para a lista **Rótulos Aplicados**.
3. Clique em **OK**.

Para remover rótulos do objeto

1. Selecione um ou mais rótulos na lista **Rótulos Aplicados**.
2. Clique no botão Seta à Esquerda para mover os rótulos selecionados para a lista **Rótulos Disponíveis**.
3. Clique em **OK**.

Para definir um novo rótulo e aplicá-lo ao objeto

1. Digite o nome do rótulo no campo **Novo Rótulo**.

2. Clique com o botão Seta à Direita para mover o novo rótulo para a lista **Rótulos Aplicados**.
3. Clique em **OK**.

Implementando fluxos

Para permitir que um fluxo seja usado com o aplicativo thin client VantagensIBM SPSS Modeler, ele deve ser implementado como um fluxo (arquivo **.str**) no repositório.

Nota: Não é possível implementar um fluxo que tenha mais de um nó de origem na ramificação de escoragem.

O fluxo pode tirar total proveito dos recursos de nível corporativo de IBM SPSS Serviços de colaboração e implantação. Veja o tópico [“Armazenando e implementando objetos do repositório”](#) na página 206 para obter mais informações.

Para implementar o fluxo atual (método do menu Arquivo)

1. No menu principal, clique em:

Arquivo > Loja > Implementar

2. Escolha o tipo de implementação e conclua o restante da caixa de diálogo conforme necessário.
3. Clique em **Implementar como Fluxo** para implementar o fluxo para usar com VantagensIBM SPSS Modeler ou IBM SPSS Serviços de colaboração e implantação.
4. Clique em **Armazenamento**. Para obter mais informações, clique em **Ajuda**.
5. Continue a partir de “Concluindo o processo de implementação.”

Para implementar o fluxo atual (método do menu Ferramentas)

1. No menu principal, clique em:

Ferramentas > Propriedades do Fluxo > Implementação

2. Escolha o tipo de implementação, conclua o restante da guia Implementação conforme necessário e clique em **Armazenamento**. Veja o tópico [“Opções de Implementação do Fluxo”](#) na página 217 para obter mais informações.

Concluindo o processo de implementação

1. Especifique configurações de conexão para o repositório, caso seja necessário. Consulte o tópico [“Conectando ao repositório”](#) na página 206 para obter informações adicionais. Para porta, senha e outros detalhes de conexão específicos, entre em contato com o administrador do sistema local.
2. Na caixa de diálogo Repositório: escolha a pasta onde deseja armazenar o objeto, especifique qualquer outra informação que deseja gravar, e clique em **Loja**. Veja o tópico [“Configurando propriedades do objeto”](#) na página 207 para obter mais informações.

Opções de Implementação do Fluxo

A guia Implementação na caixa de diálogo Opções de Fluxo permite especificar opções para implementar o fluxo.

Quando você implementa um fluxo, ele é armazenado no repositório como um arquivo com a extensão **.str**.

A implantação de um fluxo permite que você aproveita a funcionalidade adicional disponível com IBM SPSS Serviços de colaboração e implantação, como acesso multiusuário, pontuação automatizada, atualização de modelo e análise de Challenger de Campeão.

Nota: Os nós de modelagem Regras de Associação, STP e TCM não suportam os passos Avaliação de Modelo ou Champion Challenger no IBM SPSS Serviços de colaboração e implantação.

Na guia Implementação, também é possível visualizar a descrição de fluxo que o IBM SPSS Modelador cria para o fluxo. Veja o tópico “[Descrições de fluxo](#)” na página 49 para obter mais informações.

Tipo de implementação. Escolha como deseja implementar o fluxo. Todos os fluxos requerem um nó de escoragem designado para que possam ser implementados; as opções e os requisitos adicionais dependem do tipo de implementação.

Nota: Os nós de modelagem Regras de Associação, STP e TCM não suportam os passos Avaliação de Modelo ou Champion Challenger no IBM SPSS Serviços de colaboração e implantação.

- <none>. O fluxo não será implementado para o repositório. Todas as opções ficam desativadas, exceto a visualização da descrição do fluxo.
- **Somente Escoragem.** O fluxo é implementado no repositório quando você clica no botão **Armazenamento**. Dados podem ser escorados usando o nó designado no campo **Nó de Escoragem**.
- **Atualização de Modelo.** O mesmo que Somente Escoragem mas, além disso, pode ser atualizado no repositório usando os objetos que você designar nos campos **Nó de Modelagem** e **Nugget de Modelo**. Note que a atualização automática do modelo não é suportada por padrão em IBM SPSS Serviços de colaboração e implantação, portanto você deve escolher este tipo de implementação se você deseja usar este recurso ao executar um fluxo do repositório. Consulte “[Atualizar modelo](#)” na página 220 para obter mais informações.

Nó de pontuação. Selecione um gráfico, uma saída ou um nó de exportação para identificar a ramificação de fluxo a ser usada para escoragem dos dados. Embora o fluxo possa realmente conter qualquer número válido de ramificações, modelos e nós terminais, uma, e somente uma, ramificação de escoragem deve ser designada para propósitos de implementação. Esse é o requisito mais básico para implementar qualquer fluxo.

Parâmetros de pontuação. Permite especificar parâmetros que podem ser modificados quando a ramificação de escoragem é executada. Consulte o “[Parâmetros de escoragem e modelagem](#)” na página 218 para obter mais informações.

Modelagem do nó. Para atualização de modelo, especifica o nó de modelagem usado para gerar novamente ou atualizar o modelo no repositório. Deve ser um nó de modelagem do mesmo tipo que o especificado para **Nugget do Modelo**.

Parâmetros de Construção de Modelo. Permite especificar parâmetros que podem ser modificados quando o nó de modelagem é executado. Consulte o “[Parâmetros de escoragem e modelagem](#)” na página 218 para obter mais informações.

Nugget do Modelo. Para atualização de modelo, especifica o nugget do modelo que será atualizado ou gerado novamente cada vez que o fluxo for atualizado no repositório (geralmente como parte de uma tarefa planejada). O modelo deve estar localizado na ramificação de escoragem. Embora possa existir vários modelos na ramificação de escoragem, somente um pode ser designado. Observe que quando o fluxo é criado inicialmente, ele pode ser efetivamente um modelo de item temporário que é atualizado ou gerado novamente conforme novos dados são disponibilizados.

Confira. Clique neste botão para verificar se este é um fluxo válido para implementação. Todos os fluxos devem ter um nó de escoragem designado para que possam ser implementados. Mensagens de erro serão exibidas se essas condições não forem satisfeitas.

Loja. Implementa o fluxo se ele for válido. Se não, uma mensagem de erro é exibida. Clique no botão **Fix**, corrido o erro e tente novamente.

Descrição do Fluxo de Visualização. Permite visualizar o conteúdo da descrição do fluxo que o IBM SPSS Modelador cria para o fluxo. Consulte “[Descrições de fluxo](#)” na página 49 para obter mais informações.

Nota: Os nós de modelagem Regras de Associação, STP e TCM não suportam os passos Avaliação de Modelo ou Champion Challenger no IBM SPSS Serviços de colaboração e implantação.

Parâmetros de escoragem e modelagem

Durante a implementação de um fluxo no IBM SPSS Serviços de colaboração e implantação, é possível escolher quais parâmetros podem ser visualizados ou editados cada vez que o modelo é atualizado ou

escorado. Por exemplo, você pode especificar os valores máximo e mínimo ou algum outro valor que possa estar sujeito à mudança cada vez que uma tarefa for executada.

1. Para deixar um parâmetro visível para que ele possa ser visualizado ou editado após o fluxo ser implementado, selecione-o na lista na caixa de diálogo Parâmetros de Escoragem.

A lista de parâmetros disponíveis é definida na guia Parâmetros na caixa de diálogo de propriedades do fluxo. Consulte o tópico “[Configurando parâmetros de sessão e fluxo](#)” na página 45 para obter informações adicionais.

A ramificação de escoragem

Se você estiver implementando um fluxo, uma ramificação do fluxo deverá ser designada como **ramificação de escoragem** (ou seja, aquela que contém o nó de escoragem). Quando você designa uma ramificação como ramificação de escoragem, ela é destacada na tela de fluxo, como na ligação do modelo com o nugget na ramificação de escoragem. Essa representação visual é particularmente útil em fluxos complexos com diversas ramificações nas quais a ramificação de escoragem pode não ser imediatamente óbvia.

Nota: Somente uma ramificação de fluxo pode ser designada como ramificação de escoragem.

Se o fluxo já tinha uma ramificação de escoragem definida, a ramificação recém-designada a substituirá como ramificação de escoragem. É possível configurar a cor da indicação de ramificação de escoragem por meio de uma opção Cor Customizada. Consulte o tópico “[Configurando opções de exibição](#)” na página 238 para obter informações adicionais.

É possível mostrar ou ocultar a indicação de ramificação de escoragem por meio do botão da barra de ferramentas Mostrar/Ocultar Marcação de Fluxo.



Figura 18. Botão da barra de ferramentas Mostrar/Ocultar Marcação de Fluxo

Identificando a ramificação de escoragem para implementação

É possível designar a ramificação de escoragem a partir do menu pop-up de um nó terminal ou do menu Ferramentas. Se você usar o menu pop-up, o nó de escoragem será configurado automaticamente na guia Implementação das propriedades do fluxo.

Para designar uma ramificação como a ramificação de escoragem (menu pop-up)

1. Conecte o nugget do modelo a um nó terminal (recebimento de dados do nó de saída ou de processamento do nugget).
2. Clique com o botão direito no nó terminal.
3. No menu, clique em **Usar como Ramificação de Escoragem**.

Para designar uma ramificação como ramificação de escoragem (menu Ferramentas)

1. Conecte o nugget do modelo a um nó terminal (recebimento de dados do nó de saída ou de processamento do nugget).
2. No menu principal, clique em:

Ferramentas > Propriedades do Fluxo > Implementação

3. Na lista **Tipo de Implementação**, clique em **Somente Escoragem** ou **Atualização de Modelo** conforme necessário. Consulte o tópico “[Opcões de Implementação do Fluxo](#)” na página 217 para obter mais informações.
4. Clique no campo **Nó de Escoragem** e selecione um nó terminal da lista.
5. Clique em **OK**.

Atualizar modelo

Atualização de modelo é o processo de reconstrução de um modelo existente em um fluxo usando dados mais novos. O fluxo em si não muda no repositório. Por exemplo, o tipo de algoritmo e as configurações específicas do fluxo permanecem iguais, mas o modelo é retreinado nos novos dados e atualizados, caso a nova versão do modelo funcione melhor que a antiga.

Somente um nugget do modelo em um fluxo pode ser configurado--isso é conhecido como **modelo de atualização**. Se você clicar na opção **Atualização de Modelo** na guia Implementação das propriedades do fluxo (consulte “[Opções de Implementação do Fluxo](#)” na página 217), o nugget do modelo designado nesse momento se tornará o modelo de atualização. Também é possível designar um modelo como modelo de atualização no menu pop-up de um nugget do modelo. O nugget já deve estar na ramificação de escoragem para que isso seja possível.

Se você desativar o status do "modelo de atualização" de um nugget, isso será equivalente à configuração do tipo de implementação do fluxo para Somente Escoragem, e a guia Implementação da caixa de diálogo de propriedades do fluxo será atualizada de acordo. É possível ativar e desativar esse status por meio da opção **Usar como Modelo de Atualização** no menu pop-up do nugget na ramificação de escoragem atual.

A remoção da ligação do modelo de um nugget na ramificação de escoragem também remove o status "modelo de atualização" do nugget. É possível desfazer a remoção da ligação do modelo por meio do menu Editar ou da barra de ferramentas; isso também restabelece o status do "modelo de atualização" do nugget.

Como Atualizar Modelo é selecionada

Assim como a ramificação de escoragem, a ligação com um modelo de atualização também é destacada no fluxo. O nugget do modelo escolhido como o modelo de atualização e, portanto, a ligação que é destacada, depende de quantos nuggets estão no fluxo.

Modelo Único no Fluxo

Se houver um único nugget do modelo vinculado na ramificação de escoragem quando ele for identificado como tal, esse nugget se tornará o modelo de atualização para o fluxo.

Vários Modelos no Fluxo

Se houver mais de um nugget vinculado no fluxo, o modelo de atualização será escolhido como tal.

Se um nugget do modelo tiver sido definido na guia Implementação da caixa de diálogo de propriedades do fluxo e também no fluxo, ele se tornará o modelo de atualização.

Se nenhum nugget tiver sido definido na guia Implementação, ou se tiver sido definido, mas não na ramificação de escoragem, o nugget mais próximo do nó terminal se tornará o modelo de atualização.

Se você cancelar subsequente a seleção de todas as ligações de modelo como ligações de atualização, somente a ramificação de escoragem será destacada, não as ligações. O tipo de implementação é configurado para Somente Escoragem.

Nota: É possível escolher configurar uma das ligações para o status Substituir, mas não o outro. Nesse caso, o nugget do modelo escolhido como o modelo de atualização é aquele que tem uma ligação de atualização e que está mais perto do nó terminal quando a ramificação de escoragem é designada.

Nenhum Modelo no Fluxo

Se não houver modelos no fluxo, ou somente modelos sem ligações de modelo, o tipo de implementação será configurado para Somente Escoragem.

Verificando erros em uma ramificação de escoragem

Ao designar a ramificação de pontuação, ela é verificada em busca de erros.

Se um erro for localizado, a ramificação de escoragem será destacada na cor do erro de ramificação de escoragem e uma mensagem de erro será exibida. É possível configurar a cor do erro por meio de uma

opção Cor Customizada. Consulte o tópico “[Configurando opções de exibição](#)” na página 238 para obter informações adicionais.

Se um erro for localizado, continue da seguinte forma:

1. Corrija o erro de acordo com o conteúdo da mensagem de erro.
2. No menu principal, clique em:

Ferramentas > Propriedades do Fluxo > Implementação

e clique em **Verificar**.

3. Se for necessário, repita esse processo até que nenhum erro seja localizado.

Capítulo 13. Exportando para aplicativos externos

Sobre Exportando para aplicativos externos

IBM SPSS Modelador fornece inúmeros mecanismos para você exportar o processo de mineração de dados inteiro para aplicativos externos, de modo que o trabalho realizado para preparar dados e construir modelos possa ser usado em sua vantagem fora do IBM SPSS Modelador também.

A seção anterior mostrou como é possível implementar fluxos em um repositório do IBM SPSS Serviços de colaboração e implantação para aproveitar a vantagem de seu acesso multiusuário, planejamento de tarefas e outras variáveis. Da forma semelhante, os fluxos do IBM SPSS Modelador também podem ser usados com:

- VantagensIBM SPSS Modeler
- Aplicativos que podem importar e exportar arquivos em formato PMML

Para obter mais informações sobre como usar fluxos com VantagensIBM SPSS Modeler, consulte [“Abrindo um Fluxo em VantagensIBM SPSS Modeler”](#) na página 223.

Para obter informações sobre como exportar e importar modelos como arquivos PMML, possibilitando o compartilhamento de modelos com quaisquer outros aplicativos que suportam esse formato, consulte [“Importando e exportando modelos como PMML”](#) na página 223.

Abrindo um Fluxo em VantagensIBM SPSS Modeler

Os fluxos do IBM SPSS Modelador podem ser usados com o aplicativo thin client VantagensIBM SPSS Modeler. Embora seja possível criar aplicativos customizados completamente dentro do VantagensIBM SPSS Modeler, também é possível usar um fluxo já criado no IBM SPSS Modelador como a base de um fluxo de trabalho de aplicativo.

Para abrir um fluxo no VantagensIBM SPSS Modeler:

1. Implemente o fluxo no repositório do IBM SPSS Serviços de colaboração e implantação, certificando-se de clicar na opção **Implementar como Fluxo**. Consulte o tópico [“Implementando fluxos”](#) na página 217 para obter mais informações.
2. Clique no botão da barra de ferramentas Abrir no VantagensIBM SPSS Modeler, ou no menu principal, clique em:

Arquivo > Abrir em IBM SPSS Modeler Advantage

1. Especifique configurações de conexão para o repositório, caso seja necessário. Consulte o tópico [“Conectando ao repositório”](#) na página 206 para obter informações adicionais. Para porta, senha e outros detalhes de conexão específicos, entre em contato com o administrador do sistema local.

Nota: O servidor de repositório também deve ter o software VantagensIBM SPSS Modeler instalado.

1. No diálogo Repositório: Loja o diálogo, escolha a pasta onde deseja armazenar o objeto, especifique qualquer outra informação que deseja gravar, e clique no botão **Loja**. Veja o tópico [“Configurando propriedades do objeto”](#) na página 207 para obter mais informações.

Isso ativa o VantagensIBM SPSS Modeler com o fluxo já aberto. O fluxo é fechado no IBM SPSS Modelador.

Importando e exportando modelos como PMML

PMML, ou Predictive Model Markup Language, é um formato XML para descrever modelos estatísticos e mineração de dados, incluindo entradas para os modelos, transformações usadas para preparar dados para mineração de dados e os parâmetros que definem os modelos em si. IBM SPSS Modelador pode

importar e exportar PMML, possibilitando o compartilhamento de modelos com outros aplicativos que suportem esse formato, como IBM SPSS Estatísticas .

Para obter mais informações sobre PMML, consulte o website Grupo de Mineração de Dados (<http://www.dmg.org>).

Para Exportar um Modelo

A exportação de PMML é suportada para a maioria dos tipos de modelo gerados no IBM SPSS Modelador. Veja o tópico “[Tipos de modelo que suportam PMML](#)” na página 224 para obter mais informações.

1. Clique com o botão direito em um nugget do modelo na paleta de modelos. (Alternativamente, dê um clique duplo em um nugget do modelo na tela e selecione o menu Arquivo.)
2. No menu, clique em **Exportar PMML**.
3. Na caixa de diálogo Exportar (ou Salvar), especifique um diretório de resposta e um nome exclusivo para o modelo.

Nota:

Você pode alterar opções para exportação da PMML na caixa de diálogo Opções do Usuário. No menu principal, clique em:

Ferramentas > Opções > Opções do Usuário

e clique na guia PMML.

Consulte o tópico “[Configurando opções de exportação de PMML](#)” na página 239 para obter informações adicionais.

Para Importar um Modelo Salvo como PMML

Modelos exportados como PMML do IBM SPSS Modelador ou ouro aplicativo podem ser importados na paleta de modelos. Veja o tópico “[Tipos de modelo que suportam PMML](#)” na página 224 para obter mais informações.

1. Na paleta de modelos, clique com o botão direito do mouse na paleta e selecione **Importar PMML** no menu.
2. Selecione o arquivo para importar e especifique opções para rótulos de variáveis conforme necessário.
3. Clique em **Abrir**.

Usar rótulos de variáveis se presentes no modelo. O PMML pode especificar nomes de variáveis e rótulos de variáveis (como ID do Referenciador para *RefID*) para variáveis no dicionário de dados. Selecione essa opção para usar rótulos de variáveis se eles estiverem presentes no PMML exportado originalmente.

Se você tiver selecionado a opção de rótulo de variáveis, mas não houver rótulos de variáveis no PMML, os nomes de variáveis serão usados como normalmente.

Tipos de modelo que suportam PMML

Exportação de PMML

IBM SPSS Modeler modelos. Os modelos a seguir criados no IBM SPSS Modelador podem ser exportados como PMML 4.3:

- Árvore C&R
- QUEST
- CHAID
- Rede neural
- C5.0

- Regressão Logística
- Genlin
- SVM
- A priori
- Carma
- K-Médias
- Kohonen
- TwoStep
- TwoStep-AS
- GLMM (PMML é exportado para todos os modelos GLMM, mas o PMML tem apenas efeitos fixos)
- Lista de Decisão
- Cox
- Sequência (escoragem para modelos PMML de Sequência não é suportada)
- Árvores Aleatórias
- Árvore-AS
- Linear
- AS linear
- Regressão
- Logística
- GLE
- LSVM
- KNN
- Regras de associação

Modelos nativos de banco de dados. Para os modelos gerados usando algoritmos nativos de banco de dados, a exportação PMML não está disponível. Modelos criados usando Serviços de Análise da Microsoft ou Oracle Data Miner não podem ser exportados.

Importação de PMML

IBM SPSS Modelador pode importar e escorar modelos PMML gerados pelas versões atuais de todos os produtos IBM SPSS Estatísticas , incluindo modelos exportados do IBM SPSS Modelador, bem como PMML de transformação ou modelo gerado pelo IBM SPSS Estatísticas 17.0 ou posterior. Essencialmente, isso significa qualquer PMML que o mecanismo de escoragem pode escorar com as exceções a seguir:

- A Priori, CARMA, Detecção de Anomalias, Sequência e Modelos de Regra de Associação não podem ser importados.
- Modelos PMML não podem ser navegados após importação no IBM SPSS Modelador, mesmo que possam ser usados na escoragem. (Observe que isso inclui modelos que foram exportados do IBM SPSS Modelador com os quais começar. Para evitar essa limitação, exporte o modelo como um arquivo de modelo gerado [*.gm] em vez de PMML.)
- A validação limitada ocorre na importação, mas a validação integral é executada na tentativa de escorar o modelo. Assim, é possível que a importação seja bem-sucedida, mas a escoragem falhará ou produzirá resultados incorretos.

Nota: Para PMML de terceiros importados no IBM SPSS Modelador, IBM SPSS Modelador tentará escorar PMML válido que pode ser reconhecido e escorado. Mas não é garantido que todo PMML escorará ou que escorará da mesma maneira que o aplicativo que o gerou.

Capítulo 14. Projetos e relatórios

Introdução aos Projetos

Um **projeto** é um grupo de arquivos relacionados a uma tarefa de mineração de dados. Os projetos incluem fluxos de dados, gráficos, modelos gerados, relatórios e mais alguma coisa que você tenha criado no IBM SPSS Modelador. À primeira vista, pode parecer que os projetos do IBM SPSS Modelador sejam simplesmente uma maneira de organizar a saída, mas são realmente capazes de muito mais. Usando os projetos, é possível:

- Anotar cada objeto no arquivo de projeto.
- Usar a metodologia CRISP-DM para orientar seus esforços na mineração de dados. Os projetos também contêm um Sistema de Ajuda CRISP-DM que fornece detalhes e exemplos reais sobre mineração de dados com CRISP-DM.
- Incluir objetos não IBM SPSS Modelador no projeto, como a exibição de slides do PowerPoint usados para apresentar os objetivos da mineração de dados ou White Papers sobre os algoritmos que você planeja usar.
- Produzir relatórios de atualização abrangentes e simples com base em suas anotações. Esses relatórios podem ser gerados em HTML para publicação fácil na intranet de sua organização.

Nota: se a área de janela do projeto não estiver visível na janela do IBM SPSS Modelador, clique em **Projeto** no menu Visualizar.

Os objetos incluídos em um projeto podem ser visualizados de duas maneiras: **Visualização de classes** e **Visualização de CRISP-DM**. Qualquer coisa que for incluída em um projeto será incluída em ambas as visualizações, e é possível alternar entre as visualizações para criar a organização que funcione melhor.

Visualização CRISP-DM

Suportando o processo padrão de vários segmentos de mercados para mineração de dados (CRISP-DM), os projetos do IBM SPSS Modelador fornecem uma maneira não proprietária e comprovada pela indústria de organizar as partes de seus esforços de mineração de dados. CRISP-DM usa seis fases para descrever o processo do início (reunindo necessidades de negócios) ao fim (implementando seus resultados). Mesmo que algumas fases normalmente não envolvam trabalho no IBM SPSS Modelador, a área de janela do projeto inclui todas as seis fases para que você tenha uma localização central para armazenar e controlar todos os materiais associados ao projeto. Por exemplo, a fase Entendimento de Negócios geralmente envolve reunir necessidades e se encontrar com colegas para determinar objetivos em vez de trabalhar com dados no IBM SPSS Modelador. A área de janela do projeto permite armazenar suas observações feitas nessas reuniões na pasta *Entendimento de Negócios* para futura referência e inclusão em relatórios.

A visualização CRISP-DM na área de janela do projeto também é equipada com seu próprio sistema de Ajuda para lhe guiar pelo ciclo de vida de mineração de dados. No IBM SPSS Modelador, essa ajuda pode ser acessada clicando na **Ajuda do CRISP-DM** no menu Ajuda.

Nota: Se a área de janela do projeto não estiver visível na janela, clique em **Projeto** no menu Visualizar.

Configurando a fase do projeto padrão

Objetos incluídos em um projeto são incluídos em uma fase padrão do CRISP-DM. Isso significa que é necessário organizar objetos manualmente de acordo com a fase de mineração de dados em que eles foram usados. É sábio configurar a pasta padrão para a fase na qual você está trabalhando atualmente.

Para selecionar qual fase usar como padrão:

1. Na visualização CRISP-DM, clique com o botão direito na pasta para a fase para configurar como padrão.

2. No menu, clique em **Configurar como Padrão**.

A pasta padrão é exibida em negrito.

Visualização Classes

A visualização Classes na área de janela do projeto organiza seu trabalho no IBM SPSS Modelador categoricamente por tipos de objetos criados. Objetos salvos podem ser incluídos em qualquer uma das categorias a seguir:

- Fluxos
- Nós
- Modelos
- Tabelas, gráficos e relatórios
- Outros (arquivos não IBM SPSS Modelador, como exibição de slides ou White Papers relevantes para seu trabalho de mineração de dados)

A inclusão de objetos na visualização Classes também os inclui na pasta da fase padrão na visualização CRISP-DM.

Nota: Se a área de janela do projeto não estiver visível na janela, clique em **Projeto** no menu Visualizar.

Construindo um projeto

Essencialmente, um projeto é um arquivo contendo referências a todos os arquivos associados ao projeto. Isso significa que os itens do projeto são salvos individualmente e como uma referência no arquivo de projeto (.cpj). Devido a essa estrutura referencial, observe o seguinte:

- Primeiro os itens do projeto devem ser salvos individualmente antes de serem incluídos em um projeto. Se um item não estiver salvo, será solicitado que você o salve antes de incluí-lo no projeto atual.
- Objetos atualizados individualmente, como fluxos, também são atualizados no arquivo de projeto.
- Mover ou excluir objetos manualmente (como fluxos, nós e objetos de saída) do sistema de arquivos renderizará ligações no arquivo de projeto inválido.

Criando um novo projeto

Projetos novos são fáceis de criar na janela IBM SPSS Modelador. É possível começar construindo um, se não houver nenhum aberto, ou é possível fechar um projeto existente e começar do zero.

No menu principal, clique em:

Arquivo > Projeto > Novo Projeto ...

Incluindo em um projeto

Após você ter criado ou aberto um projeto, é possível incluir objetos, como fluxos de dados, nós e relatórios, usando diversos métodos.

Incluindo objetos a partir de gerenciadores

Usando os gerenciadores no canto superior direito da janela do IBM SPSS Modelador, é possível incluir fluxos ou saída.

1. Selecione um objeto, como uma tabela ou um fluxo, de uma das guias do gerenciador.

2. Clique com o botão direito e clique em **Incluir no Projeto**.

Se o objeto tiver sido salvo anteriormente, ele será automaticamente incluído na pasta de objetos apropriada (na visualização Classes) ou na pasta de fase padrão (na visualização CRISP-DM).

3. Alternativamente, é possível arrastar e soltar objetos dos gerenciadores para a área de janela do projeto.

Nota: Talvez seja solicitado que você salve o objeto primeiro. Durante o salvamento, certifique-se de que **Incluir Arquivo no Projeto** esteja selecionada na caixa de diálogo Salvar. Isso incluirá o objeto automaticamente no projeto após você salvá-lo.

Incluindo nós da tela

É possível incluir nós individuais da tela de fluxo usando a caixa de diálogo Salvar.

1. Selecione um nó na tela.
2. Clique com o botão direito e clique em **Salvar Nó**. Alternativamente, na menu principal, clique em:
Editar > Nó > Salvar Nó ...
3. Na caixa de diálogo Salvar, selecione **Incluir Arquivo no Projeto**.
4. Crie um nome para o nó e clique em **Salvar**.

Isso salva o arquivo e o inclui no projeto. Nós são incluídos na pasta *Nós* na visualização Classes e na pasta de fase padrão na visualização CRISP-DM.

Incluindo arquivos externos

É possível incluir uma grande variedade de objetos não IBM SPSS Modelador em um projeto. Isso é útil quando você está gerenciando o processo de mineração de dados inteiro dentro do IBM SPSS Modelador. Por exemplo, é possível armazenar ligações para dados, notas, apresentações e gráficos em um projeto. Na visualização CRISP-DM, arquivos externos podem ser incluídos na pasta de sua escolha. Na visualização Classes, arquivos externos podem ser salvos somente na pasta *Outro*.

Para incluir arquivos externos em um projeto:

1. Arraste arquivos da área de trabalho para o projeto.
2. Clique com o botão direito na pasta de resposta na visualização CRISP-DM ou Classes.
3. No menu, clique em **Incluir na Pasta**.
4. Selecione um arquivo na caixa de diálogo e clique em **Abrir**.

Isso incluirá uma referência ao objeto selecionado dentro dos projetos do IBM SPSS Modelador.

Transferindo Projetos para o IBM SPSS Repositório de serviços de colaboração e implantação

É possível transferir um projeto inteiro, incluindo todos os arquivos do componente, para o IBM SPSS Repositório de serviços de colaboração e implantação em um passo. Nenhum objeto que já está no local de resposta será movido. Essa variável trabalha de maneira inversa: é possível transferir projetos inteiros do IBM SPSS Repositório de serviços de colaboração e implantação para seu sistema de arquivos local.

Transferindo um Projeto

Certifique-se de que o projeto que deseja transferir esteja aberto na área de janela do projeto.

Para transferir um projeto:

1. Clique com o botão direito na pasta do projeto raiz e clique em **Transferir Projeto**.
2. Se for solicitado, efetue login no IBM SPSS Repositório de serviços de colaboração e implantação.
3. Especifique a nova localização para o projeto e clique em **OK**.

Configurando propriedades do projeto

É possível customizar a documentação e o conteúdo do projeto usando a caixa de diálogo de propriedades do projeto. Para acessar as propriedades do projeto:

1. Clique com o botão direito em um objeto ou pasta na área de janela do projeto e clique em **Propriedades do Projeto**.

2. Clique na guia **Projeto** para especificar informações básicas sobre o projeto.

Criado. Mostra a data de criação do projeto (não editável).

Sumarização. É possível inserir uma sumarização de seu projeto de mineração de dados que será exibida no relatório do projeto.

Conteúdo. Lista o tipo e o número de componentes referenciados pelo arquivo do projeto (não editável).

Salvar objeto não salvo como. Especifica se objetos não salvos devem ser salvos no sistema de arquivos local ou armazenados no repositório. Veja o tópico “[Sobre o IBM SPSS Repositório de serviços de colaboração e implantação](#)” na página 205 para obter mais informações.

Atualizar referências do objeto durante o carregamento do projeto. Selecione esta opção para atualizar as referências do projeto para seus componentes. *Nota:* Os arquivos incluídos em um projeto não são salvos no arquivo de projeto em si. Em vez disso, uma referência aos arquivos é armazenada no projeto. Isso significa que mover ou excluir um arquivo removerá o objeto do projeto.

Anotando um projeto

A área de janela do projeto fornece inúmeras maneiras de anotar seus esforços de mineração de dados. Anotações em nível de projeto normalmente são usadas para controlar objetivos e decisões de uma "perspectiva completa", enquanto anotações de pasta ou nó fornecem detalhes adicionais. A guia Anotações fornece espaço suficiente para você documentar detalhes em nível de projeto, como a exclusão de dados com dados omissos irrecuperáveis ou hipóteses promissoras formadas durante a exploração de dados.

Para anotar um projeto:

1. Selecione a pasta do projeto na visualização CRISP-DM ou Classes.
2. Clique com o botão direito na pasta e clique em **Propriedades do Projeto**.
3. Clique na guia **Anotações**.
4. Insira palavras-chave e texto para descrever o projeto.

Propriedades e anotações de pastas

Pastas do projeto individuais (nas visualizações CRISP-DM e Classes) podem ser anotadas. Na visualização CRISP-DM, essa pode ser uma maneira extremamente efetiva de documentar os objetivos de sua organização para cada fase da mineração de dados. Por exemplo, usando a ferramenta de anotação para a pasta *Entendimento dos Negócios*, é possível incluir documentação como "O objetivo de negócios para este estudo é reduzir a perda de clientes entre clientes de alto valor." Esse texto poderia ser incluído automaticamente no relatório do projeto selecionando a opção **Incluir no Relatório**.

Para anotar uma pasta:

1. Selecione uma pasta na área de janela do projeto.
2. Clique com o botão direito na pasta e clique em **Propriedades da Pasta**.

Na visualização CRISP-DM, pastas são anotadas com uma sumarização do propósito de cada fase, bem como uma orientação sobre como concluir tarefas de mineração de dados relevantes. É possível remover ou editar qualquer uma dessas anotações.

Nome. Essa área exibe o nome do campo selecionado.

Texto de dica de ferramenta. Crie Dicas de ferramenta personalizadas que serão exibidas quando você passar o mouse sobre uma pasta do projeto. Isso é útil na visualização CRISP-DM, por exemplo, para fornecer uma visão geral rápida dos objetivos de cada fase ou para marcar o status de uma fase, como "Em Andamento" ou "Concluído."

Campo de anotação. Use este campo para anotações mais longas que possam ser intercaladas no relatório do projeto. A visualização CRISP-DM inclui uma descrição de cada fase da mineração de dados na anotação, mas você pode ficar à vontade para personalizá-la para seu próprio projeto.

Incluir no relatório. Para incluir a anotação em relatórios, selecione **Incluir no Relatório**.

Propriedades do Objeto

É possível visualizar propriedades do objeto e escolher se você deseja incluir objetos individuais no relatório do projeto. Para acessar propriedades do objeto:

1. Clique com o botão direito em um objeto na área de janela do projeto.
2. No menu, clique em **Propriedades do Objeto**.

Nome. Essa área lista o nome do objeto salvo.

Caminho. Essa área lista a localização do objeto salvo.

Incluir no relatório. Selecione esta opção para incluir os detalhes do objeto em um relatório gerado.

Encerrando um projeto

Quando você sai do IBM SPSS Modelador ou abre um novo projeto, o arquivo de projeto existente (.cpj) é encerrado.

Alguns arquivos associados ao projeto (como fluxos, nós ou gráficos) podem continuar abertos. Se desejar deixar esses arquivos abertos, responda **Sim** à mensagem **... Deseja salvar e fechar esses arquivos?**

Se você modificar e salvar quaisquer arquivos associados após o encerramento de um projeto, essas versões atualizadas serão incluídas no projeto na próxima vez que você abri-lo. Para evitar esse comportamento, remova o arquivo do projeto ou salve-o sob um nome de arquivo diferente.

Gerando um relatório

Uma das variáveis mais úteis dos projetos é a capacidade de gerar relatórios com base nos itens e nas anotações do projeto. Isso é um componente crítico da mineração de dados efetiva, conforme discutido em toda metodologia CRISP-DM. É possível gerar um relatório diretamente em um dos vários tipos de arquivos ou para uma janela de saída na tela para visualização imediata. Aqui, é possível imprimir, salvar ou visualizar o relatório em um navegador da web. É possível distribuir relatórios salvos para outras pessoas em sua organização.

Geralmente relatórios são gerados a partir de arquivos de projeto várias vezes durante o processo de mineração de dados para distribuição para aqueles envolvidos no projeto. O relatório escolhe informações sobre os objetos referenciados a partir do arquivo de projeto, bem como quaisquer anotações criadas. É possível criar relatórios com base em sua visualização Classes ou visualização CRISP-DM.

Para gerar um relatório:

1. Selecione a pasta do projeto na visualização CRISP-DM ou Classes.
2. Clique com o botão direito na pasta e clique em **Relatório do Projeto**.
3. Especifique as opções de relatório e clique em **Gerar Relatório**.

As opções na caixa de diálogo de relatório fornecem várias maneiras de você gerar o tipo de relatório necessário:

Nome de saída. Especifique o nome da janela de saída se escolher enviar a saída do relatório para a tela. É possível especificar um nome personalizado ou deixar o IBM SPSS Modelador nomear a janela automaticamente para você.

Saída para tela. Selecione esta opção para gerar e exibir o relatório em uma janela de saída. Observe que você tem a opção de exportar o relatório para vários tipos de arquivo a partir da janela de saída.

Saída para arquivo. Selecione esta opção para gerar e salvar o relatório como um arquivo do tipo especificado na lista Tipo de Arquivo.

Nome do arquivo. Especifique um nome do arquivo para o relatório gerado. Os arquivos são salvos por padrão para o diretório IBM SPSS Modelador **|bin|**. Use o botão de reticências (...) para especificar uma localização diferente.

Tipo de arquivo. Os tipos de arquivo disponíveis são:

- **Documento HTML.** O relatório é salvo como um único arquivo HTML. Se seu relatório contiver gráficos, eles serão salvos como arquivos PNG e referenciados pelo arquivo HTML. Durante a publicação do seu relatório na Internet, certifique-se de fazer upload do arquivo HTML e de quaisquer imagens referenciadas por ele.
- **Documento de texto.** O relatório é salvo como um único arquivo de texto. Se seu relatório contiver gráficos, somente as referências de caminho e nome do arquivo serão incluídas no relatório.
- **Documento do Microsoft Word.** O relatório é salvo como um único documento com quaisquer gráficos integrados diretamente ao documento.
- **Documento do Microsoft Excel.** O relatório é salvo como uma única planilha com quaisquer gráficos integrados diretamente à planilha.
- **Documento do Microsoft PowerPoint.** Cada fase é mostrada em um novo slide. Quaisquer gráficos serão integrados diretamente aos slides do PowerPoint.
- **Objeto de saída.** Quando aberto no IBM SPSS Modelador, esse arquivo (.cou) é igual à opção **Saída para tela** no grupo **Formato do Relatório**.

Nota: Para exportar para um arquivo do Microsoft Office, deve-se ter o aplicativo correspondente instalado.

Título. Especifique um título para o relatório.

Estrutura de relatório. Selecione **CRISP-DM** ou **Classes**. A visualização CRISP-DM fornece um relatório de status com uma sinopse "geral", bem como detalhes sobre cada fase da mineração de dados. A visualização **Classes** é uma visualização baseada em objeto mais apropriada para o rastreamento interno de dados e fluxos.

Autor. O nome de usuário padrão é exibido, mas é possível alterá-lo.

Inclusões de relatório. Selecione um método para incluir objetos no relatório. Selecione **todas as pastas e objetos** para incluir todos os itens incluídos no arquivo de projeto. Também é possível incluir itens dependendo se **Incluir no Relatório** estiver selecionada nas propriedades do objeto. Alternativamente, para verificar itens não relatados, é possível escolher incluir somente itens marcados para exclusão (em que **Incluir no Relatório** não é marcada).

Selecionar. Esta opção permite fornecer atualizações de projeto selecionando somente **itens recentes** no relatório. Alternativamente, é possível controlar problemas mais antigos e talvez não resolvidos configurando parâmetros para **itens antigos**. Selecione **todos os itens** para indeferir tempo como um parâmetro para o relatório.

Ordenar por. É possível selecionar uma combinação das seguintes características do objeto para ordená-las da seguinte forma:

- **Tipo.** Agrupe objetos por tipo.
- **Nome.** Organize objetos em ordem alfabética.
- **Data de inclusão.** Ordene objetos usando a data em que eles foram incluídos no projeto.

Salvando e exportando relatórios gerados

Um relatório gerado para a tela é exibido em uma nova janela de saída. Quaisquer gráficos incluídos no relatório são exibidos como imagens sequenciais.

Terminologia do relatório

O número total de nós em cada fluxo é listado dentro do relatório. Os números são mostrados sob os seguintes títulos, que usam a terminologia do IBM SPSS Modelador, e não a terminologia do CRISP-DM:

- **Leitores de dados.** Nós de origem.

- **Gravadores de dados.** Nós de exportação.
- **Construtores de modelo.** Nós de construção ou modelagem.
- **Aplicadores de modelo.** Modelos gerados, também conhecidos como nuggets.
- **Construtores de saída.** Nós de gráfico ou saída.
- **Outro.** Quaisquer outros nós relacionados ao projeto. Por exemplo, aqueles que estão disponíveis na guia Operações de Campo ou Operações de Registro na Paleta de Nós.

Para salvar um relatório:

1. No menu Arquivo, clique em **Salvar**.
2. Especifique um nome de arquivo.

O relatório é salvo como um objeto de saída.

Para exportar um relatório:

3. No menu Arquivo, clique em **Exportar** e o tipo de arquivo para o qual deseja exportar.
4. Especifique um nome de arquivo.

O relatório é salvo no formato que você escolheu.

É possível exportar para os seguintes tipos de arquivo:

- HTML
- Texto
- Microsoft Word
- Microsoft Excel
- Microsoft PowerPoint

Nota: Para exportar para um arquivo do Microsoft Office, deve-se ter o aplicativo correspondente instalado.

Use os botões na parte superior da janela para:

- Imprimir o relatório.
- Visualizar o relatório como HTML em um navegador da web externo.

Capítulo 15. Customizando o IBM SPSS Modelador

Customizando as opções IBM SPSS Modelador

Há inúmeras operações que podem ser executadas para customizar o IBM SPSS Modelador para suas necessidades. Principalmente, essa customização consiste na definição de opções específicas do usuário, como alocação de memória, diretórios padrão e o uso do som e da cor. Também é possível customizar a Paleta de Nós localizada na parte inferior da janela IBM SPSS Modelador.

Configurando IBM SPSS Modelador opções

Há várias maneiras de customizar e configurar opções para o IBM SPSS Modelador:

- Configurar opções de sistema, como uso de memória e locale, clicando em **Opções do sistema** no menu **Ferramentas> Opções** .
- Configure opções de usuário, como fontes de exibição e cores, clicando em **Opções do usuário** no menu **Ferramentas> Opções** .
- Especifique o local de aplicativos que funcionam com IBM SPSS Modelador , clicando em **Ajudar Aplicativos** no menu **Ferramentas> Opções** .
- Especifique os diretórios padrão usados no IBM SPSS Modelador clicando em **Configurar Diretório** ou **Configurar Diretório do Servidor** no menu Arquivo.

Também é possível configurar opções que se apliquem a alguns ou todos os seus fluxos. Consulte o tópico ["Configurando opções para fluxos"](#) na página 37 para obter informações adicionais.

Opções de Sistema

Você pode especificar o idioma preferido ou o locale para IBM SPSS Modelador clicando em **Opções do sistema** no menu **Ferramentas> Opções** . Aqui você também pode configurar o uso máximo de memória para ModeladorSPSS, e especificar com que frequência salvar automaticamente os fluxos. Observe que as mudanças feitas nessa caixa de diálogo não entrarão em vigor até você reiniciar o ModeladorSPSS.

Memória máxima. Selecione para impor um limite em megabytes no uso de memória IBM SPSS Modelador. Em algumas plataformas, o ModeladorSPSS limita o tamanho do seu processo para reduzir a taxa em computadores com recursos limitados ou cargas pesadas. Se você estiver lidando com grandes quantidades de dados, isso pode causar um erro de "falta de memória". É possível facilitar a carga de memória especificando um novo limite.

Por exemplo, tentar exibir uma árvore de decisão muito grande pode causar um erro de memória. Neste caso, recomendamos que você aumente a memória para o alto valor como 4096Mb. Em casos como esses, nos quais é provável que você esteja processando grandes quantidades de dados, após você aumentar a concessão de memória, encerre o ModeladorSPSS e inicie-o a partir de uma linha de comandos para assegurar que a quantidade máxima de memória seja usada durante o processamento dos seus dados.

Para iniciar a partir de uma linha de comandos (assumindo que ModeladorSPSS esteja instalado no local padrão), em uma janela do Prompt de Comando, digite o seguinte:

```
C:\Program Files\IBM\SPSS\Modeler\19.0.0\bin\modeleircclient.exe" -J-Xss4096M
```

Nota: O limite máximo de memória que você pode especificar é 58368Mb devido a uma restrição Java.

Usar código de idioma do sistema. Esta opção fica selecionada por padrão e configurada para inglês (Estados Unidos). Cancele sua seleção para especificar outro idioma a partir da lista de idiomas e códigos de idioma disponíveis.

Intervalo de salvamento automático de fluxo (minutos). Especifique com que frequência você quer ModeladorSPSS salvar fluxos automaticamente. O máximo é 60 minutes, o mínimo é 1 minute e a inadimplência é 5 minutes.

Gerenciando memória

Além da configuração **Máximo de Memória** especificada na caixa de diálogo Opções do Sistema, há várias maneiras de se otimizar o uso de memória:

- Ajuste a opção **Máximo de Membros para Campos Nominais** na caixa de diálogo de propriedades do fluxo. Essa opção especifica um número máximo de membros para campos nominais após o qual o nível de medição do campo se torna *Sem Tipo*. Consulte o tópico [“Configurando opções gerais para fluxos” na página 38](#) para obter mais informações.
- Force o IBM SPSS Modelador a liberar memória clicando no canto inferior direito da janela, onde a memória que o IBM SPSS Modelador está usando e a quantidade alocada são exibidas (xxMB / xxMB). Um clique nessa região a deixa com uma tonalidade mais escura, após o qual as figuras de alocação de memória caem. Após a região retornar para sua cor regular, o IBM SPSS Modelador liberou toda memória possível.

Configurando diretórios padrão

É possível especificar o diretório padrão usado para navegadores de arquivo e saída selecionando **Configurar Diretório** ou **Configurar Diretório do Servidor** no menu Arquivo.

- **Configurar Diretório.** É possível usar essa opção para configurar o diretório ativo. O diretório ativo padrão é baseado no caminho da instalação para sua versão do IBM SPSS Modelador ou a partir do caminho da linha de comandos usado para ativar o IBM SPSS Modelador. No modo local, o diretório ativo é o caminho usado para todas as operações do lado do cliente e arquivos de saída (se eles forem referenciados com caminhos relativos).
- **Configurar Diretório do Servidor.** A opção Configurar Diretório do Servidor no menu Arquivo é ativada sempre que há uma conexão do servidor remoto. Use essa opção para especificar o diretório padrão para todos os arquivos do servidor e arquivos de dados especificados para entrada ou saída. O diretório do servidor padrão é `$CLEO/data`, em que `$CLEO` é o diretório no qual a versão de Servidor do IBM SPSS Modelador está instalada. Usando a linha de comandos, também é possível substituir esse padrão usando o flag `-server_directory` com o argumento de linha de comandos `modelerclient`.

Configurando opções do usuário

Você pode definir opções gerais para IBM SPSS Modelador selecionando **Opções do usuário** a partir do menu **Ferramentas> Opções**. Essas opções se aplicam a todos os fluxos usados no IBM SPSS Modelador.

Os tipos de opções a seguir podem ser configurados clicando na guia correspondente:

- Opções de notificação, como sobrescrição de modelo e mensagens de erro.
- Opções de exibição, como gráfico e cores do plano de fundo.
- Opções de exibição de cores de sintaxe.
- Opções de exportação de PMML usadas durante a exportação de modelos para Predictive Model Markup Language (PMML).
- Informações do usuário ou autor, como nome, iniciais e endereço de email. Essas informações podem ser exibidas na guia Anotações para nós e para outros objetos que você criar.
- Alternando entre o modo tradicional e o modo Analytic Server.

Para configurar opções específicas do fluxo, como separadores decimais, formatos de data e hora, otimização, layout de fluxo e scripts de fluxo, use a caixa de diálogo Propriedades do Fluxo disponível nos menus Arquivo e Ferramentas.

Configurando opções de notificação

Usando a guia Notificações da caixa de diálogo Opções de Usuário, é possível configurar diversas opções referentes à ocorrência e ao tipo de avisos e janelas de confirmação no IBM SPSS Modelador. Também é possível especificar o comportamento das guias Saídas e Modelos na área de janela de gerenciadores quando novos modelos e saída são gerados.

Mostrar diálogo de feedback de execução de fluxo Seleccione para exibir uma caixa de diálogo que inclua um indicador de progresso quando o fluxo estiver em execução por três segundos. A caixa de diálogo também inclui detalhes dos objetos de saída criados pelo fluxo.

- **Fechar diálogo na conclusão** Por padrão, a caixa de diálogo fecha quando o fluxo termina a execução. Limpe essa caixa de seleção se desejar que a caixa de diálogo permaneça visível quando o fluxo terminar.

Avisar quando um nó sobrescrever um arquivo Seleccione para avisar com uma mensagem de erro quando operações de nó sobrescreverem um arquivo existente.

Avisar quando um nó sobrescrever uma tabela de base de dados Seleccione para avisar com uma mensagem de erro quando operações de nó sobrescreverem uma tabela de base de dados existente.

Notificações de som

Use a lista para especificar se você receberá notificações de som quando ocorrer um evento ou erro. Há inúmeros sons disponíveis. Use o botão Play (altofalante) para reproduzir um som selecionado. Use o botão de reticências (...) para navegar e selecionar um som.

Nota: Os arquivos .wav usados para criar sons no IBM SPSS Modelador são armazenados no diretório / media/sounds de sua instalação.

- **Silenciar todos os sons** Seleccione para desativar a notificação de som para todos os eventos.

Notificações visuais

As opções nesse grupo são usadas para especificar o comportamento das guias Saídas e Modelos na área de janela de gerenciadores na parte superior direita da exibição quando novos itens são gerados. Seleccione **Novo Modelo** ou **Nova Saída** na lista para especificar o comportamento da guia correspondente.

A opção a seguir está disponível para **Novo Modelo**:

Substituir modelo anterior Se selecionada (padrão), sobrescreve um modelo existente desse fluxo na guia Modelos e na tela de fluxo. Se essa caixa estiver desmarcada, o modelo será incluído nos modelos existentes na guia e na tela. Observe que essa configuração é sobreescrita pela configuração de substituição de modelo em uma ligação de modelo.

A opção a seguir está disponível para **Nova Saída**:

Avisar quando saídas excederem [n] Seleccione se deseja exibir um aviso quando o número de itens na guia Saídas exceder uma quantidade pré-especificada. A quantidade padrão é 20; no entanto, é possível alterá-la se necessário.

As opções a seguir estão disponíveis em todos os casos:

Guia Selecionar Escolha se você alternará para a guia Saídas ou Modelos quando o objeto correspondente for gerado enquanto o fluxo é executado.

- Seleccione **Sempre** para alternar para a guia correspondente na área de janela de gerenciadores.
- Seleccione **Se Gerado por Fluxo Atual** para alternar para a guia correspondente somente para objetos gerados pelo fluxo atualmente visível na tela.
- Seleccione **Nunca** para impedir que o software alterne para a guia correspondente para notificá-lo de saídas ou modelos gerados.

Guia Atualizar Selecione se deseja atualizar a guia Saídas ou Modelos na área de janela de gerenciadores quando novas saídas ou modelos tiverem sido gerados.

- Selecione **Se Não Selecionado** para atualizar a guia correspondente (se ela ainda não estiver selecionada) sempre que novos objetos forem gerados na área de janela de gerenciadores.
- Selecione **Nunca** para impedir que o software atualize a guia correspondente para notificá-lo de objetos gerados.

Rolar paleta para tornar visível (somente **Novo Modelo**). Selecione se deseja rolar automaticamente a guia Modelos na área de janela de gerenciadores para deixar o modelo mais recente visível.

- Selecione **Sempre** para ativar a rolagem.
- Selecione **Se Gerado por Fluxo Atual** para rolar somente para objetos gerados pelo fluxo atualmente visível na tela.
- Selecione **Nunca** para impedir que o software role automaticamente a guia Modelos.

Abrir Janela (somente **Nova Saída**). Selecione se deseja abrir automaticamente uma janela de saída na geração.

- Selecione **Sempre** para sempre abrir uma nova janela.
- Selecione **Se Gerado por Fluxo Atual** para abrir uma nova janela para saída gerada pelo fluxo atualmente visível na tela.
- Selecione **Nunca** para impedir que o software abra as novas janelas automaticamente para a saída gerada.

Para reverter para as configurações padrão do sistema para essa guia, clique em **Valores Padrão**.

Configurando opções de exibição

Usando a guia Exibição da caixa de diálogo Opções de Usuário, é possível configurar opções para a exibição de fontes e cores no IBM SPSS Modelador.

Mostrar diálogo de boas-vindas na inicialização. Selecione para fazer com que a caixa de diálogo de boas-vindas seja exibida na inicialização. A caixa de diálogo bem-vindo tem opções para ativar o tutorial de exemplos do aplicativo, abrir um fluxo de demonstração ou um fluxo ou projeto existente ou criar um novo fluxo.

Mostrar fluxo e marcações SuperNode. Se selecionado, faz com que a marcação (se houver) nos fluxos e SuperNodes seja exibida por padrão. A marcação inclui comentários de fluxo, ligações de modelo e destaque de ramificações de escoragem.

Fontes & Cores padrão (efetivo no reinício). Opções nesta caixa de controle são usadas para especificar o design de tela IBM SPSS Modelador, esquema de cores e o tamanho das fontes exibidas. As opções que você seleciona aqui não entram em vigor até você fechar e reiniciar o IBM SPSS Modelador.

- **Aparência.** Selecione um esquema de cores padrão e design de tela. É possível escolher entre:
 - **SPSS Standard**, o design padrão.
 - **SPSS Clássico**, um design familiar para os usuários de versões anteriores do ModeladorSPSS.
 - **Windows**, um design do Windows que pode ser útil para um contraste aumentado na tela de fluxo e nas paletas.
 - **Carbon Analytics**, um design moderno com ícones e cores sonâmbicos.
- **Tamanho da fonte padrão para nós.** Especifique um tamanho de fonte a ser usado nas paletas do nó e para nós que são exibidos na tela do fluxo.
- **Especificar fonte de largura fixa.** Para selecionar uma fonte de largura fixa, e fonte associada **Tamanho** para uso em scripts de expressão scripting e CLEM, selecione esta caixa de seleção. A fonte padrão é Monospace plain; clique em **Alterar ...** para exibir uma lista de outras fontes que você pode selecionar.

Nota: É possível configurar o tamanho dos ícones de nó para um fluxo na área de janela Layout da guia Opções da caixa de diálogo de propriedades do fluxo. A partir do menu principal, escolha **Ferramentas > Propriedades do Fluxo > Opções > Layout**.

Cores Customizadas. Esta tabela lista as cores selecionadas atualmente que são usadas para vários itens de exibição. Para cada um dos itens que são listados na tabela, é possível alterar a cor atual dando um clique duplo na linha correspondente na coluna **Cor** e selecionando uma cor da lista. Para especificar uma cor personalizada, role para a parte inferior da lista e clique no **Color ... :NONE**.

Ordenar a Ordem de Cor da Categoria. Esta tabela lista as cores selecionadas atualmente que são usadas para exibição em gráficos recém-criados. A ordem das cores reflete a ordem em que elas são usadas no gráfico. Por exemplo, se um campo nominal usado como uma sobreposição de cor contiver quatro valores exclusivos, somente as quatro primeiras cores listadas aqui serão usadas. Para cada um dos itens que são listados na tabela, é possível alterar a cor atual dando um clique duplo na linha correspondente na coluna **Cor** e selecionando uma cor da lista. Para especificar uma cor personalizada, role para a parte inferior da lista e clique no **Color ... :NONE**. As mudanças feitas aqui não afetam gráficos criados anteriormente.

Para reverter para as configurações padrão do sistema para essa guia, clique em **Valores Padrão**.

Configurando opções de exibição de sintaxe

Usando a guia Sintaxe da caixa de diálogo Opções do Usuário, é possível configurar opções para os atributos de fonte e exibir cores nos scripts que você criar no IBM SPSS Modelador.

Destaque da Sintaxe. Esta tabela lista as cores atualmente selecionadas usadas para vários itens de sintaxe, incluindo fonte e janela na qual ela é exibida. Para cada um dos itens listados na tabela, é possível alterar a cor clicando na lista suspensa correspondente na linha e selecionando uma cor da lista. Além disso, para itens de fonte, é possível escolher incluir ênfase para negrito e itálico.

Visualizar. Esta tabela mostra uma exibição de sintaxe de exemplo que usa os atributos de cores e fontes selecionados na tabela **Destaque da Sintaxe**. Esta visualização é atualizada assim que você altera qualquer seleção.

Clique em **Valores Padrão** para reverter para as configurações padrão do sistema para esta guia.

Configurando opções de exportação de PMML

Na guia PMML, é possível controlar como o IBM SPSS Modelador exporta modelos para Predictive Model Markup Language (PMML). Veja o tópico [“Importando e exportando modelos como PMML”](#) na página 223 para obter mais informações.

Exportar PMML. Aqui é possível configurar variações de PMML que funcionam melhor com seu aplicativo de resposta.

- Selecione **Com Extensões** para permitir extensões de PMML para casos especiais nos quais não há um PMML padrão equivalente. Observe que na maioria dos casos isso produzirá o mesmo resultado que o PMML padrão.
- Selecione **Como padrão PMML ...** para exportar PMML que adere o mais próximo possível do padrão PMML.

Opções de PMML Padrão. Quando o **Como padrão PMML ...** opção é selecionada, você pode escolher uma de duas maneiras válidas para exportar modelos de regressão linear e logística:

- Como ' **Modelos PMML <GeneralRegression>**
- Como **PMML < Regressão> modelos**

Para obter mais informações sobre PMML, consulte o website Grupo de Mineração de Dados em <http://www.dmg.org>.

Configurando informações do usuário

Informações do Usuário / Autor. As informações inseridas aqui podem ser exibidas na guia Anotações dos nós e outros objetos que você criar.

Definindo o Modo

Configurações do Modo Modelador. Na guia **Mode** , você pode escolher entre os seguintes modos:

- **Modo tradicional SPSS Modeler** mostra todos os nós e expressões disponíveis na interface do usuário.
- **O modo Analytic Server** mostra apenas os nós e expressões suportados por Servidor analítico. Mas note que alguns nós e algumas expressões do CLEM ainda serão exibidos mesmo que eles não sejam *totalmente* suportados por Servidor analítico. A tabela a seguir fornece informações gerais sobre quais nós são suportados, parcialmente suportados, e não suportados por Servidor analítico.

Veja a documentação [Servidor analítico](#) para obter mais informações sobre Servidor analítico.

Note que se você configurar ModeladorSPSS para mostrar nós de banco de dados na paleta de Modelagem de Banco de Dados, eles não serão impactados ao alternar modos. Os nós do banco de dados serão sempre exibidos. Se você usar a integração do IBM Db2 for z/OS ou IBM Netezza , em alguns casos, esses nós poderão desaparecer da paleta de Modelagem de Banco de Dados após alternar para o modo Servidor analítico . Se isso acontecer, vá em **Ferramentas > Opções > Ajudar Aplicativos** e redefinir as caixas de seleção.

Tabela 38. Suporte ao nó			
Tipo de nó (nome da paleta)	Suportado por Servidor analítico	Parcialmente suportado por Servidor analítico	Não suportado pelo Analytic Server
Origens	<ul style="list-style-type: none">• Servidor analítico nó de origem		<ul style="list-style-type: none">• Banco de dados• Var. Arquivo• Arquivo fixo• Arquivo de estatísticas• Data Collection• IBM Cognos• Importação de TWC• Importação do TM1• Arquivo SAS• Excel• Entrada do Usuário XML• Ger Sim• Visualização de dados• Geo-espacial• Object Storage• Importação de extensão• Importação de R• SNA (Análise de Difusos e Análise de Grupo)

Tabela 38. Suporte ao nó (continuação)

Tipo de nó (nome da paleta)	Suportado por Servidor analítico	Parcialmente suportado por Servidor analítico	Não suportado pelo Analytic Server
Operações de Registro	<ul style="list-style-type: none"> • Selecionar • Ordenar • Saldo • Distinguir • Agregado de RFM • Anexar • TS de fluxo • Transformação De Extensão • Fluxo de TCM 	<ul style="list-style-type: none"> • Amostra (Somente suporta Random% para método Simples. Complexo não é suportado.) • Mesclar (apenas suporta junção por chaves e condições) • Agregado (o 1º quantil, o 3º quantil e a mediana não são compatíveis com o site Servidor analítico) 	<ul style="list-style-type: none"> • Space-Time-Boxes • Otimização de CPLEX
Operações de Campo	<ul style="list-style-type: none"> • Tipo • Filtro • Derivar • Preenchedor • Reclassificar • Combinação • SetToFlag • Reestruturar • Reordenar campo • Reprojetar • Intervalos de tempo 	<ul style="list-style-type: none"> • Auto Data Prep (só suporta transformação) • O binning (método de binning equalFreq não é suportado quando Ties setting Keep in current é selecionado) • Análise RFM (método de binning Tiles não é suportado quando Ties configurando Keep in current" é selecionado) • Partição (não suportada por Servidor analítico , a menos que um campo exclusivo seja usado para designar repetições de linhas para partições) 	<ul style="list-style-type: none"> • Tornar anônimo • Histórico • Transpor

Tabela 38. Suporte ao nó (continuação)

Tipo de nó (nome da paleta)	Suportado por Servidor analítico	Parcialmente suportado por Servidor analítico	Não suportado pelo Analytic Server
Gráficos	<ul style="list-style-type: none"> • Gráfico • Multigráficos • Gráfico de tempo • Distribuição • Histograma • Coleção • Web • Avaliação • Visualização de mapa • E-Plot (Beta) • t-SNE 	<ul style="list-style-type: none"> • Grafboard (apenas suporta a função de modo de agregação para campos tendo um nível de medição de discreta, nominal, ordinal ou bandeira) 	
Modelagem	<ul style="list-style-type: none"> • Séries temporais • TCM • Isotonic-AS • Árvores Aleatórias • Árvore-AS • AS linear • GLE • LSVM • STP • TwoStep-AS • Regras de associação • XGBoost-AS • K-Means-AS 	<ul style="list-style-type: none"> • Classificador Automático • Auto Numérico (os dois nós apenas suportam split, e um campo com uma função split deve ser fornecido ao usar a opção Auto Classifier Run on Analytic Server (splits ativado)) • Extensão (o modelo de construção de sintaxe R não é suportado por Servidor analítico) • Os seguintes nós só suportam split e PSM: <ul style="list-style-type: none"> – Árvore C&R – Linear – Rede neural – CHAID – Quest 	<ul style="list-style-type: none"> • Cluster automático • Lista de Decisão • C5.0 • Regressão • PCA/Fator • Seleção de Variáveis • Discriminante • Logística • GenLin • GLMM • Rede Bayes • A priori • Carma • Sequência • K-Médias • Kohonen • TwoStep • Anomalia • KNN • R • Floresta Aleatória • Os nós a seguir têm asl mas apenas leitura asl: <ul style="list-style-type: none"> – Cox – SVM – SLM

Tabela 38. Suporte ao nó (continuação)

Tipo de nó (nome da paleta)	Suportado por Servidor analítico	Parcialmente suportado por Servidor analítico	Não suportado pelo Analytic Server
Saída	<ul style="list-style-type: none"> • Tabela • Matriz • Análise • Auditoria de dados • Transformar • Estatísticas • Médias • Relatório • Configurar globais 	<ul style="list-style-type: none"> • Saída de extensão • Saída De Sintaxe R 	<ul style="list-style-type: none"> • Avaliação Sim • Ajuste de simulação • Saída de R
Exportar	<ul style="list-style-type: none"> • Servidor analítico Nó Exportar 	<ul style="list-style-type: none"> • Exportação de extensão • Exportação De Sintaxe R 	<ul style="list-style-type: none"> • Banco de dados • Arquivo simples • Exportação de estatísticas • Data Collection • Excel • IBM Cognos Exportar • Exportação do TM1 • SAS • Exportação XML • Object Storage • Exportação de R
IBM SPSS Statistics			<ul style="list-style-type: none"> • Arquivo de estatísticas • Transformação estatística • Modelo estatístico • Saída do Statistics • Exportação de estatísticas
IBM SPSS Text Analytics	<ul style="list-style-type: none"> • Análise de link de texto • Mineração de texto • nó de linguagem 		<ul style="list-style-type: none"> • Lista de Arquivos • web feed • Análise de link de texto • Traduzir • Mineração de texto • File Viewer

Tabela 38. Suporte ao nó (continuação)

Tipo de nó (nome da paleta)	Suportado por Servidor analítico	Parcialmente suportado por Servidor analítico	Não suportado pelo Analytic Server
Python			<ul style="list-style-type: none">• SMOTE• SVM de classe única• Árvore XGBoost• XGBoost Linear• t-SNE• Floresta Aleatória• HDBSCAN
Spark	Todos suportados		

Customizando a Paleta de Nós

Fluxos são construídos usando nós. A Paleta de Nós na parte inferior da janela do IBM SPSS Modelador contém todos os nós que podem ser usados na construção do fluxo. Veja o tópico “[Paleta de nós](#)” na [página 13](#) para obter mais informações.

É possível reorganizar a Paleta de Nós de duas maneiras:

- Customizar o Gerenciador de Paleta. Consulte o tópico “[Customizando o Gerenciador de Paleta](#)” na [página 244](#) para obter informações adicionais.
- Alterar como as guias da paleta contendo subpaletas são exibidas na Paleta de Nós. Consulte o tópico “[Criando uma subpaleta](#)” na [página 246](#) para obter informações adicionais.

Customizando o Gerenciador de Paleta

O Gerenciador de Paleta pode ser customizado para acomodar seu uso do IBM SPSS Modelador. Por exemplo, se você analisar frequentemente dados de séries temporais a partir de um banco de dados, talvez você queira se certificar de que o nó de Origem da Base de Dados, nó de Intervalos de Tempo, nó de Séries Temporais e nó de Gráfico de Plotagem de Tempo estejam disponíveis juntos em uma única guia da paleta. O Gerenciador de Paleta permite fazer esses ajustes facilmente criando suas guias da paleta personalizadas na Paleta de Nós.

O Gerenciador de Paleta permite executar várias tarefas:

- Controlar quais guias da paleta são mostradas na Paleta de Nós abaixo da tela de fluxo.
- Alterar a ordem em que as guias da paleta são mostradas na Paleta de Nós.
- Criar e editar suas próprias guias da paleta e quaisquer subpaletas associadas.
- Edite as seleções de nó padrão em suas guias.

Para acessar o Gerenciador de Paleta, no menu Ferramentas, clique em **Gerenciar Paletas**.

Nome da Paleta. Cada guia da paleta disponível, independentemente de ser ou não mostrada na Paleta de Nós, é listada. Isso inclui quaisquer guias da paleta que você criou. Veja o tópico “[Criando uma guia de paleta](#)” na [página 245](#) para obter mais informações.

Não. de nós. O número de nós exibido em cada guia da paleta. Um número mais alto aqui significa que você pode achar mais conveniente criar subpaletas para dividir os nós na guia. Veja o tópico “[Criando uma subpaleta](#)” na [página 246](#) para obter mais informações.

Exibido? Selecione este campo para exibir a guia da paleta na Paleta de Nós. Veja o tópico “[Exibindo guias de paleta na Paleta de Nós](#)” na [página 245](#) para obter mais informações.

Subpaletas. Para selecionar subpaletas para exibir em uma guia da paleta, destaque o **Nome da Paleta** necessário e clique nesse botão para exibir a caixa de diálogo Subpaletas. Consulte o tópico “[Criando uma subpaleta](#)” na página 246 para obter mais informações.

Restaurar Padrões. Para remover completamente todas as mudanças e adições feitas nas paletas e subpaletas e retornar para as configurações de paleta padrão, clique nesse botão.

Criando uma guia de paleta

Para criar uma guia de paleta customizada:

1. No menu Ferramentas, abra o Gerenciador de Paleta.
2. À direita do *Shown?* coluna, clique no botão Adicionar Paleta; a caixa de diálogo Criar / Editar Paleta é exibida.
3. Digite um **Nome de Paleta** exclusivo.
4. Na área **Nós Disponíveis**, selecione o nó a ser incluído na guia da paleta.
5. Clique no botão Seta à Direita Incluir Nó para mover o nó destacado para the área **Nós Selecionados**. Repita até ter incluído todos os nós desejados.

Após você ter incluído todos os nós necessários, é possível alterar a ordem em que eles são exibidos na guia de paleta:

6. Use os botões de seta simples para mover um nó uma linha para cima ou para baixo.
7. Use os botões de seta de linha para mover um nó para a parte inferior ou parte superior da lista.
8. Para remover um nó de uma paleta, destaque o nó e clique no botão Excluir à direita da área **Nós Selecionados**.

Exibindo guias de paleta na Paleta de Nós

Pode haver opções disponíveis dentro do IBM SPSS Modelador que você nunca usa; nesse caso, é possível usar o Gerenciador de Paleta para ocultar as guias contendo esses nós.

Para selecionar quais guias devem ser mostradas na Paleta de Nós:

1. No menu Ferramentas, abra o Gerenciador de Paleta.
2. Usando as caixas de seleção no *Shown?* coluna, selecione se deve incluir ou ocultar cada guia de paleta.

Para remover permanentemente uma guia de paletas da Paleta do Nós, destaque o nó e clique no botão Excluir à direita da coluna *Mostrado?* coluna. Após a exclusão, uma guia da paleta não pode ser recuperada.

Nota: Não é possível excluir a paleta padrão fornecida com o IBM SPSS Modelador, exceto a guia Favoritos.

Alterando a ordem de exibição na Paleta de Nós

Após ter selecionado quais guias da paleta deseja exibir, é possível alterar a ordem em que elas são exibidas na Paleta de Nós:

1. Use os botões de seta simples para mover uma guia da paleta uma linha para cima ou para baixo. Movê-las para cima as move para a esquerda da Paleta de Nós e vice-versa.
2. Use os botões de seta de linha para mover uma guia da paleta para cima ou para baixo na lista. Aquelas na parte superior da lista serão mostradas à esquerda da Paleta de Nós.

Exibindo subpaletas em uma guia de paleta

Da mesma maneira que é possível controlar quais guias de paleta são exibidas na Paleta de Nós, é possível controlar quais subpaletas estão disponíveis a partir da guia de paleta pai.

Para selecionar várias subpaletas para exibição em uma guia da paleta:

1. No menu Ferramentas, abra o Gerenciador de Paleta.
2. Selecione a paleta requerida.
3. Clique no botão Subpaletas; a caixa de diálogo Subpaletas é exibida.
4. Usando as caixas de seleção no *Shown?* coluna, selecione se deve incluir cada subpaleta na aba paleta. A subpaleta **Todos** é sempre mostrada e não pode ser excluída.
5. Para remover permanentemente uma subpaleta da guia de paletas, destaque a subpaleta e clique no botão Excluir à direita da coluna *Mostrado?* coluna.

Nota: Não é possível excluir as subpaletas padrão fornecidas com a guia de paleta Modelagem.

Alterando a ordem de exibição na guia de paleta

Após você ter selecionado quais subpaletas deseja exibir, é possível alterar a ordem em que elas são exibidas na guia de paleta pai:

1. Use os botões de seta simples para mover uma subpaleta uma linha para cima ou para baixo.
2. Use os botões de seta de linha para mover uma subpaleta para a parte inferior ou parte superior da lista.

As subpaletas que você cria são exibidas na Paleta do Nó quando você seleciona sua guia de paleta pai. Consulte o tópico [“Alterando uma visualização da guia paleta”](#) na página 246 para obter informações adicionais.

Criando uma subpaleta

Como é possível incluir qualquer nó existente nas guias de paleta personalizadas que você cria, é provável que você selecionará mais nós do que o que pode ser facilmente exibido na tela sem rolagem. Para evitar essa rolagem, é possível criar subpaletas nas quais você coloca os nós escolhidos na guia da paleta. Por exemplo, se você criasse uma guia da paleta contendo os nós que devem ser usados com mais frequência para criar seus fluxos, seria possível criar quatro subpaletas que dividiriam as seleções por nó de origem, operações de campo, modelagem e saída.

Nota: Só é possível selecionar nós de subpaleta entre aqueles incluídos na guia da paleta pai.

Para criar uma subpaleta:

1. No menu Ferramentas, abra o Gerenciador de Paleta.
2. Selecione a paleta na qual deseja incluir subpaletas.
3. Clique no botão Subpaletas; a caixa de diálogo Subpaletas é exibida.
4. À direita do *Shown?* coluna, clique no botão Add Sub Palette; a caixa de diálogo Create / Edit Sub Palette é exibida.
5. Digite um **Nome de Subpaleta** exclusivo.
6. Na área **Nós Disponíveis**, selecione o nó a ser incluído na subpaleta.
7. Clique no botão Seta à Direita Incluir Nó para mover um nó selecionado para a área **Nós Selecionados**.
8. Quando tiver incluído os nós necessários, clique em **OK** para retornar para a caixa de diálogo Subpaletas.

As subpaletas que você cria são exibidas na Paleta do Nó quando você seleciona sua guia de paleta pai. Consulte o tópico [“Alterando uma visualização da guia paleta”](#) na página 246 para obter informações adicionais.

Alterando uma visualização da guia paleta

Devido ao grande número de nós disponíveis no IBM SPSS Modelador, eles podem não estar todos visíveis em telas menores, sem rolar para a esquerda ou para a direita na Paleta de Nós; isso é perceptível principalmente na guia Paleta de Modelagem. Para reduzir a necessidade de rolagem, é possível escolher a exibição somente dos nós contidos em uma subpaleta (onde disponível). Veja o tópico [“Criando uma subpaleta”](#) na página 246 para obter mais informações.

Para alterar os nós mostrados em uma guia paleta, selecione a guia paleta e, em seguida, no menu à esquerda, selecione a exibição de todos os nós ou apenas daqueles que estão em uma subpaleta específica.

Capítulo 16. Considerações de desempenho para fluxos e nós

É possível projetar seus fluxos para maximizar o desempenho organizando os nós na configuração mais eficiente, ativando caches de nó quando apropriado e prestando atenção em outras considerações conforme detalhado nesta seção.

Com exceção das considerações discutidas aqui, melhorias de desempenho adicionais e mais substanciais normalmente podem ser obtidas com o uso efetivo de seu banco de dados, principalmente por meio de otimização de SQL.

Ordem dos nós

Mesmo quando você não está usando otimização de SQL, a ordem dos nós em um fluxo pode afetar o desempenho. O objetivo geral é minimizar o processamento do recebimento de dados; portanto, quando você tiver nós que reduzem a quantidade de dados, coloque-os próximo do início do fluxo. ServidorIBM SPSS Modeler pode aplicar algumas regras de reordenação automaticamente durante a compilação para apresentar certos nós quando isso for comprovadamente seguro. (Essa variável fica ativada por padrão. Verifique com seu administrador do sistema para se certificar de que ela esteja ativada em sua instalação.)

Ao usar otimização de SQL, você deseja maximizar sua disponibilidade e eficiência. Como a otimização é parada quando o fluxo contém uma operação que não pode ser executada no banco de dados, é melhor agrupar operações otimizadas por SQL no início do fluxo. Essa estratégia mantém mais do processamento no banco de dados e, então, menos dados são executados no IBM SPSS Modelador.

As operações a seguir podem ser executadas na maioria dos bancos de dados. Tente agrupá-las no *início* do fluxo:

- Mesclar por chave (junção)
- Selecionar
- Agregado
- Ordenar
- Amostra
- Anexar
- Operações distintas no modo *incluir*, no qual todos os campos são selecionados
- Operações de preenchimento
- Operações básicas de derivação usando manipulação de sequência ou aritmética padrão (dependendo de quais operações são suportadas pelos bancos de dados)
- Configurar para flag

As operações a seguir não podem ser executadas na maioria dos bancos de dados. Elas devem ser colocadas no fluxo *após* as operações na lista precedente:

- Operações em quaisquer dados que não são do banco de dados, como arquivos simples
- Mesclar por ordem
- Saldo
- Operações distintas no modo *descartar* ou onde somente um subconjunto de campos é selecionado como distinto
- Qualquer operação que requer acesso a dados de registros além daqueles sendo processados
- Derivações de campo de contagem e estado

- Operações do nó de histórico
- Operações envolvendo funções @" (séries temporais)
- Modos de verificação de tipo *Avisar* e *Interromper*
- Construção, aplicação e análise de modelo

Nota: Árvores de decisão, conjuntos de regras, regressão linear e modelos gerados por fator podem gerar SQL e, portanto, ser enviadas de volta para o banco de dados.

- Saída de dados para qualquer lugar além do mesmo banco de dados que está processando os dados

Caches de nó

Para otimizar a execução do fluxo, é possível configurar um *cache* em um nó não terminal. Quando você configura um cache em um nó, o cache é preenchido com os dados que são passados pelo nó na próxima vez que você executa o fluxo de dados. Daí em diante, os dados são lidos do cache (que é armazenado em disco em um diretório temporário) e não da origem de dados.

O armazenamento em cache é mais útil após uma operação demorada, como uma ordenação, mesclagem ou agregação. Por exemplo, suponha que você tenha um nó de origem configurado para ler dados de vendas de um banco de dados e um nó agregado que resuma as vendas por localização. É possível configurar um cache no nó agregado em vez de no nó de origem, pois você deseja que o cache armazene os nós agregados, e não o conjunto de dados inteiro.

Nota: O armazenamento em cache nos nós de origem, que simplesmente armazena uma cópia dos dados originais como ele é lido em IBM SPSS Modelador, não irá melhorar o desempenho na maioria das circunstâncias.

Nós com o armazenamento em cache ativado são exibidos com um ícone de documento pequeno no canto superior direito. Quando os dados são armazenados em cache no nó, o ícone de documento fica verde.

Para ativar um cache

1. Na tela de fluxo, clique com o botão direito no nó e clique em **Cache** no menu.
2. No submenu de armazenamento em cache, clique em **Ativar**.
3. É possível desativar o cache clicando com o botão direito no nó e clicando em **Desativar** no submenu de armazenamento em cache.

Nós de armazenamento em cache em um banco de dados

Para fluxos executados em um banco de dados, dados podem ser armazenados em cache no midstream para uma tabela temporária no banco de dados em vez de para o sistema de arquivo. Quando combinado com otimização de SQL, isso pode resultar em ganhos significativos no desempenho. Por exemplo, a saída de um fluxo que mescla várias tabelas para criar uma visualização de mineração de dados pode ser armazenada em cache e reutilizada conforme necessário. Ao gerar SQL automaticamente para todos os nós de recebimento de dados, o desempenho pode ser melhorado.

Para aproveitar a vantagem do armazenamento em cache do banco de dados, a otimização de SQL e o armazenamento em cache do banco de dados devem estar ativados. Observe que as configurações de otimização do Servidor substituem as do Cliente. Consulte o tópico [“Configurando opções de otimização para fluxos”](#) na página 40 para obter informações adicionais.

Com o armazenamento em cache do banco de dados ativado, basta clicar com o botão direito em qualquer nó não terminal para armazenar dados em cache nesse ponto e, automaticamente, o cache será criado diretamente no banco de dados na próxima vez que o fluxo for executado. Se o armazenamento em cache do banco de dados ou a otimização de SQL não estiverem ativados, o cache será gravado no sistema de arquivos.

Nota: Os bancos de dados a seguir suportam tabelas temporárias para fins de armazenamento em cache: Db2, Oracle, SQL Server e Teradata. Outros bancos de dados, como o Netezza, usarão uma tabela normal para armazenamento em cache do banco de dados. O código SQL pode ser customizado para bancos de dados específicos - entre em contato com os Serviços para obter assistência.

Desempenho: Nós de Processamento

Ordem. O nó de Ordem deve ler o conjunto inteiro de dados de entrada para que ele possa ser ordenado. Os dados são armazenados na memória até um certo limite, e o excesso é estendido para o disco. O algoritmo de ordenação é um algoritmo de combinação: dados são lidos na memória até um certo limite e ordenados usando um algoritmo de ordenação rápida híbrido. Se todos os dados se ajustarem à memória, a ordenação foi concluída. Caso contrário, um algoritmo de ordenação por mesclagem será aplicado. Os dados ordenados são gravados em arquivo e o chunk de dados seguinte é lido na memória, ordenado e gravado em disco. Isso é repetido até que todos os dados tenham sido lidos; em seguida, os chunks ordenados são mesclados. A mesclagem pode requerer aprovações repetidas para os dados armazenados em disco. No uso de pico, o nó de Ordem terá duas cópias completas dos dados configurados no disco: ordenados e não ordenados.

O tempo de execução geral do algoritmo está na ordem de $N \log(N)$, em que N é o número de registros. A ordenação na memória é mais rápida que a mesclagem do disco, portanto, o tempo de execução pode ser reduzido pela alocação de mais memória para a ordenação. O algoritmo aloca a si mesmo uma fração de RAM física controlada pela opção de configuração ServidorIBM SPSS Modeler *Memória de uso de memória*. Para aumentar a memória usada para ordenação, forneça mais RAM física ou aumente esse valor. Observe que quando a proporção de memória usada excede o conjunto de trabalhos do processo, de modo que parte da memória seja paginada para o disco, o desempenho é degradado, pois o padrão de acesso à memória do algoritmo de ordenação na memória é aleatório e pode causar paginação excessiva. O algoritmo de ordenação é usado por vários nós além do nó de Ordem, mas as mesmas regras de desempenho se aplicam.

Categorização. O nó de Categorização lê o conjunto de dados de entrada inteiro para calcular os limites de categoria antes de alocar registros para categorias. O conjunto de dados é armazenado em cache enquanto os limites são calculados; depois ele é varrido novamente para alocação. Quando o método de categorização é *largura fixa* ou *média+desvio padrão*, o conjunto de dados é armazenado em cache diretamente no disco. Esses métodos têm um tempo de execução linear e requerem espaço em disco suficiente para armazenar o conjunto de dados inteiro. Quando o método de categorização é *ranqueamentos* ou *ladrilhos*, o conjunto de dados é ordenado usando o algoritmo de ordenação descrito anteriormente e o conjunto de dados ordenado é usado como o cache. A ordenação dá a esses métodos um tempo de execução de $M \times N \log(N)$, em que M é o número de campos categorizados e N é o número de registros; isso requer espaço em disco igual a duas vezes o tamanho do conjunto de dados.

A geração de um nó Derivar baseada em categorias geradas melhorará o desempenho em passagens subsequentes. Operações de derivação são mais rápidas que categorização.

Mesclagem por Chave (Junção). O nó de Mesclagem, quando o método de mesclagem é *chaves* (equivalente a uma junção de banco de dados), ordena cada um de seus conjuntos de dados de entrada por campos-chave. Essa parte do procedimento tem um tempo de execução de $M \times N \log(N)$, em que M é o número de entradas e N é o número de registros na maior entrada; isso requer espaço em disco suficiente para armazenar todos os seus conjuntos de dados de entrada, mais uma segunda cópia do maior conjunto de dados. O tempo de execução da mesclagem em si é proporcional ao tamanho do conjunto de dados de saída, o que depende da frequência de chaves correspondentes. No pior dos casos, em que a saída é um produto Cartesiano das entradas, o tempo de execução pode se aproximar de NM . Isso é raro-a maioria das junções tem muitas chaves menos correspondentes. Se um conjunto de dados for relativamente maior que o(s) outro(s), ou se os dados recebidos já estiverem ordenados por uma campo-chave, é possível melhorar o desempenho desse nó usando a guia Otimização.

Agregado. Quando a opção *Chaves são contíguas* não está configurada, esse nó lê (mas não armazena) seu conjunto de dados de entrada inteiro antes de produzir qualquer saída agregada. Nas situações mais extremas, em que o tamanho dos dados agregados atinge um limite (determinado pela opção de configuração ServidorIBM SPSS Modeler *Memory usage multiplier*), o restante do conjunto de dados é ordenado e processado como se a opção *Keys are contiguous* fosse definida. Quando essa opção é configurada, nenhum dado é armazenado porque os registros de saída agregados são produzidos conforme os dados de entrada são lidos.

Distinto. O nó Distinto armazena todos os campos-chave exclusivos no conjunto de dados de entrada; em casos em que todos os campos são campos-chave e todos os registros são exclusivos, ele armazena

o conjunto de dados inteiro. Por padrão, o nó Distinto ordena os dados nos campos-chave e depois seleciona (ou descarta) o primeiro registro distinto de cada grupo. Para conjuntos de dados menores com um baixo número de chaves distintas, ou aqueles que foram pré-ordenados, é possível escolher opções para melhorar a velocidade e a eficiência do processamento.

Tipo. Em algumas ocorrências, o nó Tipo armazena em cache os dados de entrada durante a leitura de valores; o cache é usado para processamento de recebimento de dados. O cache requer espaço em disco suficiente para armazenar o conjunto de dados inteiro, mas acelera o armazenamento.

Avaliação. O nó de Avaliação deve ordenar os dados de entrada para calcular ladrilhos. A ordem é repetida para cada modelo avaliado, pois os escores e a ordem de registro subsequente são diferentes em cada caso. O tempo de execução é $M \times N \times \log(N)$, em que M é o número de modelos e N é o número de registros.

Desempenho: Nós de Modelagem

Rede Neural e Kohonen. Algoritmos de treinamento de rede neural (incluindo o algoritmo Kohonen) ignoram muitos dados de treinamento. Os dados são armazenados na memória até um certo limite, e o excesso é estendido para o disco. O acesso aos dados de treinamento a partir do disco é caro porque o método de acesso é aleatório, o que pode levar à atividade excessiva do disco. É possível desativar o uso do armazenamento em disco para esses algoritmos, forçando todos os dados a serem armazenados na memória, selecionando a opção **Otimizar para Velocidade** na guia Modelo da caixa de diálogo do nó. Observe que se a quantidade de memória necessária para armazenar os dados for maior que o conjunto de trabalhos do processo do servidor, parte dela será paginada para o disco e o desempenho sofrerá de acordo.

Quando **Otimizar para Memória** está ativada, a porcentagem de RAM física é alocada para o algoritmo de acordo com o valor da opção de configuração do ServidorIBM SPSS Modeler, *Porcentagem Limite de Memória de Modelagem*. Para usar mais memória para treinamento de redes neurais, forneça mais RAM ou aumente o valor dessa opção, mas observe que a configuração de um valor muito alto causará paginação.

O tempo de execução dos algoritmos de rede neural depende do nível de precisão necessário. É possível controlar o tempo de execução configurando uma condição de parada na caixa de diálogo do nó.

K-médias. O algoritmo de clusterização k-médias tem as mesmas opções para controlar o uso de memória que os algoritmos de rede neural. No entanto, o desempenho nos dados armazenados em disco é melhor, pois o acesso aos dados é sequencial.

Desempenho: CLEM Expressões

As funções em sequência do CLEM (“funções @”) que reconsideram o fluxo de dados devem armazenar dados suficientes para satisfazer o lookback mais longo. Para operações cujo grau de lookback não tem limites, todos os valores do campo devem ser armazenados. Uma operação sem limites é aquela na qual o valor de compensação não é um número inteiro literal; por exemplo, @OFFSET(Sales, Month). O valor de compensação é o nome do campo Month, cujo valor é desconhecido até ser executado. O servidor deve salvar todos os valores do campo Vendas para assegurar resultados precisos. Onde um limite máximo for conhecido, é necessário fornecê-lo como um argumento adicional; por exemplo, @OFFSET(Sales, Month, 12). Essa operação instrui o servidor a armazenar no máximo os 12 valores mais recentes de Sales. As funções em sequência, limitadas ou não, quase sempre inibem a geração de SQL.

Capítulo 17. Acessibilidade em IBM SPSS Modelador

Visão geral da Acessibilidade em IBM SPSS Modelador

IBM SPSS Modelador fornece suporte de acessibilidade para todos os usuários, bem como suporte específico para usuários com deficiência visual e outras deficiências funcionais. Esta seção descreve as variáveis e os métodos para se trabalhar usando aprimoramentos de acessibilidade, como leitores de tela e atalhos de teclado.

Tipos de suporte à acessibilidade

Independentemente de você ter uma deficiência visual ou depender do teclado para manipulação, há uma grande variedade de métodos alternativos para o uso desse kit de ferramentas de mineração de dados. Por exemplo, é possível construir fluxos, especificar opções e ler saídas, tudo sem usar o mouse. Os atalhos de teclado disponíveis estão listados nos tópicos a seguir. Além disso, o IBM SPSS Modelador fornece suporte extensivo para leitores de tela, como JAWS para Windows. Também é possível otimizar o esquema de cores para fornecer contraste adicional. Esses tipos de suporte serão discutidos nos tópicos a seguir.

Acessibilidade para deficiência visual

Há inúmeras propriedades que podem ser especificadas no IBM SPSS Modelador que aprimorarão sua capacidade de usar o software.

Exibir opções

É possível selecionar cores para a exibição de gráficos. Também é possível escolher usar suas configurações específicas do Windows para o próprio software. Isso pode ajudar a aumentar o contraste visual.

1. Para configurar opções de exibição no menu Ferramentas, clique em **Opções do Usuário**.
2. Clique na guia **Exibir**. As opções nessa guia incluem o esquema de cores do software, cores de gráficos e tamanhos de fontes para nós.

Nota: O leitor de tela não é capaz de ler gráficos, portanto, estes não são acessíveis aos usuários com deficiência visual.

Uso de Sons para Notificação

Ativando ou desativando sons, é possível controlar a maneira como você é alertado para operações específicas no software. Por exemplo, é possível ativar sons para eventos, como criação e exclusão de nó ou geração de nova saída ou modelos.

1. Para configurar opções de notificação, no menu Ferramentas, clique em **Opções do Usuário**.
2. Clique na tarefa **Notificações**.

Controlar Ativação Automática de Novas Janelas

A guia Notificações na caixa de diálogo Opções do Usuário também é usada para controlar se uma saída recém-gerada, como tabelas e gráficos, é ativada em uma janela separada. Talvez seja mais fácil para você desativar essa opção e abrir uma janela de saída somente quando necessário.

1. Para configurar essas opções, no menu Ferramentas, clique em **Opções do Usuário**.
2. Clique na tarefa **Notificações**.
3. Na caixa de diálogo, selecione **Nova Saída** na lista no grupo **Notificações Visuais**.

4. Em **Abrir Janela**, selecione **Nunca**.

Tamanho do nó

Nós podem ser exibidos usando um tamanho padrão ou pequeno. Talvez você queira ajustar esses tamanhos para de acordo com suas necessidades.

1. Para configurar opções de tamanho de nó, no menu Arquivo, clique em **Propriedades do Fluxo**.
2. Clique na guia **Layout**.
3. Na lista **Tamanho do Ícone**, selecione **Padrão**.

Acessibilidade para usuários cegos

O suporte para usuários cegos é predominantemente dependente do uso de um leitor de tela, como JAWS for Windows. Para otimizar o uso de um leitor de tela com IBM SPSS Modelador, é possível especificar inúmeras configurações.

Exibir opções

Leitores de tela tendem a ter um desempenho melhor quando o contraste visual é maior na tela. Se você já tiver uma configuração do Windows de alto contraste, é possível escolher usar essas configurações do Windows para o software em si.

1. Para configurar opções de exibição no menu Ferramentas, clique em **Opções do Usuário**.
2. Clique na guia **Exibir**.

Nota: O leitor de tela não é capaz de ler gráficos, portanto, estes não são acessíveis aos usuários cegos.

Uso de Sons para Notificação

Ativando ou desativando sons, é possível controlar a maneira como você é alertado para operações específicas no software. Por exemplo, é possível ativar sons para eventos, como criação e exclusão de nó ou geração de nova saída ou modelos.

1. Para configurar opções de notificação, no menu Ferramentas, clique em **Opções do Usuário**.
2. Clique na tarefa **Notificações**.

Controlar Ativação Automática de Novas Janelas

A guia Notificações na caixa de diálogo Opções do Usuário também é usada para controlar se uma saída recém-gerada é ativada em uma janela separada. Talvez seja mais fácil para você desativar essa opção e abrir uma janela de saída conforme necessário.

1. Para configurar essas opções, no menu Ferramentas, clique em **Opções do Usuário**.
2. Clique na tarefa **Notificações**.
3. Na caixa de diálogo, selecione **Nova Saída** na lista no grupo **Notificações Visuais**.
4. Em **Abrir Janela**, selecione **Nunca**.

Teclado de acessibilidade

A funcionalidade do produto está acessível a partir do teclado. No nível mais básico, é possível pressionar Alt mais a tecla apropriada para ativar menus de janela (como Alt+F para acessar o menu Arquivo) ou pressionar a tecla Tab para rolar pelos controles da caixa de diálogo. No entanto, há alguns problemas especiais relacionados a cada uma das janelas de menu do produto e sugestões úteis para navegação nas caixas de diálogo.

Esta seção cobrirá os destaques do teclado de acessibilidade, da abertura de um fluxo ao uso de caixas de diálogo de nó e ao trabalho com saída. Além disso, listas de atalhos de teclado são fornecidas para uma navegação muito mais eficiente.

Atalhos para navegação na janela principal

Você faz a maioria do seu trabalho de mineração de dados na janela principal do IBM SPSS Modelador. A área principal é chamada **tela de fluxo** e é usada para construir e executar fluxos de dados. A parte inferior da janela contém as **paletas de nó**, que contêm todos os nós disponíveis. As paletas são organizadas nas guias correspondentes ao tipo de operação de mineração de dados para cada grupo de nós. Por exemplo, nós usados para trazer dados para o IBM SPSS Modelador são agrupados na guia Origens, e nós usados para derivar, filtrar ou digitar campos são agrupados na guia Ops do Campo (abreviação de Operações do Campo).

O lado direito da janela contém várias ferramentas para gerenciar fluxos, saída e projetos. A metade superior direita contém os **gerenciadores** e tem três guias que são usadas para gerenciar fluxos, saída e modelos gerados. É possível acessar esses objetos selecionando a guia e um objeto na lista. A metade inferior direita contém a **área de janela do projeto**, que permite organizar seu trabalho em projetos. Há duas guias nessa área refletindo duas visualizações diferentes de um projeto. A **visualização Classes** ordena objetos do projeto por tipo, enquanto a **visualização CRISP-DM** ordena objetos por fase de mineração de dados relevante, como Preparação de Dados ou Modelagem. Esses vários aspectos da janela do IBM SPSS Modelador são discutidos em todo o Sistema de Ajuda e no Guia do Usuário.

A seguir está uma tabela dos atalhos usados para você se mover dentro da janela principal do IBM SPSS Modelador e construir fluxos. Atalhos para caixas de diálogo e saída são listados nos tópicos que seguem. Observe que essas teclas de atalho estão disponíveis somente na janela principal.

Tabela 39. Atalhos da Janela Principal

Tecla de Atalho	Função
Ctrl+F5	Move o foco para as paletas de nó.
Ctrl+F6	Move o foco para a tela de fluxo.
Ctrl+F7	Move o foco para a área de janela de gerenciadores.
Ctrl+F8	Move o foco para a área de janela do projeto.

Tabela 40. Atalhos de Fluxo e Nó

Tecla de Atalho	Função
Ctrl+N	Cria uma nova tela de fluxo em branco.
Ctrl+O	Exibe a caixa de diálogo Abrir, onde é possível selecionar e abrir um fluxo existente.
Ctrl+teclas de número	Move o foco para a guia correspondente em uma janela ou área de janela. Por exemplo, dentro de uma janela ou área de janela com guias, Ctrl+1 move para a primeira guia, começando da esquerda, Ctrl+2 para a segunda, etc.
Ctrl+Seta para Baixo	Usado na paleta do nó para mover o foco de uma guia da paleta para o primeiro nó sob essa guia.
Ctrl+Seta para Cima	Usado na paleta do nó para mover o foco de um nó para sua guia da paleta.
Inserir	Quando um nó é selecionado na paleta do nó (incluindo modelos refinados na paleta de modelos gerada), esse pressionamento de tecla inclui o nó na tela de fluxo. O pressionamento de Enter quando um nó já está selecionado na tela abre a caixa de diálogo para esse nó.
Ctrl+Enter	Quando um nó é selecionado na paleta, inclui esse nó na tela de fluxo sem selecioná-lo e move o foco para o primeiro nó na paleta.

Tabela 40. Atalhos de Fluxo e Nô (continuação)

Tecla de Atalho	Função
Alt+Enter	Quando um nó é selecionado na paleta, inclui esse nó na tela de fluxo e o seleciona, enquanto move o foco para o primeiro nó na paleta.
Shift+Barra de espaço	Quando um nó ou comentário está em foco na paleta, alterna entre a seleção e o cancelamento da seleção do nó ou comentário. Se algum outro nó ou comentário também estiver selecionado, isso faz com que a seleção seja cancelada.
Ctrl+Shift+Barra de Espaço	Quando um nó ou comentário está em foco no fluxo, ou um nó ou comentário está em foco na paleta, alterna entre a seleção e o cancelamento da seleção do nó ou comentário. Isso não afeta nenhum outro nó ou comentário selecionado.
Seta para Esquerda/Direita	Se a tela de fluxo estiver em foco, move o fluxo inteiro horizontalmente na tela. Se uma guia da paleta estiver em foco, faz ciclos entre as guias. Se um nó da paleta estiver em foco, se move entre os nós na paleta.
Seta para Cima/Baixo	Se a tela de fluxo estiver em foco, move o fluxo inteiro verticalmente na tela. Se um nó da paleta estiver em foco, se move entre os nós na paleta. Se uma subpaleta estiver em foco, se move entre outras subpaletas para essa guia da paleta.
Alt+Seta para Esquerda/Direita	Move nós e comentários selecionados na tela de fluxo horizontalmente na direção da tecla de seta.
Alt+Seta para Cima/Baixo	Move nós e comentários selecionados na tela de fluxo verticalmente na direção da tecla de seta.
Ctrl+A	Seleciona todos os nós em um fluxo.
Ctrl+Q	Quando um nó está em foco, seleciona-o e o recebimento de dados de todos os nós e cancela a seleção do envio de dados de todos os nós.
Ctrl+W	Quando um nó selecionado está em foco, cancela sua seleção e o recebimento de dados de todos os nós selecionados.
Ctrl+Alt+D	Duplica um nó selecionado.
Ctrl+Alt+L	Quando um nugget do modelo é selecionado no fluxo, abra uma caixa de diálogo Inserir para permitir que você carregue um modelo salvo de um arquivo .nod no fluxo.
Ctrl+Alt+R	Exibe a guia Anotações para um nó selecionado, permitindo renomear o nó.
Ctrl+Alt+U	Cria um nó de origem Entrada do Usuário.
Ctrl+Alt+C	Ativa ou desativa o cache para um nó.
Ctrl+Alt+F	Limpa o cache para um nó.
Guia	Na tela de fluxo, faz ciclos por todos os nós de origem e comentários no fluxo atual. Em uma paleta do nó, se move entre os nós na paleta. Em uma subpaleta selecionada, se move para o primeiro nó na subpaleta.

Tabela 40. Atalhos de Fluxo e Nô (continuação)

Tecla de Atalho	Função
Shift+Tab	Executa a mesma operação que Tab, mas na ordem inversa.
Ctrl+Tab	Com o foco na área de janela dos gerenciadores ou área de janela do projeto, move o foco para a tela de fluxo. Com o foco em uma paleta do nó, move o foco entre um nó e sua guia da paleta.
Qualquer tecla alfabética	Com o foco em um nó no fluxo atual, direciona o foco e faz um ciclo para o próximo nó cujo nome começa com a tecla pressionada.
F1	Abra o Sistema de Ajuda em um tópico relevante para o foco.
F2	Inicia o processo de conexão para um nó selecionado na tela. Use a tecla tab para mover para o nó necessário na tela e pressione Shift+Barra de Espaço para concluir a conexão.
F3	Exclui todas as conexões para o nó selecionado na tela.
F6	Move o foco entre a área de janela dos gerenciadores, área de janela do projeto e paletas de nó.
F10	Abre o menu Arquivo.
Shift+F10	Abre o menu pop-up para o nó ou fluxo.
Excluir	Exclui um nó selecionado da tela.
Esc	Fecha um menu pop-up ou caixa de diálogo.
Ctrl+Alt+X	Expande um SuperNode.
Ctrl+Alt+Z	Aumenta o zoom em um SuperNode.
Ctrl+Alt+Shift+Z	Diminui o zoom de um SuperNode.
Ctrl+E	Com o foco na tela de fluxo, executa o fluxo atual.

Inúmeras teclas de atalho padrão também são usadas no IBM SPSS Modelador, como Ctrl+C para copiar. Veja o tópico [“Usando teclas de atalho”](#) na página 18 para obter mais informações.

Atalhos para caixas de diálogo e tabelas

Várias teclas de atalho e de leitor de tela são úteis quando você está trabalhando com caixas de diálogo, tabelas e tabelas em caixas de diálogo. Uma lista completa de atalhos de teclado e leitor de tela especiais é mostrada a seguir.

Tabela 41. Atalhos da Caixa de Diálogo e do Construtor de Expressões

Tecla de Atalho	Função
Alt+4	Usado para indeferir todas as janelas de saída ou caixas de diálogo abertas. A saída pode ser recuperada da guia Saídas na área de janela de gerenciadores.
Ctrl+End	Com o foco em qualquer controle no construtor de expressões, move o ponto de inserção para o final da expressão.
Ctrl+1	No construtor de expressões, move o foco para o controle de edição de expressão.
Ctrl+2	No construtor de expressões, move o foco para a lista de funções.

Tabela 41. Atalhos da Caixa de Diálogo e do Construtor de Expressões (continuação)

Tecla de Atalho	Função
Ctrl+3	No construtor de expressões, move o foco para a lista de campos.

Atalhos de Tabela

Atalhos de tabela são usados para tabelas de saída, bem como controles de tabela em caixas de diálogo para nós, como Tipo, Filtro e Mesclagem. Normalmente, você usará a tecla tab para se mover entre células da tabela e a Ctrl+Tab para sair do controle de tabela. *Nota:* Ocasionalmente, um leitor de tela pode não começar a ler imediatamente o conteúdo de uma célula. O pressionamento das teclas de seta uma ou duas vezes reconfigurará o software e iniciará o discurso.

Tabela 42. Atalhos de Tabela

Tecla de Atalho	Função
Ctrl+W	Para tabelas, lê a descrição simples da linha A selecionada. Por exemplo, "Os valores da linha 2 selecionada são sexo, flag, m/f, etc."
Ctrl+Alt+W	Para tabelas, lê a descrição detalhada da linha A selecionada. Por exemplo, "Os valores da linha 2 selecionada são campo = sexo, tipo = flag, sexo = m/f, etc."
Ctrl+D	Para tabelas, lê a Descrição simples da área selecionada. Por exemplo, "Seleção é uma linha por seis colunas."
Ctrl+Alt+D	Para tabelas, fornece a Descrição detalhada da área selecionada. Por exemplo, "Seleção é uma linha por seis colunas. As colunas selecionadas são Campo, Tipo, Omissio. A linha selecionada é 1."
Ctrl+T	Para tabelas, fornece uma descrição simples das colunas selecionadas. Por exemplo, "Campos, Tipo, Omissio."
Ctrl+Alt+T	Para tabelas, fornece uma descrição detalhada das colunas selecionadas. Por exemplo, "Colunas selecionadas são Campos, Tipo, Omissio."
Ctrl+R	Para tabelas, fornece o número de Registros na tabela.
Ctrl+Alt+R	Para tabelas, fornece o número de Registros na tabela, bem como nomes de coluna.
Ctrl + I	Para tabelas, lê as Informações , ou conteúdo, da célula para a célula que está em foco.
Ctrl+Alt+I	Para tabelas, lê a descrição detalhada das Informações da célula (nome da coluna e conteúdo da célula) para a célula que está em foco.
Ctrl+G	Para tabelas, fornece informações de seleção Gerais simples .
Ctrl+Alt+G	Para tabelas, fornece informações de seleção Gerais detalhadas .
Ctrl+Q	Para tabelas, fornece uma alternância Rápida das células da tabela. Ctrl+Q lê as descrições detalhadas, como "Sexo=Feminino", conforme você se move pela tabela usando as teclas de seta. A seleção de Ctrl+Q novamente alterna para descrições simples (conteúdo da célula).

Tabela 42. Atalhos de Tabela (continuação)

Tecla de Atalho	Função
F8	Para tabelas, quando o foco está na tabela, configura o foco para o cabeçalho da coluna.
Barra de espaços	Para tabelas, quando o foco está no cabeçalho da coluna, permite a ordenação da coluna.

Atalhos para comentários

Ao trabalhar com comentários na tela, é possível usar os atalhos a seguir.

Tabela 43. Atalhos de comentário

Tecla de Atalho	Função
Alt+C	Altera a variável mostrar/ocultar comentário.
Alt+M	Insere um novo comentário se os comentários estiverem sendo exibidos atualmente; mostra os comentários se eles estiverem atualmente ocultos.
Guia	Na tela de fluxo, faz ciclos por todos os nós de origem e comentários no fluxo atual.
Inserir	Quando um comentário está em foco, indica o início da edição.
Alt+Enter ou Ctrl+Tab	Termina a edição e salva as mudanças na edição.
Esc	Cancela a edição. As mudanças feitas durante a edição são perdidas.
Alt+Shift+Seta para Cima	Reduz a altura da área de texto em uma célula de grade (ou um pixel) se "ajustar à grade" estiver ativado (ativado / desativado).
Alt+Shift+Seta para Baixo	Aumenta a altura da área de texto em uma célula de grade (ou um pixel) se "ajustar à grade" estiver ativado (ativado / desativado).
Alt+Shift+Seta para Esquerda	Reduz a largura da área de texto em uma célula de grade (ou um pixel) se "ajustar à grade" estiver ativado (ativado / desativado).
Alt+Shift+Seta para Direita	Aumenta a largura da área de texto em uma célula de grade (ou um pixel) se "ajustar à grade" estiver ativado (ativado / desativado).

Atalhos para o Visualizador de Cluster e Visualizador de Modelo

Há teclas de atalho disponíveis para navegação nas janelas Visualizador de Cluster e Visualizador de Modelo.

Tabela 44. Atalhos Gerais - Visualizador de Cluster e Visualizador de Modelo

Tecla de Atalho	Função
Guia	Move o foco para o próximo controle de tela.
Shift+Tab	Move o foco para o controle de tela anterior.

Tabela 44. Atalhos Gerais - Visualizador de Cluster e Visualizador de Modelo (continuação)

Tecla de Atalho	Função
Seta para baixo	Se uma lista suspensa estiver em foco, abre a lista ou move para o próximo item na lista. Se um menu estiver em foco, move para o próximo item no menu. Se um gráfico miniatura estiver em foco, move para o próximo no conjunto (ou para o primeiro se o último estiver em foco).
Seta para cima	Se uma lista suspensa estiver aberta, move para o item anterior na lista. Se um menu estiver em foco, move para o item anterior no menu. Se um gráfico miniatura estiver em foco, move para o anterior no conjunto (ou para o último se a primeira miniatura estiver em foco).
Inserir	Fecha uma lista suspensa aberta ou faz uma seleção em um menu aberto.
F6	Altera o foco entre as áreas de janela esquerda e direita da janela.
Setas para Esquerda e Direita	Se uma guia estiver em foco, move para a guia anterior ou seguinte. Se um menu estiver em foco, move para o menu anterior ou seguinte.
Alt+ <i>letra</i>	Seleciona o botão ou menu com essa letra sublinhada em seu nome.
Esc	Fecha uma lista suspensa ou menu aberto.

Somente Visualizador de Cluster

O Visualizador de Cluster tem uma visualização Clusters que contém uma grade "cluster por variáveis".

Para escolher a visualização Clusters em vez de a visualização Sumarização do Modelo:

1. Pressione Tab repetidamente até o botão **Visualização** ser selecionado.
2. Pressione Seta para Baixo duas vezes para selecionar **Clusters**.

Aqui, é possível selecionar uma célula individual dentro da grade:

3. Pressione Tab repetidamente até chegar ao último ícone na barra de ferramentas de visualização.



Figura 19. Ícone Mostrar Árvore de Visualização

4. Pressione Tab mais uma vez, depois a barra de espaço, depois uma tecla de seta.

Os atalhos de teclado a seguir agora estão disponíveis:

Tabela 45. Atalhos do Visualizador de Cluster

Tecla de Atalho	Função
Tecla de seta	Move o foco entre células individuais na grade. A exibição da distribuição de células na área de janela direita muda conforme o foco é movido.
Ctrl+, (vírgula)	Seleciona ou cancela a seleção da coluna inteira na grade em que a célula está em foco. Para incluir uma coluna na seleção, use as teclas de seta para navegar para uma célula nessa coluna e pressione Ctrl+, novamente.
Guia	Move o foco para fora da grade e para o próximo controle de tela.
Shift+Tab	Move o foco para fora da grade e de volta para o controle de tela anterior.
F2	Entra no modo de edição (somente células de descrição e rótulo).
Inserir	Salva mudanças de edição e sai do modo de edição (somente células de descrição e rótulo).
Esc	Sai do modo de edição sem salvar as mudanças (somente células de descrição e rótulo).

Exemplo de teclas de atalho: Construindo Fluxos

Para deixar um processo de construção de fluxo mais claro para os usuários dependentes de teclado ou em um leitor de tela, veja a seguir um exemplo de como construir um fluxo sem usar o mouse. Neste exemplo, você construirá um fluxo contendo um nó de Arquivo Variável, um nó de Derivação e um nó de Histograma usando os passos a seguir:

- 1. Inicie o IBM SPSS Modelador.** Quando o IBM SPSS Modelador for iniciado pela primeira vez, o foco estará na guia Favoritos da paleta do nó.
- 2. Ctrl+Seta para Baixo.** Move o foco da guia para o corpo da guia.
- 3. Seta para a Direita.** Move o foco para o nó de Arquivo Variável.
- 4. Barra de Espaço.** Seleciona o nó de Arquivo Variável.
- 5. Ctrl+Enter.** Inclui o nó de Arquivo Variável na tela de fluxo. Essa combinação de teclas também mantém a seleção no nó de Arquivo Variável para que o próximo nó incluído seja conectado a ele.
- 6. Tab.** Move o foco de volta para a paleta do nó.
- 7. Seta para a Direita 4 vezes.** Move para o nó de Derivação.
- 8. Barra de Espaço.** Seleciona o nó de Derivação.
- 9. Alt+Enter.** Inclui o nó de Derivação na tela e move a seleção para o nó de Derivação. Agora esse nó está pronto para ser conectado ao próximo nó incluído.
- 10. Tab.** Move o foco de volta para a paleta do nó.
- 11. Seta para a Direita 5 vezes.** Move o foco para o nó de Histograma na paleta.
- 12. Barra de Espaço.** Seleciona o nó de Histograma.
- 13. Inserir.** Inclui o nó no fluxo e move o foco para a tela de fluxo.

Continue com o próximo exemplo ou salve o fluxo se quiser tentar o próximo exemplo depois.

Exemplo de teclas de atalho: Editando Nós

Neste exemplo, você usará o fluxo para construir o exemplo anterior. O fluxo consiste em um nó de Arquivo Variável, um nó de Derivação e um nó de Histograma. As instruções começam com o foco no terceiro nó no fluxo, o nó de Histograma.

1. **Ctrl+Seta para Esquerda 2 vezes.** Move o foco de volta para o nó de Arquivo Variável.
2. **Inserir.** Abre a caixa de diálogo Arquivo Variável. Tabule para o campo Arquivo e digite um nome e um caminho de arquivo de texto para selecionar tal arquivo. Pressione Ctrl+Tab para navegar para a parte inferior da caixa de diálogo, tabule para o botão OK e pressione Enter para fechar a caixa de diálogo.
3. **Ctrl+Seta para a Direita.** Direcciona o foco para o segundo nó, um nó de Derivação.
4. **Inserir.** Abre a caixa de diálogo do nó de Derivação. Tabule para selecionar campos e especificar condições de derivação. Pressione Ctrl+Tab para navegar para o botão OK e pressione Enter para fechar a caixa de diálogo.
5. **Ctrl+Seta para a Direita.** Direcciona o foco para o terceiro nó, um nó de Histograma.
6. **Inserir.** Abre a caixa de diálogo do nó de Histograma. Tabule para selecionar campos e especificar opções de gráfico. Para listas suspensas, pressione Setas para Baixo para abrir a lista e para destacar um item da lista; em seguida, pressione Enter para selecionar o item da lista. Tabule para o botão OK e pressione Enter para fechar a caixa de diálogo.

Nesse ponto, é possível incluir nós adicionais ou executar o fluxo atual. Lembre-se das dicas a seguir ao construir fluxos:

- Ao conectar nós manualmente, use F2 para criar o ponto de início de conexão, tabule para se mover para o terminal e, em seguida, use Shift+Barra de Espaço para finalizar a conexão.
- Use F3 para destruir todas as conexões para um nó selecionado na tela.
- Após ter criado um fluxo, use Ctrl+E para executar o fluxo atual.

Uma lista completa de teclas de atalho está disponível. Consulte o tópico [“Atalhos para navegação na janela principal”](#) na página 255 para obter informações adicionais.

Usando um leitor de tela

Inúmeros leitores de tela estão disponíveis no mercado. IBM SPSS Modelador é configurado para suportar JAWS for Windows usando Java Access Bridge, que é instalado com IBM SPSS Modelador. Se você tiver o JAWS instalado, basta ativar o JAWS antes de ativar o IBM SPSS Modelador para usar esse produto.

Nota: É recomendado ter pelo menos 6 GB de espaço para executar JAWS com ModeladorSPSS.

Devido à natureza da representação gráfica exclusiva do IBM SPSS Modelador do processo de mineração de dados, diagramas e gráficos são melhores usados visualmente. No entanto, é possível que você entenda e tome decisões baseadas na saída e nos modelos visualizados textualmente usando um leitor de tela.

Nota: Com computadores clientes de 64 bits, algumas variáveis de tecnologia assistiva não funcionam. Isso porque o Java Access Bridge não foi projetado para uma operação de 64 bits.

Usando o Arquivo do Dicionário IBM SPSS Modelador

Um arquivo de dicionário do IBM SPSS Modelador (*Awt.JDF*) está disponível para inclusão com JAWS. Para usar esse arquivo:

1. Navegue para o subdiretório */accessibility* da instalação do IBM SPSS Modelador e copie o arquivo de dicionário (*Awt.JDF*).
2. Copie-o no diretório com seus scripts JAWS.

Você já pode ter um arquivo denominado *Awt.JDF* em sua máquina se tiver outros aplicativos JAVA em execução. Nesse caso, talvez você não possa usar esse arquivo de dicionário sem editar manualmente o arquivo de dicionário.

Usando um leitor de tela com saída HTML

Durante a visualização da saída exibida como HTML dentro do IBM SPSS Modelador usando um leitor de tela, você pode encontrar algumas dificuldades. Inúmeros tipos de saída são afetados, incluindo:

- Saída visualizada na guia Avançado para os nós Regressão, Regressão Logística e Fator/PCA
- Saída do nó do relatório

Em cada uma dessas janelas ou caixas de diálogo, há uma ferramenta na barra de ferramentas que pode ser usada para ativar a saída em seu navegador padrão, o que fornece suporte ao leitor de tela padrão. É possível usar o leitor de tela para transmitir as informações de saída.

Acessibilidade na janela de árvore interativa

A exibição padrão de um modelo de árvore de decisão na janela Árvore Interativa pode causar problemas para leitores de tela. Para acessar uma versão acessível, nos menus Árvore Interativa, clique em:

View > Janela Acessível

Isso exibe uma visualização semelhante ao mapa de árvore padrão, mas que o JAWS possa ler corretamente. É possível mover para cima, para baixo, para direita ou para esquerda usando as teclas de seta padrão. Conforme navega na janela acessível, o foco na janela Árvore Interativa é movido de acordo. Use a Barra de Espaço para alterar a seleção ou Ctrl+Barra de Espaço para estender a seleção atual.

Dicas de uso

Há várias dicas para tornar o ambiente do IBM SPSS Modelador mais acessível. A seguir estão sugestões gerais para o trabalho no IBM SPSS Modelador.

- **Sair das caixas de texto estendidas.** Use Ctrl+Tab para sair das caixas de texto estendidas. Note que Ctrl + Tab também é usado para sair de controles de tabela.
- **Usar a tecla tab em vez de teclas de seta.** Ao selecionar opções para uma caixa de diálogo, use a tecla tab para se mover entre os botões de opções. As teclas de seta não funcionarão nesse contexto.
- **Listas suspensas.** Em uma lista suspensa para caixas de diálogo, você pode usar a tecla Escape ou a barra de espaço para selecionar um item e, em seguida, fechar a lista. Também é possível usar a tecla Escape para fechar as listas suspensas que não são fechadas quando você tabula para outro controle.
- **Status de execução.** Quando você está executando um fluxo em um banco de dados grande, o JAWS pode ficar para trás na leitura do status do fluxo para você. Pressione a tecla Ctrl periodicamente para atualizar o relato de status.
- **Usar paletas de nó.** Quando você entra em uma guia das paletas de nó pela primeira vez, às vezes o JAWS lê "caixa de grupo" em vez de o nome do nó. Nesse caso, é possível usar Ctrl+Seta para a Direita e depois Ctrl+Seta para a Esquerda para reconfigurar o leitor de tela e ouvir o nome do nó.
- **Ler menus.** Ocionalmente, quando você abre um menu pela primeira vez, o JAWS pode não ler o primeiro item de menu. Se você suspeitar que isso aconteceu, use a Tecla para Baixo e depois a Tecla para Cima para ouvir o primeiro item no menu.
- **Menus em cascata.** JAWS não lê o primeiro nível de um menu em cascata. Se você ouvir uma quebra na fala ao se mover pelo menu, pressione a tecla Seta para a Direita para ouvir os itens de menu filhos.

Além disso, se você tiver o IBM SPSS Modelador de análise de texto instalado, as sugestões a seguir podem deixar a interface do ambiente de trabalho interativa mais acessível.

- **Entrar nas caixas de diálogo.** Talvez seja necessário pressionar a tecla tab para direcionar o foco para o primeiro controle ao entrar na caixa de diálogo.
- **Sair das caixas de texto estendidas.** Use Ctrl+Tab para sair das caixas de texto estendidas e se mover para o próximo controle. Note que Ctrl + Tab também é usado para sair de controles de tabela.
- **Digitar a primeira letra para localizar o elemento na lista de árvores.** Quando você procura um elemento na área de janela de categorias, área de janela de resultados extraídos ou árvore de

biblioteca, é possível digitar a primeira letra do elemento quando o foco estiver na área de janela. Isso irá selecionar a próxima ocorrência de um elemento começando com a letra inserida.

- **Listas suspensas.** Em uma lista suspensa para caixas de diálogo, você pode usar a barra de espaço para selecionar um item e, em seguida, fechar a lista.

Interferência de outro software

Durante o teste do IBM SPSS Modelador com leitores de tela, como JAWS, nossa equipe de desenvolvimento descobriu que o uso de um Systems Management Server (SMS) dentro de sua organização pode interferir na capacidade do JAWS de ler aplicativos baseados em Java, como IBM SPSS Modelador. A desativação do SMS corrigirá essa situação. Visite o website da Microsoft para obter mais informações sobre o SMS.

JAWS e Java

Versões diferentes do JAWS fornecem níveis variados de suporte para aplicativos de software baseados em Java. Embora o IBM SPSS Modelador trabalhe com todas as versões recentes de JAWS, algumas versões podem ter pequenos problemas quando usadas com sistemas baseados em Java. Visite o website do JAWS for Windows em <http://www.FreedomScientific.com>.

Usando Gráficos em IBM SPSS Modelador

Exibições visuais de informações, como histogramas, gráficos de avaliação, multigráficos e gráficos de dispersão, são difíceis de interpretar com um leitor de tela. Observe, no entanto, que gráficos da web e distribuições podem ser visualizados usando a sumarização textual disponível na janela de saída.

Capítulo 18. Suporte a Unicode

Suporte a Unicode em IBM SPSS Modelador

IBM SPSS Modelador é totalmente ativado por Unicode para IBM SPSS Modelador e ServidorIBM SPSS Modeler. Isso possibilita a troca de dados com outros aplicativos que suportam Unicode, incluindo bancos de dados multi-idiomas, sem qualquer perda de informações que possa ser causada pela conversão em ou de um esquema de codificação específico de um código de idioma.

- IBM SPSS Modelador armazena dados Unicode internamente e pode ler e gravar dados multi-idiomas armazenados como Unicode em bancos de dados sem perda.
- IBM SPSS Modelador pode ler e gravar arquivos de texto codificados em UTF-8. A importação e a exportação de arquivo de texto serão padronizadas para a codificação do código de idioma, mas suportarão UTF-8 como alternativa. Essa configuração pode ser especificada nos nós de importação e exportação de arquivo, ou a codificação padrão pode ser alterada na caixa de diálogo de propriedades do fluxo. Veja o tópico [“Configurando opções gerais para fluxos”](#) na página 38 para obter mais informações.
- Arquivos de dados de Estatísticas, SAS e Texto armazenados no locale-encoding serão convertidos em UTF-8 na importação e volta novamente na exportação. Durante a gravação em um arquivo, se houver caracteres Unicode que não existem no conjunto de caracteres de código de idioma, eles serão substituídos e um aviso será exibido. Isso deve ocorrer somente quando os dados tiverem sido importados de uma origem de dados que suporta Unicode (um arquivo de texto de banco de dados ou UTF-8) e contenha caracteres de um código de idioma diferente ou de vários códigos de idioma ou conjuntos de caracteres.
- Imagens do Editor de soluçõesIBM SPSS Modeler são codificadas em UTF-8 e são totalmente móveis entre plataformas e códigos de idioma.

Sobre Unicode

O objetivo do padrão Unicode é fornecer uma maneira consistente de codificar um texto multilíngue para que ele possa ser facilmente compartilhado entre bordas, códigos de idioma e aplicativos. O padrão Unicode, agora na versão 4.0.1, define um conjunto de caracteres que é um superconjunto de todos os conjuntos de caracteres de uso comum no mundo atualmente e designa a cada caractere um nome exclusivo e um ponto de código. Os caracteres e seus pontos de código são idênticos aqueles do Conjunto de Caracteres Universais (UCS) definido pelo ISO-10646. Para obter mais informações, consulte a [Página inicial do Unicode](#).

Avisos

Estas informações foram desenvolvidas para produtos e serviços oferecidos nos EUA. Este material pode estar disponível na IBM em outros idiomas. Entretanto, poderá ser necessário ter uma cópia do produto ou uma versão do produto nesse idioma para acessá-lo.

É possível que a IBM não ofereça os produtos, serviços ou recursos discutidos nesta publicação em outros países. Consulte um representante IBM local para obter informações sobre produtos e serviços disponíveis atualmente em sua área. Qualquer referência a produtos, programas ou serviços IBM não significa que apenas produtos, programas ou serviços IBM possam ser utilizados. Qualquer outro produto, programa ou serviço, funcionalmente equivalente, poderá ser utilizado em substituição daqueles, desde que não infrinja nenhum direito de propriedade intelectual da IBM. Entretanto, a avaliação e verificação da operação de qualquer produto, programa ou serviço não IBM são de responsabilidade do Cliente.

A IBM pode ter patentes ou solicitações de patentes pendentes relativas a assuntos tratados nesta publicação. O fornecimento desta publicação não lhe garante direito algum sobre tais patentes. Pedidos de licença devem ser enviados, por escrito, para:

*Gerência de Relações Comerciais e Industriais da IBM Brasil
IBM Brasil Ltda
Botafogo
Rio de Janeiro, RJ
CEP 22290-240*

Para pedidos de licença relacionados a informações de DBCS (Conjunto de Caracteres de Byte Duplo), entre em contato com o Departamento de Propriedade Intelectual da IBM em seu país ou envie pedidos de licença, por escrito, para:

*Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan*

INTERNATIONAL BUSINESS MACHINES CORPORATION FORNECE ESTA PUBLICAÇÃO “NO ESTADO EM QUE SE ENCONTRA”, SEM GARANTIA DE NENHUM TIPO, SEJA EXPRESSA OU IMPLÍCITA, INCLUINDO, MAS A ELAS NÃO SE LIMITANDO, AS GARANTIAS IMPLÍCITAS (OU CONDIÇÕES) DE NÃO INFRAÇÃO, COMERCIALIZAÇÃO OU ADEQUAÇÃO A UM DETERMINADO PROPÓSITO. Alguns países não permitem a exclusão de garantias expressas ou implícitas em certas transações; portanto, essa disposição pode não se aplicar ao Cliente.

Essas informações podem conter imprecisões técnicas ou erros tipográficos. São feitas alterações periódicas nas informações aqui contidas; tais alterações serão incorporadas em futuras edições desta publicação. A IBM pode, a qualquer momento, aperfeiçoar e/ou alterar os produtos e/ou programas descritos nesta publicação, sem aviso prévio.

Quaisquer referências nessas informações em sites não IBM são fornecidas por conveniência apenas e não de modo a servir como endosso para esses websites. Os materiais contidos nesses websites não fazem parte dos materiais desse produto IBM e a utilização desses websites é de inteira responsabilidade do Cliente.

A IBM pode usar ou distribuir quaisquer informações que você fornecer da forma como julgar apropriado, sem incorrer qualquer obrigação ao cliente.

Licenciados deste programa que desejam obter informações sobre este assunto com objetivo de permitir: (i) a troca de informações entre programas criados independentemente e outros programas (incluindo este) e (ii) a utilização mútua das informações trocadas, devem entrar em contato com:

*Gerência de Relações Comerciais e Industriais da IBM Brasil
IBM Brasil Ltda
Botafogo
Rio de Janeiro, RJ
CEP 22290-240*

Tais informações podem estar disponíveis, sujeitas a termos e condições apropriadas, incluindo em alguns casos o pagamento de uma taxa.

O programa licenciado descrito nesta publicação e todo o material licenciado disponível são fornecidos pela IBM sob os termos do Contrato com o Cliente IBM, do Contrato Internacional de Licença do Programa IBM ou de qualquer outro contrato equivalente.

Os dados de desempenho e exemplos de clientes citados são apresentados para propósitos ilustrativos somente. Os resultados do desempenho real podem variar, dependendo das configurações específicas e condições operacionais.

As informações a respeito de produtos não IBM foram obtidas de fornecedores dos próprios produtos, seus anúncios publicados ou outras fontes disponíveis publicamente. A IBM não testou esses produtos e não pode confirmar a precisão de seu desempenho ou de sua compatibilidade, nem quaisquer reclamações relacionadas aos produtos não IBM. Perguntas sobre os recursos de produtos não IBM devem ser endereçadas aos fornecedores desses produtos.

Instruções a respeito de direção ou intento futuro da IBM estão sujeitas a alteração ou retirada sem aviso, e representam somente metas ou objetivos.

Estas informações contêm exemplos de dados e relatórios utilizados nas operações diárias de negócios. Para ilustrá-los da forma mais completa possível, os exemplos incluem nomes de indivíduos, empresas, marcas e produtos. Todos esses nomes são fictícios e quaisquer semelhanças a pessoas reais ou empresas é meramente coincidência.

Marcas comerciais

IBM, o logotipo IBM e ibm.com são marcas comerciais ou marcas registradas da International Business Machines Corp., registradas em várias jurisdições no mundo todo. Outros nomes de produto e de serviço podem ser marcas registradas da IBM ou de outras empresas. Uma lista atual de marcas comerciais da IBM está disponível na web em "Copyright and trademark information" em www.ibm.com/legal/copytrade.shtml.

Adobe, o logotipo Adobe, PostScript e o logotipo PostScript são marcas registradas ou marcas comerciais da Adobe Systems Incorporated nos Estados Unidos e/ou em outros países.

Intel, o logotipo Intel, Intel Inside, o logotipo Intel Inside, Intel Centrino, o logotipo Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium e Pentium são marcas comerciais ou marcas registradas da Intel Corporation ou de suas subsidiárias nos Estados Unidos e em outros países.

Linux é uma marca registrada da Linus Torvalds nos Estados Unidos, e/ou em outros países.

Microsoft, Windows, Windows NT e o logotipo Windows são marcas comerciais da Microsoft Corporation nos Estados Unidos e/ou em outros países.

UNIX é uma marca registrada da The Open Group nos Estados Unidos e em outros países.

Java e todas as marcas registradas e logotipos baseados em Java são marcas ou marcas registradas da Oracle e/ou suas afiliadas.

Termos e condições para documentação do produto

Permissões para uso dessas publicações são concedidas mas estão sujeitas aos termos e condições a seguir.

Aplicabilidade

Esses termos e condições estão completando quaisquer termos para uso do website IBM.

Uso pessoal

O Cliente pode reproduzir estas publicações para uso pessoal e não comercial, contanto que todos os avisos do proprietário sejam preservados. Você pode não distribuir, exibir ou fazer trabalhos derivados dessas publicações, ou de qualquer parte delas, sem o consentimento expresso da IBM.

Uso comercial

É possível reproduzir, distribuir e exibir estas publicações exclusivamente dentro de sua empresa desde que todos os avisos do proprietário sejam preservados. Você pode não fazer trabalhos derivados dessas publicações, ou reproduzir, distribuir ou exibir estas publicações ou qualquer parte delas fora de sua empresa, sem o consentimento expresso da IBM.

Direitos

Exceto quando expressamente concedida nessa permissão, nenhuma outra propriedade intelectual, licenças ou direitos são concedidos, nem expressos ou implícitos, às publicações ou quaisquer informações, dados, software ou outra propriedade intelectual contida nela.

A IBM reserva-se ao direito de retirar as permissões concedidas aqui sempre que, de acordo com seus critérios, o uso das publicações for prejudicial ao seu interesse ou, conforme determinado pela IBM, as instruções acima não estiverem sendo seguidas adequadamente.

Você não pode fazer download, exportar ou re-exportar essas informações, exceto em conformidade total com todas as leis e regulamentações aplicáveis, incluindo todas as leis e regulamentações de exportação dos Estados Unidos.

IBM NÃO APRESENTA QUALQUER GARANTIA SOBRE O CONTEÚDO DESSAS PUBLICAÇÕES. AS PUBLICAÇÕES SÃO FORNECIDAS "COMO ESTÃO" E SEM GARANTIA DE QUALQUER TIPO, EXPRESSAS OU IMPLÍCITAS, INCLUINDO MAS NÃO SE LIMITANDO A GARANTIAS IMPLÍCITAS DE COMERCIALIZAÇÃO, NÃO INFRAÇÃO E ADEQUAÇÃO A UM DETERMINADO PROPÓSITO.

Index

Special Characters

área
 funções espaciais [178](#)
árvores de decisão
 acessibilidade [263](#)
ícones
 configurando opções [18](#), [42](#)

A

acessibilidade
 dicas no IBM SPSS Modeler [263](#)
 exemplo [261](#), [262](#)
 variáveis no IBM SPSS Modeler [253](#)
agrupamento
 controlando largura da coluna para texto quebrado [131](#)
agrupando linhas e colunas [126](#)
ajuste de escala
 tabelas dinâmicas [131](#), [133](#)
ajuste de escala de fluxos para visualização [18](#)
alinhamento
 saída [110](#), [138](#)
análise baseada em árvore
 aplicativos típicos [21](#)
análise do Champion Challenger [206](#), [217](#)
Anotações do
 convertendo em comentários [55](#)
 pasta [230](#)
 projeto [230](#)
anotando
 fluxos [51](#), [56](#)
 nós [51](#), [56](#)
aplicando zoom [15](#)
aplicativos [21](#)
aplicativos de mineração de dados [22](#)
aplicativos típicos [21](#)
aprendizado de máquina [21](#)
armazenando objetos no Repositório do IBM SPSS Collaboration and Deployment Services [207](#)
arquivo de dicionário [262](#)
arquivos
 incluindo um arquivo de texto no Visualizador [112](#)
arquivos BMP
 exportando gráficos [115](#), [121](#)
arquivos de dados
 vários arquivos de dados abertos [138](#)
arquivos de dados de texto
 codificação [265](#)
arquivos de Estatísticas
 codificação [265](#)
arquivos de fluxo de backup
 restaurando [56](#)
arquivos de log
 exibindo SQL gerado [42](#)
arquivos de saída
 salvando [57](#)

arquivos do PowerPoint [228](#)
arquivos EPS
 exportando gráficos [115](#), [122](#)
arquivos JPEG
 exportando gráficos [115](#), [121](#)
arquivos PNG
 exportando gráficos [115](#), [121](#)
arquivos PostScript (contidos)
 exportando gráficos [115](#), [122](#)
arquivos SAS
 codificação [265](#)
Arquivos TIFF
 exportando gráficos [115](#), [121](#)
atalkhos
 teclado [18](#), [254](#), [255](#), [257](#), [259](#)
 uso geral [60](#)
atalkhos do teclado [254](#), [255](#), [257](#), [259](#)
ativar nós [33](#)
atributo [21](#)
atualização de modelo [217](#)
atualizando modelos [220](#)
atualizar
 nós de origem [38](#)
automação [147](#)
avisos
 configurando opções [237](#)

B

banco de dados
 funções [156](#)
barra de ferramentas [15](#)
bloqueando nós [36](#)
bloqueando objetos do Repositório do IBM SPSS Collaboration and Deployment Services [214](#)
bordas
 exibindo bordas ocultas [136](#)
botão do meio do mouse
 simulando [18](#), [32](#)
box plots [68](#)

C

cabeçalhos [123](#)
cache
 ativando [236](#)
 configurando um cache [34](#)
 limpando [35](#), [38](#)
 salvando [35](#)
caixa de diálogo bem-vindo [238](#)
caixa de diálogo inicialização [238](#)
caixas de diálogo
 exibindo nomes de variáveis [138](#)
 exibindo rótulos de variáveis [138](#)
 ordem de exibição de variável [138](#)
camadas
 criando [128](#)

camadas (*continued*)
em tabelas dinâmicas 128
exibição 128, 129
impressão 122, 131, 133

Campos
em expressões do CLEM 159
visualizando valores 159

campos de hora
convertendo 194

campos de modelo 60

campos essenciais 58, 60

campos obrigatórios 60

camundongo
usando no IBM SPSS Modeler 18, 32

caractere de barra invertida em expressões do CLEM 164

caracteres 163, 164

caracteres especiais
removendo das sequências de caracteres 151

carregando
estados 58
nós 58

casas decimais
formatos de exibição 40

Catálogo de gráficos
Editor de Dados 63

células em tabelas dinâmicas
formatos 132
larguras 136
mostrando 129
ocultando 129
selecionando 136

centralizando saída 110, 138

Classes do 15, 227, 228

classificação
linhas da tabela dinâmica 127

CLEM
construindo expressões 155
exemplos 147
expressões 150, 163
funções 156
idioma 163
introdução 20, 147
tipos de dados 163–165
verificando expressões 160

cliente
diretório padrão 236

close_to
funções espaciais 178

close_to funtion 178

codificação 38, 265

codificação de texto 38

codificação padrão de fluxo 38

codificação UTF-8 38, 265

colando saída em outros aplicativos 113

colar 15

colunas
alterando largura em tabelas dinâmicas 136
selecionando em tabelas dinâmicas 136

comentários
atâlhos do teclado 259
em nós e fluxos 51
listando todos em um fluxo 55

Compilador Java 264

concatenando sequências de caracteres 171

Condições 150

conexão única 8

conexão única, IBM SPSS Collaboration and Deployment Services Repository 205, 206

conexões
cluster de servidores 9
no IBM SPSS Analytic Server 10
para o IBM SPSS Modeler Server 8, 9
para o repositório do IBM SPSS Collaboration and Deployment Services 206

configuração da página
cabeçalhos e rodapés 123
tamanho do gráfico 123

conformidade 508 253

conjuntos 38

conjuntos de múltiplas categorias
em expressões do CLEM 154

conjuntos de múltiplas dicotomias
em expressões do CLEM 154

conjuntos de múltiplas respostas
em expressões do CLEM 154, 159

conjuntos de regras
avaliando 38

Construtor de Expressões
acessando 155
utilização 155
visão geral 155

Construtor de gráfico
layout 64
termos 64

Construtor de Gráfico
galeria 64

controlando o número de linhas para exibição 131

convenções 170

coordenadas geoespaciais
formato de exibição 43
selecionando sistemas 44

Coordenador de processos 9

COP 9

copiando e colando saída em outros aplicativos 113

copiar 15

copiar especial 113

cor do plano de fundo 133

cores
configurando 238

cores de linha alternativas
tabelas dinâmicas 132

cores de script
configurando 239

cores em tabelas dinâmicas
bordas 132

credenciais
para o Re却itório do IBM SPSS Collaboration and Deployment Services 207

criação de paleta customizada
criação de subpaleta 246

CRISP-DM
visualização de projetos 227

cruzes
funções espaciais 178

customização da guia paleta 246

D

dados
 visualização 36
dados de mapeamento 60
dados ruidosos 22
datas
 convertendo 194
 manipulando 194
dentro
 funções espaciais 178
desativar nós 33, 34
desbloqueando objetos do Repositório do IBM SPSS Collaboration and Deployment Services 214
descoberta do conhecimento 21
descrições de fluxo 49, 50
desfazer 15
dias de sobreposição 39
dicas
 para acessibilidade 263
 uso geral 60
diretório
 padrão 236
diretório temporário 11
distância
 funções espaciais 178
distribuição f
 funções de probabilidade 178
distribuição normal
 funções de probabilidade 178
distribuição qui-quadrado
 funções de probabilidade 178
distribuição t
 funções de probabilidade 178
dividindo tabelas
 controlando quebras de tabela 136
Documentação 3
DTD 223

E

e operador 175
Editor de Dados
 Catálogo de gráficos 63
 opções de estatísticas descritivas 140
 vários arquivos de dados abertos 138
efetuando login no IBM SPSS Modeler Server 8
em branco 142, 152
erro na visualização de renderização
 memória insuficiente 235
erro no estouro da capacidade 235
espaço em branco
 removendo das sequências de caracteres 151, 181
espaços
 removendo das sequências de caracteres 151, 181
estados
 carregando 58
 salvando 57
estrutura de tópicos
 alterando níveis 111
 expandindo 111
 no Visualizador 110
 reduzindo 111
excluindo saída 110

executando fluxos 50

exemplos
 Guia de Aplicativos 3
 visão geral 4
exemplos de aplicativos 3
exportando
 descrições de fluxo 50
 PMML 223, 224
exportando gráficos 115, 121, 122
exportando saída
 formato do Excel 115, 118
 formato do PowerPoint 115
 formato do Word 115, 117
 formato HTML 115
 Formato PDF 115, 119
 HTML 116
 relatório da web 116
expressões 163
expressões do CLEM
 o desempenho 252

F

fator 263
ferramenta de mapeamento de dados 58, 59
fluxo 12
fluxos
 ajuste de escala para visualização 18
 anotando 51, 56
 armazenando no Repositório do IBM SPSS Collaboration and Deployment Services 209
 arquivos de backup 56
 carregando 58
 conectando nós 32
 desativando nós 33
 efetuando bypass em nós 33
 em execução 50
 incluindo comentários 51
 incluindo em projetos 228
 incluindo nós 32, 34
 opções 37–40, 42, 43
 opções de implementação 217
 prédio 31
 renomeando 49, 56
 salvando 56
 sistema de coordenadas geoespaciais 44
 visualizando tempos de execução 44
fluxos de dados
 prédio 31
fontes
 na área de janela da estrutura de tópicos 111
formato de exibição de moeda 40
formato do Excel
 exportando saída 115, 118
 workbook de exportação 115
formato do PowerPoint
 exportando saída 115
 workbook de exportação 115
formato do Word
 exportando saída 115, 117
 tabelas grandes 115
 workbook de exportação 115
formatos de coordenadas geoespaciais 43
formatos de datas 39, 165, 166

formatos de exibição
casas decimais 40
científico 40
coordenadas geoespaciais 43, 44
moeda 40
números 40
símbolo de agrupamento 40
formatos de exibição de número 40
formatos de horário 39, 165, 166
função @BLANK 145, 170, 201
função @DIFF 194
função @FIELD 145, 202
função @FIELDS_BETWEEN 145, 153, 202
função @FIELDS_MATCHING 145, 153, 202
função @INDEX 194
função @LAST_NON_BLANK 194, 201
função @MAX 194
função @MEAN 194
função @MIN 194
função @MULTI_RESPONSE_SET 154, 202
função @NULL 145, 170, 201
função @OFFSET
 considerações de desempenho 252
função @PARTITION_FIELD 202
função @PREDICTED 202
função @SDEV 194
função @SINCE 194
função @SUM 194
função @TARGET 202
função @TESTING_PARTITION 202
função @THIS 194
função @TODAY 187
função @TRAINING_PARTITION 202
função @VALIDATION_PARTITION 202
função abs 176
função aleatória 180
função allbutfirst 181
função allbutlast 181
função alphabefore 181
função arccos 177
função arccosh 177
função arcsin 177
função arcsinh 177
função arctan 177
função arctan2 177
função arctanh 177
função area 178
função cdf_chisq 178
função cdf_f 178
função cdf_normal 178
função cdf_t 178
função cos 177
função cosh 177
função count_equal 153, 173
função count_greater_than 153, 173
função count_less_than 153, 173
função count_non_nulls 173
função count_not_equal 153, 173
função count_nulls 145, 153, 173
função count_substring 181
função crosses 178
função date_before 173
função datetime_date 171
função de subsequência 181
função DIFF 194
função distance 178
função div 176
função do membro 173
função endstring 181
função exponential 176
função first_index 154, 173
função first_non_null 154, 173
função first_non_null_index 154, 173
função fracof 176
função hasendstring 181
função hasmidstring 181
função hasstartstring 181
função hassubstring 181
função INDEX 194
função integer_bitcount 179
função integer_leastbit 179
função integer_length 179
função intof 176
função is_date 170
função is_datetime 170
função is_integer 170
função is_number 170
função is_real 170
função is_string 170
função is_time 170
função is_timestamp 170
função isalphacode 181
função isendstring 181
função islowercode 181
função ismidstring 181
função isnumbercode 181
função isstartstring 181
função issubstring 181
função issubstring_count 181
função issubstring_lim 181
função isuppercode 181
função last_index 154, 173
função LAST_NON_BLANK 194
função last_non_null 154, 173
função last_non_null_index 154, 173
função length 181
função locchar 181
função locchar_back 181
função log 176
função log10 176
função lowertoupper 181
função matches 181
Função max 173
Função MAX 194
função max_index 154, 173
função max_n 153, 173
função MEAN 194
função mean_n 153, 176
Função min 173
Função MIN 194
função min_index 154, 173
função min_n 153, 173
função mod 176
função negate 176
função num_points 178
função OFFSET 194
função oneof 180
função overlap 178

função pi [177](#)
 função power (exponencial) [176](#)
 função random0 [180](#)
 função rem [176](#)
 função replace [181](#)
 função replicate [181](#)
 função round [176](#)
 função SDEV [194](#)
 função sdev_n [153, 176](#)
 função sign [176](#)
 função sin [177](#)
 função SINCE [194](#)
 função sinh [177](#)
 função skipchar [181](#)
 função skipchar_back [181](#)
 função soundex [187](#)
 função soundex_difference [187](#)
 função sqrt [176](#)
 função startstring [181](#)
 função stripchar [181](#)
 função strmember [181](#)
 função subscrs [181](#)
 função substring_between [181](#)
 Função SUM [194](#)
 função sum_n [153, 176](#)
 função tan [177](#)
 função tanh [177](#)
 função testbit [179](#)
 função THIS [194](#)
 função time_before [173](#)
 função to_date [171, 187](#)
 função to_dateline [187](#)
 função to_datetime [171](#)
 função to_integer [171](#)
 função to_number [171](#)
 função to_real [171](#)
 função to_string [171](#)
 função to_time [171, 187](#)
 função to_timestamp [171, 187](#)
 função trim [181](#)
 função trim_start [181](#)
 função trimend [181](#)
 função undef [201](#)
 função unicode_char [181](#)
 função unicode_value [181](#)
 função uppertolower [181](#)
 função value_at [154, 173](#)
 função within [178](#)
 funções
 @BLANK [145](#)
 @FIELD [155, 202](#)
 @GLOBAL_MAX [200](#)
 @GLOBAL_MEAN [200](#)
 @GLOBAL_MIN [200](#)
 @GLOBAL_SDEV [200](#)
 @GLOBAL_SUM [200](#)
 @PARTITION [202](#)
 @PREDICTED [155, 202](#)
 @TARGET [155, 202](#)
 banco de dados [156](#)
 em expressões do CLEM [156](#)
 exemplos [147](#)
 funções definidas pelo usuário (UDFs) [156](#)
 manipulando valores omissos [145](#)

funções bit a bit [179](#)
 funções de banco de dados
 em expressões do CLEM [156](#)
 funções definidas pelo usuário (UDFs) [156](#)
 funções de comparação [173](#)
 funções de conversão [171](#)
 funções de data
 date_anterior [173, 187](#)
 date_days_diferença [187](#)
 date_in_anos [187](#)
 date_in_days [187](#)
 date_in_meses [187](#)
 date_in_semanas [187](#)
 date_months_diferença [187](#)
 date_weeks_diferença [187](#)
 date_years_diferença [187](#)
 função @TODAY [187](#)
 funções de data e hora [165, 166](#)
 funções de data/hora
 data_adicionar_dias [187](#)
 data_adicionar_semanas [187](#)
 date_add_hours [187](#)
 date_add_minutes [187](#)
 date_add_months [187](#)
 date_add_seconds [187](#)
 date_add_years [187](#)
 datetime_date [187](#)
 datetime_day [187](#)
 datetime_day_name [187](#)
 datetime_day_short_name [187](#)
 datetime_hour [187](#)
 datetime_in_seconds [187](#)
 datetime_minute [187](#)
 datetime_month [187](#)
 datetime_month_name [187](#)
 datetime_month_short_name [187](#)
 datetime_now datetime_second [187](#)
 datetime_time [187](#)
 datetime_timestamp [187](#)
 datetime_weekday [187](#)
 datetime_year [187](#)
 funções de distribuição [178](#)
 funções de horário
 time_anterior [173, 187](#)
 time_hours_diferença [187](#)
 time_in_horas [187](#)
 time_in_mins [187](#)
 time_in_secs [187](#)
 time_mins_diferença [187](#)
 time_secs_diferença [187](#)
 funções de informação [170](#)
 funções de probabilidade [178](#)
 funções de sequência [194](#)
 funções de sequências de caracteres [181](#)
 funções definidas pelo usuário (UDFs) [156](#)
 Funções do CLEM
 aleatório [180](#)
 bit a bit [179](#)
 comparação [173](#)
 conversão [171](#)
 data/hora [187](#)
 em branco e nulos [201](#)
 espaciais [178](#)
 funções especiais [202](#)

Funções do CLEM (*continued*)

globais 200
informações 170
lista de disponíveis 168
lógico 175
numérico 176
probabilidade 178
sequência 181, 194
trigonométricas 177
valores omissos 145
funções espaciais 178
funções especiais 202
funções globais 200
funções if, then, else 175
funções lógicas 175
funções numéricas 176
funções trigonométricas 177

G

geração de SQL
criação de log 42
visualizando 42
gerenciadores 14
girando rótulos 127
Gráfico de resumo em pontos 94
gráfico de tabela 137
Gráfico Q-Q 90
gráficos
3D 64
Árvore 101
Avaliação 75
barra 65
barra de erros 65, 74
bolha 69
box plot 68
castiçal 69
coluna 65
criando a partir de tabelas dinâmicas 137
Curva matemática 84
customizar 71
dendrograma 72
Eixos Y duplos 72
Empacotamento circular 70
exportando 115
gráfico de dispersão 94
gráfico de pontos 94
Gráfico de resumo em pontos 94
Gráfico Q-Q 90
histograma 78
incluindo em projetos 228
linha 80
Linha de projeção 94
mapa 81, 83
Mapa de árvore 103
mapa de calor 77
Matriz de série 97
modelos 104
Multigráfico 85
múltiplos eixos 72
nuvem de palavras 103
ocultando 109
painei 107
paralelo 87

gráficos (*continued*)

Pareto 88
pirâmide populacional 78
Pirâmide populacional 90
pizza 88
Polígono de frequência 78
radar 92
relacionamento 93
salvando a saída 57
sunburst 97
t-SNE 98
Theme River 100
tramas do tempo 99
várias séries 86
Gráficos 3D 64
gráficos candlestick (candelabro) 69
gráficos customizados 71
Gráficos de árvore 101
Gráficos de avaliação 75
gráficos de axes múltiplas 72
gráficos de barras 65
gráficos de barras 3-D 65
gráficos de barras de 74
gráficos de barras de erro 65
gráficos de barras em cluster 65
Gráficos de barras empilhadas 65
gráficos de bolhas 69
gráficos de colunas 65
Gráficos de curva de matemática 84
gráficos de dendrograma 72
gráficos de dispersão
1-D 94
3-D 94
agrupadas 94
matriz 94, 96
parcelas de ponto 94
simples 94
sobreposição 94
gráficos de eixos duplos 72
gráficos de embalagem círculo 70
Gráficos de explosão solar 97
gráficos de linha agrupados 80
gráficos de linhas
Linha de projeção 94, 96
tramas do tempo 99
gráficos de mapas 81
Gráficos de mapas de utilização 77
Gráficos de matriz de série 97
gráficos de multigráficos 85
gráficos de nuvem de 103
Gráficos de Pareto 88
Gráficos de pirâmide populacional 90
gráficos de pizza 88
gráficos de queda de linha 94, 96
gráficos de radar 92
Gráficos de relacionamento 93
Gráficos de treemap 103
gráficos do rio tema 100
gráficos duplos y-eixos 72
gráficos multilinhas 80
gráficos paralelos 87
gráficos t-SNE 98
graus
unidades de medida 40

Guia Pasta de Trabalho
movendo saída [110](#)

H

histogramas [78](#)
HTML
exportando saída [115, 116](#)
workbook de exportação [115](#)

I

IBM SPSS Analytic Server
Conexão [10](#)
Conexões Múltiplas [10](#)
IBM SPSS Collaboration and Deployment Services [206](#)
IBM SPSS Collaboration and Deployment Services
Repository
armazenando objetos [207](#)
bloqueando e desbloqueando objetos [214](#)
conexão com [206](#)
conexão única [206](#)
credenciais [207](#)
excluindo objetos e versões [214](#)
pastas [214, 215](#)
procurando [207](#)
procurando no [212](#)
propriedades do objeto [215](#)
recuperando objetos [211](#)
transferindo projetos para [229](#)
IBM SPSS Modeler
dicas e atalhos [60](#)
Documentação [3](#)
executando a partir da linha de comandos [7](#)
introdução [7](#)
opções [235](#)
recursos de acessibilidade [253](#)
visão geral [7, 235](#)
IBM SPSS Modeler Advantage [206](#)
IBM SPSS Modeler Server
ID do Usuário [8](#)
nome de domínio (Windows) [8](#)
nome do host [8, 9](#)
número da porta [8, 9](#)
Senha [8](#)
ID do Usuário
IBM SPSS Modeler Server [8](#)
idioma
alterando o idioma de saída [128](#)
opções [235](#)
implementação [206](#)
importando
PMML [223, 224](#)
impressão
ajuste de escala das tabelas [131, 133](#)
cabeçalhos e rodapés [123](#)
camadas [122, 131, 133](#)
controlando quebras de tabela [136](#)
espaço entre itens de saída [123](#)
fluxos [18, 34](#)
gráficos [122](#)
números de página [123](#)
saída de texto [122](#)

impressão (*continued*)
tabelas dinâmicas [122](#)
tamanho do gráfico [123](#)
visualização de impressão [122](#)
incluindo
em um projeto [228](#)
incluindo conexões do IBM SPSS Modeler Server [9](#)
incluindo rótulos de grupo [126](#)
inserindo rótulos de grupo [126](#)
introdução
IBM SPSS Modeler [7, 235](#)

J

janela Árvore Interativa
acessibilidade [263](#)
janela principal [12](#)
JAWS [253, 262, 264](#)
justificação
saída [110, 138](#)

L

largura da coluna
controlando a largura máxima [131](#)
controlando a largura padrão [139](#)
controlando largura para texto quebrado [131](#)
tabelas dinâmicas [136](#)
legenda [134](#)
leitores de tela
exemplo [261, 262](#)
linha de comandos
iniciando o IBM SPSS Modeler [7](#)
linhas
selecionando em tabelas dinâmicas [136](#)
linhas de grade
tabelas dinâmicas [136](#)
Lista [163, 164](#)
listando todos os comentários para um fluxo [55](#)
localidade
opções [235](#)
localizar e substituir
Documentos do visualizador [112](#)
locatário
IBM SPSS Analytic Server [10](#)

M

maiúsculas e minúsculas [21](#)
manipulando em branco
Funções do CLEM [201](#)
mapeando campos [58](#)
memória
erro no estouro da capacidade [235](#)
gerenciando [235, 236](#)
memória insuficiente [235](#)
mensagens
exibindo SQL gerado [42](#)
mensagens de erro [44](#)
metarquivos
exportando gráficos [115](#)
métodos de seleção

métodos de seleção (*continued*)
 selecionando linhas e colunas nas tabelas dinâmicas 136
mineração de dados
 Estratégia 23
 exemplos de aplicativos 30
minimizando 17
modelagem
 ramificação 51
modelo de processo CRISP-DM 23, 24
modelos
 armazenando no Repositório do IBM SPSS Collaboration and Deployment Services 211
 atualizar 220
 exportando 239
 gráficos 104
 incluindo em projetos 228
 substituindo 237
modelos PMML
 regressão linear 239
 regressão logística 239
mostrando
 legenda 134
 linhas ou colunas 129
 notas de rodapé 134
 resultados 109
 rótulos de dimensão 129
 títulos 129
movendo linhas e colunas 126

N

navegação
 atâlhos do teclado 254

Nó agregado
 o desempenho 251

Nó classificação
 o desempenho 251

nó de auditoria de dados
 usar na exploração 21
 usar na mineração de dados 22

Nó de Avaliação
 o desempenho 251

Nó de Categorização
 o desempenho 251

Nó de mesclagem
 o desempenho 251

nó de preenchimento
 valores omissos 145

nó de rede neural
 conjuntos grandes 38

nó de Rede Neural
 o desempenho 252

nó de regra de construção
 carregando 58

Nó distinto
 o desempenho 251

nó do arquivo de cache
 carregando 58

Nó k-médias
 conjuntos grandes 38
 o desempenho 252

Nó Kohonen
 conjuntos grandes 38

Nó Kohonen (*continued*)
 o desempenho 252
nó Qualidade
 valores omissos 142
nó Seleção de Variável
 valores omissos 142
nó Tipo
 o desempenho 251
 valores omissos 145
nome de domínio (Windows)
 IBM SPSS Modeler Server 8
nome do host
 IBM SPSS Modeler Server 8, 9
nomeando nós e fluxos 56
nomes de fluxo 56
nomes de nó 56
Nomes de variáveis
 em caixas de diálogo 138
nós
 anotando 51, 56
 armazenando no Repositório do IBM SPSS Collaboration and Deployment Services 210
 ativando 33
 bloqueando 36
 carregando 58
 conectando em um fluxo 32
 configurando opções 34
 criação de paleta customizada 245
 criação de subpaleta customizada 246
 customização da guia paleta 246
 desativando em um fluxo 33
 desativar 33, 34
 duplicando 34
 editando 34
 efetuando bypass em um fluxo 33
 excluindo 32
 excluindo conexões 34
 exibindo na paleta 245
 incluindo 32, 34
 incluindo comentários em 51
 incluindo em projetos 228
 introdução 31
 o desempenho 251, 252
 ordem dos 249
 procurando 48
 removendo da paleta 245
 salvando 56
 tempos de execução 44
 visualização de dados 36
 visualizando dados 36
nós de modelagem
 customização da guia paleta de modelagem 246
 o desempenho 252
nós de origem
 atualizar 38
 mapeamento de dados 59
nós de processamento
 o desempenho 251
nós de saída 31
nós terminais 31
notação científica
 formato de exibição 40
 suprimindo na saída 138
notas de rodapé

notas de rodapé (*continued*)
marcadores 131
numerando novamente 135
notificações
configurando opções 237
Novos recursos 5
nuggets
definido 14
nuggets do modelo 51
nulos 152
num_points
funções espaciais 178
numeração de página 123
número da porta
IBM SPSS Modeler Server 8, 9
números 152, 163
números inteiros 163

O

o desempenho
dos nós de modelagem 252
dos nós de processamento 251
expressões do CLEM 252
O que há de novo em SPSS Modeler 5
objetos
Propriedades 231
objetos de saída
armazenando no Repositório do IBM SPSS Collaboration and Deployment Services 210
ocultando
legenda 134
linhas e colunas 129
notas de rodapé 134
resultados do procedimento 110
rótulos de dimensão 129
títulos 129
opções
aparência da tabela dinâmica 139
estatísticas descritivas no Editor de Dados 140
exibição 238
geral 138
para IBM SPSS Modeler 235
PMML 239
propriedades de fluxo 37–40, 42–44
rótulos de saída 140
sintaxe 239
Usuário 236
Visualizador 138
opções de implementação 217
opções do usuário 236
operador de igual 173
operador maior que 173
operador menor que 173
operador não igual 173
operador not 175
operador or 175
operadores
associando sequências de caracteres 171
em expressões do CLEM 156
ordem de exibição 126

P

padrão
fase do projeto 227
painel
gráficos 107
palavras-chave
anotando nós 56
paleta de modelos 211
paleta de modelos gerados 14
paletas
customizando 244
Parâmetros
construção de modelo 218
em expressões do CLEM 159
fluxo 46
pontuação 218
prompts de tempo de execução 46
Sessão de 46
tipo 46
parâmetros de fluxo 46
parâmetros de sessão 46
parar execução 15
pastas, Repositório do IBM SPSS Collaboration and Deployment Services 214, 215
PDF
exportando saída 115, 119
workbook de exportação 115
período 38
pirâmides da população 78
planilha
exportando 115
plots de dispersão de matriz 94, 96
PMML
exportando modelos 223, 224
importando modelos 223, 224
opções de exportação 239
polígonos de frequência 78
pontuação
ramificação 51, 219, 220
PowerPoint
exportando saída como PowerPoint 118
precedência 166
precedência do operador 166
Preferências de visualização global 104
procurando
por nós em um fluxo 48
procurando conexões no COP 9
procurando objetos no Repositório do IBM SPSS Collaboration and Deployment Services 212
procurar e substituir
Documentos do visualizador 112
programação visual 12
projetos
anotando 230
armazenando no Repositório do IBM SPSS Collaboration and Deployment Services 210
Configurando propriedades 229
configurando uma pasta padrão 227
criando nova 228
fechando 231
gerando relatórios 231
incluindo objetos 228

projetos (*continued*)
no Repositório do IBM SPSS Collaboration and Deployment Services [229](#)
prédio [228](#)
propriedades da pasta [230](#)
propriedades do objeto [231](#)
Visualização Classes [228](#)
Visualização CRISP-DM [227](#)
prompts, tempo de execução [46](#)
Propriedades
fases do relatório [231](#)
para fluxos de dados [37](#)
pasta do projeto [230](#)
Tabelas [131](#)
tabelas dinâmicas [130](#)
propriedades da célula [133](#)
propriedades de fluxo
Analytic Server [43](#)
propriedades do objeto, Repositório do IBM SPSS Collaboration and Deployment Services [215](#)

Q

quebras de tabela [136](#)

R

radianos
unidades de medida [40](#)
ramificações, modelagem e escoragem [51, 219, 220](#)
reais [163](#)
recorte [15](#)
recuperando objetos do Repositório do IBM SPSS Collaboration and Deployment Services [211](#)
Recurso de produção
usando sintaxe de comando a partir do arquivo de diário [138](#)
redimensionamento [17](#)
registros
valores ausentes do sistema [143](#)
valores omissos [142](#)
regressão [263](#)
regressão linear
exportar como PMML [239](#)
regressão logística
exportar como PMML [239](#)
relatório ativo do Cognos [116](#)
relatório da web
exportando saída [116](#)
relatórios
Configurando propriedades [231](#)
gerando [231](#)
incluindo em projetos [228](#)
salvando a saída [57](#)
removendo o mapeamento de campos [58](#)
removendo rótulos de grupo [126](#)
renomeando
fluxos [49](#)
nós [56](#)
reordenando linhas e colunas [126](#)
resultado interativo [114](#)
rodapés [123](#)
rolagem

rolagem (*continued*)
configurando opções [42](#)
rótulo
excluindo [126](#)
exibição [38](#)
inserindo rótulos de grupo [126](#)
valor [223](#)
Variável [223](#)
rótulos de grupo [126](#)
rótulos de valor
em tabelas dinâmicas [140](#)
na área de janela da estrutura de tópicos [140](#)
rótulos de variáveis
em caixas de diálogo [138](#)
em tabelas dinâmicas [140](#)
na área de janela da estrutura de tópicos [140](#)
rótulos de versão, objeto do Repositório do IBM SPSS Collaboration and Deployment Services [216](#)
rótulos, objeto do Repositório do IBM SPSS Collaboration and Deployment Services Repository [216](#)

S

saída
alinhamento [110, 138](#)
alterando o idioma de saída [128](#)
centralizando [110, 138](#)
colando em outros aplicativos [113](#)
copiando [110](#)
criptografia [124](#)
excluindo [110](#)
exportando [115](#)
interativo [114](#)
mostrando [109](#)
movendo [110](#)
ocultando [109](#)
salvando [124](#)
Visualizador [109](#)
saída HTML
leitor de tela [263](#)
salvando
estados [57](#)
fluxos [56](#)
nós [56](#)
objetos de saída [57](#)
objetos múltiplos [57](#)
salvando a pasta de trabalho
formato de texto [115](#)
formato do Excel [115](#)
formato do PowerPoint [115](#)
formato do Word [115](#)
formato HTML [115](#)
Formato PDF [115](#)
HTML [115](#)
salvando a saída
formato de texto [115, 120](#)
formato do Excel [115, 118](#)
formato do PowerPoint [115, 118](#)
formato do Word [115, 117](#)
formato HTML [115](#)
Formato PDF [115, 119](#)
HTML [115, 116](#)
relatório da web [116](#)
salvando gráficos

salvando gráficos (continued)
arquivos BMP [115](#), [121](#)
arquivos EMF [115](#)
arquivos EPS [115](#), [122](#)
arquivos JPEG [115](#), [121](#)
arquivos PICT [115](#)
arquivos PNG [121](#)
arquivos PostScript [122](#)
Arquivos TIFF [121](#)
metarquivos [115](#)
script [20](#), [147](#)
seleção da paleta do nó [245](#)
seleção do sistema de coordenadas geoespaciais [44](#)
Senha
 IBM SPSS Analytic Server [10](#)
 IBM SPSS Modeler Server [8](#)
Sequências
 correspondência [151](#)
 manipulando em expressões do CLEM [151](#)
 substituindo [151](#)
serviço de mapa [83](#)
Servidor
 criação de log em [8](#)
 diretório padrão [236](#)
 incluindo conexões [9](#)
 procurando servidores no COP [9](#)
símbolo de agrupamento
 formatos de exibição de número [38](#)
Símbolo decimal
 formatos de exibição de número [38](#)
sistema
 opções [235](#)
sistema de medição [138](#)
sobreposição
 funções espaciais [178](#)
SPLOMs [94](#)
subpaleta
 criação [246](#)
 exibindo na guia de paleta [245](#)
 removendo da guia de paleta [245](#)
substituindo modelos [237](#)
Supporte a Unicode [265](#)

T

- **Tabelas**
 - alinhamento [133](#)
 - controlando quebras de tabela [136](#)
 - cor do plano de fundo [133](#)
 - fontes [133](#)
 - incluindo em projetos [228](#)
 - margens [133](#)
 - propriedades da célula [133](#)
 - salvando a saída [57](#)
 - **tabelas dinâmicas**
 - agrupando linhas e colunas [126](#)
 - ajuste da largura da coluna padrão [139](#)
 - ajuste de escala para ajustar página [131](#), [133](#)
 - alinhamento [133](#)
 - alterando a aparência [130](#)
 - alterando a ordem de exibição [126](#)
 - aparência padrão para novas tabelas [139](#)
 - bordas [132](#)
 - camadas [128](#)

tabelas dinâmicas (*continued*)
classificando linhas [127](#)
colando como tabelas [113](#)
colando em outros aplicativos [113](#)
controlando o número de linhas para exibição [131](#)
controlando quebras de tabela [136](#)
cor do plano de fundo [133](#)
cores de linha alternativas [132](#)
criando gráficos a partir de tabelas [137](#)
desagrupando linhas ou colunas [126](#)
desfazendo mudanças [128](#)
dinamização [125](#), [126](#)
editando [125](#)
excluindo rótulos de grupo [126](#)
exibindo bordas ocultas [136](#)
exportando como HTML [115](#)
fontes [133](#)
formatos de célula [132](#)
girando rótulos [127](#)
idioma [128](#)
impressão de camadas [122](#)
imprimindo tabelas grandes [136](#)
inserindo linhas e colunas [127](#)
inserindo rótulos de grupo [126](#)
larguras da célula [136](#)
legenda [134](#)
linhas de grade [136](#)
manipulando [125](#)
margens [133](#)
mostrando e ocultando células [129](#)
movendo linhas e colunas [126](#)
notas de rodapé [134](#), [135](#)
ocultando [109](#)
Propriedades [130](#)
propriedades da célula [133](#)
propriedades de nota de rodapé [131](#)
propriedades gerais [131](#)
renderizar tabelas mais rapidamente [139](#)
rótulos de valor [127](#)
rótulos de variáveis [127](#)
selecionando linhas e colunas [136](#)
tabelas dinâmicas rápidas [139](#)
tabelas legadas [138](#)
tela de continuação [133](#)
transpondo linhas e colunas [126](#)
usando ícones [126](#)
tabelas dinâmicas rápidas [139](#)
tabelas grandes
colando no Microsoft Word [113](#)
tabelas legadas [138](#)
TableLooks
aplicando [130](#)
criando [130](#)
tamanhos
na estrutura de tópicos [111](#)
teclas de acesso rápido [18](#)
tela [12](#)
tela de fluxo
configurações [42](#)
tempos de execução, visualizando [44](#)
Texto
exportadora de workbook como texto [115](#)
exportando saída como texto [115](#), [120](#)
incluindo no Visualizador [112](#)

Texto (*continued*)
 incluindo um arquivo de texto no Visualizador [112](#)
texto de continuação
 para tabelas dinâmicas [133](#)
texto do rótulo vertical [127](#)
tipo de implementação [217](#)
tipos de dados
 em parâmetros [46](#)
títulos
 incluindo no Visualizador [112](#)
ToolTips
 anotando nós [56](#)
tramas de dispersão agrupadas [94](#)
Tramas de dispersão de 3-D [94](#)
tramas de dispersão de sobreposição [94](#)
tramas do tempo [99](#)
transpondo linhas e colunas [126](#)

U

URL
 IBM SPSS Analytic Server [10](#)

V

vaga
 estados [58](#)
 fluxos [58](#)
 modelos [58](#)
 nós [58](#)
 projetos [228](#)
 saída [58](#)
valores
 incluindo em expressões do CLEM [159](#)
 visualizando a partir de uma auditoria de dados [159](#)
valores ausentes do sistema
 em registros [143](#)
valores de data/hora [152](#)
valores globais
 em expressões do CLEM [159](#)
valores omissos
 em registros [142](#)
 expressões do CLEM [145](#)
 sistema [143](#)
várias sessões do IBM SPSS Modeler [11](#)
variáveis
 ordem de exibição nas caixas de diálogo [138](#)
vários arquivos de dados abertos [138](#)
vários gráficos de séries [86](#)
verificando expressões do CLEM [160](#)
vírgula [38](#)
visualização
 dados do nó [36](#)
Visualizador
 alterando a fonte da estrutura de tópicos [111](#)
 alterando níveis da estrutura de tópicos [111](#)
 alterando tamanhos da estrutura de tópicos [111](#)
 área de janela da estrutura de tópicos [109](#)
 área de janela de resultados [109](#)
 espaço entre itens de saída [123](#)
 estrutura de tópicos [110](#)
 excluindo saída [110](#)
 exibindo nomes de variáveis [140](#)

Visualizador (*continued*)
 exibindo rótulos de valor [140](#)
 exibindo rótulos de variáveis [140](#)
 exibindo valores de dados [140](#)
 expandido a estrutura de tópicos [111](#)
 localizar e substituir informações [112](#)
 ocultando resultados [109](#)
 opções de exibição [138](#)
 procurar e substituir informações [112](#)
 reduzindo a estrutura de tópicos [111](#)
 salvando documento [124](#)

W

workbook de exportação
 formato do Excel [115](#)
 formato do PowerPoint [115](#)
 formato do Word [115](#)
 formato HTML [115](#)
 Formato PDF [115](#)

IBM.[®]