

IBM SPSS Custom Tables 26

IBM

注释

请先阅读 第 15 页的『通知』 中的信息，然后再使用本信息及其支持的产品。

产品信息

此版本适用于 IBM® SPSS Statistics26.0.0 以及所有后续发行版和修订版，直至在新版本中另有说明为止。

目录

定制表.	1	通知.	15
"定制表"界面	1	商标.	16
表构建器界面	1		
构建表	1	索引.	19
定制表: "检验统计"	6		
样本文件.	7		

定制表

SPSS® Statistics Standard Edition 或"定制表"选项中包含以下定制表功能。

"定制表"界面

表构建器界面

定制表使用简单的拖放表构建器界面，使您可以在选择变量和选项时预览表。它还具有典型对话框所不具备的灵活性，能够更改窗口大小以及窗口中窗格的大小。

构建表

在"定制表"界面中选择将在表中显示的变量和汇总测量。

分析 > 表 > 定制表

变量列表。数据文件中的变量显示在对话框的左侧窗格中。定制表会对变量的两个不同测量级别加以区别，并根据测量级别的不同分别处理：

分类。包含有限数量的不同值或类别（例如，性别或宗教）的数据。分类变量可以是字符串（字母数值）变量或使用数值代码表示类别的数值变量（例如，0 = *male*, 1 = *female*）。这种数据也称为定性数据。分类变量既可以是名义变量，也可以是顺序变量

- **名义 (Nominal).** 当变量值表示不具有内在等级的类别时，该变量可以作为名义变量；例如，雇员任职的公司部门。名义变量的示例包括地区、邮政编码和宗教信仰。
- **有序 (Ordinal).** 当变量值表示带有某种内在等级的类别时，该变量可以作为有序变量；例如，从十分不满意到十分满意的服务满意度水平。有序变量的示例包括表示满意度或可信度的态度分数和优先选择评分。

分类变量定义表中的类别（行、列和层），缺省汇总统计为计数（每个类别中的个案数）。例如，分类性别变量的缺省表将只显示男性人数和女性人数。

刻度。以区间或比率刻度度量的数据，其中数据值既表示值的顺序，也表示值之间的距离。例如，72,195 美元的薪金比 52,398 美元的薪金高，这两个值之间的距离是 19,797 美元。也称为定量或连续数据。

刻度变量通常在分类变量的类别中进行摘要，缺省汇总统计为平均值。例如，性别类别的缺省收入表将显示男性的平均收入和女性的平均收入。

也可以对刻度变量自身进行摘要，无需使用分类变量定义组。这主要用于多个刻度变量的堆积摘要。

多重响应集

此外，定制表还支持一种称为**多重响应集**的特殊变量。多响应集不是通常意义上真正的变量。它们不在数据编辑器中显示，也不能由其他过程识别。多响应集使用多个变量记录对问题的答复，其中响应者可以给出多个答案。多响应集以与分类变量相同的方式对待，可以对分类变量执行的大多数操作也可以对多响应集执行。

变量列表中每个变量旁的图标标识变量类型。

类别。 在变量列表中选择分类变量后，将在“变量信息”窗格中显示为该变量定义类别。如果在表中使用该变量，还将在画布窗格中显示这些类别。如果变量没有已定义的类别，那么“变量信息”窗格和画布窗格将显示两个占位符类别：类别 1 和类别 2。

显示在表构建器中的已定义类别基于**值标签**，即指定给不同数据值的描述性标签（例如，数值 0 和 1 的值标签分别为 *male* 和 *female*）。您可以在数据编辑器的“变量信息”窗格中定义值标签。

画布窗格。 通过将变量拖放到画布窗格的行和列上可以构建表。画布窗格显示要创建的表的预览。画布窗格不显示单元格中的实际数据值，但会提供十分准确的最终表的布局视图。对于分类变量，如果数据文件包含尚未定义值标签的唯一值，则实际表可能会比预览包含更多的类别。

构建表的基本规则和限制

- 对于分类变量，汇总统计基于统计源维度中最里面的变量。
- 分类变量的缺省统计源维度（行或列）基于这些变量拖放到画布窗格中的顺序。例如，如果首先将变量拖动到行托盘，则行维度是缺省的统计源维度。
- 刻度变量只能在行或列维度中最里面变量的类别中进行摘要。（可以将刻度变量放在表的任何级别，但将在最里面的级别对该变量进行摘要。）
- 刻度变量不能在其他刻度变量中进行摘要。可以对多个刻度变量的摘要进行堆积，也可以对分类变量的类别中的刻度变量进行摘要。不能将一个刻度变量嵌套在另一个刻度变量中，也不能将行维度中的一个刻度变量放在列维度中的另一个刻度变量中。
- 如果活动数据集中的任何变量包含 12,000 个以上的已定义值标签，则不能使用表构建器创建表。如果不需要在表中包含超过此限制的变量，可以定义和应用排除这些变量的变量集。如果需要在表中包含任何具有 12,000 个以上的已定义值标签的变量，则可以使用 CTABLES 命令语法生成这些表。

构建表

1. 从菜单中选择：

分析 > 表 > 定制表

2. 将一个或多个变量拖放到画布窗格的行和/或列区域。
3. 单击**创建**以创建表。

要从画布窗格删除变量

1. 在画布窗格中选择（单击）变量。
2. 单击右键并从下拉菜单中选择**删除变量**。

嵌套变量

与交叉制表相似，嵌套可以显示两个分类变量之间的关系，区别在于使用嵌套时一个变量嵌套在同一维度中的另一个变量内部。例如，可以在行维度中的 *Age category* 内部嵌套 *Gender*，从而显示每个年龄组中的男性和女性人数。

也可以在分类变量中嵌套刻度变量。例如，可以在 *Gender* 中嵌套 *Income*，以显示男性和女性收入的独立平均值（中位数或其他汇总测量）。

嵌套变量

1. 将一个分类变量拖放到画布窗格的行或列区域中。
2. 将一个分类或标度变量拖放到分类行或列变量的顶部。
3. 从菜单中选择**嵌套在所有变量上方**、**嵌套在左侧**或**嵌套在右侧**。

表 1. 嵌套分类变量

变量 1	变量 2	汇总统计
类别 1	类别 1	12
	类别 2	34
	类别 3	56
类别 2	类别 1	12
	类别 2	34
	类别 3	56

注：定制表不会显示分层拆分文件处理。要取得与分层拆分文件相同的结果，可将拆分文件变量放在表的最外嵌套层中。

编辑统计

"编辑统计"窗格允许：

- 在表中添加和移去汇总统计。

"编辑统计"窗格中可用的统计（和其他选项）取决于统计源变量的测量级别。统计源（统计所基于的变量）由以下因素确定：

- **测量级别。**如果表（或堆积表中的表部分）包含标度变量，那么统计基于该标度变量。
- **变量选择顺序。**分类变量的缺省统计源维度（行或列）基于将这些变量拖放到画布窗格上的顺序。例如，如果首先将变量拖动到行区域，则行维度是缺省的统计源维度。
- **嵌套。**对于分类变量，统计基于统计源维度中最里面的变量。

分类变量的汇总统计：适用于分类变量的基本统计有计数和百分比。也可以为总计和小计指定定制汇总统计。这些定制汇总统计包括集中趋势度量（例如平均值和中位数）及可能适于一些序数分类变量的离差测量（例如标准差）。

计数。每个表单元格中的个案数或多响应集的响应数。如果加权有效，此值就是加权计数。

- 如果加权有效，该值就是加权计数。
- 对于全局数据集加权（数据 > 对个案进行加权...），加权计数都是相同的。

未加权的计数。每个表单元格中的未加权的个案数。仅当加权生效时，此统计才与计数有区别。

调整后计数。在有效基数权重计算中使用的调整后计数。如果不使用有效基数权重变量，调整后计数与计数将相同。

行百分比。每一行中的百分比。子表的每一行中的百分比（简单百分比）的总和为 100%。通常仅在具有分类列 变量时，行百分比才有用。

列百分比。每一列中的百分比。子表的每一列中的百分比（简单百分比）的总和为 100%。通常仅在具有分类行 变量时，列百分比才有用。

子表百分比。每个单元格中的百分比都基于子表。所有单元格百分比都基于相同的个案总数且总和为 100%。在嵌套表中，位于最里面的嵌套级别之前的变量定义子表。例如，在一个嵌套表中，*Marital status* 嵌套在 *Gender* 中，而后者又嵌套在 *Age category* 中，则由 *Gender* 定义子表。

表百分比。 每个单元格中的百分比基于整个表。所有单元格百分比都基于相同的个案总数且总和为 100%（简单百分比）。

置信区间

- 置信限制上限和下限可用于计数、百分比、平均值、中位数、百分位和求和。
- 标签中的文本字符串"&[Confidence Level]"在表中列标签中包含置信度。
- 标准错误可用于计数、百分比、平均值和求和。
- 置信区间和标准错误不可用于多响应集。

水平 这是置信区间的置信度，以百分比表示。值必须大于 0 且小于 100。

多响应集(U)

多响应集可以具有基于个案数、响应数或计数的百分比。请参阅『多重响应集的汇总统计』主题以获取更多信息。

百分比基数： 可以用三种不同的方法计算百分比，这些方法由计算基数中的缺失值处理确定：

简单百分比。 百分比基于表中所使用的个案数且总和始终为 100%。如果从表中排除了某一类别，也将排除该类别中的个案。在任何情况下都将从基数中排除具有系统缺失值的个案。如果从表中排除用户缺失类别（缺省设置），也会排除具有用户缺失值的个案；相反，如果在表中包含用户缺失类别，也会包含具有用户缺失值的个案。任何其名称中不包含有效 N 或总计 N 的百分比都是简单百分比。

总计 N 百分比。 将向简单百分比基数中添加具有系统缺失值和用户缺失值的个案。百分比的总和可能小于 100%。

有效 N 百分比。 即使在表中包含用户缺失类别，也会从简单百分比基数中移去具有用户缺失值的个案。

注：任何情况下都将从基数中排除手动排除类别（用户缺失类别除外）中的个案。

多重响应集的汇总统计： 以下附加汇总统计可用于多响应集。

列/行/层响应 %。 基于响应数的百分比。

列/行/层响应 %（基准：计数）。 响应数是分子，总数是分母。

列/行/层计数 %（基准：响应）。 计数是分子，总响应数是分母。

列/行/层响应 %。 跨子表的百分比。基于响应数的百分比。

分层列/行响应 %（基准：计数）。 跨子表的百分比。响应数是分子，总数是分母。

分层列/行响应 %（基准：响应）。 跨子表的百分比。计数是分子，总响应数是分母。

响应。 响应计数。

子表/表响应 %。 基于响应数的百分比。

子表/表响应 %（基准：计数）。 响应数是分子，总数是分母。

子表/表计数 %（基准：响应）。 计数是分子，总响应数是分母。

标度变量和分类定制总计的汇总统计：除了适用于分类变量的计数和百分比之外，还为刻度变量提供了以下汇总统计，并且这些汇总统计还作为分类变量的定制总计和小计摘要。这些汇总统计不适用于多响应集或字符串（字母数值）变量。

平均值。 算术平均值；总和除以个案数。

中位数。 一个值，大于该值和小于该值的个案数各占一半，第 50 个百分位。

众数。 出现频率最高的值。如果存在出现频率相等的值，则显示最小值。

最小值。 最小（最低）值。

最大值。 最大（最高）值。

缺失。 缺失值（用户和系统缺失值）计数。

百分位数。 可以包含第 5 个、第 25 个、第 75 个、第 95 个和/或第 99 个百分位。

范围。 最大值和最小值之差。

标准差。 对围绕平均值的离差的测量。在正态分布中，68% 的个案在平均值的一个标准差范围内，95% 的个案在平均值的两个标准差范围内。例如，在正态分布（方差的平方根）中，如果平均年龄为 45，标准差为 10，则 95% 的个案将处于 25 到 65 之间。

和。 值的总和。

合计百分比。 基于总和的百分比。适用于行和列（在子表中）、所有行和列（跨子表）、层、子表和整个表。

总计 N。 无缺失值、用户缺失值和系统缺失值的计数。不包含手动排除的类别（用户缺失类别除外）中的个案。

调整后总计 N。 在有效基数权重计算中使用的调整后总计 N。如果不使用有效基数权重变量（"选项"选项卡），调整后总计 N 与总计 N 将相同。此统计信息不能用于多个响应集。

有效 N。 无缺失值的计数。不包含手动排除的类别（用户缺失类别除外）中的个案。

调整后有效 N。 在有效基数权重计算中使用的调整后有效 N。如果不使用有效基数权重变量（"选项"选项卡），调整后有效 N 与有效 N 将相同。此统计信息不能用于多个响应集。

方差。 对围绕平均值的离差的测量，值等于与平均值的差的平方和除以个案数减一。度量方差的单位是变量本身的单位的平方（标准差的平方）。

置信区间

- 置信限制上限和下限可用于计数、百分比、平均值、中位数、百分位和求和。
- 标签中的文本字符串"&[Confidence Level]"在表中列标签中包含置信度。
- 标准错误可用于计数、百分比、平均值和求和。
- 置信区间和标准错误不可用于多响应集。

水平 这是置信区间的置信度，以百分比表示。值必须大于 0 且小于 100。

堆积表

堆积变量所定义每个表部分均被视为一个单独表，并将相应地为其计算汇总统计。

分类和总计

定制表允许您：

- 重新排序类别。
- 插入总计。
- 对于未定义值标签的变量，只能排列类别和插入总计。

访问类别和总计选项

1. 将某一分类变量或多响应集拖放到画布窗格中。
2. 右键单击画布窗格上的该变量，然后从弹出菜单中选择一个类别或总计选项。

对类别排序

1. 右键单击画布窗格上的变量，从弹出菜单中选择**对类别排序**，然后选择排序方法：
 - 按值
 - 按标签
 - 按计数
 - 按下级

总计

1. 右键单击画布窗格上的变量，从弹出菜单中选择**显示总计**，然后选择显示总计的位置：
 - 高于类别
 - 低于类别

如果选定变量嵌套在另一个变量中，则将为每个小计插入总计。

定制表："检验统计"

"检验统计"功能提供定制表的显著性检验。



这些检验不适用于类别标签移出其缺省表维度的表，也不适用于已计算类别。

列均值和列比例检测

列均值检验可用于刻度变量。列比例检验可用于分类变量。

比较列均值

针对列均值相等性的成对检验。表必须在列中具有一个分类变量，且具有一个作为行最内部级别的刻度变量。该表必须包含平均值作为汇总统计。

针对有序分类变量，从所有类别或仅比较的类别估计差异。针对多响应变量，均值检测的差异始终仅基于比较的类别。

比较列的比例(P)

针对列比例相等性的成对检验。在列和行中，表必须至少具有一个分类变量。表必须包含计数或列百分比。

显著性水平

列均值和列比例检验的显著性级别。

- 值必须大于 0 且小于 1。

- 如果指定两个显著性级别，那么会使用大写字母来标识小于或等于较小级别的显著性值。将使用小写字母标识小于或等于较大级别的显著性值。
- 如果选择使用 **APA 样式下标**，那么会忽略第二个值。

针对多个比较调整 p 值

Bonferroni 纠正针对族系误差率 (FWER) 进行调整。**Benjamini-Hochberg** 方法为伪发现率 (FDR) 调整。此方法较 Bonferroni 纠正更为保守。

标识显著性差异

针对列均值和列比例检验，可在单独表或主表中显示显著性结果。

在单独表中

显著性检验结果显示在单独的表中。如果两个值显著不同，则较大值对应的单元格将显示键，此键标识较小值所在列。

显示显著性值

显著性值显示在单元格中每个键值后括号中。仅当显著性结果显示在单独表中，此选项才可用。

在主表中

显著性检验结果显示在主表中。表中每个列类别使用字母关键字进行标识。对于每一显著性对，具有较小列均值或比例类别关键字将显示在具有较大列均值或比例的类别中。

- 将光标悬停在透视表中列标签单元格中的关键字上时，表中具有该显著性关键字的所有单元格都会突出显示。如果是在列维度中具有多个变量的表，只有该子表中的单元格会突出显示。
- 要选择表（或子表）中具有相同显著性关键字的所有单元格，请右键单击列标签单元格，然后选择 **选择 > 选择具有此显著性关键字的所有单元格**。

使用 APA 样式下标

使用 APA 样式格式（使用下标字母）标识显著性差异。如果两个值显著不同，这些值将显示不同的下标字母。这些下标并非脚注。当此选项有效时，在当前表格外观中定义脚注样式将被覆盖，并且脚注将显示为上标数字。要选择同一行中具有相同显著性关键字的所有单元格，请右键单击一个具有显著性关键字的单元格，然后选择 **选择具有类似显著性的单元格**

独立性检验 (卡方验证)(T)

读表执行独立性卡方检验，该表的行和列中至少同时有一个分类变量。

使用小计代替小计类别(U)

每个小计将替换其类别以用于显著性检验。否则，只有隐藏其小计类别的小计才替换其类别以用于检验。

检验中包含多重响应变量(I)

多响应集类别包含在显著性检验中。否则，多响应集不包含在显著性检验中。

样本文件

随产品一起安装的样本文件可以在安装目录的 *Samples* 子目录中找到。以下每种语言的 "Samples" 子目录中都有一个独立的文件夹：英语、法语、德语、意大利语、日语、韩语、波兰语、俄语、简体中文、西班牙语和繁体中文。

并非所有样本文件均提供此处的全部语言版本。如果样本文件未提供某种语言的版本，则相应语言文件夹中包含该样本文件的英语版本。

描述

以下是对在整个文档的各种示例中使用的样本文件的简要描述。

- **accidents.sav**。 该假设数据文件涉及某保险公司，该公司正在研究给定区域内汽车事故的年龄和性别风险因子。每个个案对应一个年龄类别和性别类别的交叉分类。
- **adl.sav**。 该假设数据文件涉及在确定针对脑卒中患者的建议治疗类型的优点方面的举措。医师将女性脑卒中患者随机分配到两组中的一组。第一组患者接受标准的物理治疗，而第二组患者则接受附加的情绪治疗。在进行治疗的三个月时间里，将为每个患者进行一般日常生活行为的能力评分并作为原始变量。
- **advert.sav**。 该假设数据文件涉及某零售商在检查广告支出与销售业绩之间的关系方面的举措。为此，他们收集了过去的销售数字以及相关的广告成本。
- **aflatoxin.sav**。 该假设数据文件涉及对谷物的黄曲霉毒素的检测，该毒素的浓度会因谷物产量的不同（不同谷物之间及同种谷物之间）而有较大变化。谷物加工机从 8 个谷物产量的每一个中收到 16 个样本并以十亿分之几 (PPB) 为单位来测量黄曲霉毒素的水平。
- **anorectic.sav**。 在研究厌食/暴食行为的标准症状参照时，研究人员¹对 55 名已知存在进食障碍的青少年进行了调查。其中每名患者每年都将进行四次检查，因此总观测数为 220。在每次观测期间，将对这些患者按 16 种症状逐项评分。但 71 号和 76 号患者的症状得分均在时间点 2 缺失，47 号患者的症状得分在时间点 3 缺失，因此有效观测数为 217。
- **anticonvulsants.sav**。 医学研究者可以用广义线性混合模型来判定一种新型抗惊厥药能否降低患者癫痫发作的比率。对同一患者重复测量通常是正相关的，因此适合使用具有随机效应的混合模型。目标字段为发作次数，其值为正整数，因此可能适合使用具有泊松分布和对数关联的广义线性混合模型。
- **bankloan.sav**。 该假设数据文件涉及某银行在降低贷款拖欠率方面的举措。该文件包含 850 位过去和潜在客户的财务和人口统计信息。前 700 个个案是以前曾获得贷款的客户。剩下的 150 个个案是潜在客户，银行需要按高或低信用风险对他们进行分类。
- **bankloan_binning.sav**。 该假设数据文件包含 5,000 位过去客户的财务和人口统计信息。
- **bankloan_cs.sav**。 该假设数据文件涉及某银行在确定指示可能会在贷款时违约的人员的特征并随后使用这些特征确定良好和不良信用风险方面的举措。
- **bankloan_cs_noweights.sav**。 该假设数据文件涉及某银行在确定指示可能会在贷款时违约的人员的特征并随后使用这些特征确定良好和不良信用风险方面的举措。抽样权重不包含在文件中。
- **behavior.sav**。 在一个经典示例中²，52 名学生被要求以 10 点的标度对 15 种情况和 15 种行为的组合进行评价，该 10 点的标度从 0 = "极得体"到 9 = "极不得体"。平均值在个人值之上，值被视为相异性。
- **behavior_ini.sav**。 该数据文件包含 *behavior.sav* 的二维解的初始配置。
- **brakes.sav**。 该假设数据文件涉及某生产高性能汽车盘式制动器的工厂的质量控制。该数据文件包含对 8 台专用机床中每一台的 16 个盘式制动器的直径测量。盘式制动器的目标直径为 322 毫米。
- **breakfast.sav**。 在一项经典研究中³，21 名 Wharton School MBA 学生及其配偶被要求按照喜好程度顺序对 15 种早餐食品进行评价，从 1 = 他们的喜好根据六种不同的情况加以记录，从"全部喜欢"到"只带饮料的快餐"。
- **breakfast-overall.sav**。 该数据文件只包含早餐食品喜好的第一种情况，即"全部喜欢"。
- **broadband_1.sav**。 该假设数据文件包含各地区订制了全国宽带服务的客户的数量。该数据文件包含 4 年期间 85 个地区每月的订户数量。

1. Van der Ham, T., J. J. Meulman, D. C. Van Strien, and H. Van Engeland. 1997. Empirically based subgrouping of eating disorders in adolescents: A longitudinal perspective. *British Journal of Psychiatry*, 170, 363-368.

2. Price, R. H., and D. L. Bouffard. 1974. Behavioral appropriateness and situational constraints as dimensions of social behavior. *Journal of Personality and Social Psychology*, 30, 579-586.

3. Green, P. E., and V. Rao. 1972. *Applied multidimensional scaling*. Hinsdale, Ill.: Dryden Press.

- **broadband_2.sav**。该数据文件和 *broadband_1.sav* 一样，但包含另外三个月的数据。
- **cable_survey.sav**。电视、电话和互联网服务的电缆提供商的高级管理人员希望了解有关潜在客户的更多信息。他们对其服务区域内的 2000 人进行了一项调查，调查问卷的问题包括他们 (1) 是否未采用此服务、(2) 是否向其他提供商订阅了此服务或 (3) 是否采用了公司的服务（对于三项服务中的每项）。此外，该调查还收集了一些人口统计信息，例如性别、年龄类别（4 个级别）、教育类别（3 个级别）、收入类别（3 个级别）、居住类型类别（4 个级别）、在当前地址的居住年限类别（3 个级别）、家庭成员数等。
- **car_insurance_claims.sav**。在别处被提出和分析的⁴关于汽车损坏赔偿的数据集。平均理赔金额可以当作其具有伽玛分布来建模，通过使用逆关联函数将因变量的平均值与投保者年龄、车辆类型和车龄的线性组合关联。提出理赔的数量可以作为刻度权重。
- **car_sales.sav**。该数据文件包含假设销售估计值、订价以及各种品牌和型号的车辆的物理规格。订价和物理规格可以从 *edmunds.com* 和制造商处获得。
- **car_sales_uprepared.sav**。这是 *car_sales.sav* 的修改版本，不包含字段的任何已转换版本。
- **carpet.sav**。在常用示例⁵中，有一家公司想要销售一种新型的地毯清洁用品。该公司希望调查以下会对消费者偏好产生影响的五个因素：包装设计、品牌名称、价格、优质家用品标志、以及退货保证。包装设计有三个因子级别，每个因子级别因刷体位置而不同；有三个品牌名称（*K2R*、*Glory* 和 *Bissell*）；有三个价格水平；最后两个因素各有两个级别（有或无）。十名消费者对这些因素所定义的 22 个特征进行了排序。变量优选包含对每个概要文件的平均等级的排序。低等级与高偏好相对应。此变量反映了对每个概要文件的偏好的总体度量。
- **carpet_prefs.sav**。该数据文件所基于的示例和在 *carpet.sav* 中所描述的一样，但它还包含从 10 位消费者的每一位中收集到的实际排列顺序。消费者被要求按照从最喜欢到最不喜欢的顺序对 22 个产品概要文件进行排序。*carpet_plan.sav* 中定义了变量 *PREF1* 到 *PREF22* 包含相关特征的标识。
- **catalog.sav**。该数据文件包含某编目公司出售的三种产品的假设每月销售数据。同时还包含 5 个可能的预测变量的数据。
- **catalog_seasfac.sav**。除添加了一组从“季节性分解”过程中计算出来的季节性因素和附带的日期变量外，该数据文件和 *catalog.sav* 是相同的。
- **cellular.sav**。该假设数据文件涉及某便携式电话公司在减少客户流失方面的举措。客户流失倾向分被应用到帐户，分数范围从 0 到 100。得到 50 分或更高分数的帐户可能会更换提供商。
- **ceramics.sav**。该假设数据文件涉及某制造商在确定新型优质合金是否比标准合金具有更高的耐热性方面的举措。每个个案代表对一种合金的单独检验；个案中会记录合金的耐热极限。
- **cereal.sav**。该假设数据文件涉及一份 880 人参与的关于早餐喜好的民意调查，该调查记录了参与者的年龄、性别、婚姻状况以及生活方式是否积极（根据他们是否每周至少做两次运动）。每个个案代表一个单独的响应者。
- **clothing_defects.sav**。这是关于某服装厂的质量控制过程的假设数据文件。检验员要对工厂中每次大批量生产的服装进行抽样检测并清点不合格的服装的数量。
- **coffee.sav**。这是关于六种冰咖啡的认知品牌形象⁶的数据文件。对于 23 种冰咖啡特征属性中的每种属性，人们选择了由该属性所描述的所有品牌。为保密起见，六种品牌用 AA、BB、CC、DD、EE 和 FF 来表示。
- **contacts.sav**。该假设数据文件涉及一组公司计算机销售代表的联系方式列表。根据这些销售代表所在的公司部门及其公司的等级来对每个联系方式进行分类。同时还记录了最近一次的销售量、最近一次销售距今的时间和所联系公司的规模。

4. McCullagh, P., and J. A. Nelder. 1989. *Generalized Linear Models*, 2nd ed. London: Chapman & Hall.

5. Green, P. E., and Y. Wind. 1973. *Multiaattribute decisions in marketing: A measurement approach*. Hinsdale, Ill.: Dryden Press.

6. Kennedy, R., C. Riquier, and B. Sharp. 1996. Practical applications of correspondence analysis to categorical data in market research. *Journal of Targeting, Measurement, and Analysis for Marketing*, 5, 56-70.

- **credit_card.sav**。这是一项对信用卡使用情况的假设研究，它关注两年时间内每个主体其主卡的月度支出以及按交易类型（杂货、零售、娱乐、旅游和其他）细分的开支。数据集中的每条记录都对应于指定月份的支出及交易类型，因此对每个主体收集的数据需要 2 年 × 12 个月（每年）× 5 类交易 = 120 条记录。
- **creditpromo.sav**。该假设数据文件涉及某百货公司在评价最新信用卡促销的效果方面的举措。为此，随机选择了 500 位持卡人。其中一半收到了宣传关于在接下来的三个月内降低消费利率的广告。另一半收到了标准的季节性广告。
- **cross_sell.sav**。一家邮寄公司拥有一家书店和一家 CD 俱乐部。每个月他们都会为俱乐部会员举行一次特别优惠活动。该公司希望根据总书籍采购量、CD 采购量以及为俱乐部会员提供的优惠类型为月度总特殊优惠采购活动创建模型。二阶最小平方回归适用于这种情况，因为在特别优惠活动中支出的费用是未花费在书籍或 CD 上的费用，因此在响应和这两个预测变量之间存在反馈循环。
- **customer_dbase.sav**。该假设数据文件涉及某公司在使用数据仓库中的信息来为最有可能回应的客户提供特惠商品方面的举措。随机选择客户群的子集并为其提供特惠商品，同时记录下他们的回应。
- **customer_information.sav**。该假设数据文件包含客户邮寄信息，如姓名和地址。
- **customer_subset.sav**。来自 *customer_dbase.sav* 的拥有 80 个个案的子集。
- **debate.sav**。该假设数据文件涉及在某政治辩论前后对该辩论的参与者所做的调查的成对回答。每个个案对应一个单独的响应者。
- **debate_aggregate.sav**。该假设数据文件汇总了 *debate.sav* 中的回答。每个个案对应一个辩论前后的偏好的交叉分类。
- **demo.sav**。这是关于购物客户数据库的假设数据文件，用于寄出每月的商品。将记录客户对商品是否有回应以及各种人口统计信息。
- **demo_cs_1.sav**。该假设数据文件涉及某公司在汇编调查信息数据库方面的举措的第一步。每个个案对应不同的城市，并记录地区、省、区和城市标识。
- **demo_cs_2.sav**。该假设数据文件涉及某公司在汇编调查信息数据库方面的举措的第二步。每个个案对应来自第一步中所选城市的不同的家庭单元格，并记录地区、省、区、市、子区和单元格标识。还包括设计前两个阶段的抽样信息。
- **demo_cs.sav**。该假设数据文件包含用复杂抽样设计收集的调查信息。每个个案对应不同的家庭单元格，并记录各种人口统计和抽样信息。
- **diabetes_costs.sav**。这是假设数据文件，其中包含保险公司维护的有关患有糖尿病的投保人的信息。每个个案对应一个不同投保人。
- **dietstudy.sav**。该假设数据文件包含对 "Stillman diet"⁷ 的研究结果。每个个案对应一个单独的主体，并记录其在实行饮食方案前后的体重（磅）以及甘油三酸酯的水平（毫克/100 毫升）。
- **dmdata.sav**。该假设数据文件包含直销公司的人口统计学和购买信息。*dmdata2.sav* 包含收到测试邮件的一部分联系人的信息，而 *dmdata3.sav* 包含其余未收到测试邮件的联系人信息。
- **dvdplayer.sav**。这是关于开发新的 DVD 播放器的假设数据文件。营销团队用原型收集了焦点小组数据。每个个案对应一个单独的被调查用户，并记录他们的人口统计信息及其对原型问题的回答。
- **Employee data.sav**。该假设数据文件包含特定于员工的信息（教育水平、雇佣类别、当前薪资、先前工作经验等）。
- **german_credit.sav**。该数据文件取自加州大学欧文分校的 Repository of Machine Learning Databases⁸ 中的 "German credit" 数据集。

7. Rickman, R., N. Mitchell, J. Dingman, and J. E. Dalen. 1974. Changes in serum cholesterol during the Stillman Diet. *Journal of the American Medical Association*, 228:, 54-58.

8. Blake, C. L., and C. J. Merz. 1998. "UCI Repository of machine learning databases." Available at <http://www.ics.uci.edu/~mllearn/MLRepository.html>.

- **grocery_1month.sav**。该假设数据文件是在数据文件 *grocery_coupons.sav* 的基础上加上了每周购物"累计", 所以每个个案对应一个单独的客户。所以, 一些每周更改的变量消失了, 而且现在记录的消费金额是为期四周的研究过程中的消费金额之和。
- **grocery_coupons.sav**。该假设数据文件包含由重视顾客购物习惯的杂货连锁店收集的调查数据。对每位顾客调查四周, 每个个案对应一个单独的顾客周, 并记录有关顾客购物地点和方式的信息 (包括那一周里顾客在杂货上的消费金额)。
- **guttman.sav**。Bell⁹创建了一个表, 用来阐释可能的社会群体。Guttman¹⁰引用了该表的一部分, 其中包括五个变量, 用于描述以下七个理论社会群体的社会交往、对群体的归属感、成员的物理亲近度以及关系正式性: 观众 (比如在足球比赛现场的人们)、听众 (比如在剧院或听课堂讲座的人们)、公众 (比如报纸或电视观众)、组织群体 (与观众类似但具有紧密的关系)、初级群体 (关系密切)、次级群体 (自发组织) 及现代社区 (因在物理上亲近而导致关系松散并需要专业化服务)。
- **health_funding.sav**。该假设数据文件包含关于保健基金 (每 100 人的金额)、发病率 (每 10,000 人的比率) 以及保健提供商拜访率 (每 10,000 的比率) 的数据。每个个案代表不同的城市。
- **hivassay.sav**。该假设数据文件涉及某药物实验室在开发用于检测 HIV 感染的快速化验方面的举措。化验结果为八个加深的红色阴影, 如果有更深的阴影则表示感染的可能性很大。用 2,000 份血液样本来进行实验室试验, 其中一半受到 HIV 感染而另一半没有受到感染。
- **hourlywagedata.sav**。该假设数据文件涉及在政府机关和医院工作的具有不同经验水平的护士的时薪。
- **insurance_claims.sav**。该假设数据文件涉及某保险公司, 该公司希望构建一个模型来标记具有潜在欺骗性的可疑理赔。每个个案代表一次单独的理赔。
- **insure.sav**。该假设数据文件涉及某保险公司, 该公司正在研究指示客户是否会根据 10 年的人寿保险合同提出理赔的风险因子。数据文件中的每个个案代表一副根据年龄和性别进行匹配的合同, 其中一份记录了一次理赔而另一份则没有。
- **judges.sav**。该假设数据文件涉及经过训练的裁判 (加上一个体操爱好者) 对 300 次体操表演给出的分数。每行代表一次单独的表演; 裁判们观看相同的表演。
- **kinship_dat.sav**。Rosenberg 和 Kim¹¹ 开始分析 15 个亲属关系项 (伯母、兄弟、表兄妹、女儿、父亲、孙女、祖父、祖母、孙子、母亲、侄子或外甥、侄女或外甥女、姐妹、儿子和叔叔)。他们让四组大学生 (两组女同学, 两组男同学) 根据相似程度将各项排序。他们让其中的两组同学 (一组女同学, 一组男同学) 进行了两次排序, 第二次排序和第一次排序采取的标准不同。这样, 一共得到了六组"源"。每个源对应一个 15 × 15 的近似值矩阵, 该矩阵单元格中的值等于源中的人数减去该源中划分对象的次数。
- **kinship_ini.sav**。该数据文件包含 *kinship_dat.sav* 的三维解的初始配置。
- **kinship_var.sav**。该数据文件包含自变量 *gender*、*gener(ation)* 和 *degree (of separation)*, 这些变量可用于解释 *kinship_dat.sav* 的解的维数。具体而言, 它们可用来将解的空间限制为这些变量的线性组合。
- **marketvalues.sav**。该数据文件涉及 1999–2000 年期间 Algonquin, Ill. 地区新住宅开发中的房屋销售情况。这些销售情况属于公共记录。
- **nhis2000_subset.sav**。美国健康访问调查 (NHIS) 是针对美国全体公民的大型人口调查。该调查对美国的具有全国代表性的家庭样本进行了面对面的访问, 并获取了每个家庭的成员的健康行为和健康状态的人口统计信息和观察数据。该数据文件包含取自 2000 年调查信息的子集。国家健康统计中心。2000 年美国健康访问调查。公用数据文件和文档。 ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHIS/2000/。2003 年发布。

9. Bell, E. H. 1961. *Social foundations of human behavior: Introduction to the study of sociology*. New York: Harper & Row.

10. Guttman, L. 1968. A general nonmetric technique for finding the smallest coordinate space for configurations of points. *Psychometrika*, 33, 469-506.

11. Rosenberg, S., and M. P. Kim. 1975. The method of sorting as a data-gathering procedure in multivariate research. *Multivariate Behavioral Research*, 10, 489-502.

- **ozone.sav**。 这些数据包含了用来根据其余变量预测臭氧浓度的六个气象变量的 330 个观察值。在以前的研究人员中,¹² 和 ¹³ 发现了这些变量之间的非线性, 这妨碍了标准回归方法。
- **pain_medication.sav**。 该假设数据文件包含用于治疗慢性关节炎疼痛的抗炎药的临床试验结果。我们感兴趣的是该药见效的时间以及它和现有药物的比较。
- **patient_los.sav**。 该假设数据文件包含被医院确诊为疑似心肌梗塞(即 MI 或"心脏病发作")的患者的治疗记录。每个个案对应一位单独的患者, 并记录与其住院期有关的一些变量。
- **patlos_sample.sav**。 该假设数据文件包含在治疗心肌梗塞(即 MI 或"心脏病发作")期间收到溶解血栓剂的患者的样本治疗记录。每个个案对应一位单独的患者, 并记录与其住院期有关的一些变量。
- **poll_cs.sav**。 该假设数据文件涉及民意测验专家在确定正式立法前公众对法案的支持水平方面的举措。个案对应注册的选民。每个个案记录选民居住的县、镇、区。
- **poll_cs_sample.sav**。 该假设数据文件包含在 *poll_cs.sav* 中列出的选民的样本。该样本是根据 *poll_csplan* 中指定的设计来选取的, 而且该数据文件记录包含概率和样本权重。请注意, 由于该抽样计划使用与大小成正比 (PPS) 方法, 因此, 还有一个文件 (*poll_jointprob.sav*) 包含联合选择概率。在选取了样本之后, 对应于选民人群统计信息及其对提交法案的意见的附加变量将被收集并添加到数据文件。
- **property_assess.sav**。 该假设数据文件涉及某县资产评估员在利用有限的资源不断更新资产价值评估方面的举措。个案对应过去一年中县里所出售的资产。数据文件中的每个个案记录资产所在的镇、最后评估资产的评估员、该次评估距今的时间、当时的估价以及资产的出售价格。
- **property_assess_cs.sav**。 该假设数据文件涉及某州资产评估员在利用有限的资源不断更新资产价值评估方面的举措。个案对应该州的资产。数据文件中的每个个案记录资产所在的县、镇和区, 最后一次评估距今的时间以及当时的估价。
- **property_assess_cs_sample.sav**。 该假设数据文件包含在 *property_assess_cs.sav* 中列出的资产的样本。该样本是根据 *property_assess_csplan* 中指定的设计来选取的, 而且该数据文件记录包含概率和样本权重。在选取了样本之后, 附加变量 *Current value* 将被收集并添加到数据文件。
- **recidivism.sav**。 该假设数据文件涉及某政府执法机构在了解其管辖区域内的屡犯率方面的举措。每个个案对应先前的一名罪犯, 并记录其人口统计信息和第一次犯罪的详细资料; 如果在第一次被捕后两年内又第二次被捕, 则还将记录两次被捕间隔的时间。
- **recidivism_cs_sample.sav**。 该假设数据文件涉及某政府执法机构在了解其管辖区域内的屡犯率方面的举措。每个个案对应第一次被捕后于 2003 年 6 月期间释放的一名先前的罪犯, 并记录其人口统计信息、第一次犯罪的详细资料以及第二次被捕的数据(如果于 2006 年 6 月底之前第二次被捕)。罪犯是根据 *recidivism_cs_csplan* 中指定的抽样计划从抽样的部门中选择的; 由于该抽样计划使用与大小成正比 (PPS) 方法, 因此, 还有一个文件 (*recidivism_cs_jointprob.sav*) 包含联合选择概率。
- **rfm_transactions.sav**。 此假设数据文件包含购买交易数据, 即每笔交易的购买日期、购买商品和消费金额。
- **salesperformance.sav**。 这是关于评估两个新的销售培训课程的假设数据文件。60 名员工被分成 3 组且都接受标准的培训。另外, 组 2 接受技术培训; 组 3 接受实践教程。在培训课程结束时, 对每名员工进行测试并记录他们的分数。数据文件中的每个个案代表一名单独的受训者, 并记录其被分配到的组以及测验的分数。
- **satisf.sav**。 该假设数据文件涉及某零售公司在 4 个商店位置所进行的满意度调查。总共对 582 位客户进行了调查, 每个个案代表一位单独客户的回答。
- **screws.sav**。 该数据文件包含关于螺钉、螺栓、螺母和图钉的特征的信息¹⁴。

12. Breiman, L., and J. H. Friedman. 1985. Estimating optimal transformations for multiple regression and correlation. *Journal of the American Statistical Association*, 80, 580-598.

13. Hastie, T., and R. Tibshirani. 1990. *Generalized additive models*. London: Chapman and Hall.

14. Hartigan, J. A. 1975. *Clustering algorithms*. New York: John Wiley and Sons.

- **shampoo_ph.sav** 这是关于某发制品厂的质量控制的假设数据文件。在规定的的时间间隔对六批独立输出的产品进行检测并记录它们的 pH 值。目标范围是 4.5–5.5。
- **ships.sav**。在别处被提出和分析的¹⁵关于波浪对货船造成的损坏的数据集。在给定了船的类型、建造工期和服务期后，可以根据以泊松比率发生来为事件计数建模。在因子交叉分类构成的表格中，每个单元格的汇总服务月数提供遇到风险的值。
- **site.sav**。该假设数据文件涉及某公司在为扩展业务而选择新址方面的举措。该公司聘请了两名顾问分别对选址进行评估，除了提供长期报告外，他们还要以"前景颇佳"、"前景良好"或"前景不佳"来对每个选址进行总结。
- **smokers.sav**。该数据文件摘自 1998 年全国家庭药物滥用调查并且是美国家庭的概率样本。(http://dx.doi.org/10.3886/ICPSR02934) 因此，分析该数据文件的第一步应该是对数据进行加权，从而反映总体趋势。
- **stocks.sav** 该假设数据文件包含某一年的股票价格和成交量。
- **stroke_clean.sav**。该假设数据文件包含某医学数据库在经过 Statistics Base Edition 中的过程清理后的状态。
- **stroke_invalid.sav**。该假设数据文件包含某医学数据库的初始状态及一些数据输入错误。
- **stroke_survival**。此假设数据文件涉及正在研究结束缺血性中风后复元计划的患者存活时间的研究人员面临着很多挑战。中风后，记录心肌梗塞、缺血性中风或出血性中风的发生及其时间。样本为左侧截断，因为只包含在中风后管理的复元计划结束后存活的患者。
- **stroke_valid.sav**。该假设数据文件包含在使用"验证数据"过程检查值后，某医学数据库的状态。它仍包含潜在异常个案。
- **survey_sample.sav**。此数据文件包含调查数据，包括人口统计学数据和各种态度测量。它基于 1998 NORC 综合社会调查的变量子集，但某些数据值已经过修改，并添加了其他虚拟变量以供演示用途。
- **tcm_kpi.sav**。这是一个假设数据文件，其中包含公司的每周关键绩效指标的值。它还包含相同时间段内多个可控制指标的每周数据。
- **tcm_kpi_upd.sav**。该数据文件和 *tcm_kpi.sav* 一样，但包含另外四周的数据。
- **telco.sav**。该假设数据文件涉及某电信公司在减少客户群中的客户流失方面的举措。每个个案对应一个单独的客户，并记录各类人口统计和服务用途信息。
- **telco_extra.sav**。该数据文件与 *telco.sav* 数据文件类似，但删除了 "tenure" 和经对数转换的客户消费变量，代替它们的是标准化的对数转换客户消费变量。
- **telco_missing.sav**。该数据文件是 *telco.sav* 数据文件的子集，但某些人口统计数据值已被缺失值替换。
- **testmarket.sav**。该假设数据文件涉及某快餐连锁店为其菜单添加新项目的计划。有三种可能的促销新产品的活动，所以会在多个随机选择的地点引入新的项目。在每个地点采用不同的促销方式，并记录新项目前四周的每周销售情况。每个个案对应单独地点的一周。
- **testmarket_1month.sav**。该假设数据文件是在数据文件 *testmarket.sav* 的基础上加上了每周销售"累计"，所以每个个案对应一个单独的地点。所以，一些每周更改的变量消失了，而且现在记录的销售是为期四周的研究过程中的销售之和。
- **tree_car.sav**。该假设数据文件包含人口统计和车辆购买价格数据。
- **tree_credit.sav**。该假设数据文件包含人口统计和银行贷款历史数据。
- **tree_missing_data.sav** 该假设数据文件包含具有大量缺失值的人口统计和银行贷款历史数据。
- **tree_score_car.sav**。该假设数据文件包含人口统计和车辆购买价格数据。

15. McCullagh, P., and J. A. Nelder. 1989. *Generalized Linear Models*, 2nd ed. London: Chapman & Hall.

- **tree_textdata.sav**。这是一个只有两个变量的样本数据文件，主要打算在指定测量级别和值标签之前显示变量的缺省状态。
- **tv-survey.sav**。该假设数据文件涉及由某电视演播室进行的一项关于是否要继续制作一档成功的节目的调查。906 位响应者被问及他们在各种情况下是否会收看该节目。每行代表一位单独的响应者；每列代表一种单独的情况。
- **ulcer_recurrence.sav**。此文件包含某项研究的部分信息，该研究旨在比较两种用来防止溃疡复发的治疗的功效。它提供了区间数据的优秀示例并且已在别处被提出和分析¹⁶。
- **ulcer_recurrence_recoded.sav**。该文件重新组织 *ulcer_recurrence.sav* 中的信息以允许为研究的每个区间的事件概率建模而不是简单地研究结束事件概率建模。它已在别处被提出和分析¹⁷。
- **verd1985.sav**。该数据文件涉及某项调查¹⁸。该调查记录了 15 个主体对 8 个变量的响应。需要处理的变量被分成 3 个集。数据集 1 包含 年龄 和 婚姻；数据集 2 包含 宠物 和 新闻；数据集 3 包含 音乐 和 居住。宠物被刻度化为多名义而年龄被刻度化为有序；所有其他变量都被刻度化为单名义。
- **virus.sav**。该假设数据文件涉及某因特网服务提供商 (ISP) 在确定病毒对其网络的影响方面的举措。他们从发现病毒到威胁得以遏制这段时间内跟踪其网络上受感染的电子邮件的流量的（近似）百分比。
- **wheeze_steubenville.sav**。这是关于空气污染对儿童健康影响的纵向研究的一个子集¹⁹。这些数据包含儿童的气喘状况的重复二分类测量（这些儿童来自 Steubenville, Ohio, 年龄为 7 到 10 岁），以及母亲在研究的第一年中是否为吸烟者的固定记录。
- **workprog.sav**。该假设数据文件涉及一份尝试为弱势群体提供较好的工作的政府工作计划。文件后还有一个潜在计划参与者的样本，其中一些参与者是被随机选择来参加该计划的，而其他参与者则不是。每个个案代表一位单独的计划参与者。
- **worldsales.sav** 该假设数据文件包含按不同大洲和产品列出的销售收入。

16. Collett, D. 2003. *Modelling survival data in medical research*, 2 ed. Boca Raton: Chapman & Hall/CRC.

17. Collett, D. 2003. *Modelling survival data in medical research*, 2 ed. Boca Raton: Chapman & Hall/CRC.

18. Verdegaal, R. 1985. *Meer sets analyse voor kwalitatieve gegevens (in Dutch)*. Leiden: Department of Data Theory, University of Leiden.

19. Ware, J. H., D. W. Dockery, A. Spiro III, F. E. Speizer, and B. G. Ferris Jr.. 1984. Passive smoking, gas cooking, and respiratory health of children living in six cities. *American Review of Respiratory Diseases*, 129, 366-374.

通知

本信息是为在美国提供的产品和服务编写的。本资料的其他语言版本可以从 IBM 获取。但是，您可能需要拥有该语言的产品副本或产品版本才能访问这些资料。

IBM 可能在其他国家或地区不提供本文中讨论的产品、服务或功能特性。有关您当前所在区域的产品和服务的信息，请向您当地的 IBM 代表咨询。任何对 IBM 产品、程序或服务的引用并非意在明示或暗示只能使用 IBM 的产品、程序或服务。只要不侵犯 IBM 的知识产权，任何同等功能的产品、程序或服务，都可以代替 IBM 产品、程序或服务。不过，用户应自行负责评估和验证任何非 IBM 产品、程序或服务的工作情况。

IBM 公司可能已拥有或正在申请与本文档内容有关的各项专利。您获得该文档并不意味着授予您任何这些专利许可。您可以将书面许可查询函件发送至：

*IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US*

有关双字节 (DBCS) 信息的许可查询，请与您所在国家或地区的 IBM 知识产权部门联系，或用书面方式将查询寄往：

*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"按现状"提供本出版物，不附有任何种类的（无论是明示的还是暗含的）保证，包括但不限于暗含的有关非侵权、适销和适用于某种特定用途的保证。某些管辖区域在某些交易中不允许免除明示或暗含的保证。因此本条款可能不适用于您。

本信息可能含有技术误差或排版错误。此处的信息会定期进行更改；这些更改会体现在本出版物的新版本中。IBM 可以随时对本资料中描述的产品和/或程序进行改进和/或更改，而不另行通知。

本信息中对任何非 IBM Web 站点的引用都只是为了方便起见才提供的，不以任何方式充当对那些 Web 站点的保证。那些 Web 站点中的资料不是 IBM 产品资料的一部分，使用那些 Web 站点带来的风险将由您自行承担。

IBM 可以按它认为适当的任何方式使用或分发您所提供的任何信息而无须对您承担任何责任。

本程序的被许可方如果要了解有关程序的信息以达到如下目的：(i) 允许在独立创建的程序和其他程序（包括本程序）之间进行信息交换，以及 (ii) 允许对已经交换的信息进行相互使用，请与下列地址联系：

*IBM Director of Licensing
IBM Corporation*

North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US

此类信息的提供应遵照相关条款和条件，其中包括在某些情况下支付适当费用。

本资料中描述的许可程序及其所有可用的许可资料均由 IBM 依据 IBM 客户协议、IBM 国际软件许可协议或任何同等协议中的条款提供。

所引用的性能数据和客户示例只用于阐述说明。根据具体配置和操作条件，实际性能结果可能有所不同。

涉及非 IBM 产品的信息可从这些产品的供应商、其出版说明或其他可公开获得的资料中获取。IBM 没有对这些产品进行测试，也无法确认其性能的精确性、兼容性或任何其他关于非 IBM 产品的声明。有关非 IBM 产品性能的问题应当向这些产品的供应商提出。

有关 IBM 未来方向或意向的声明均可能未经通知即变更或撤销，并且仅代表目标和目的。

本信息包含日常业务运营中使用的数据和报告的示例。为了尽可能详尽地对其进行说明，示例中包含了人员的姓名、公司、品牌和产品的名称。所有这些名称都是虚构的，如果与实际人员或公司企业有任何类似则纯属巧合。

版权许可：

本信息包括源语言形式的样本应用程序，这些样本说明不同操作平台上的编程方法。如果是为按照在编写样本程序的操作平台上的应用程序编程接口 (API) 进行应用程序的开发、使用、经销或分发为目的，您可以任何形式对这些样本程序进行复制、修改、分发，而无须向 IBM 付费。这些示例并未在所有条件下作全面测试。因此，IBM 不能担保或暗示这些程序的可靠性、可维护性或功能。本样本程序仍然是"按现状"提供的，不附有任何种类的保证。对于因使用样本程序所引起的任何损害，IBM 概不负责。

凡这些实例程序的每份拷贝或其任何部分或任何衍生产品，都必须包括如下版权声明：

© IBM 2019. 此部分代码是根据 IBM Corp. 公司的样本程序衍生出来的。

© Copyright IBM Corp. 1989 - 20019. All rights reserved.

商标

IBM、IBM 徽标和 ibm.com 是 International Business Machines Corp., 在全球许多管辖区域注册的商标或注册商标。其他产品和服务名称可能是 IBM 或其他公司的商标。Web 站点 www.ibm.com/legal/copytrade.shtml 上的"Copyright and trademark information"部分中提供了 IBM 商标的最新列表。

Adobe、Adobe 徽标、PostScript 和 PostScript 徽标是 Adobe Systems Incorporated 在美国和/或其他国家或地区的注册商标或商标。

Intel、Intel 徽标、Intel Inside、Intel Inside 徽标、Intel Centrino、Intel Centrino 徽标、Celeron、Intel Xeon、Intel SpeedStep、Itanium 和 Pentium 是 Intel Corporation 或其子公司在美国和其他国家或地区的商标或注册商标。

Linux 是 Linus Torvalds 在美国、其他国家或地区或两者的注册商标。

Microsoft、Windows、Windows NT 和 Windows 徽标是 Microsoft Corporation 在美国、其他国家或地区或两者的商标。

UNIX 是 The Open Group 在美国和其他国家或地区的注册商标。

Java 和所有基于 Java 的商标和徽标是 Oracle 及/或其分支机构的商标和注册商标。

索引

[B]

- 百分比
 - 多响应集 4
 - 在定制表中 3, 4
- 标准差
 - 设定表格 5
- 表
 - 设定表格 1

[C]

- 测量级别
 - 在定制表中更改 1
- 拆分文件处理
 - 定制表 2

[D]

- 定制表
 - 百分比值 3, 4
 - 拆分文件处理 2
 - 重新排列类别 6
 - 多响应集 1
 - 多响应集的百分比 4
 - 分类变量 1
 - 分类变量的值标签 1
 - 更改测量级别 1
 - 汇总统计 3, 4, 5
 - 检验统计 6
 - 刻度变量 1
 - 控制显示的小数位数 3
 - 排除类别 6
 - 如何构建表 2
 - 显示格式 3
 - 小计 6
 - 已计算类别 6
 - 总计 6
- 对类别进行重新排序
 - 设定表格 6
- 多响应集
 - 百分比值 4

[F]

- 范围
 - 设定表格 5
- 方差
 - 设定表格 5

[J]

- 检验统计
 - 设定表格 6

[P]

- 排除类别
 - 设定表格 6
- 平均值
 - 设定表格 5

[S]

- 删除类别
 - 设定表格 6

[X]

- 显著性检验
 - 设定表格 6
- 小计
 - 设定表格 6
- 小数位
 - 控制定制表中显示的小数位数 3

[Y]

- 样本文件
 - 位置 7
- 有效 N
 - 设定表格 5

[Z]

- 中间值
 - 设定表格 5
- 众数
 - 设定表格 5
- 总和
 - 设定表格 5
- 总计
 - 设定表格 6
- 最大值
 - 设定表格 5
- 最小值
 - 设定表格 5



Printed in China