

*IBM SPSS Categories 29*



**Nota**

Prima di utilizzare queste informazioni e il prodotto che supportano, leggere le informazioni in [“Informazioni particolari” a pagina 45](#).

**Informazioni sul prodotto**

Questa edizione si applica alla versione 29, release 0, modifica 1 di IBM® SPSS Statistics e a tutte le release e modifiche successive se non diversamente indicato nelle nuove edizioni.

© **Copyright International Business Machines Corporation .**

---

# Indice

<b>Capitolo 1. Categorie.....</b>	<b>1</b>
Introduzione alle procedure di scaling ottimali per i dati categoriali.....	1
Cosa È Scaling Ottimale?.....	1
Perché Utilizzare Scaling Ottimale?.....	1
Livello di Scaling ottimale e Livello di misurazione.....	1
Quale Procedura è Best for Your Application?.....	4
Aspetto Rapporto in Diagrammi Di Scaling Ottimali.....	9
Regressione Categoriale (CATREG).....	10
Definire Scale in Regressione categoriale.....	11
Categorical Regression Discretization.....	11
Valori Mancanti di regressione categoriale.....	12
Opzioni Di Regressione Categoriale.....	12
Regolarizzazione Regressione Categoriale.....	13
Output Regression Categoriale.....	13
Regressione Categoriale Salva.....	14
Categorical Regression Transformation Plotti.....	14
Funzioni Aggiuntive Comando CATREG.....	14
Categorical Principal Components Analysis (CATPCA).....	15
Definire Scale e Peso in CATPCA.....	16
Categorical Principal Components Analysis Discretization.....	16
Categorical Principal Components Analysis Missing Valori.....	17
Categorical Principal Components Analysis Opzioni.....	17
Categorical Principal Components Analysis Output.....	19
Categorical Principal Components Analysis Save.....	19
Categorical Principal Components Analysis Object Plotti.....	19
Categorical Principal Components Analysis Categorie di analisi.....	20
Categorical Principal Components Analisi Caricamento Lotti.....	20
Categorical Principal Components Analysis Bootstrap.....	20
Funzioni Aggiuntive di comando CATPCA.....	21
Analisi della correlazione canonica non lineare (OVERALS).....	21
Definire intervallo e scala.....	23
Definisci intervallo.....	23
Opzioni Di Analisi Di Correlazione Canoniche Non Lineari.....	23
OVERALS Comando Funzioni Aggiuntive.....	24
Analisi delle corrispondenze.....	24
Definire intervallo di righe in Corrispondenza Analisi.....	26
Definire l'intervallo di colonne in Analisi della corrispondenza.....	26
Modello di analisi della corrispondenza.....	26
Statistiche di analisi della corrispondenza.....	27
Grafici di analisi della corrispondenza.....	28
CORRISPONDENZA Comando Funzioni Aggiuntive.....	28
Analisi delle corrispondenze multiple.....	29
Definire Peso variabile in Analisi Corrispondenza Multipla.....	29
Più Discretizzazione Analisi Corrispondenza.....	30
Più Analisi Corrispondenza valori mancanti.....	30
Opzioni Analisi di corrispondenza multipla.....	30
Output Analisi di corrispondenza multipla.....	31
Multiplo Analisi di corrispondenza Salva.....	32
Multi Corrispondenza oggetto di analisi di corrispondenza.....	32
Più Corrispondenza di Analisi Corrispondenza.....	32
MULTIPLE CORRISPONDENZA Comando Funzioni aggiuntive.....	33

Scaling multidimensionale (PROXSCAL).....	33
Prossimità in Matrici attraverso le Colonne.....	34
Prossimità in Colonne.....	34
Prossimità in una Colonna.....	35
Crea prossimità dai dati.....	35
Crea Misura da Dati.....	35
Definire un modello di Scaling Multidimensionale.....	36
Restrizioni Scaling Multidimensionali.....	37
Scaling multidimensionale: Opzioni.....	37
Grafici Scaling multidimensionali, Versione 1.....	37
Grafici Scaling multidimensionali, Versione 2.....	38
Output di Scaling multidimensionale.....	38
PROXSCAL Command Aggiuntive Features.....	39
Unfolding multidimensionale (PREFSCAL).....	39
Definire un modello di unidimensazione multidimensionale.....	40
Limitazioni Unfolding Multidimensionali.....	41
Opzioni di sfolding multidimensionali.....	41
Grafici di sfolding multidimensionali.....	42
Output unificato multidimensionale.....	43
Funzioni Aggiuntive di comando PREFSCAL.....	43
<b>Informazioni particolari.....</b>	<b>45</b>
Marchi.....	46
<b>Indice analitico.....</b>	<b>49</b>

---

# Capitolo 1. Categorie

Le seguenti funzioni categorie sono incluse in SPSS Statistiche Professional Edition o l'opzione Categorie.

## Introduzione alle procedure di scaling ottimali per i dati categoriali

Le procedure delle categorie utilizzano una scalatura ottimale per analizzare i dati difficili o impossibili per le procedure statistiche standard da analizzare. Questo capitolo descrive cosa ogni procedura fa, le situazioni in cui ogni procedura è più appropriata, le relazioni tra le procedure, e i rapporti di queste procedure ai loro omologhi statistici standard.

*Nota:* queste procedure e la loro implementazione in IBM SPSS Statistics sono state sviluppate dal Data Theory Scaling System Group (DTSS), composto da membri dei dipartimenti di Educazione e Psicologia, Facoltà di Scienze sociali e comportamentali, Università Leiden.

### Cosa È Scaling Ottimale?

L'idea alla base di una scaletta ottimale è quella di assegnare quantificazioni numeriche alle categorie di ogni variabile, permettendo così di utilizzare le procedure standard per ottenere una soluzione sulle variabili quantificate.

I valori di scala ottimali vengono assegnati a categorie di ogni variabile in base al criterio di ottimizzazione della procedura in uso. A differenza delle etichette originali delle variabili nominali o ordinali nell'analisi, questi valori di scala hanno proprietà metriche.

Nella maggior parte delle procedure di Categorie, la quantificazione ottimale per ogni variabile scalata si ottiene attraverso un metodo iterativo chiamato **alternando minimi quadrati** in cui, dopo che le attuali quantificazioni vengono utilizzate per trovare una soluzione, le quantificazioni vengono aggiornate utilizzando tale soluzione. Le quantificazioni aggiornate vengono poi utilizzate per trovare una nuova soluzione, che viene utilizzata per aggiornare le quantificazioni, e così via, fino ad arrivare a qualche criterio che segnali il processo di arresto.

### Perché Utilizzare Scaling Ottimale?

I dati categoriali si trovano spesso nella ricerca di marketing, ricerca indagine e ricerca nelle scienze sociali e comportamentali. Infatti molti ricercatori si occupano quasi esclusivamente di dati categoriali.

Mentre gli adattamenti dei modelli più standard esistono appositamente per analizzare i dati categoriali, spesso non si eseguono bene per i datasets che caratterizzano:

- Troppe poche osservazioni
- Troppe variabili
- Troppi valori per variabile

Quantificando le categorie, le tecniche di scaling ottimali evitano problemi in queste situazioni. Inoltre, sono utili anche quando le tecniche specializzate sono appropriate.

Piuttosto che interpretare le stime dei parametri, l'interpretazione dell'output di scaling ottimale si basa spesso su display grafici. Le tecniche di scaling ottimali offrono eccellenti analisi esplorative, che completano bene altri modelli IBM SPSS Statistics. Restringendo il focus della tua indagine, visualizzare i tuoi dati attraverso una scalatura ottimale può costituire la base di un'analisi che i centri sull'interpretazione dei parametri del modello.

### Livello di Scaling ottimale e Livello di misurazione

Questo può essere un concetto molto confuso quando si utilizzano le procedure di Categorie prime. Quando si specifica il livello, si specifica non il livello in cui le variabili sono *misurate* ma il livello

in cui sono *scalate*. L'idea è che le variabili da quantificare possano avere relazioni non lineari indipendentemente da come si misurano.

Per Categorie si trovano tre livelli fondamentali di misurazione:

- Il livello **nominale** implica che i valori di una variabile rappresentino categorie non ordinate. Esempi di variabili che potrebbero essere nominali sono regione, area codice zip, affiliazione religiosa e categorie a scelta multipla.
- Il livello **ordinale** implica che i valori di una variabile rappresentino categorie ordinate. Esempi includono scale di assetto che rappresentano il grado di soddisfazione o i punteggi di valutazione della fiducia e della preferenza
- Il livello **numerico** implica che i valori di una variabile rappresentino categorie ordinate con una metrica significativa in modo che i confronti a distanza tra le categorie siano appropriati. Esempi includono l'età in anni e il reddito in migliaia di dollari.

Si supponga, ad esempio, che le variabili *region*, *job* e *age* siano codificate come riportato nella seguente tabella.

Codice di area	Valore della regione	Codice lavoro	Valore del lavoro	Età
1	Nord	1	intern	20
2	Sud	2	REP VENDITE	22
3	Est	3	Manager	25
4	Ovest			27

I valori mostrati rappresentano le categorie di ogni variabile. *Regione* sarebbe una variabile nominale. Ci sono quattro categorie di *regione*, senza ordinamento intrinseco. I valori da 1 a 4 rappresentano semplicemente le quattro categorie; lo schema di codifica è completamente arbitrario. *Lavoro*, invece, potrebbe essere ipotizzata una variabile ordinale. Le categorie originali formano una progressione da intern a manager. Codici più grandi rappresentano un lavoro superiore sulla scala aziendale. Tuttavia, solo le informazioni sull'ordine sono note - non si può dire nulla sulla distanza tra le categorie adiacenti. Al contrario, *età* potrebbe essere ipotizzata una variabile numerica. Nel caso di *età*, le distanze tra i valori sono intrinsecamente significative. La distanza tra 20 e 22 è la stessa della distanza tra 25 e 27, mentre la distanza tra 22 e 25 è maggiore di una di queste.

## Selezione del livello di scaling ottimale

È importante capire che non ci sono proprietà intrinseche di una variabile che predefinisce automaticamente quale livello di scaling ottimale si deve specificare per esso. Puoi esplorare i tuoi dati in qualsiasi modo che abbia un senso e rende più facile l'interpretazione. Analizzando una variabile a livello numerico a livello ordinale, ad esempio, l'utilizzo di una trasformazione non lineare può consentire una soluzione in meno dimensioni.

I seguenti due esempi illustrano come il livello di misurazione "ovvio" potrebbe non essere il livello di scaling ottimale. Supponga che un tipo variabile oggetti in fasce di età. Anche se l'età può essere scaldata come variabile numerica, può essere vero che per le persone più giovani di 25 la sicurezza ha un rapporto positivo con l'età, mentre per le persone più vecchie di 60 la sicurezza ha un rapporto negativo con l'età. In questo caso potrebbe essere meglio trattare l'età come variabile nominale.

Come un altro esempio, una variabile che ordina le persone per preferenza politica appare essenzialmente nominale. Tuttavia, se ordinate le parti da sinistra politica a destra politica, potreste volere la quantificazione delle parti per rispettare questo ordine utilizzando un livello ordinale di analisi.

Anche se non ci sono proprietà predefinite di una variabile che lo rendono esclusivamente un livello o un altro, ci sono alcune linee guida generali per aiutare l'utente novello. Con una quantificazione singola - nominale, di solito non si conosce l'ordine delle categorie ma si vuole che l'analisi ne imponga uno.

Se l'ordine delle categorie è noto, si dovrebbe provare una quantificazione ordinale. Se le categorie sono inordinabili, si potrebbe provare una quantificazione multipla - nominale.

## Grafici di trasformazione

I diversi livelli in cui ogni variabile può essere scalato impongono diverse limitazioni alle quantificazioni. I tracciati di trasformazione illustrano il rapporto tra le quantificazioni e le categorie originali derivanti dal livello di scaling ottimale selezionato. Ad esempio, una trama di trasformazione lineare determina quando una variabile viene trattata come numerica. Variabili trattate come risultato ordinale in una trama di trasformazione non decrescente. I tracciati di trasformazione per le variabili trattate nominalmente che sono a forma di U (o l'inverso) visualizzano una relazione quadratica. Le variabili nominali potrebbero anche generare trame di trasformazione senza evidenti tendenze modificando completamente l'ordine delle categorie. La seguente figura visualizza una trama di trasformazione del campione.

I tracciati di trasformazione sono particolarmente adatti a determinare quanto esegue bene il livello di scaling ottimale selezionato. Se diverse categorie ricevono analoghe quantificazioni, collassare queste categorie in una sola categoria può essere giustificata. In alternativa, se una variabile trattata come nominale riceve quantificazioni che visualizzano una tendenza crescente, una trasformazione ordinale può comportare un adattamento simile. Se quella tendenza è lineare, può essere opportuno un trattamento numerico. Tuttavia, se le categorie collassanti o la modifica dei livelli di scaling è giustificata, l'analisi non cambierà in modo significativo.

## Codici Categoria

Alcune cure dovrebbero essere prese quando si codificano variabili categoriali perché alcuni schemi di codifica possono produrre output indesiderato o analisi incomplete. I possibili schemi di codifica per *lavoro* vengono visualizzati nella seguente tabella.

Categoria	A	B	C	D
intern	1	1	5	1
REP VENDITE	2	2	6	5
Manager	3	7	7	3

Alcune Procedure di Categorie richiedono che venga definita la gamma di ogni variabile utilizzata. Qualsiasi valore esterno a questa gamma viene trattato come un valore mancante. Il valore minimo della categoria è sempre 1. Il valore di categoria massimo è fornito dall'utente. Questo valore non è il *numero* delle categorie per una variabile - è il valore della categoria *più grande*. Ad esempio, nella tabella, lo schema A ha un valore di categoria massimo di 3 e lo schema B ha un valore di categoria massimo di 7, eppure entrambi i regimi codificano le stesse tre categorie.

L'intervallo variabile determina quali categorie saranno omesse dall'analisi. Eventuali categorie con codici esterni all'intervallo definito sono omesse dall'analisi. Questo è un metodo semplice per omettere categorie ma può provocare analisi indesiderate. Una categoria massima definita erroneamente può omettere *valide* categorie dall'analisi. Ad esempio, per il regime B, definire il valore della categoria massima in essere 3 indica che *lavoro* ha categorie codificate da 1 a 3; la categoria *manager* viene trattata come mancante. Poiché in realtà nessuna categoria è stata codificata 3, la terza categoria nell'analisi non contiene casi. Se si volesse omare tutte le categorie manager, questa analisi sarebbe appropriata. Tuttavia, se i gestori devono essere inclusi, la categoria massima deve essere definita come 7, e i valori mancanti devono essere codificati con valori superiori a 7 o inferiori a 1.

Per le variabili trattate come nominali o ordinarie, la gamma delle categorie non influiscono sui risultati. Per le variabili nominali, è importante solo l'etichetta e non il valore associato a tale etichetta. Per le variabili ordinarie, l'ordine delle categorie è conservato nelle quantificazioni; i valori di categoria stessi non sono importanti. Tutti gli schemi di codifica che comportano lo stesso ordinamento di categoria avranno risultati identici. Ad esempio, i primi tre schemi della tabella sono funzionalmente equivalenti se

*lavoro* viene analizzato a livello ordinale. L'ordine delle categorie è identico in questi schemi. Lo schema D, invece, inverte la seconda e la terza categoria e darà risultati diversi rispetto agli altri schemi.

Sebbene molti schemi di codifica per una variabile siano funzionalmente equivalenti, si preferiscono schemi con piccole differenze tra codici perché i codici hanno un impatto sulla quantità di output prodotto da una procedura. Sono valide tutte le categorie codificate con valori compresi tra 1 e il massimo definito dall'utente. Se una qualsiasi di queste categorie è vuota, le relative quantificazioni saranno di sistema - mancante o 0, a seconda della procedura. Sebbene nessuna di queste assegnazioni influenzi le analisi, l'output viene prodotto per queste categorie. Così, per il regime B, *lavoro* ha quattro categorie che ricevono valori mancanti di sistema. Per il regime C ci sono anche quattro categorie che ricevono indicatori mancanti di sistema. Al contrario, per il regime A non ci sono quantificazioni mancanti di sistema. Utilizzare numeri interi consecutivi come codici per le variabili trattate come risultati nominali o ordinali in molto meno output senza influenzare i risultati.

Gli schemi di codifica per le variabili trattate come numeriche sono più ristretti rispetto al caso ordinale. Per queste variabili sono importanti le differenze tra le categorie consecutive. La seguente tabella visualizza tre schemi di codifica per *age*.

Tabella 3. Schemi di codifica alternativi per l'età			
Categoria	A	B	C
20	20	1	1
22	22	3	2
25	25	6	3
27	27	8	4

Qualsiasi ricovero di variabili numeriche deve preservare le differenze tra le categorie. Utilizzare i valori originali è un metodo per garantire la conservazione delle differenze. Tuttavia, questo può comportare in molte categorie avere indicatori mancanti di sistema. Ad esempio, lo schema A impiega i valori osservati originali. Per tutte le procedure di Categorie tranne che per Corrispondenza Analisi, il valore di categoria massimo è di 27 e il valore minimo di categoria è fissato a 1. Le prime 19 categorie sono vuote e ricevono indicatori mancanti di sistema. L'output può diventare rapidamente piuttosto ingombrante se la categoria massima è molto maggiore di 1 e ci sono molte categorie vuote tra 1 e il massimo.

Per ridurre la quantità di output, è possibile effettuare il ricovero. Tuttavia, nel caso numerico, la struttura automatica Recode non deve essere utilizzata. La codifica agli interi consecutivi determina differenze di 1 tra tutte le categorie consecutive e, di conseguenza, tutte le quantificazioni saranno ugualmente distanziate. Le caratteristiche metriche ritenute importanti quando si trattano una variabile come numerica vengono distrutte recuperando numeri interi consecutivi. Ad esempio, lo schema C nella tabella corrisponde al ricovero automatico *age*. La differenza tra le categorie 22 e 25 è cambiata da tre a una e le quantificazioni rispecchieranno quest' ultima differenza.

Uno schema di ricovero alternativo che preserva le differenze tra le categorie è quello di sottrarre il valore di categoria più piccolo da ogni categoria e aggiungere 1 ad ogni differenza. Lo schema B deriva da questa trasformazione. Il valore di categoria più piccolo, 20, è stato sottratto da ogni categoria e 1 è stato aggiunto ad ogni risultato. I codici trasformati hanno un minimo di 1 e tutte le differenze sono identiche ai dati originali. Il valore massimo della categoria è ora 8, e le quantificazioni zero prima della prima quantificazione non zero sono tutte eliminate. Eppure, le quantificazioni non zero corrispondenti a ciascuna categoria derivanti dal regime B sono identiche alle quantificazioni dal regime A.

## Quale Procedura è Best for Your Application?

Le tecniche incarnite in quattro di queste procedure (Analisi della corrispondenza, analisi di corrispondenza multipla, Analisi dei componenti principali categoriali e Analisi Nonlineare Canonica Correlazione) rientrano nell'area generale di analisi dei dati multivariati nota come **riduzione della quota**. Cioè le relazioni tra variabili sono rappresentate in poche dimensioni - dicono due o tre - il più spesso possibile. Questo consente di descrivere strutture o modelli nelle relazioni che sarebbero troppo difficili da sfatare nella loro originale ricchezza e complessità. Nelle applicazioni di ricerca di mercato queste

tecniche possono essere una forma di **mappatura percettiva**. Un grande vantaggio di queste procedure è che essi allungano i dati con diversi livelli di scaling ottimale.

Categorical Regression descrive la relazione tra una variabile di risposta categoriale e una combinazione di variabili predittive categoriali. L'influenza di ogni variabile predittiva sulla variabile di risposta è descritta dal corrispondente peso di regressione. Come nelle altre procedure, i dati possono essere analizzati con diversi livelli di scaling ottimale.

Multidimensionale Scaling e Multidimensionale Unfolding descrivono le relazioni tra gli oggetti in uno spazio basso-dimensionale, utilizzando le prossimità tra gli oggetti.

Di seguito le brevi linee guida per ciascuna delle procedure:

- Utilizzare Regressione Catoriale per prevedere i valori di una variabile dipendente categoriale da una combinazione di variabili indipendenti categoriali.
- Utilizzare Categorical Principal Components Analysis per rendere conto dei modelli di variazione in un unico insieme di variabili di livelli di scaling ottimali misti.
- Utilizzare Analisi di correlazione canonica non lineare per valutare l'estensione a cui sono correlati due o più insiemi di variabili di livelli di scaling misti ottimali.
- Utilizzare Analisi Corrispondenza per analizzare tabelle di contingenza a due vie o dati che possono essere espressi come tabella a due vie, come ad esempio la preferenza del marchio o i dati di scelta sociometrica.
- Utilizzare Più Analisi Corrispondenza per analizzare una matrice dati multivariata categoriale quando si è disposti a non assumere supponenza che tutte le variabili siano analizzate a livello nominale.
- Utilizza Scaling Multidimensionale per analizzare i dati di prossimità per trovare una rappresentazione di almeno quadrati di una singola serie di oggetti in uno spazio basso-dimensionale.
- Utilizza Unfolding Multidimensionale per analizzare i dati di prossimità per trovare una rappresentazione di almeno quadrati di due insiemi di oggetti in uno spazio basso-dimensionale.

## Regressione categoriale

L'utilizzo di Regression categoriale è più appropriato quando l'obiettivo della tua analisi è quello di prevedere una variabile dipendente (risposta) da una serie di variabili indipendenti (predittore). Come per tutte le procedure di scaling ottimali, i valori di scala sono assegnati a ciascuna categoria di ogni variabile tale che questi valori siano ottimali rispetto alla regressione. La soluzione di una regressione categoriale massimizza la correlazione quadrata tra la risposta trasformata e la combinazione ponderata dei predittori trasformati.

**Relazione ad altre procedure di categorie.** La regressione categoriale con scaling ottimale è paragonabile ad un'ottimale analisi di correlazione canonica con due set, uno dei quali contiene solo la variabile dipendente. In quest' ultima tecnica, la somiglianza dei set deriva confrontando ogni insieme ad una variabile sconosciuta che si trova da qualche parte tra tutti i set. Nella regressione categoriale, si valuta direttamente la somiglianza della risposta trasformata e la combinazione lineare di predittori trasformati.

**Relazione alle tecniche standard.** Nella regressione lineare standard, le variabili categoriali possono essere recuperate come variabili indicatore o essere trattate nella stessa moda delle variabili di livello dell'intervallo. Nel primo approccio, il modello contiene un'intercettazione e una pendenza separate per ogni combinazione dei livelli delle variabili categoriali. Ne deriva un gran numero di parametri da interpretare. Nel secondo approccio viene stimato un solo parametro per ogni variabile. Tuttavia, la natura arbitraria delle codifiche di categoria rende impossibile le generalizzazioni.

Se alcune variabili non sono continue, sono disponibili analisi alternative. Se la risposta è continua e i predittori sono categoriali, l'analisi della varianza è spesso impiegata. Se la risposta è categoriale e i predittori sono continui, può essere appropriata la regressione logistica o l'analisi discriminante. Se la risposta e i predittori sono entrambi categoriali, spesso vengono utilizzati modelli loglineari.

La regressione con scaling ottimale offre tre livelli di scaling per ogni variabile. Le combinazioni di questi livelli possono rendere conto di una vasta gamma di relazioni nonlineari per le quali qualsiasi metodo

singolo "standard" è mal adatto. Di conseguenza, scaling ottimale offre una maggiore flessibilità rispetto agli approcci standard con una complessità aggiunta minimale.

Inoltre, le trasformazioni non lineari dei predittori di solito riducono le dipendenze tra i predittori. Se si confronta gli autovalori della matrice di correlazione per i predittori con gli autovalori della matrice di correlazione per i predittori ottimamente scalati, quest' ultimo set sarà di solito meno variabile rispetto all'ex. In altre parole, nella regressione categoriale, scaling ottimale rende più piccoli gli autovalori più grandi della matrice di correlazione predittiva e i più piccoli autovalori.

## **Analisi Componenti Principali Categoriali**

L'utilizzo di Categorie Principal Components Analysis è più appropriato quando si desidera rendere conto dei modelli di variazione in un unico insieme di variabili di livelli di scaling ottimali misti. Questa tecnica tenta di ridurre la dimensionalità di una serie di variabili pur rappresentando quanto più possibile la variazione. I valori di scala sono assegnati a ciascuna categoria di ogni variabile in modo che questi valori siano ottimali rispetto alla soluzione dei componenti principali. Gli oggetti nell'analisi ricevono i punteggi dei componenti in base ai dati quantificati. I traccianti dei punteggi dei componenti rivelano modelli tra gli oggetti nell'analisi e possono rivelare oggetti insoliti nei dati. La soluzione di un'analisi dei componenti principal categoriali massimizza le correlazioni dei punteggi oggetto con ciascuna delle variabili quantificate per il numero di componenti (dimensioni) specificato.

Un'importante applicazione di componenti principali categoriali è quella di esaminare i dati di preferenza, in cui gli intervistati classificano o valutano una serie di articoli rispetto alle preferenze. Nella consueta IBM SPSS Statistics configurazione dei dati, le righe sono individui, le colonne sono misurazioni per gli articoli e i punteggi su righe sono punteggi di preferenza (su una scala da 0 a 10, ad esempio), rendendo la riga di dati - condizionale. Per i dati di preferenza, si può voler trattare gli individui come variabili. Utilizzando la procedura Transpose è possibile recepire i dati. I rateri diventano le variabili e tutte le variabili sono dichiarate ordinali. Non c'è obiezione a usare più variabili di oggetti in CATPCA.

**Relazione ad altre procedure di categorie.** Se tutte le variabili sono dichiarate più nominali, l'analisi dei componenti principali categoriali produce un'analisi equivalente ad un'analisi di corrispondenza multipla eseguita sulle stesse variabili. Così, l'analisi dei componenti principali categoriali può essere considerata un tipo di analisi di corrispondenza multipla in cui alcune variabili sono dichiarate ordinali o numeriche.

**Relazione alle tecniche standard.** Se tutte le variabili sono scalate sul livello numerico, l'analisi dei componenti principali categoriali è equivalente all'analisi dei componenti principali standard.

Più in generale, l'analisi dei componenti principali categoriali è un'alternativa per calcolare le correlazioni tra scale non numeriche e analizzarle utilizzando un approccio principal standard o un approccio di analisi dei fattori. L'uso ingenuo del consueto coefficiente di correlazione di Pearson come misura di associazione per i dati ordinali può portare a bias non banali in stima delle correlazioni.

## **Analisi della correlazione canonica non lineare (OVERALS)**

Nonlinear Canonical Correlazione Analysis è una procedura molto generale con molte applicazioni diverse. L'obiettivo dell'analisi di correlazione canonica non lineare è quello di analizzare le relazioni tra due o più insiemi di variabili invece che tra le variabili stesse, come nelle principali analisi componenti. Ad esempio, si possono avere due serie di variabili, in cui una serie di variabili potrebbe essere elementi di sfondo demografico su una serie di intervistati e una seconda serie potrebbe essere la risposta ad una serie di elementi di assetto. I livelli di scaling nell'analisi possono essere qualsiasi mix di nominale, ordinale e numerico. L'analisi di correlazione canonica a scaling ottimale determina la somiglianza tra i set confrontando contemporaneamente le variabili canoniche da ciascuna serie a una serie di compromessi assegnati agli oggetti.

**Relazione ad altre procedure di categorie.** Se ci sono due o più serie di variabili con una sola variabile per set, l'analisi di correlazione canonica di scaling ottimale equivale a un'analisi dei componenti principali di scaling ottimale. Se tutte le variabili in un'analisi one - variabile - per - set sono molteplici, l'analisi di correlazione canonica a scala ottimale è equivalente a più analisi di corrispondenza. Se ci sono due insiemi di variabili, una delle quali contiene una sola variabile, l'analisi della correlazione canonica a scaling ottimale equivale a regressione categoriale con scaling ottimale.

**Relazione alle tecniche standard.** L'analisi di correlazione canonica standard è una tecnica statistica che trova una combinazione lineare di una serie di variabili e una combinazione lineare di una seconda serie di variabili che sono massimamente correlate. Dato questo insieme di combinazioni lineari, l'analisi di correlazione canonica può trovare successive serie indipendenti di combinazioni lineari, indicate come variabili canoniche, fino ad un numero massimo pari al numero di variabili nel set più piccolo.

Se ci sono due insiemi di variabili nell'analisi e tutte le variabili sono definite per essere numeriche, l'analisi di correlazione canonica a scala ottimale è equivalente a un'analisi di correlazione canonica standard. Sebbene IBM SPSS Statistics non abbia una procedura di analisi di correlazione canonica, molte delle statistiche rilevanti possono essere ottenute da analisi multivariate di varianza.

L'analisi di correlazione canonica a scaling ottimale ha varie altre applicazioni. Se si dispone di due insiemi di variabili e uno dei set contiene una variabile nominale dichiarata come singola nominale, i risultati di analisi della correlazione canonica a scaling ottimale possono essere interpretati in modo simile all'analisi di regressione. Se si considera la variabile per essere più nominale, l'analisi di scaling ottimale è un'alternativa all'analisi discriminante. Raggruppare le variabili in più di due serie fornisce una varietà di modi per analizzare i tuoi dati.

## Analisi delle corrispondenze

L'obiettivo dell'analisi della corrispondenza è quello di rendere i bigrafici per le tabelle di corrispondenza. In una tabella di corrispondenza si ipotizza che le variabili di riga e colonne rappresentino categorie non ordinate; pertanto, viene sempre utilizzato il livello di scaling ottimale nominale. Entrambe le variabili sono ispezionate solo per le loro informazioni nominali. Cioè l'unica considerazione è il fatto che alcuni oggetti sono nella stessa categoria mentre altri no. Non si assume nulla circa la distanza o l'ordine tra le categorie della stessa variabile.

Un utilizzo specifico dell'analisi della corrispondenza è l'analisi delle tabelle di contingenza a due vie. Se una tabella dispone di righe attive  $r$  e di colonne attive  $c$ , il numero di dimensioni nella soluzione di analisi della corrispondenza è il minimo di  $r$  meno 1 o  $c$  meno 1, qualunque sia meno. In altre parole, si potrebbero rappresentare perfettamente le categorie di righe o le categorie di colonne di una tabella di contingenza in uno spazio di dimensioni. Praticamente parlando, però, si vorrebbe rappresentare le categorie di fila e colonne di un tavolo a due vie in uno spazio basso-dimensionale, diciamo due dimensioni, per il motivo che i tracciati bidimensionali sono più facilmente comprensibili delle rappresentazioni spaziali multidimensionali.

Quando si utilizzano meno del numero massimo di dimensioni possibili, le statistiche prodotte nell'analisi descrivono come le categorie di righe e colonne siano rappresentate nella rappresentazione a basso dimensionale. A condizione che la qualità di rappresentazione della soluzione bidimensionale sia buona, è possibile esaminare trame dei punti di riga e dei punti di colonna per apprendere quali categorie della variabile di riga sono simili, quali categorie della variabile colonna sono simili e quali categorie di righe e colonne sono simili tra loro.

**Relazione ad altre procedure di categorie.** L'analisi della corrispondenza semplice è limitata alle tabelle a due vie. Se ci sono più di due variabili di interesse, è possibile combinare le variabili per creare variabili di interazione. Ad esempio, per le variabili *regione*, *lavoro* e *età*, è possibile combinare *regione* e *lavoro* per creare una nuova variabile *rilavoro* con le 12 categorie mostrate nella seguente tabella. Questa nuova variabile forma una tabella a due vie con *age* (12 righe, 4 colonne), che può essere analizzato nell'analisi della corrispondenza.

Tabella 4. Combinazioni di regione e lavoro			
Codice categoria	Definizione categoria	Codice categoria	Definizione categoria
1	Nord, intern	7	Est, intern
2	Nord, rep di vendita	8	Est, rep di vendita
3	Nord, gestore	9	Est, gestore
4	Sud, intern	10	Ovest, intern
5	Sud, rep di vendita	11	West, rep di vendita

Tabella 4. Combinazioni di regione e lavoro (Continua)

Codice categoria	Definizione categoria	Codice categoria	Definizione categoria
6	Sud, gestore	12	West, manager

Un accorciamento di questo approccio è che qualsiasi coppia di variabili può essere combinata. Possiamo combinare *lavoro* e *età*, ottenendo un'altra variabile di 12 categoria. Oppure possiamo combinare *regione* e *età*, che si traduce in una nuova variabile a 16 categoria. Ognuna di queste variabili di interazione forma una tabella a due vie con la variabile rimanente. Le analisi di corrispondenza di queste tre tabelle non daranno risultati identici, eppure ognuno è un approccio valido. Inoltre, se ci sono quattro o più variabili, si possono costruire tavole a due vie che confrontino una variabile interazione con un'altra variabile di interazione. Il numero di tabelle possibili da analizzare può diventare abbastanza grande, anche per alcune variabili. È possibile selezionare una di queste tabelle da analizzare, oppure è possibile analizzarle tutte. In alternativa, la procedura di Analisi Corrispondenza Multipla può essere utilizzata per esaminare tutte le variabili contemporaneamente senza la necessità di costruire variabili di interazione.

**Relazione alle tecniche standard.** La procedura di Crosstabs può essere utilizzata anche per analizzare le tabelle di contingenza, con l'indipendenza come focus comune nelle analisi. Tuttavia, anche in piccoli tavoli, rilevare la causa delle partenze dall'indipendenza può essere difficile. L'utilità dell'analisi della corrispondenza risiede nella visualizzazione di tali modelli per tabelle a due vie di qualsiasi dimensione. Se c'è un'associazione tra le variabili di fila e colonne - ovvero se il valore chi - quadrato è significativo - l'analisi della corrispondenza può aiutare a rivelare la natura del rapporto.

## Analisi delle corrispondenze multiple

Più Correspondence Analysis cerca di produrre una soluzione in cui gli oggetti all'interno della stessa categoria sono tracciati vicini e gli oggetti in diverse categorie sono tracciati molto distanti. Ogni oggetto è il più vicino possibile ai punti di categoria delle categorie che si applicano all'oggetto. In questo modo le categorie suddividono gli oggetti in sottogruppi omogenei. Le variabili sono considerate omogenee quando classificano gli oggetti nelle stesse categorie negli stessi sottogruppi.

Per una soluzione unidimensionale, l'analisi di corrispondenza multipla assegna valori di scala ottimali (quantificazioni di categoria) a ciascuna categoria di ciascuna variabile in modo che complessivamente, in media, le categorie abbiano la massima diffusione. Per una soluzione bidimensionale, l'analisi di corrispondenza multipla trova una seconda serie di quantificazioni delle categorie di ogni variabile non correlate al primo set, tentando nuovamente di massimizzare la diffusione, e così via. Poiché le categorie di una variabile ricevono altrettanti scorci come ci sono dimensioni, le variabili nell'analisi si ipotizza essere più nominali in livello di scaling ottimale.

L'analisi di corrispondenza multipla assegna anche i punteggi agli oggetti nell'analisi in modo che le quantificazioni di categoria siano le medie, o i centroidi, dei punteggi oggetto degli oggetti di quella categoria.

**Relazione ad altre procedure di categorie.** Più analisi di corrispondenza è nota anche come analisi omogenea o dual scaling. Dà risultati comparabili, ma non identici, all'analisi della corrispondenza quando ci sono solo due variabili. L'analisi della corrispondenza produce un output unico che sintetizza l'adattamento e la qualità della rappresentazione della soluzione, comprese le informazioni di stabilità. Così, l'analisi della corrispondenza è di solito preferibile a più analisi di corrispondenza nel caso di due variabili. Un'altra differenza tra le due procedure è che l'input a più analisi di corrispondenza è una matrice dati, dove le righe sono oggetti e le colonne sono variabili, mentre l'input per l'analisi della corrispondenza può essere la stessa matrice dati, una matrice di prossimità generale o una tabella di contingenza congiunta, ovvero una matrice aggregata in cui sia le righe che le colonne rappresentano categorie di variabili. L'analisi di corrispondenza multipla può anche essere pensata come analisi componenti principali dei dati scalati al livello nominale multiplo.

**Relazione alle tecniche standard.** Un'analisi di corrispondenza multipla può essere pensata come l'analisi di un tavolo di contingenza multistrada. Le tabelle di contingenza multimodo possono essere analizzate anche con la procedura Crosstabs, ma Crosstabs fornisce statistiche di riepilogo separate per ogni categoria di ogni variabile di controllo. Con più analisi di corrispondenza è spesso possibile sintetizzare il rapporto tra tutte le variabili con una singola trama bidimensionale. Un utilizzo avanzato

di più analisi di corrispondenza è quello di sostituire i valori di categoria originali con i valori di scala ottimali dalla prima dimensione ed eseguire un'analisi multivariata secondaria. Dal momento che l'analisi di corrispondenza multipla sostituisce le etichette di categoria con valori di scala numerica, molte diverse procedure che richiedono dati numerici possono essere applicate dopo l'analisi di corrispondenza multipla. Ad esempio, la procedura di analisi di Fattore produce una prima componente principale che equivale alla prima dimensione dell'analisi di corrispondenza multipla. I punteggi dei componenti nella prima dimensione sono uguali ai punteggi dell'oggetto e i loadini dei componenti al quadrato sono pari alle misure di discriminazione. La seconda dimensione di analisi della corrispondenza multipla, tuttavia, non è pari alla seconda dimensione dell'analisi dei fattori.

## Scaling multidimensionale

L'utilizzo di Scaling Multidimensionale è più appropriato quando l'obiettivo della tua analisi è quello di trovare la struttura in una serie di misure a distanza tra una singola serie di oggetti o casi. Ciò si compie assegnando le osservazioni a specifiche località in uno spazio concettuale basso-dimensionale in modo che le distanze tra i punti nello spazio corrispondono al dato (dis) analogie il più possibile. Il risultato è una rappresentazione meno quadrante degli oggetti in quello spazio basso-dimensionale che, in molti casi, ti aiuterà a comprendere ulteriormente i tuoi dati.

**Relazione ad altre procedure di categorie.** Quando si hanno dati multivariati da cui si creano le distanze e che poi si analizzano con scaling multidimensionale, i risultati sono simili all'analisi dei dati che utilizzano le analisi dei componenti principali categoriali con la normalizzazione del principal object. Questo tipo di PCA è anche noto come analisi delle coordinate principali.

**Relazione alle tecniche standard.** La Procedura di scaling multidimensionale (PROXSCAL) offre diversi miglioramenti alla procedura di scaling disponibile nell'opzione Statistiche Base (ALSCAL). PROXSCAL offre un algoritmo accelerato per alcuni modelli e consente di inserire limitazioni allo spazio comune. Inoltre, PROXSCAL tenta di minimizzare lo stress grezzo normalizzato piuttosto che S - stress (indicato anche come **ceppo**). Lo stress grezzo normalizzato è generalmente preferito perché è una misura basata sulle distanze, mentre lo S - stress si basa sulle distanze quadrati.

## Unfolding multidimensionale

L'utilizzo di Unfolding Multidimensionale è più appropriato quando l'obiettivo della tua analisi è quello di trovare la struttura in una serie di misure a distanza tra due insiemi di oggetti (indicati come oggetti di fila e colonne). Ciò si compie assegnando le osservazioni a specifiche località in uno spazio concettuale basso-dimensionale in modo che le distanze tra i punti nello spazio corrispondono al dato (dis) analogie il più possibile. Il risultato è una rappresentazione meno quadrante degli oggetti di fila e di colonna in quello spazio basso-dimensionale che, in molti casi, ti aiuterà a comprendere ulteriormente i tuoi dati.

**Relazione ad altre procedure di categorie.** Se i tuoi dati sono costituiti da distanze tra una singola serie di oggetti (una matrice quadrata, simmetrica), utilizzare Scaling Multidimensionale.

**Relazione alle tecniche standard.** La Procedura di sfolding multidimensionale (PREFSCAL) offre diversi miglioramenti alla funzionalità di sfolding disponibile nell'opzione Statistiche Base (tramite ALSCAL). PREFSCAL consente di inserire limitazioni allo spazio comune; inoltre, PREFSCAL tenta di minimizzare una misura di stress penalizzata che lo aiuta ad evitare le soluzioni degenerate (a cui gli algoritmi più anziani sono inclini).

## Aspetto Rapporto in Diagrammi Di Scaling Ottimali

Rapporto di aspetto in lotti di scaling ottimali è isotropo. In una trama bidimensionale, la distanza che rappresenta una unità nella dimensione 1 è pari alla distanza che rappresenta una unità nella dimensione 2. Se si modifica la gamma di una dimensione in una trama bidimensionale, il sistema cambia la dimensione dell'altra dimensione per mantenere uguali le distanze fisiche. Il rapporto di aspetto isotropo non può essere sovrascritto per le procedure di scaling ottimali.

## Regressione Categoriale (CATREG)

---

**La regressione categoriale** quantifica i dati categoriali assegnando valori numerici alle categorie, determinando un'equazione di regressione lineare ottimale per le variabili trasformate. La regressione categoriale è nota anche con l'acronimo CATREG, per la regressione *gattoegizia rig*.

L'analisi di regressione lineare standard comporta la minimizzazione della somma delle differenze quadrati tra una variabile di risposta (dipendente) e una combinazione ponderata di variabili predittiva (indipendente). Le variabili sono tipicamente quantitative, con dati categoriali (nominali) recuperabili a variabili binarie o di contrasto. Di conseguenza, le variabili categoriali servono a separare gruppi di casi, e la tecnica stima serie separate di parametri per ciascun gruppo. I coefficienti stimati riflettono il modo in cui le modifiche dei predittori influiscono sulla risposta. La previsione della risposta è possibile per qualsiasi combinazione di valori predittori.

Un approccio alternativo comporta la regressione della risposta sui valori predittivi categoriali stessi. Di conseguenza, si stima un coefficiente per ogni variabile. Tuttavia, per le variabili categoriali, i valori di categoria sono arbitrari. Codificando le categorie in modi diversi si rendono diversi i coefficienti, rendendo difficili i confronti tra le analisi delle stesse variabili.

CATREG estende l'approccio standard scalando contemporaneamente variabili nominali, ordinali e numeriche. La procedura quantifica le variabili categoriali in modo che le quantificazioni riflettano caratteristiche delle categorie originali. La procedura tratta le variabili categoriali quantificate allo stesso modo delle variabili numeriche. Utilizzando trasformazioni non lineari permettono di analizzare le variabili ad una varietà di livelli per trovare il modello di adattamento migliore.

**Esempio.** La regressione categoriale potrebbe essere utilizzata per descrivere come la soddisfazione del lavoro dipende dalla categoria lavoro, dalla regione geografica e dalla quantità di viaggio. Potreste scoprire che alti livelli di soddisfazione corrispondono ai manager e ai viaggi bassi. L'equazione di regressione risultante potrebbe essere utilizzata per prevedere la soddisfazione del lavoro per qualsiasi combinazione delle tre variabili indipendenti.

**Statistiche e grafici.** Frequenze, coefficienti di regressione, tabella ANOVA, storia dell'iterazione, quantificazioni di categoria, correlazioni tra predittori non trasformati, correlazioni tra predittori trasformati, tracciati residui e complotti di trasformazione.

Considerazioni Sui Dati di regressione categoriale

**Dati.** CATREG opera sulle variabili indicatore di categoria. Gli indicatori di categoria dovrebbero essere interi positivi. È possibile utilizzare la finestra di dialogo Discretization per convertire variabili di valore frazionale e variabili stringa in numeri interi positivi.

**Ipotesi.** È consentita una sola variabile di risposta, ma il numero massimo di variabili predittore è di 200. I dati devono contenere almeno tre casi validi e il numero di casi validi deve superare il numero di variabili predittore più uno.

**Procedure correlate.** CATREG equivale ad analisi di correlazione canonica categoriale con scaling ottimale (OVERALS) con due insiemi, uno dei quali contiene una sola variabile. Scalare tutte le variabili a livello numerico corrisponde all'analisi di regressione multipla standard.

Per Ottenere una Regressione categoriale

1. Dai menu, scegliere:

**Analizzare > Regressione > Scaling Ottimale (CATREG) ...**

2. Selezionare la variabile dipendente e indipendente variabile.

3. Fare clic su **OK**.

Facoltativamente, modificare il livello di scala per ogni variabile.

## Definire Scale in Regressione categoriale

È possibile impostare il livello di scaling ottimale per le variabili dipendenti e indipendenti. Per impostazione predefinita, vengono scalati come sprighe monotoniche di secondo grado (ordinale) con due nodi interni. Inoltre è possibile impostare il peso per le variabili di analisi.

**Livello di scaling ottimale.** È inoltre possibile selezionare il livello di scala per quantificare ogni variabile.

- **Spline Ordinale.** L'ordine delle categorie della variabile osservata viene conservato nella variabile scalata in modo ottimale. I punti di categoria saranno su una retta (vettore) attraverso l'origine. La trasformazione risultante è un polinomio monotono liscio del grado prescelto. I pezzi sono specificati dal numero specificato dall'utente e dal posizionamento determinato dalla procedura dei nodi interni.
- **Spline Nominale.** Le uniche informazioni nella variabile osservata che si conservano nella variabile ottimamente scalata è il raggruppamento di oggetti in categorie. L'ordine delle categorie della variabile osservata non è conservato. I punti di categoria saranno su una retta (vettore) attraverso l'origine. La trasformazione risultante è un polinomio liscio, possibilmente nonmonotonico, del grado prescelto. I pezzi sono specificati dal numero specificato dall'utente e dal posizionamento determinato dalla procedura dei nodi interni.
- **Ordinale.** L'ordine delle categorie della variabile osservata viene conservato nella variabile scalata in modo ottimale. I punti di categoria saranno su una retta (vettore) attraverso l'origine. La trasformazione risultante si adatta meglio alla trasformazione ordinale spline ma è meno liscia.
- **Nominale.** Le uniche informazioni nella variabile osservata che si conservano nella variabile ottimamente scalata è il raggruppamento di oggetti in categorie. L'ordine delle categorie della variabile osservata non è conservato. I punti di categoria saranno su una retta (vettore) attraverso l'origine. La trasformazione risultante si adatta meglio alla trasformazione nominale spline ma è meno liscia.
- **Numerico.** Le categorie sono trattate come ordinate e ugualmente distanziate (livello di intervallo). L'ordine delle categorie e le distanze uguali tra i numeri di categoria della variabile osservata sono conservati nella variabile ottimamente scalata. I punti di categoria saranno su una retta (vettore) attraverso l'origine. Quando tutte le variabili sono a livello numerico, l'analisi è analoga all'analisi dei componenti principali standard.

## Categorical Regression Discretization

Dialogo Discretization consente di selezionare un metodo di ricovero delle variabili. Le variabili a valore frazionale sono raggruppate in sette categorie (o nel numero di valori distinti della variabile se questo numero è inferiore a sette) con una distribuzione approssimativamente normale, salvo diversa indicazione. Le variabili stringa vengono sempre convertite in numeri interi positivi assegnando gli indicatori di categoria in base all'ordine alfanumerico crescente. La discretizzazione per le variabili di stringa si applica a questi interi. Altre variabili sono lasciate sole per impostazione predefinita. Le variabili discretizzate vengono poi utilizzate nell'analisi.

**Metodo.** Scegliere tra raggruppamento, graduatoria e moltiplicarsi.

- **Raggruppamento.** Recupera in un numero specificato di categorie o recupera per intervallo.
- **Classificazione.** La variabile è discretizzata ordinando i casi.
- **Moltiplicando.** I valori correnti della variabile sono standardizzati, moltiplicati per 10, arrotondati e hanno una costante aggiunta in modo che il valore più basso discretizzato sia di 1.

**Raggruppamento.** Le seguenti opzioni sono disponibili quando le variabili discretizing si raggruppano:

- **Numero di categorie.** Specificare una serie di categorie e se i valori della variabile devono seguire una distribuzione approssimativamente normale o uniforme in quelle categorie.
- **Gli intervalli di uguale.** Le variabili sono recuperate in categorie definite da questi intervalli ugualmente dimensionati. È necessario specificare la lunghezza degli intervalli.

## Valori Mancanti di regressione categoriale

La finestra di dialogo Valori mancanti consente di scegliere la strategia per la gestione dei valori mancanti nelle variabili di analisi e le variabili supplementari.

**Strategia.** Scegli di escludere gli oggetti con valori mancanti (cancellazione listwise) o di imputare i valori mancanti (trattamento attivo).

- **Escludere gli oggetti con valori mancanti su questa variabile.** Gli oggetti con valori mancanti sulla variabile selezionata sono esclusi dall'analisi. Questa strategia non è disponibile per le variabili supplementari.
- **Assegna i valori mancanti.** Gli oggetti con valori mancanti sulla variabile selezionata hanno quei valori imputati. È possibile scegliere il metodo di imputazione. Selezionare **Modalità** per sostituire i valori mancanti con la categoria più frequente. Quando ci sono più modalità, viene utilizzato quello con l'indicatore di categoria più piccolo. Selezionare **Extra category** per sostituire i valori mancanti con la stessa quantificazione di una categoria extra. Ciò implica che gli oggetti con un valore mancante su questa variabile sono considerati appartenenti alla stessa categoria (extra).

## Opzioni Di Regressione Categoriale

La finestra di dialogo Opzioni consente di selezionare lo stile di configurazione iniziale, specificare l'iterazione e i criteri di convergenza, selezionare gli oggetti supplementari e impostare l'etichettatura dei tracciati.

**Oggetti integrali.** Questo consente di specificare gli oggetti che si desidera trattare come supplenti. Semplicemente digitare il numero di un oggetto supplementare (o specificare una gamma di casi) e fare clic su **Aggiungi**. Non si possono ponderare gli oggetti integrali (i pesi specificati vengono ignorati).

**Configurazione iniziale.** Se nessuna variabile viene trattata come nominale, selezionare la configurazione **Numerico**. Se almeno una variabile viene trattata come nominale, selezionare la configurazione **Random**.

In alternativa, se almeno una variabile ha un livello di scala ordinale o spline, il consueto algoritmo di adattamento modello può determinare una soluzione subottimale. La scelta di **Più sistemazioni sistematiche** con tutti i possibili modelli di segno da testare troverà sempre la soluzione ottimale, ma il tempo di elaborazione necessario aumenta rapidamente come numero di variabili ordinali e spline nell'aumento del dataset. È possibile ridurre il numero di modelli di prova specificando una percentuale di perdita di soglia di varianza, dove sarà esclusa la soglia superiore, i modelli più di segno. Con questa opzione, ottenere la soluzione ottimale non è garantita, ma diminuisce la possibilità di ottenere una soluzione subottimale. Inoltre, se non si trova la soluzione ottimale, la possibilità che la soluzione subottimale sia molto diversa dalla soluzione ottimale è diminuita. Quando vengono richieste più sistemazioni sistematiche, i segni dei coefficienti di regressione per ogni inizio vengono scritti in un file di dati IBM SPSS Statistics esterno o in un dataset nella sessione corrente. Per ulteriori informazioni, consultare la sezione [“Regressione Categoriale Salva”](#) a pagina 14.

I risultati di una precedente esecuzione con più iniziati sistematici consentono di **Utilizza segni fissi per i coefficienti di regressione**. I segni (indicati da 1 e -1) devono essere in una riga del dataset o file specificato. Il numero di partenza a valore intero è il numero di caso della riga in questo file che contiene i segni da utilizzare.

**Criteri.** È possibile specificare il numero massimo di iterazioni che la regressione può percorrere nei suoi calcoli. È inoltre possibile selezionare un valore di criterio di convergenza. La regressione interrompe l'iterazione se la differenza di adattamento totale tra le ultime due iterazioni è inferiore al valore di convergenza o se si raggiunge il numero massimo di iterazioni.

**Etichetta Tracciati Per.** Consente di specificare se le variabili e le etichette di valore o i nomi e i valori variabili verranno utilizzati nei tracciati. È inoltre possibile specificare una lunghezza massima per le etichette.

## Regolarizzazione Regressione Catoriale

**Metodo.** I metodi di regolarizzazione possono migliorare l'errore predittivo del modello riducendo la variabilità nelle stime del coefficiente di regressione restringendo le stime verso il 0. Il Lasso e Elastic Net ridurranno alcune stime di coefficiente a esattamente 0, fornendo così una forma di selezione variabile. Quando viene richiesto un metodo di regolarizzazione, il modello regolarizzato e i coefficienti per ogni valore del coefficiente di penalità sono scritti in un file di dati IBM SPSS Statistics esterno o dataset nella sessione corrente. Per ulteriori informazioni, consultare la sezione [“Regressione Catoriale Salva”](#) a pagina 14.

- **Regressione Ridge** Ridge regressione restringe i coefficienti introducendo un termine di penalità pari alla somma dei coefficienti quadrati volte a un **coefficiente di penalità**. Questo coefficiente può spaziare da 0 (nessuna penalità) a 1; la procedura cercherà il valore "migliore" della penalità se si specifica un intervallo e un incremento.
- **Lazo.** Il Termine di pena del lazo si basa sulla somma dei coefficienti assoluti, e la specifica di un coefficiente di penalità è simile a quella della regressione di Ridge; tuttavia, il Lasso è più intensivo computazionale.
- **net Elastic.** L'Elastic Net combina semplicemente le pene di regressione di Lasso e Ridge e cercherà la griglia di valori specificati per trovare i coefficienti di penalità di regressione "best" Lasso e Ridge. Per un determinato paio di pene di regressione di Lasso e Ridge, l'Elastic Net non è molto più computazionalmente costoso del Lasso.

**Visualizza i grafici di regolarizzazione.** Si tratta di appezzamenti dei coefficienti di regressione contro la penalità di regolarizzazione. Quando si ricerca una gamma di valori per il coefficiente di penalità "migliore", fornisce una visione di come i coefficienti di regressione cambiano su quella gamma.

**Plotti Elastici Net.** Per il metodo Elastic Net, i tracciati di regolarizzazione separati sono prodotti da valori della penalità di regressione di Ridge. **Tutti i tracciati possibili** utilizzano ogni valore nella gamma determinato dai valori minimi e massimi di penalità di regressione Ridge specificati. **Per alcune sanzioni Ridge** consente di specificare un sottoinsieme dei valori nella gamma determinati dal minimo e massimo. Basta digitare il numero di un valore di penalità (o specificare una serie di valori) e fare clic su **Aggiungi**.

## Output Regression Catoriale

La finestra di dialogo Output consente di selezionare le statistiche da visualizzare nell'output.

**Tabelle.** Consente di creare tabelle per:

- **Più R.** Include  $R^2$ , regolato  $R^2$ , e regolato  $R^2$  prendendo in considerazione il scaling ottimale.
- **ANOVA.** Questa opzione include la regressione e le somme residue di quadrati, quadrati medi e  $F$ . Vengono visualizzate due tabelle ANOVA: una con gradi di libertà per la regressione uguale al numero di variabili predittiva e una con gradi di libertà per la regressione che prende in considerazione il scaling ottimale.
- **Coefficienti.** Questa opzione dà tre tabelle: una tabella Coefficienti che include beta, errore standard dei beta, valori  $t$  e significato; un tavolo Coefficienti - Ottimale Scaling con l'errore standard delle scommesse che prendono in considerazione i gradi di scaling ottimali della libertà; e una tabella con la correlazione zero - ordine, parte e parziale, la misura di importanza relativa di Pratt per i predittori trasformati, e la tolleranza prima e dopo la trasformazione.
- **Cronologia iterazioni.** Per ogni iterazione, inclusi i valori di partenza per l'algoritmo, vengono mostrati i molteplici errori  $R$  e regressione. L'aumento di più  $R$  è elencato a partire dalla prima iterazione.
- **Correlazioni di variabili originali.** Viene visualizzata una matrice che mostra le correlazioni tra le variabili non trasformate.
- **Correlazioni delle variabili trasformate.** Viene visualizzata una matrice che mostra le correlazioni tra le variabili trasformate.
- **Modelli e coefficienti regolarizzati.** Visualizza i valori di penalità, R-square e i coefficienti di regressione per ogni modello regolarizzato. Se viene specificato un metodo di ricampionamento o se vengono specificati oggetti integrali (casi di test), visualizza anche l'errore di previsione o il test MSE.

**Ricampionamento.** I metodi di ricampionamento ti danno una stima dell'errore di previsione del modello.

- **Crossvalidation.** La convalida incrociata divide il campione in vari sottocampioni o occorrenze. Vengono poi generati modelli di regressione categoriale, esclusi i dati di ciascun sottocampione a turno. Il primo modello si basa su tutti i casi eccetto quelli contenuti nella prima occorrenza del campione, il secondo modello si basa su tutti i casi eccetto quelli contenuti nella seconda occorrenza del campione e così via. Per ogni modello, l'errore di previsione è stimato applicando il modello al sottocampione escluso nel generarlo.
- **.632 Bootstrap.** Con il bootstrap le osservazioni vengono tracciate casualmente dai dati con sostituzione, ripetendo questo processo diverse volte per ottenere un numero di campioni di bootstrap. Un modello è idoneo per ogni campione di bootstrap. L'errore di previsione per ogni modello è stimato applicando il modello montato ai casi non nel campione di bootstrap.

**Categoria Quantificazioni.** Vengono visualizzate le tabelle che mostrano i valori trasformati delle variabili selezionate.

**Statistiche descrittive.** Vengono visualizzate le tabelle che mostrano le frequenze, i valori mancanti e le modalità delle variabili selezionate.

## Regressione Categoriale Salva

La finestra di dialogo Salva consente di salvare i valori previsti, i residui e i valori trasformati al dataset attivo e / o salvare i dati discretizzati, i valori trasformati, i modelli e i coefficienti regolarizzati e i segni dei coefficienti di regressione ad un file di dati IBM SPSS Statistics esterno o dataset nella sessione corrente.

- I Datasets sono disponibili durante la sessione corrente ma non sono disponibili nelle sessioni successive a meno che non si salvino esplicitamente come file di dati. I nomi di dataset devono aderire a regole di denominazione variabili.
- I nomi di filenomi o di dataset devono essere diversi per ogni tipo di dati salvato.

I modelli e i coefficienti regolarizzati vengono salvati ogni qualvolta viene selezionato un metodo di regolarizzazione sulla finestra di dialogo *Regolarizzazione*. Per impostazione predefinita, la procedura crea un nuovo dataset con un nome univoco, ma è possibile ovviamente specificare un nome di propria scelta o scrivere su un file esterno.

I segni dei coefficienti di regressione vengono salvati ogni volta che vengono utilizzati più sistemi sistematici come configurazione iniziale sulla finestra di dialogo *Opzioni*. Per impostazione predefinita, la procedura crea un nuovo dataset con un nome univoco, ma è possibile ovviamente specificare un nome di propria scelta o scrivere su un file esterno.

## Categorical Regression Transformation Plotti

La finestra di dialogo Plotti consente di specificare le variabili che produrranno la trasformazione e i tracciati residui.

**Grafici di trasformazione.** Per ognuna di queste variabili, le quantificazioni di categoria sono tracciate contro i valori di categoria originali. Le categorie vuote appaiono sull'asse orizzontale ma non influenzano i calcoli. Queste categorie sono identificate mediante interruzioni nella linea che collegano le quantificazioni.

**Grafici residui.** Per ognuna di queste variabili, i residui (calcolati per la variabile dipendente previsti da tutte le variabili predittive tranne la variabile predittiva in questione) sono tracciati contro gli indicatori di categoria e le quantificazioni di categoria ottimali moltiplicate per la beta contro gli indicatori di categoria.

## Funzioni Aggiuntive Comando CATREG

È possibile personalizzare la regressione categoriale se incollare le selezioni in una finestra di sintassi e modificare la sintassi del comando CATREG risultante. Il linguaggio della sintassi dei comandi consente inoltre di:

- Specificare i rootnomi per le variabili trasformate quando si salvano al dataset attivo (con il comando SAVE).

Vedere *Command Syntax Reference* per informazioni dettagliate sulla sintassi.

## Categorical Principal Components Analysis (CATPCA)

---

Questa procedura quantifica contemporaneamente le variabili categoriali riducendo al contempo la dimensionalità dei dati. L'analisi dei componenti principali categoriali è nota anche con l'acronimo CATPCA, per l'analisi dei componenti principali *cat*.

L'obiettivo dell'analisi dei componenti principali è quello di ridurre un insieme originale di variabili in una serie più piccola di componenti non correlati che rappresentano la maggior parte delle informazioni trovate nelle variabili originali. La tecnica è più utile quando un gran numero di variabili proibisce un'interpretazione efficace delle relazioni tra oggetti (soggetti e unità). Riducendo la dimensionalità, si interpretano alcuni componenti piuttosto che un gran numero di variabili.

L'analisi dei componenti principali standard presuppone relazioni lineari tra variabili numeriche. D'altra parte, l'approccio di scaling ottimale permette di scalare le variabili a diversi livelli. Le variabili categoriali sono ottimamente quantificate nella dimensionalità specificata. Di conseguenza, le relazioni non lineari tra variabili possono essere modellate.

**Esempio.** Le analisi dei componenti principali categoriali potrebbero essere utilizzate per visualizzare graficamente il rapporto tra categoria lavoro, divisione lavoro, regione, quantità di viaggio (alto, medio e basso) e soddisfazione del lavoro. Si potrebbe scoprire che due dimensioni rappresentano una grande quantità di varianza. La prima dimensione potrebbe separare la categoria di lavoro da regione, mentre la seconda dimensione potrebbe separare la divisione lavoro dalla quantità di viaggio. Potresti anche scoprire che l'alta soddisfazione del lavoro è correlata a un mezzo di viaggio medio.

**Statistiche e grafici.** Frequenze, valori mancanti, livello di scaling ottimale, modalità, varianza contabilizzate da coordinate centroide, coordinate vettoriali, totale per variabile e per dimensione, loadings per variabili vettoriali, quantificazioni e coordinate di categoria, cronologia delle iterazioni, correlazioni delle variabili trasformate e autovalori della matrice di correlazione, correlazioni delle variabili originali e autovalori della matrice di correlazione, punteggi oggetto, complotti di categoria, complotti di categoria, complotti di trasformazione, tracciati residui, tracciati di centrocampo, tracciati di oggetti, bilotti, triplotti e complotti dei lotti di componenti.

Categorical Principal Components Analysis Data Considerazioni

**Dati.** I valori variabili stringa vengono sempre convertiti in numeri interi positivi per ordine alfanumerico crescente. Valori mancanti definiti dall'utente, valori mancanti di sistema e valori inferiori a 1 sono considerati dispersi; è possibile recuperare o aggiungere una costante a variabili con valori inferiori a 1 per renderli non mancanti.

**Ipotesi.** I dati devono contenere almeno tre casi validi. L'analisi si basa su dati interi positivi. L'opzione di discretizzazione categorizza automaticamente una variabile a valore frazionato raggruppando i propri valori in categorie con una distribuzione più vicina alla "normale" e convertirà automaticamente i valori delle variabili stringa in numeri interi positivi. È possibile specificare altri schemi di discretizzazione.

**Procedure correlate.** Scalare tutte le variabili a livello numerico corrisponde all'analisi dei componenti principali standard. Le funzioni di plotting alternativo sono disponibili utilizzando le variabili trasformate in un'analisi dei componenti principali lineari standard. Se tutte le variabili hanno più livelli di scaling nominali, l'analisi dei componenti principali categoriali è identica a più analisi di corrispondenza. Se le serie di variabili sono di interesse, si deve utilizzare l'analisi di correlazione canonica categoriale (non lineare).

Per Ottenere un Analisi Componenti Principali Categorie

1. Dai menu, scegliere:  
**Analizzare > Riduzione della quota > Scaling ottimale ...**
2. Selezionare **Alcune variabili non più nominali**.
3. Selezionare **Un set**.
4. Fare clic su **Definisci**.

5. Selezionare almeno due variabili di analisi e specificare il numero di dimensioni nella soluzione.

6. Fare clic su **OK**.

È possibile specificare facoltativamente le variabili integrative, che sono inserite nella soluzione trovata, o le variabili di etichettatura per i tracciati.

## Definire Scale e Peso in CATPCA

È possibile impostare il livello di scaling ottimale per le variabili di analisi e le variabili supplementari. Per impostazione predefinita, vengono scalati come sprighe monotoniche di secondo grado (ordinale) con due nodi interni. Inoltre è possibile impostare il peso per le variabili di analisi.

**Peso variabile.** Puoi scegliere di definire un peso per ogni variabile. Il valore specificato deve essere un intero positivo. Il valore assunto è 1.

**Livello di scaling ottimale.** È inoltre possibile selezionare il livello di scala da utilizzare per quantificare ogni variabile.

- **Spline ordinale.** L'ordine delle categorie della variabile osservata viene conservato nella variabile scalata in modo ottimale. I punti di categoria saranno su una retta (vettore) attraverso l'origine. La trasformazione risultante è un polinomio monotono liscio del grado prescelto. I pezzi sono specificati dal numero specificato dall'utente e dal posizionamento determinato dalla procedura dei nodi interni.
- **Spline nominale.** Le uniche informazioni nella variabile osservata che si conservano nella variabile ottimamente scalata è il raggruppamento di oggetti in categorie. L'ordine delle categorie della variabile osservata non è conservato. I punti di categoria saranno su una retta (vettore) attraverso l'origine. La trasformazione risultante è un polinomio liscio, possibilmente nonmonotonico, del grado prescelto. I pezzi sono specificati dal numero specificato dall'utente e dal posizionamento determinato dalla procedura dei nodi interni.
- **Più nominale.** Le uniche informazioni nella variabile osservata che si conservano nella variabile ottimamente scalata è il raggruppamento di oggetti in categorie. L'ordine delle categorie della variabile osservata non è conservato. I punti di categoria saranno nel centroide degli oggetti nelle categorie particolari. *Più* indica che sono ottenute diverse serie di quantificazioni per ogni dimensione.
- **Ordinale.** L'ordine delle categorie della variabile osservata viene conservato nella variabile scalata in modo ottimale. I punti di categoria saranno su una retta (vettore) attraverso l'origine. La trasformazione risultante si adatta meglio alla trasformazione ordinale spline ma è meno liscia.
- **Nominale.** Le uniche informazioni nella variabile osservata che si conservano nella variabile ottimamente scalata è il raggruppamento di oggetti in categorie. L'ordine delle categorie della variabile osservata non è conservato. I punti di categoria saranno su una retta (vettore) attraverso l'origine. La trasformazione risultante si adatta meglio alla trasformazione nominale spline ma è meno liscia.
- **Numerico.** Le categorie sono trattate come ordinate e ugualmente distanziate (livello di intervallo). L'ordine delle categorie e le distanze uguali tra i numeri di categoria della variabile osservata sono conservati nella variabile ottimamente scalata. I punti di categoria saranno su una retta (vettore) attraverso l'origine. Quando tutte le variabili sono a livello numerico, l'analisi è analoga all'analisi dei componenti principali standard.

## Categorical Principal Components Analysis Discretization

Dialogo Discretization consente di selezionare un metodo di ricovero delle variabili. Le variabili valutate frazionali sono raggruppate in sette categorie (o nel numero di valori distinti della variabile se questo numero è inferiore a sette) con una distribuzione approssimativamente normale, a meno che non sia specificato diversamente. Le variabili stringa vengono sempre convertite in numeri interi positivi assegnando gli indicatori di categoria in base all'ordine alfanumerico crescente. La discretizzazione per le variabili di stringa si applica a questi interi. Altre variabili sono lasciate sole per impostazione predefinita. Le variabili discretizzate vengono poi utilizzate nell'analisi.

**Metodo.** Scegliere tra raggruppamento, graduatoria e moltiplicarsi.

- **Raggruppamento.** Recupera in un numero specificato di categorie o recupera per intervallo.

- **Classificazione.** La variabile è discretizzata ordinando i casi.
- **Moltiplicando.** I valori correnti della variabile sono standardizzati, moltiplicati per 10, arrotondati, e hanno una costante aggiunta tale che il valore più basso discretizzato è di 1.

**Raggruppamento.** Le seguenti opzioni sono disponibili quando si sono discretizing variabili per raggruppamento:

- **Numero di categorie.** Specificare una serie di categorie e se i valori della variabile devono seguire una distribuzione approssimativamente normale o uniforme in quelle categorie.
- **Gli intervalli di uguale.** Le variabili sono recuperate in categorie definite da questi intervalli ugualmente dimensionati. È necessario specificare la lunghezza degli intervalli.

## Categorical Principal Components Analysis Missing Valori

Utilizzare la finestra di dialogo Valori mancanti per scegliere la strategia per la gestione dei valori mancanti nelle variabili di analisi e le variabili supplementari.

**Strategia.** Scegliere di escludere valori mancanti (trattamento passivo), imputare i valori mancanti (trattamento attivo) oppure escludere gli oggetti con valori mancanti (eliminazione listwise).

- **Escludi i valori mancanti; per le correlazioni, assegna dopo la quantificazione** Gli oggetti con valori mancanti sulla variabile selezionata non contribuiscono all'analisi per questa variabile. Se tutte le variabili sono date trattamento passivo, allora gli oggetti con valori mancanti su tutte le variabili sono trattati come integrali. Se le correlazioni sono specificate nella finestra di dialogo Output, allora (dopo l'analisi) i valori mancanti sono imputati con la categoria più frequente, o la modalità, della variabile per le correlazioni delle variabili originali. Per le correlazioni delle variabili ottimamente scalate è possibile scegliere il metodo di imputazione.
  - **Modalità.** Sostituire i valori mancanti con la modalità della variabile ottimamente scalata.
  - **Categoria Extra.** Sostituire i valori mancanti con la quantificazione di una categoria extra. Questa impostazione implica che gli oggetti con un valore mancante su questa variabile siano considerati appartenenti alla stessa categoria (extra).
  - **Categoria casuale.** Imputare ogni valore mancante su una variabile con il valore quantificato di un diverso numero di categoria casuale in base alle frequenze marginali delle categorie della variabile.
- **Assegna i valori mancanti.** Gli oggetti con valori mancanti sulla variabile selezionata hanno quei valori imputati. È possibile scegliere il metodo di imputazione.
  - **Modalità.** Sostituire i valori mancanti con la categoria più frequente. Quando ci sono più modalità, viene utilizzato quello con l'indicatore di categoria più piccolo.
  - **Categoria Extra.** Sostituire i valori mancanti con la stessa quantificazione di una categoria extra. Questa impostazione implica che gli oggetti con un valore mancante su questa variabile siano considerati appartenenti alla stessa categoria (extra).
  - **Categoria casuale.** Sostituire ogni valore mancante su una variabile con un numero di categoria casuale diverso in base alle frequenze marginali delle categorie.
- **Escludere gli oggetti con valori mancanti su questa variabile.** Gli oggetti con valori mancanti sulla variabile selezionata sono esclusi dall'analisi. Questa strategia non è disponibile per le variabili supplementari.

## Categorical Principal Components Analysis Opzioni

La finestra di dialogo Opzioni fornisce controlli per selezionare la configurazione iniziale, specificare iterazione e criteri di convergenza, selezionare un metodo di normalizzazione, scegliere il metodo per etichettare i tracciati e specificare gli oggetti integrativi.

**Oggetti supplementari.** Specificare il numero di caso dell'oggetto, oppure i numeri di primo e ultimo caso di una gamma di oggetti, che si desidera effettuare integrazioni e quindi fare clic su **Aggiungi**. Se un oggetto viene specificato come supplementare, allora i pesi dei casi vengono ignorati per quell'oggetto.

**Metodo di normalizzazione.** È possibile specificare una delle cinque opzioni per la normalizzazione dei punteggi degli oggetti e delle variabili. In ogni analisi può essere utilizzato un solo metodo di normalizzazione.

- **Principal variabile.** Questa opzione ottimizza l'associazione tra le variabili. Le coordinate delle variabili nello spazio oggetto sono i loadings dei componenti (correlazioni con componenti principali, come dimensioni e punteggi oggetto). Questo metodo è utile quando si è interessati principalmente alla correlazione tra le variabili.
- **Oggetto principale.** Questa opzione ottimizza le distanze tra gli oggetti. Questo metodo è utile quando si è interessati principalmente a differenze o analogie tra gli oggetti.
- **Simmetrico.** Utilizzare questa opzione di normalizzazione se si è interessati principalmente alla relazione tra oggetti e variabili.
- **Independent.** Utilizzare questa opzione di normalizzazione se si desidera esaminare le distanze tra oggetti e correlazioni tra variabili separatamente.
- **Personalizzato.** È possibile specificare qualsiasi valore reale nell'intervallo chiuso  $[-1, 1]$ . Un valore di 1 è uguale al metodo Object Principal. Un valore di 0 è uguale al metodo Symmetrical. Un valore di -1 è uguale al metodo della variabile Principal. Specificando un valore superiore a -1 e inferiore a 1, è possibile diffondere l'autovalore su entrambi gli oggetti e le variabili. Questo metodo è utile per realizzare una biplot o triplot su misura.

**Criteri.** È possibile specificare il numero massimo di iterazioni la procedura può passare nelle sue elaborazioni. È inoltre possibile selezionare un valore di criterio di convergenza. L' algoritmo interrompe l' iterazione se la differenza di adattamento totale tra le ultime due iterazioni è inferiore al valore di convergenza o se si raggiunge il numero massimo di iterazioni.

**Etichetta Tracciati Per.** È possibile specificare se le variabili e le etichette di valore o i nomi e i valori variabili sono utilizzati nei tracciati. È inoltre possibile specificare una lunghezza massima per le etichette.

**Dimensioni della trama.** È possibile controllare le dimensioni visualizzate nell'output.

- **Visualizza tutte le dimensioni nella soluzione.** Tutte le dimensioni nella soluzione vengono visualizzate in una matrice di dispersione.
- **Limitare il numero di dimensioni.** Le dimensioni visualizzate sono limitate a coppie tracciate. Se si limita le dimensioni, è necessario selezionare le dimensioni più basse e più alte da tracciare. La dimensione più bassa può spaziare da 1 al numero di dimensioni nella soluzione meno 1 ed è tracciata contro dimensioni superiori. Il valore di quota più alto può spaziare da 2 al numero di dimensioni nella soluzione e indica la dimensione più alta da utilizzare nel tracciamento delle coppie di dimensioni. Questa specifica si applica a tutti i tracciati multidimensionali richiesti.

**Rotazione.** È possibile selezionare un metodo di rotazione per ottenere risultati ruotati.

Nota: questi metodi di rotazione non sono disponibili se si seleziona **Perform bootstrap** nella finestra di dialogo **Bootstrap**.

- **Varimax.** Un metodo di rotazione ortogonale che minimizza il numero di variabili che hanno alti caricamenti su ogni componente. Semplifica l'interpretazione dei componenti.
- **Quartimax.** Un metodo di rotazione che minimizza il numero di componenti necessari per spiegare ogni variabile. Semplifica l'interpretazione delle variabili osservate.
- **Equamax.** Un metodo di rotazione che è una combinazione del metodo Varimax, che semplifica i componenti, e il metodo Quartimax, che semplifica le variabili. Il numero di variabili che caricano altamente su un componente e il numero di componenti necessari per spiegare una variabile sono minimizzati.
- **Oblimin.** Un metodo per la rotazione obliqua (non ortogonale). Quando il delta equivale a 0, i componenti sono più obliqui. Poiché il delta diventa più negativo, i componenti diventano meno obliqui. I valori positivi permettono una correlazione di componente aggiuntiva. Il valore di **Delta** deve essere inferiore o uguale a 0,8.
- **Promax.** Una rotazione obliqua (non ortogonale), che consente di correlare i componenti. Può essere calcolato più rapidamente di una rotazione Oblimin diretta, quindi è utile per i grandi datasets. La

quantità di correlazione (obliquità) consentita è controllata dal parametro kappa. Il valore di **Kappa** deve essere maggiore o uguale a 1 e meno 10.000.

**Configurazione.** È possibile leggere i dati di un file che contiene le coordinate di una configurazione. La prima variabile nel file contiene le coordinate per la prima dimensione. La seconda variabile contiene le coordinate per la seconda dimensione, e così via.

- **iniziale.** La configurazione nel file specificato viene utilizzata come punto di partenza dell'analisi.
- **Fisso.** La configurazione nel file specificato viene utilizzata per adattarsi alle variabili. Le variabili che sono montate devono essere selezionate come variabili di analisi, ma perché la configurazione è fissa, sono trattate come variabili integrali (quindi non devono essere selezionate come variabili supplementari).

## Categorical Principal Components Analysis Output

La finestra di dialogo Output controlla la visualizzazione dei risultati.

**Punteggi oggetto.** Visualizza i punteggi dell'oggetto e presenta le seguenti opzioni:

- **Incluse Categorie Di.** Visualizza gli indicatori di categoria delle variabili di analisi selezionate.
- **Etichetta oggetto di etichetta per.** Dall'elenco delle variabili specificate come variabili di etichettatura è possibile selezionarne una per etichettare gli oggetti.

**Loadings del componente.** Visualizza i loadings dei componenti per tutte le variabili che non sono stati indicati più livelli di scaling nominali. È possibile ordinare i loadings dei componenti per dimensione.

**Cronologia iterazioni.** Per ogni iterazione viene mostrata la varianza contabilizzata, la perdita e l'aumento della varianza contabilizzati.

**Correlazioni di variabili originali.** Mostra la matrice di correlazione delle variabili originali e gli autovalori di quella matrice.

**Correlazioni delle variabili trasformate.** Mostra la matrice di correlazione delle variabili trasformate (ottimamente scalate) e gli autovalori di quella matrice.

**Varianza contabilizzata.** Visualizza la quantità di varianza contabilizzata da coordinate di centroide, coordinate vettoriali e totali (centroid e coordinate vettoriali combinate) per variabile e per dimensione.

**Quantificazioni di categoria.** Fornisce le quantificazioni e le coordinate di categoria per ogni dimensione delle variabili selezionate.

**Statistiche Descrittive.** Visualizza le frequenze, il numero dei valori mancanti e la modalità delle variabili selezionate.

## Categorical Principal Components Analysis Save

I controlli nella finestra di dialogo Salva salvano i dati discretizzati, i punteggi degli oggetti, i valori trasformati e altri risultati al dataset attivo, un nuovo dataset nella sessione corrente o un file esterno.

- I Datasets sono disponibili durante la sessione corrente ma non sono disponibili nelle sessioni successive a meno che non si salvino esplicitamente come file di dati. I nomi di dataset devono aderire a regole di denominazione variabili.
- I nomi di file o di dataset devono essere diversi per ogni tipo di dati salvato.
- Se si salvano i punteggi oggetto o i valori trasformati nel dataset attivo, è possibile specificare il numero di più dimensioni nominali.
- Le opzioni nel gruppo **Bootstrap Confidence Ellipses** sono disponibili solo se si seleziona **Perform bootstrap** nella finestra di dialogo Bootstrap.

## Categorical Principal Components Analysis Object Plotti

La finestra di dialogo Oggetto e Variabile Grafici consente di specificare i tipi di grafici desiderati e le variabili per le quali verranno prodotti i grafici.

**Punti oggetto.** Viene visualizzata una trama dei punti oggetto.

**Oggetti e variabili (biplot).** I punti oggetto vengono tracciati con la tua scelta delle coordinate variabili -- loadings o centroidi variabili.

**Oggetti, loadings e centroidi (triplot).** I punti oggetto sono tracciati con i centroidi di più variabili di livello - scaling nominale e i loadings di componenti di altre variabili.

**Biplot e Triplot Variabili.** È possibile scegliere di utilizzare tutte le variabili per i bigrafici e i trigrafici o selezionare un sottoinsieme.

**Oggetti di etichetta.** Puoi scegliere di avere oggetti etichettati con le categorie di variabili selezionate (puoi scegliere valori indicatori di categoria o etichette di valore nella finestra di dialogo Opzioni) o con i relativi numeri di caso. Una trama viene prodotta per variabile se si seleziona **Variabile**.

## Categorical Principal Components Analysis Categorie di analisi

La finestra di dialogo Categoria di categoria consente di specificare i tipi di grafici desiderati e le variabili per le quali verranno prodotti i grafici.

**Pagine di categoria.** Per ogni variabile selezionata viene tracciata una trama del centroide e delle coordinate vettoriali. Per le variabili con più livelli di scaling nominali, le categorie sono nei centroidi degli oggetti nelle categorie particolari. Per tutti gli altri livelli di scaling le categorie sono su un vettore attraverso l'origine.

**Plotti di categoria comuni.** Si tratta di una trama unica delle coordinate centroide e vettoriali di ogni variabile selezionata.

**Grafici di trasformazione.** Visualizza una trama delle quantificazioni di categoria ottimali contro gli indicatori di categoria. È possibile specificare il numero di dimensioni per le variabili con più livelli di scaling nominali; verrà generata una trama per ogni dimensione. È inoltre possibile scegliere di visualizzare i tracciati residui per ogni variabile selezionata.

**Centroidi Progetto Di.** Potrai scegliere una variabile e proiettare i suoi centroidi su variabili selezionate. Le variabili con più livelli di scaling nominali non possono essere selezionate su progetto. Quando questa trama viene richiesta, viene visualizzata anche una tabella con le coordinate dei centrocampisti proiettati.

## Categorical Principal Components Analisi Caricamento Lotti

La Finestra di dialogo Caricamento Lotti controlla le variabili che sono incluse nella trama, visualizzazione dei centroidi nella trama dei loadings e visualizzazione di appezzamenti di varianza contabilizzati.

**Varianza contabilizzata.** Per ogni dimensione, visualizza una trama di varianza contabilizzata.

**Visualizza i loadamenti dei componenti.** Se selezionato, viene visualizzata una trama dei loadings del componente.

**Variabili di caricamento.** Puoi scegliere di utilizzare tutte le variabili per la trama dei loadings del componente o selezionare un sottoinsieme.

**Includi i centroidi.** Le variabili con più livelli di scaling nominali non hanno loadings di componenti, ma è possibile scegliere di includere i centroidi di quelle variabili nella trama. È possibile scegliere di utilizzare tutte le variabili nominali multiple o selezionare un sottoinsieme.

## Categorical Principal Components Analysis Bootstrap

La finestra di dialogo Bootstrap specifica i parametri per l'analisi bootstrap.

**Esegui bootstrap.** Esegue il ricampionamento del bootstrap.

- Se vengono richiesti complotti di loadings, categorie o punteggi dei componenti, vengono visualizzati i complotti extra. Questi grafici visualizzano i punti per il campione di dati e le stime di bootstrap. I tracciati di trasformazione includono le regioni di fiducia. Viene visualizzata anche una trama per gli autovalori.

- Se viene specificata una soluzione bidimensionale, vengono visualizzati i tracciati delle ellissi di confidenza per gli autovalori, i loadings dei componenti, i punti di categoria e i punti oggetto.
- Il ricampionamento di bootstrap non è disponibile se si specifica un metodo di rotazione nella finestra di dialogo Opzioni.

È possibile selezionare **Balanced** o **Sbilanciato** bootstrap.

**Numero di campioni.** Il numero di campioni di bootstrap utilizzati per calcolare le stime di bootstrap. Il valore deve essere un intero positivo.

**Livello di confidenza.** Il livello di confidenza delle stime di bootstrap, espresso in percentuale. Il valore deve essere un numero positivo inferiore a 100.

**Metodo Matching.** Le due alternative sono **Procrustes** e **Reflection**.

**Ellissi di fiducia.** Controlla l'area di soglia per le ellissi di fiducia nei tracciati. Per le aree superiori a (> operatore) il valore specificato, il numero di ellissi diminuisce man mano che aumenta il valore di soglia.

- Le impostazioni in questo gruppo sono disponibili solo se il numero di dimensioni specificato sulla finestra di dialogo principale è 2.
- Gli ellissi di confidenza per i tracciati di carico sono disponibili solo se **Display component loadings** viene selezionato nella finestra di dialogo Lotti Caricamento.
- Gli ellissi di confidenza per i tracciati degli oggetti sono disponibili solo se **Punti oggetto** viene selezionato nella finestra di dialogo Oggetto e Variabile Grafici.
- Le ellissi di confidenza per i complotti di categoria sono disponibili solo se una o più variabili sono specificate nell'elenco **Plotti di categoria** nella finestra di dialogo Categoria Plotti.

**Numero di punti di contorno di ellisse.** Il numero di punti di trama utilizzati per disegnare ogni ellisse di fiducia. I valori più grandi producono delle ellissi affumicate. Il valore deve essere un intero positivo inferiore o uguale a 100.

## Funzioni Aggiuntive di comando CATPCA

È possibile personalizzare l'analisi dei componenti principali categoriali se incollare le selezioni in una finestra di sintassi e modificare la sintassi del comando CATPCA risultante. Il linguaggio della sintassi dei comandi consente inoltre di:

- Specificare i rootnomi per le variabili trasformate, i punteggi degli oggetti e le approssimazioni quando si risparmia al dataset attivo (con il comando SAVE).
- Specificare una lunghezza massima per le etichette per ogni trama separatamente (con il comando PLOT).
- Specificare un elenco di variabili separato per i tracciati residui (con il comando PLOT).

Vedere *Command Syntax Reference* per informazioni dettagliate sulla sintassi.

## Analisi della correlazione canonica non lineare (OVERALS)

L'analisi di correlazione canonica non lineare corrisponde a analisi di correlazione canonica categoriale con scaling ottimale. Lo scopo di questa procedura è quello di determinare in che modo simili serie di variabili categoriali sono l'una all'altra. L'analisi di correlazione canonica non lineare è nota anche con l'acronimo OVERALS.

L'analisi di correlazione canonica standard è un'estensione di più regressione, in cui il secondo set non contiene una singola variabile di risposta ma contiene invece più variabili di risposta. L'obiettivo è quello di spiegare il più possibile della varianza nelle relazioni tra due insiemi di variabili numeriche in uno spazio a basso dimensionali. Inizialmente le variabili in ogni set sono linearmente combinate in modo che le combinazioni lineari abbiano una correlazione massima. Date queste combinazioni, le combinazioni lineari successive sono determinate che sono slegate con le combinazioni precedenti e che hanno la più grande correlazione possibile.

L'approccio di scaling ottimale espande l'analisi standard in tre modi cruciali. In primo luogo, OVERALS consente più di due serie di variabili. In secondo luogo, le variabili possono essere scalate come nominali, ordinali o numeriche. Di conseguenza, le relazioni non lineari tra variabili possono essere analizzate. Infine, invece di massimizzare le correlazioni tra i set di variabili, i set vengono confrontati con un set di compromesso sconosciuto definito dai punteggi oggetto.

**Esempio.** L'analisi di correlazione canonica categoriale con scaling ottimale potrebbe essere utilizzata per visualizzare graficamente il rapporto tra una serie di variabili contenenti la categoria di lavoro e gli anni di istruzione e un'altra serie di variabili contenenti regione di residenza e di genere. Si potrebbe scoprire che gli anni di istruzione e la regione di residenza discriminano meglio le restanti variabili. Si potrebbe anche scoprire che gli anni di istruzione discriminano al meglio la prima dimensione.

**Statistiche e grafici.** Frequenze, centroidi, storia dell'iterazione, punteggi oggetto, quantificazioni di categoria, pesi, loadings, componenti singoli e multipli, grafici di punteggi oggetto, coordinate di categoria, lotti di loadings, complotti di categoria, complotti di trasformazione.

Nonlinear Canonical Correlation Analysis Data Considerations

**Dati.** Utilizzare i numeri interi per codificare le variabili categoriali (livello di scaling nominale o ordinale). Per ridurre al minimo l'output, utilizzare numeri interi consecutivi a partire da 1 per codificare ogni variabile. Le variabili che vengono scalate a livello numerico non devono essere recuperate in numeri interi consecutivi. Per minimizzare l'output, per ogni variabile che viene scalata a livello numerico, sottrarre il valore osservato più piccolo da ogni valore e aggiungere 1. I valori frazionari vengono troncati dopo il decimale.

**Ipotesi.** Le variabili possono essere classificate in due o più serie. Le variabili nell'analisi vengono scalate come più nominali, singole nominali, ordinali o numeriche. Il numero massimo di dimensioni che vengono utilizzati nella procedura dipende dal livello di scaling ottimale delle variabili. Se tutte le variabili sono specificate come ordinale, singola nominale o numerica, il numero massimo di dimensioni è il minore dei due valori seguenti: il numero di osservazioni meno 1 o il numero totale di variabili. Tuttavia, se vengono definiti solo due insiemi di variabili, il numero massimo di dimensioni è il numero di variabili nel set più piccolo. Se alcune variabili sono più nominali, il numero massimo di dimensioni è il numero totale di più categorie nominali più il numero di variabili nominali non multiple meno il numero di variabili nominali multiple. Ad esempio, se l'analisi coinvolge cinque variabili, una delle quali è più nominale con quattro categorie, il numero massimo di dimensioni è  $(4 + 4 - 1)$  o 7. Se si specifica un numero superiore al massimo, viene utilizzato il valore massimo.

**Procedure correlate.** Se ogni insieme contiene una sola variabile, l'analisi di correlazione canonica non lineare equivale a un'analisi componenti principali con scaling ottimale. Se ciascuna di queste variabili è più nominale, l'analisi corrisponde a più analisi di corrispondenza. Se sono coinvolte due serie di variabili e una delle serie contiene una sola variabile, l'analisi è identica alla regressione categoriale con scaling ottimale.

Per Ottenere un'analisi di Correlazione Canonica Non Lineare

1. Dai menu, scegliere:

**Analizzare > Riduzione della quota > Scaling ottimale ...**

2. Selezionare **Tutte le variabili multiple** o **Alcune variabili non più nominali**.

3. Selezionare **Più serie**.

4. Fare clic su **Definisci**.

5. Definire almeno due serie di variabili. Selezionare la variabile (s) che si desidera includere nel primo set. Per passare al prossimo set, fare clic su **Avanti** e selezionare le variabili che si desidera includere nel secondo set. È possibile aggiungere ulteriori set. Clicca su **Previous** per tornare al set di variabili precedentemente definito.

6. Definire la gamma di valore e la scala di misurazione (livello di scaling ottimale) per ciascuna variabile selezionata.

7. Fare clic su **OK**.

8. Oppure:

- Selezionare una o più variabili per fornire etichette di punti per i tracciati dei punteggi degli oggetti. Ogni variabile produce una trama separata, con i punti etichettati dai valori di quella variabile. È necessario definire un intervallo per ciascuna di queste variabili di etichetta della trama. Quando si utilizza la finestra di dialogo, una singola variabile non può essere utilizzata sia nell'analisi che come variabile di etichettatura. Se si desidera etichettare la trama dei punteggi degli oggetti con una variabile che viene utilizzata nell'analisi, utilizzare la struttura Compute (disponibile dal menu Trasforma) per creare una copia di quella variabile. Usa la nuova variabile per etichettare la trama. In alternativa, può essere utilizzata la sintassi del comando.
- Specificare il numero di dimensioni che si desidera nella soluzione. In generale, scegliere come poche dimensioni come necessario per spiegare la maggior parte della variazione. Se l'analisi coinvolge più di due dimensioni, si producono complotti tridimensionali delle prime tre dimensioni. Altre dimensioni possono essere visualizzate modificando il grafico.

## Definire intervallo e scala

È necessario definire un intervallo per ogni variabile. Il valore massimo specificato deve essere un intero. I valori dei dati frazionali sono troncati nell'analisi. Un valore di categoria che si trova al di fuori della gamma specificata viene ignorato nell'analisi. Per ridurre al minimo l'output, utilizzare la struttura Recode Automatica (disponibile dal menu Trasforma) per creare categorie consecutive a partire da 1 per le variabili trattate come nominali o ordinali. Il ricovero in numeri interi consecutivi non è consigliato per le variabili che vengono scalate a livello numerico. Per ridurre al minimo l'output per le variabili trattate come numeriche, per ogni variabile, sottrarre il valore minimo da ogni valore e aggiungere 1.

È necessario selezionare anche la scalatura da utilizzare per quantificare ogni variabile.

- **Ordinale.** L'ordine delle categorie della variabile osservata è conservato nella variabile quantificata.
- **Single nominale.** Nella variabile quantificata, gli oggetti della stessa categoria ricevono lo stesso punteggio.
- **Più nominale.** Le quantificazioni possono essere diverse per ogni dimensione.
- **numerico discreto.** Le categorie sono trattate come ordinate e ugualmente distanziate. Le differenze tra i numeri di categoria e l'ordine delle categorie della variabile osservata sono conservate nella variabile quantificata.

## Definisci intervallo

È necessario definire un intervallo per ogni variabile. Il valore massimo specificato deve essere un intero. I valori dei dati frazionali sono troncati nell'analisi. Un valore di categoria che si trova al di fuori della gamma specificata viene ignorato nell'analisi. Per ridurre al minimo l'output, utilizzare la struttura Recode Automatica (disponibile dal menu Trasforma) per creare categorie consecutive a partire da 1.

È necessario definire anche un intervallo per ogni variabile utilizzato per etichettare i tracciati dei punteggi dell'oggetto. Tuttavia, le etichette per le categorie con valori di dati che sono al di fuori dell'intervallo definito per la variabile fanno apparire sui tracciati.

## Opzioni Di Analisi Di Correlazione Canoniche Non Lineari

La finestra di dialogo Opzioni consente di selezionare statistiche e complotti opzionali, salvare i punteggi degli oggetti come nuove variabili nel dataset attivo, specificare iterazione e criteri di convergenza e specificare una configurazione iniziale per l'analisi.

**Visualizza.** Le statistiche disponibili includono frequenze marginali (conteggi), centroidi, cronologia delle iterazioni, pesi e loadamenti dei componenti, quantificazioni di categoria, punteggi oggetto e statistiche di adattamento singolo e multiplo.

- *Baricentri.* Quantificazioni di categoria e medie proiettate ed effettive dei punteggi per gli oggetti (casi) inclusi in ogni insieme per i casi appartenenti alla stessa categoria della variabile.
- *Pesi e carichi di componente.* I coefficienti di regressione in ogni dimensione per ogni variabile quantificata in un insieme, dove i punteggi degli oggetti sono regressi rispetto alle variabili quantificate

e la proiezione della variabile quantificata nello spazio oggetti. Fornisce un'indicazione del contributo di ciascuna variabile alla dimensione, all'interno di ogni insieme.

- *Adattamento singolo e multiplo.* Misure della bontà di adattamento delle coordinate di una o più categorie o delle quantificazioni di categoria rispetto agli oggetti.
- *Quantificazioni di categoria.* Valori di scala ottimale assegnati alle categorie di una variabile.
- *Punteggi degli oggetti.* Punteggio ottimale assegnato a un oggetto (caso) in una determinata dimensione.

**Grafico.** È possibile produrre complotti di coordinate di categoria, punteggi oggetto, loadori di componenti, centrocampisti di categoria e trasformazioni.

**Punteggi oggetto di salvataggio.** È possibile salvare i punteggi dell'oggetto come nuove variabili nel dataset attivo. I punteggi oggetto vengono salvati per il numero di dimensioni specificate nella finestra di dialogo principale.

**Utilizza configurazione iniziale casuale.** Una configurazione iniziale casuale deve essere utilizzata se alcune o tutte le variabili sono nominali singole. Se questa opzione non viene selezionata viene utilizzata una configurazione iniziale nidificata.

**Criteri.** È possibile specificare il numero massimo di iterazioni che l'analisi di correlazione canonica non lineare può passare attraverso i suoi calcoli. È inoltre possibile selezionare un valore di criterio di convergenza. L'analisi interrompe l'iterazione se la differenza di adattamento totale tra le ultime due iterazioni è inferiore al valore di convergenza o se si raggiunge il numero massimo di iterazioni.

## OVERALS Comando Funzioni Aggiuntive

È possibile personalizzare l'analisi della correlazione canonica non lineare se incolla le selezioni in una finestra di sintassi e modificare la sintassi del comando OVERALS risultante. Il linguaggio della sintassi dei comandi consente inoltre di:

- Specificare le coppie di dimensioni da tracciare, piuttosto che tracciare tutte le dimensioni estratte (utilizzando la parola chiaveNDIM nel sottocomando PLOT).
- Specificare il numero di caratteri etichetta di valore che vengono utilizzati per etichettare i punti sui tracciati (con il comandoPLOT).
- Designare più di cinque variabili come variabili di etichettatura per i tracciati dei punteggi oggetto (con il comandoPLOT).
- Selezionare le variabili che vengono utilizzate nell'analisi come variabili di etichettatura per i tracciati dei punteggi oggetto (con il comando PLOT).
- Selezionare le variabili per fornire etichette di punti per la trama di punteggio di quantificazione (con il comando PLOT).
- Specificare il numero di casi da includere nell'analisi se non si desidera utilizzare tutti i casi nel dataset attivo (con il comando NOBSERVATIONS).
- Specificare i rootnomi per le variabili create salvando i punteggi degli oggetti (con il comando SAVE).
- Specificare il numero di dimensioni da salvare, piuttosto che salvare tutte le dimensioni estratte (con il comando SAVE).
- Scrivere le quantificazioni di categoria in un file di matrice (utilizzando il comando MATRIX).
- Produrre grafici a bassa risoluzione che possono essere più facili da leggere rispetto ai soliti complotti ad alta risoluzione (utilizzando il comando SET).
- Produrre complotti di centroid e di trasformazione per le variabili specificate solo (con il comando PLOT).

Vedere *Command Syntax Reference* per informazioni dettagliate sulla sintassi.

## Analisi delle corrispondenze

---

Uno degli obiettivi dell'analisi della corrispondenza è quello di descrivere le relazioni tra due variabili nominali in una tabella di corrispondenza in uno spazio basso-dimensionale, descrivendo

contemporaneamente le relazioni tra le categorie per ogni variabile. Per ogni variabile, le distanze tra i punti di categoria in una trama riflettono le relazioni tra le categorie con categorie simili tracciate vicine tra loro. Proiettare punti per una variabile sul vettore dall'origine a un punto di categoria per l'altra variabile descrive il rapporto tra le variabili.

Un'analisi delle tabelle di contingenza spesso include i profili di righe e colonne e i test per l'indipendenza attraverso la statistica chi quadrato. Tuttavia, il numero dei profili può essere abbastanza grande e il test chi quadrato non rivela la struttura di dipendenza. La procedura Crosstabs offre diverse misure di associazione e test di associazione ma non può rappresentare graficamente alcuna relazione tra le variabili.

L'analisi fattoriale è una tecnica standard per descrivere le relazioni tra variabili in uno spazio basso-dimensionale. Tuttavia, l'analisi dei fattori richiede dati di intervallo e il numero di osservazioni dovrebbe essere cinque volte il numero di variabili. L'analisi della corrispondenza, invece, assume variabili nominali e può descrivere le relazioni tra le categorie di ciascuna variabile, nonché il rapporto tra le variabili. Inoltre, l'analisi della corrispondenza può essere utilizzata per analizzare qualsiasi tabella di misure di corrispondenza positive.

**Esempio.** L'analisi della corrispondenza potrebbe essere utilizzata per visualizzare graficamente il rapporto tra la categoria del personale e le abitudini di fumo. Potreste scoprire che per quanto riguarda il fumo, i junior manager si differenziano dai segretari, ma i segretari non differiscono da senior manager. Si potrebbe anche scoprire che il fumo pesante è associato ai junior manager, mentre il fumo leggero è associato ai segretari.

**Statistiche e grafici.** Misure di corrispondenza, profili di righe e colonne, valori singolari, punteggi di righe e colonne, inerzia, massa, righe e colonne punteggio di confidenza statistiche, statistiche di fiducia del valore singolare, complotti di trasformazione, tracciati di punti di riga, tracciati di punti colonna e bigrafici.

Considerazioni Dati Analisi di corrispondenza

**Dati.** Le variabili categoriali da analizzare vengono scalate nominalmente. Per i dati aggregati o per una misura di corrispondenza diversa dalle frequenze, utilizzare una variabile di ponderazione con valori di similitudine positivi. In alternativa, per i dati della tabella, utilizzare la sintassi per leggere la tabella.

**Ipotesi.** Il numero massimo di dimensioni utilizzate nella procedura dipende dal numero di righe attive e dalle categorie di colonne e dal numero di vincoli di uguaglianza. Se non vengono utilizzati vincoli di uguaglianza e tutte le categorie sono attive, la dimensionalità massima è inferiore al numero di categorie per la variabile con le categorie fewest. Ad esempio, se una variabile ha cinque categorie e l'altra ha quattro, il numero massimo di dimensioni è di tre. Le categorie complementari non sono attive. Ad esempio, se una variabile ha cinque categorie, due delle quali complementari, e l'altra variabile ha quattro categorie, il numero massimo di dimensioni è di due. Trattare tutte le serie di categorie vincolate ad essere uguali come una categoria. Ad esempio, se una variabile ha cinque categorie, tre delle quali vincolate ad essere uguali, quella variabile deve essere trattata come avere tre categorie quando si determina la massima dimensionalità. Due delle categorie sono svincolate e la terza categoria corrisponde alle tre categorie vincolate. Se si specifica un numero di dimensioni superiore al massimo, viene utilizzato il valore massimo.

**Procedure correlate.** Se sono coinvolte più di due variabili, utilizzare più analisi di corrispondenza. Se le variabili devono essere scalate ordinalmente, utilizzare le analisi dei componenti principali categoriali.

Per Ottenere un'analisi di corrispondenza

1. Dai menu, scegliere:

**Analizzare > Riduzione della quota > Analisi corrispondenza ...**

2. Selezionare una variabile di riga.
3. Selezionare una variabile di colonna.
4. Definire gli intervalli per le variabili.
5. Fare clic su **OK**.

## Definire intervallo di righe in Corrispondenza Analisi

È necessario definire un intervallo per la variabile di riga. I valori minimi e massimi specificati devono essere interi. I valori dei dati frazionali sono troncati nell'analisi. Un valore di categoria che si trova al di fuori della gamma specificata viene ignorato nell'analisi.

Tutte le categorie sono inizialmente svincolate e attive. È possibile vincolare le categorie di righe a parità di altre categorie di righe oppure definire una categoria di riga come supplente.

- **Categorie devono essere uguali.** Le categorie devono avere uguali punteggi. Utilizzare i vincoli di uguaglianza se l'ordine ottenuto per le categorie è indesiderabile o controintuitivo. Il numero massimo di categorie di righe che possono essere vincolate ad essere uguali è il numero totale di categorie di righe attive meno 1. Per imporre diversi vincoli di uguaglianza su insiemi di categorie, utilizzare la sintassi. Ad esempio, utilizzare la sintassi per vincolare le categorie 1 e 2 per essere uguali e le categorie 3 e 4 ad essere uguali.
- **Categoria è integrativa.** Le categorie complementari non influenzano l'analisi ma sono rappresentate nello spazio definito dalle categorie attive. Le categorie complementari non svolgono alcun ruolo nella definizione delle dimensioni. Il numero massimo di categorie di righe supplementari è il numero totale delle categorie di righe meno 2.

## Definire l'intervallo di colonne in Analisi della corrispondenza

È necessario definire un intervallo per la variabile della colonna. I valori minimi e massimi specificati devono essere interi. I valori dei dati frazionali sono troncati nell'analisi. Un valore di categoria che si trova al di fuori della gamma specificata viene ignorato nell'analisi.

Tutte le categorie sono inizialmente svincolate e attive. È possibile vincolare le categorie di colonne a pari altre categorie di colonne oppure definire una categoria di colonne come supplente.

- **Categorie devono essere uguali.** Le categorie devono avere uguali punteggi. Utilizzare i vincoli di uguaglianza se l'ordine ottenuto per le categorie è indesiderabile o controintuitivo. Il numero massimo di categorie di colonne che possono essere vincolate ad essere uguali è il numero totale di categorie di colonne attive meno 1. Per imporre diversi vincoli di uguaglianza su insiemi di categorie, utilizzare la sintassi. Ad esempio, utilizzare la sintassi per vincolare le categorie 1 e 2 per essere uguali e le categorie 3 e 4 ad essere uguali.
- **Categoria è integrativa.** Le categorie complementari non influenzano l'analisi ma sono rappresentate nello spazio definito dalle categorie attive. Le categorie complementari non svolgono alcun ruolo nella definizione delle dimensioni. Il numero massimo di categorie di colonne supplementari è il numero totale delle categorie di colonne meno 2.

## Modello di analisi della corrispondenza

La finestra di dialogo Modello consente di specificare il numero di dimensioni, la misura della distanza, il metodo di standardizzazione e il metodo di normalizzazione.

**Dimensioni in soluzione.** Specificare il numero di dimensioni. In generale, scegliere come poche dimensioni come necessario per spiegare la maggior parte della variazione. Il numero massimo di dimensioni dipende dal numero di categorie attive utilizzate nell'analisi e dai vincoli di uguaglianza. Il numero massimo di dimensioni è il più piccolo di:

- Il numero di categorie di righe attive meno il numero di categorie di righe vincolate ad essere uguali, più il numero di serie di categoria di righe vincolate.
- Il numero di categorie di colonne attive meno il numero di categorie di colonne vincolate ad essere uguali, più il numero di serie di categoria colonne vincolate.

**Misura di distanza.** È possibile selezionare la misura della distanza tra le righe e le colonne della tabella di corrispondenza. Scegli una delle seguenti alternative:

- **Chi-quadrato.** Utilizzare una distanza di profilo ponderata, dove il peso è la massa delle righe o delle colonne. Questa misura è richiesta per l'analisi della corrispondenza standard.

- **Euclidea.** Utilizzare la radice quadrata della somma delle differenze quadrate tra coppie di righe e coppie di colonne.

**Metodo di standardizzazione.** Scegliere una delle seguenti alternative:

- **La riga e i mezzi di colonna vengono rimossi.** Sia le righe che le colonne sono centrate. Questo metodo è richiesto per l'analisi della corrispondenza standard.
- **Il mezzo della riga viene rimosso.** Solo le righe sono centrate.
- **I mezzi di colonna sono rimossi.** Solo le colonne sono centrate.
- **I totali della riga sono equalizzati e i mezzi sono rimossi.** Prima di centrare le righe, i margini di riga sono equiparati.
- **I totali di colonna sono equalizzati e i mezzi sono rimossi.** Prima di centrare le colonne, i margini delle colonne sono equiparati.

**Metodo di normalizzazione.** Scegli una delle seguenti alternative:

- **Symmetrical.** Per ogni dimensione, i punteggi di riga sono la media ponderata dei punteggi delle colonne diviso per il valore singolare corrispondente e i punteggi delle colonne sono la media ponderata dei punteggi di riga divisi per il valore singolare corrispondente. Utilizzare questo metodo se si desidera esaminare le differenze o le similitudini tra le categorie delle due variabili.
- **Principale.** Le distanze tra i punti di fila e i punti di colonna sono approssimazioni delle distanze nella tabella di corrispondenza in base alla misura a distanza selezionata. Utilizzare questo metodo se si desidera esaminare le differenze tra le categorie di entrambe o entrambe le variabili invece delle differenze tra le due variabili.
- **Principal Row.** Le distanze tra i punti di fila sono approssimazioni delle distanze nella tabella di corrispondenza in base alla misura a distanza selezionata. I punteggi di riga sono la media ponderata dei punteggi delle colonne. Utilizzare questo metodo se si desidera esaminare differenze o similitudini tra categorie della variabile di riga.
- **principal colonna.** Le distanze tra i punti di colonna sono approssimazioni delle distanze nella tabella di corrispondenza in base alla misura a distanza selezionata. I punteggi delle colonne sono la media ponderata dei punteggi di riga. Utilizzare questo metodo se si desidera esaminare differenze o similitudini tra le categorie della variabile di colonna.
- **Personalizzato.** È necessario specificare un valore compreso tra -1 e 1. Un valore di -1 corrisponde al principal della colonna. Un valore di 1 corrisponde al principal riga. Un valore di 0 corrisponde a simmetrico. Tutti gli altri valori diffondono l'inerzia su entrambi i punteggi della riga e della colonna in varia misura. Questo metodo è utile per effettuare bilotti su misura.

## Statistiche di analisi della corrispondenza

La finestra di dialogo Statistiche consente di specificare l'output numerico prodotto.

**Tabella Corrispondenza.** Una crosatura delle variabili di ingresso con totali marginali di riga e colonna.

**Panoramica dei punti di riga.** Per ogni categoria di fila, i punteggi, la massa, l'inerzia, il contributo all'inerzia della dimensione, e il contributo della dimensione all'inerzia del punto.

**Panoramica dei punti colonna.** Per ogni categoria di colonna, i punteggi, la massa, l'inerzia, il contributo all'inerzia della dimensione, e il contributo della dimensione all'inerzia del punto.

**Profili di riga.** Per ogni categoria di riga, la distribuzione attraverso le categorie della variabile di colonna.

**Profili di colonna.** Per ogni categoria di colonna, la distribuzione attraverso le categorie della variabile di riga.

**Permutazioni della tabella di corrispondenza.** La tabella di corrispondenza riorganizzata in modo tale che le righe e le colonne siano in ordine crescente secondo i punteggi sulla prima dimensione. Facoltativamente, è possibile specificare il numero di quota massima per cui verranno prodotte tabelle permutate. Viene prodotta una tabella permutata per ogni dimensione dal 1 al numero specificato.

**Statistiche di fiducia per i punti di riga.** Include deviazione standard e correlazioni per tutti i punti di riga non supplementari.

**Statistiche di fiducia per i punti Colonna.** Include deviazione standard e correlazioni per tutti i punti di colonna non supplementari.

## Grafici di analisi della corrispondenza

La finestra di dialogo Pagine consente di specificare quali complotti sono prodotti.

**Grafici a dispersione.** Produce una matrice di tutti i complotti pairwise delle dimensioni. Gli scattergrafici disponibili includono:

- **Biplot.** Produce una matrice di tracciati comuni della riga e dei punti colonna. Se viene selezionata la normalizzazione principale, la biplot non è disponibile.
- **Punti di riga.** Produce una matrice di tracciati dei punti di fila.
- **Punti di colonna.** Produce una matrice di tracciati dei punti colonna.

Facoltativamente, è possibile specificare quanti caratteri etichetta di valore utilizzare quando etichettare i punti. Questo valore deve essere un intero non negativo inferiore o uguale a 20.

**Pagine di linea.** Produce una trama per ogni dimensione della variabile selezionata. I tracciati di linea disponibili includono:

- **Categorie di righe trasformate.** Produce una trama dei valori della categoria di righe originali contro i relativi punteggi di riga.
- **Categorie di colonne trasformate.** Produce una trama dei valori di categoria della colonna originale contro i relativi punteggi della colonna.

Facoltativamente, è possibile specificare quanti caratteri etichetta di valore utilizzare quando etichettare l'asse della categoria. Questo valore deve essere un intero non negativo inferiore o uguale a 20.

**Dimensioni della trama.** Consente di controllare le dimensioni visualizzate nell'output.

- **Visualizza tutte le dimensioni nella soluzione.** Tutte le dimensioni nella soluzione vengono visualizzate in una matrice di dispersione.
- **Limitare il numero di dimensioni.** Le dimensioni visualizzate sono limitate a coppie tracciate. Se si limita le dimensioni, è necessario selezionare le dimensioni più basse e più alte da tracciare. La dimensione più bassa può spaziare da 1 al numero di dimensioni nella soluzione meno 1, ed è tracciata contro dimensioni superiori. Il valore di quota più alto può spaziare da 2 al numero di dimensioni nella soluzione e indica la dimensione più alta da utilizzare nel tracciamento delle coppie di dimensioni. Questa specifica si applica a tutti i tracciati multidimensionali richiesti.

## CORRISPONDENZA Comando Funzioni Aggiuntive

È possibile personalizzare l'analisi della corrispondenza se incollare le selezioni in una finestra di sintassi e modificare la sintassi del comando CORRESPONDENCE risultante. Il linguaggio della sintassi dei comandi consente inoltre di:

- Specificare i dati della tabella come input invece di utilizzare i dati casewise (utilizzando il comando `TABLE = ALL`).
- Specificare il numero di caratteri in etichetta di valore utilizzati per etichettare i punti per ogni tipo di matrice di dispersione o matrice di biplot (con il comando `PLOT`).
- Specificare il numero di caratteri in etichetta di valore utilizzati per etichettare i punti per ogni tipo di trama di linea (con il comando `PLOT`).
- Scrivere una matrice di punteggi di righe e colonne in un file di dati di matrice (con il comando `OUTFILE`).
- Scrivi una matrice di statistiche di fiducia (varianze e covarianze) per i valori singolari e i punteggi ad un file di dati di matrice (con il comando `OUTFILE`).
- Specificare più serie di categorie ad essere uguali (con il comando `EQUAL`).

Vedere *Command Syntax Reference* per informazioni dettagliate sulla sintassi.

## Analisi delle corrispondenze multiple

---

Analisi Multiple Corrispondenza quantifica i dati nominali (categoriali) assegnando valori numerici ai casi (oggetti) e categorie in modo che gli oggetti all'interno della stessa categoria siano vicini e gli oggetti in diverse categorie siano distanti. Ogni oggetto è il più vicino possibile ai punti di categoria delle categorie che si applicano all'oggetto. In questo modo le categorie suddividono gli oggetti in sottogruppi omogenei. Le variabili sono considerate omogenee quando classificano gli oggetti nelle stesse categorie negli stessi sottogruppi.

**Esempio.** L'Analisi di Corrispondenza Multipla potrebbe essere utilizzata per visualizzare graficamente il rapporto tra categoria lavoro, classificazione di minoranza e sesso. Si potrebbe scoprire che la classificazione minoritaria e la discriminazione di genere tra le persone ma quella categoria di lavoro no. Si potrebbe anche scoprire che le categorie latine e afro - americane sono simili tra loro.

**Statistiche e grafici.** Punteggi oggetto, misure di discriminazione, storia dell'iterazione, correlazioni di variabili originali e trasformate, quantificazioni di categoria, statistiche descrittive, tracciati di punti oggetto, bilotti, complotti, complotti di categoria, complotti di trasformazione e misure di discriminazione complotti.

Considerazioni Dati Analisi di corrispondenza multipla

**Dati.** I valori variabili stringa vengono sempre convertiti in numeri interi positivi per ordine alfanumerico crescente. Valori mancanti definiti dall'utente, valori mancanti di sistema e valori inferiori a 1 sono considerati dispersi; è possibile recuperare o aggiungere una costante a variabili con valori inferiori a 1 per renderli non mancanti.

**Ipotesi.** Tutte le variabili hanno il livello di scaling nominale multiplo. I dati devono contenere almeno tre casi validi. L'analisi si basa su dati interi positivi. L'opzione di discretizzazione categorizza automaticamente una variabile a valore frazionato raggruppando i propri valori in categorie con una distribuzione più vicina alla normale e convertirà automaticamente i valori delle variabili stringa in numeri interi positivi. È possibile specificare altri schemi di discretizzazione.

**Procedure correlate.** Per due variabili, Multiple Correspondence Analysis è analoga a Corrispondenza Analysis. Se si ritiene che le variabili possiedano proprietà ordinarie o numeriche, si dovrebbe utilizzare l'Analisi Componenti Principali di categoria. Se le serie di variabili sono di interesse, deve essere utilizzata l'analisi di Correlazione Canonica Non Lineare.

Per Ottenere un'analisi di Corrispondenza Multipla

1. Dai menu, scegliere:

**Analizzare > Riduzione della quota > Scaling ottimale ...**

2. Selezionare **Tutte le variabili multiple nominali**.

3. Selezionare **Un set**.

4. Fare clic su **Definisci**.

5. Selezionare almeno due variabili di analisi e specificare il numero di dimensioni nella soluzione.

6. Fare clic su **OK**.

È possibile specificare facoltativamente le variabili integrative, che sono inserite nella soluzione trovata, o le variabili di etichettatura per i tracciati.

## Definire Peso variabile in Analisi Corrispondenza Multipla

È possibile impostare il peso per le variabili di analisi.

**Peso variabile.** Puoi scegliere di definire un peso per ogni variabile. Il valore specificato deve essere un intero positivo. Il valore assunto è 1.

## Più Discretizzazione Analisi Corrispondenza

Dialogo Discretization consente di selezionare un metodo di ricovero delle variabili. Le variabili valutate frazionali sono raggruppate in sette categorie (o nel numero di valori distinti della variabile se questo numero è inferiore a sette) con una distribuzione approssimativamente normale, salvo diversa indicazione. Le variabili stringa vengono sempre convertite in numeri interi positivi assegnando gli indicatori di categoria in base all'ordine alfanumerico crescente. La discretizzazione per le variabili di stringa si applica a questi interi. Altre variabili sono lasciate sole per impostazione predefinita. Le variabili discretizzate vengono poi utilizzate nell'analisi.

**Metodo.** Scegliere tra raggruppamento, graduatoria e moltiplicarsi.

- **Raggruppamento.** Recupera in un numero specificato di categorie o recupera per intervallo.
- **Classificazione.** La variabile è discretizzata ordinando i casi.
- **Moltiplicando.** I valori correnti della variabile sono standardizzati, moltiplicati per 10, arrotondati e hanno una costante aggiunta in modo che il valore più basso discretizzato sia di 1.

**Raggruppamento.** Le seguenti opzioni sono disponibili quando le variabili discretizing si raggruppano:

- **Numero di categorie.** Specificare una serie di categorie e se i valori della variabile devono seguire una distribuzione approssimativamente normale o uniforme in quelle categorie.
- **Gli intervalli di uguale.** Le variabili sono recuperate in categorie definite da questi intervalli ugualmente dimensionati. È necessario specificare la lunghezza degli intervalli.

## Più Analisi Corrispondenza valori mancanti

La finestra di dialogo Valori mancanti consente di scegliere la strategia per la gestione dei valori mancanti nelle variabili di analisi e le variabili supplementari.

**Strategia di valore mancante.** Scegliere di escludere valori mancanti (trattamento passivo), impadare i valori mancanti (trattamento attivo) oppure escludere gli oggetti con valori mancanti (eliminazione listwise).

- **Escludi i valori mancanti; per le correlazioni, assegna dopo la quantificazione** Gli oggetti con valori mancanti sulla variabile selezionata non contribuiscono all'analisi per questa variabile. Se tutte le variabili sono date trattamento passivo, allora gli oggetti con valori mancanti su tutte le variabili sono trattati come integrali. Se le correlazioni sono specificate nella finestra di dialogo Output, allora (dopo l'analisi) i valori mancanti sono imputati con la categoria più frequente, o la modalità, della variabile per le correlazioni delle variabili originali. Per le correlazioni delle variabili ottimamente scalate è possibile scegliere il metodo di imputazione. Selezionare **Modalità** per sostituire i valori mancanti con la modalità della variabile ottimamente scalata. Selezionare **Extra category** per sostituire i valori mancanti con la quantificazione di una categoria extra. Ciò implica che gli oggetti con un valore mancante su questa variabile sono considerati appartenenti alla stessa categoria (extra).
- **Assegna i valori mancanti.** Gli oggetti con valori mancanti sulla variabile selezionata hanno quei valori imputati. È possibile scegliere il metodo di imputazione. Selezionare **Modalità** per sostituire i valori mancanti con la categoria più frequente. Quando ci sono più modalità, viene utilizzato quello con l'indicatore di categoria più piccolo. Selezionare **Extra category** per sostituire i valori mancanti con la stessa quantificazione di una categoria extra. Ciò implica che gli oggetti con un valore mancante su questa variabile sono considerati appartenenti alla stessa categoria (extra).
- **Escludere gli oggetti con valori mancanti su questa variabile.** Gli oggetti con valori mancanti sulla variabile selezionata sono esclusi dall'analisi. Questa strategia non è disponibile per le variabili supplementari.

## Opzioni Analisi di corrispondenza multipla

La finestra di dialogo Opzioni consente di selezionare la configurazione iniziale, specificare iterazione e criteri di convergenza, selezionare un metodo di normalizzazione, scegliere il metodo per etichettare i tracciati e specificare gli oggetti integrativi.

**Oggetti supplementari.** Specificare il numero di caso dell'oggetto (o il primo e l'ultimo numero di casi di una gamma di oggetti) che si desidera fare supplementare, quindi fare clic su **Aggiungi**. Continua fino a quando non hai specificato tutti i tuoi oggetti integrali. Se un oggetto viene specificato come supplementare, allora i pesi dei casi vengono ignorati per quell' oggetto.

**Metodo di normalizzazione.** È possibile specificare una delle cinque opzioni per la normalizzazione dei punteggi degli oggetti e delle variabili. Solo un metodo di normalizzazione può essere utilizzato in una determinata analisi.

- **Principal variabile.** Questa opzione ottimizza l'associazione tra le variabili. Le coordinate delle variabili nello spazio oggetto sono i loadings dei componenti (correlazioni con componenti principali, come dimensioni e punteggi oggetto). Ciò è utile quando si è interessati principalmente alla correlazione tra le variabili.
- **Oggetto principale.** Questa opzione ottimizza le distanze tra gli oggetti. Ciò è utile quando si è interessati principalmente a differenze o analogie tra gli oggetti.
- **Simmetrico.** Utilizzare questa opzione di normalizzazione se si è interessati principalmente alla relazione tra oggetti e variabili.
- **Independent.** Utilizzare questa opzione di normalizzazione se si desidera esaminare le distanze tra oggetti e correlazioni tra variabili separatamente.
- **Personalizzato.** È possibile specificare qualsiasi valore reale nell'intervallo chiuso  $[-1, 1]$ . Un valore di 1 è uguale al metodo Object Principal, un valore di 0 è uguale al metodo Symmetrical e un valore di -1 è uguale al metodo della variabile Principal. Specificando un valore superiore a -1 e inferiore a 1, è possibile diffondere l'autovalore su entrambi gli oggetti e le variabili. Questo metodo è utile per realizzare una biplot o triplot su misura.

**Criteri.** È possibile specificare il numero massimo di iterazioni la procedura può passare nelle sue elaborazioni. È inoltre possibile selezionare un valore di criterio di convergenza. L'algoritmo interrompe l'iterazione se la differenza di adattamento totale tra le ultime due iterazioni è inferiore al valore di convergenza o se si raggiunge il numero massimo di iterazioni.

**Etichetta Tracciati Per.** Consente di specificare se le variabili e le etichette di valore o i nomi e i valori variabili verranno utilizzati nei tracciati. È inoltre possibile specificare una lunghezza massima per le etichette.

**Dimensioni della trama.** Consente di controllare le dimensioni visualizzate nell'output.

- **Visualizza tutte le dimensioni nella soluzione.** Tutte le dimensioni nella soluzione vengono visualizzate in una matrice di dispersione.
- **Limitare il numero di dimensioni.** Le dimensioni visualizzate sono limitate a coppie tracciate. Se si limita le dimensioni, è necessario selezionare le dimensioni più basse e più alte da tracciare. La dimensione più bassa può spaziare da 1 al numero di dimensioni nella soluzione meno 1 ed è tracciata contro dimensioni superiori. Il valore di quota più alto può spaziare da 2 al numero di dimensioni nella soluzione e indica la dimensione più alta da utilizzare nel tracciamento delle coppie di dimensioni. Questa specifica si applica a tutti i tracciati multidimensionali richiesti.

**Configurazione.** È possibile leggere i dati di un file contenente le coordinate di una configurazione. La prima variabile nel file dovrebbe contenere le coordinate per la prima dimensione, la seconda variabile dovrebbe contenere le coordinate per la seconda dimensione, e così via.

- **iniziale.** La configurazione nel file specificato verrà utilizzata come punto di partenza dell'analisi.
- **Fisso.** La configurazione nel file specificato verrà utilizzata per adattarsi alle variabili. Le variabili che sono montate devono essere selezionate come variabili di analisi, ma, poiché la configurazione è fissa, sono trattate come variabili integrali (quindi non devono essere selezionate come variabili supplementari).

## Output Analisi di corrispondenza multipla

La finestra di dialogo Output consente di produrre tabelle per punteggi oggetto, misure di discriminazione, cronologia delle iterazioni, correlazioni di variabili originali e trasformate, quantificazioni di categoria per variabili selezionate e statistiche descrittive per variabili selezionate.

**Punteggi oggetto.** Visualizza i punteggi dell'oggetto, tra cui massa, inerzia e contributi e ha le seguenti opzioni:

- **Incluse Categorie Di.** Visualizza gli indicatori di categoria delle variabili di analisi selezionate.
- **Etichetta oggetto di etichetta per.** Dall'elenco delle variabili specificate come variabili di etichettatura, è possibile selezionarne una per etichettare gli oggetti.

**Misure di discriminazione.** Visualizza le misure di discriminazione per variabile e per dimensione.

**Cronologia iterazioni.** Per ogni iterazione viene mostrata la varianza contabilizzata, la perdita e l'aumento della varianza contabilizzati.

**Correlazioni di variabili originali.** Mostra la matrice di correlazione delle variabili originali e gli autovalori di quella matrice.

**Correlazioni delle variabili trasformate.** Mostra la matrice di correlazione delle variabili trasformate (ottimamente scalate) e gli autovalori di quella matrice.

**Quantificazioni di categoria e contributi.** Fornisce le quantificazioni di categoria (coordinate), tra cui massa, inerzia e contributi, per ogni dimensione della variabile (s) selezionata.

*Nota:* le coordinate e i contributi (compresa la massa e l'inerzia) vengono visualizzati in strati separati dell'output della tabella pivot, con le coordinate mostrate per impostazione predefinita. Per visualizzare i contributi, attivare (doppio clic) sulla tabella e selezionare Contributi dall'elenco a discesa Layer.

**Statistiche descrittive.** Visualizza le frequenze, il numero dei valori mancanti e la modalità della variabile (s) selezionata.

## Multiple Analyses of Correspondence Save

La finestra di dialogo Salva consente di salvare i dati discretizzati, i punteggi degli oggetti e i valori trasformati in un file di dati IBM SPSS Statistics esterno o dataset nella sessione corrente. È anche possibile salvare i valori trasformati e i punteggi degli oggetti nel dataset attivo.

- I Datasets sono disponibili durante la sessione corrente ma non sono disponibili nelle sessioni successive a meno che non si salvino esplicitamente come file di dati. I nomi di dataset devono aderire a regole di denominazione variabili.
- I nomi di file o di dataset devono essere diversi per ogni tipo di dati salvato.
- Se si salvano i punteggi oggetto o i valori trasformati nel dataset attivo, è possibile specificare il numero di più dimensioni nominali.

## Multiple Correspondence Object of Correspondence Analysis

La finestra di dialogo Oggetti Oggetto consente di specificare i tipi di tracciati desiderati e le variabili da tracciare

**Punti oggetto.** Viene visualizzata una trama dei punti oggetto.

**Oggetti e centroidi (biplot).** I punti oggetto sono tracciati con i centrocampisti variabili.

**Variabili di biplot.** È possibile scegliere di utilizzare tutte le variabili per i bigrafici o selezionare un sottoinsieme.

**Oggetti di etichetta.** Puoi scegliere di avere oggetti etichettati con le categorie di variabili selezionate (puoi scegliere valori indicatori di categoria o etichette di valore nella finestra di dialogo Opzioni) o con i relativi numeri di caso. Una trama viene prodotta per variabile se si seleziona **Variabile**.

## Multiple Correspondence Analysis of Correspondence

La finestra di dialogo Grafici Variabili consente di specificare i tipi di tracciati desiderati e le variabili da tracciare.

**Pagine di categoria.** Per ogni variabile selezionata viene tracciata una trama delle coordinate del centroide. Le categorie sono nei centroidi degli oggetti nelle categorie particolari.

**Plotti di categoria comuni.** Si tratta di una trama unica delle coordinate del centroide di ciascuna variabile selezionata.

**Grafici di trasformazione.** Visualizza una trama delle quantificazioni di categoria ottimali contro gli indicatori di categoria. È possibile specificare il numero di dimensioni; una trama verrà generata per ogni dimensione. È inoltre possibile scegliere di visualizzare i tracciati residui per ogni variabile selezionata.

**Misure di discriminazione.** Produce una trama unica delle misure di discriminazione per le variabili selezionate.

## MULTIPLE CORRISPONDENZA Comando Funzioni aggiuntive

È possibile personalizzare la tua Analisi Multipla se incollare le selezioni in una finestra di sintassi e modificare la sintassi del comando MULTIPLE CORRESPONDENCE risultante. Il linguaggio della sintassi dei comandi consente inoltre di:

- Specificare i rootnomi per le variabili trasformate, i punteggi degli oggetti e le approssimazioni quando si risparmia al dataset attivo (con il comando SAVE).
- Specificare una lunghezza massima per le etichette per ogni trama separatamente (con il comando PLOT).
- Specificare un elenco di variabili separato per i tracciati residui (con il comando PLOT).

Vedere *Command Syntax Reference* per informazioni dettagliate sulla sintassi.

## Scaling multidimensionale (PROXSCAL)

---

I tentativi di scaling multidimensionale per trovare la struttura in una serie di misure di prossimità tra gli oggetti. Questo processo è compiuto assegnando le osservazioni a specifiche località in uno spazio concettuale a basso dimensionamento tale che le distanze tra i punti nello spazio corrispondono al dato (dis) analogie il più possibile. Il risultato è una rappresentazione meno quadrante degli oggetti in quello spazio basso-dimensionale che, in molti casi, ti aiuterà a comprendere ulteriormente i tuoi dati.

**Esempio.** Scaling multidimensionale può essere molto utile nel determinare relazioni percettive. Ad esempio, quando si considera la propria immagine del prodotto, è possibile condurre un'indagine per ottenere un dataset che descriva la somiglianza percepita (o la vicinanza) del proprio prodotto a quelle dei tuoi concorrenti. Utilizzando queste variabili prossimamente e indipendenti (come il prezzo), è possibile provare a determinare quali variabili sono importanti per come le persone visualizzano questi prodotti, e si può regolare di conseguenza la propria immagine.

**Statistiche e grafici.** Cronologia delle iterazioni, misure di stress, decomposizione dello stress, coordinate dello spazio comune, distanze degli oggetti all'interno della configurazione finale, pesi singoli spazi, spazi individuali, prossimità trasformate, variabili indipendenti trasformate, tracciati di stress, spazi comuni di scatterplotti, singoli spazi di scatterplotti, singoli spazi scattergrafici, grafici di trasformazione, grafici di trasformazione di Shepard e grafici di trasformazione delle variabili indipendenti.

Considerazioni sui dati dello Scaling multidimensionale

**Dati.** I dati possono essere forniti sotto forma di matrici di prossimità o variabili che vengono convertite in matrici di prossimità. Le matrici possono essere formattate in colonne o in righe. Le prossimità possono essere trattate sul rapporto tra rapporto, intervallo, ordinale o scala di spline.

**Ipotesi.** Devono essere specificate almeno tre variabili. Il numero di dimensioni non può superare il numero di oggetti meno uno. La riduzione dimensionalità viene omessa se combinata con più inizi casuali. Se viene specificata una sola origine, tutti i modelli sono equivalenti al modello di identità; pertanto, l'analisi viene impostata per default sul modello di identità.

**Procedure correlate.** Scalare tutte le variabili a livello numerico corrisponde ad analisi di scaling multidimensionali standard.

Per Ottenere una Scala Multidimensionale

1. Dai menu, scegliere:

## Analizzare > Scala > Scaling multidimensionale (PROXSCAL) ...

In questo modo si apre la finestra di dialogo Formato dati.

### 2. Specificare il formato dei tuoi dati:

**Formato dati.** Specificare se i tuoi dati sono costituiti da misure di prossimità o vuoi creare prossimità dai dati.

**Numero di origini.** Se i tuoi dati sono prossimità, specificare se hai una singola origine o più sorgenti di misure di prossimità.

**Una origine.** Se vi è una fonte di prossimità, specificare se il proprio dataset viene formattato con le prossimità in una matrice attraverso le colonne o in una singola colonna con due variabili distinte per identificare la riga e la colonna di ogni prossimità.

- *Le prossimità sono in una matrice per colonne.* La matrice di prossimità è distribuita su un numero di colonne uguale al numero di oggetti. L'operazione porta alla finestra di dialogo Prossimità in matrici per colonne.
- *Le prossimità si trovano in una singola colonna.* La matrice di prossimità è compressa in una singola colonna o variabile. Sono necessarie due variabili aggiuntive, che identificano riga e colonna per ciascuna cella. L'operazione porta alla finestra di dialogo Prossimità in una sola colonna.

**Più origini.** Se ci sono più fonti di prossimità, specificare se il dataset viene formattato con le prossimità in matrici in pila su colonne, in più colonne con una sola sorgente per colonna o in una singola colonna.

- *Le prossimità sono in matrici sovrapposte per colonne.* Le matrici di prossimità sono distribuite su un numero di colonne pari al numero di oggetti e impilate una sopra l'altra su un numero di righe pari al numero di oggetti moltiplicati per il numero di origini. L'operazione porta alla finestra di dialogo Prossimità in matrici per colonne.
- *Le prossimità si trovano in una colonna, un'origine per colonna.* Le matrici di prossimità sono compresse in più colonne o variabili. Sono necessarie due variabili aggiuntive, che identificano riga e colonna per ciascuna cella. L'operazione porta alla finestra di dialogo Prossimità in colonne.
- *Le prossimità sono sovrapposte in una singola colonna.* Le matrici di prossimità sono compresse in una singola colonna o variabile. Sono necessarie tre variabili aggiuntive, che identificano riga, colonna e origine per ciascuna cella. L'operazione porta alla finestra di dialogo Prossimità in una sola colonna.

### 3. Fare clic su **Definisci**.

## Prossimità in Matrici attraverso le Colonne

Se si selezionano le prossimità in modello di dati matrici per una sola sorgente o più fonti nella finestra di dialogo Formato dati, quindi effettuare quanto segue:

1. Selezionare tre o più variabili di prossimità. (Assicurarsi che l'ordine delle variabili nell'elenco corrisponda all'ordine delle colonne delle prossimità.)
2. Opzionalmente, selezionare una serie di variabili di peso pari al numero di variabili di prossimità. (Assicurarsi che l'ordine dei pesi corrisponda all'ordine delle prossimità che pesano.)
3. Opzionalmente, se ci sono più fonti, selezionare una variabile di origini. (Il numero di casi in ogni variabile di prossimità dovrebbe eguagliare il numero di variabili di prossimità volte il numero di sorgenti.)

Inoltre, è possibile definire un modello per la scaling multidimensionale, porre limitazioni allo spazio comune, impostare criteri di convergenza, specificare la configurazione iniziale da utilizzare e scegliere i grafici e l'output.

## Prossimità in Colonne

Se si seleziona il modello di più colonne per più fonti nella finestra di dialogo Formato dati, procedere quindi:

1. Selezionare due o più variabili di prossimità. (Ogni variabile è ipotizza una matrice di prossimità da una sorgente separata.)
2. Selezionare una variabile di righe per definire le posizioni di riga per le prossimità in ogni variabile di prossimità.
3. Selezionare una variabile di colonne per definire le posizioni delle colonne per le prossimità in ogni variabile di prossimità. (Celle della matrice di prossimità che non vengono date una designazione a righe / a colonna sono trattate come mancanti.)
4. Opzionalmente, selezionare una serie di variabili di peso pari al numero di variabili di prossimità.

Inoltre, è possibile definire un modello per la scaling multidimensionale, porre limitazioni allo spazio comune, impostare criteri di convergenza, specificare la configurazione iniziale da utilizzare e scegliere i grafici e l'output.

## Prossimità in una Colonna

Se si seleziona il modello di una colonna per una sola origine o più sorgenti nella finestra di dialogo Formato dati, quindi effettuare quanto segue:

1. Selezionare una variabile di prossimità. (si suppone che siano una o più matrici di prossimità.)
2. Selezionare una variabile di righe per definire le posizioni di riga per le prossimità nella variabile di prossimità.
3. Selezionare una variabile di colonne per definire le posizioni delle colonne per le prossimità nella variabile di prossimità.
4. Se ci sono più fonti, selezionare una variabile di origini. (Per ogni fonte, le celle della matrice di prossimità che non hanno una designazione di riga / colonna sono trattate come mancanti.)
5. Opzionalmente, selezionare una variabile di peso.

Inoltre, è possibile definire un modello per la scaling multidimensionale, porre limitazioni allo spazio comune, impostare criteri di convergenza, specificare la configurazione iniziale da utilizzare e scegliere i grafici e l'output.

## Crea prossimità dai dati

Se si sceglie di creare prossimità dai dati nella finestra di dialogo Formato dati, procedere quindi:

1. Se si creano le distanze tra le variabili (consultare la finestra di dialogo Crea Misura dei dati), selezionare almeno tre variabili. Queste variabili verranno utilizzate per creare la matrice di prossimità (o le matrici, se ci sono più fonti). Se si creano distanze tra i casi, è necessaria una sola variabile.
2. Se ci sono più fonti, selezionare una variabile di origini.
3. Opzionalmente, scegliere una misura per la creazione di prossimità.

Inoltre, è possibile definire un modello per la scaling multidimensionale, porre limitazioni allo spazio comune, impostare criteri di convergenza, specificare la configurazione iniziale da utilizzare e scegliere i grafici e l'output.

## Crea Misura da Dati

La procedura Scaling multidimensionale utilizza dati di dissimilarità per creare una soluzione di scaling. Se i dati disponibili sono dati multivariati (valori di variabili misurate), è necessario creare dati di dissimilarità in modo da calcolare una soluzione di scaling multidimensionale. È possibile specificare i dettagli della creazione delle misure di dissimilarità a partire dai dati disponibili.

**Misura.** Consente di specificare la misura di dissimilarità per l'analisi. Selezionare un'alternativa dal gruppo Misura corrispondente al proprio tipo di dati, quindi selezionare una delle misure dall'elenco a discesa corrispondente a quel tipo di misura. Le alternative disponibili sono:

- **Intervallo.** Distanza euclidea, Distanza euclidea al quadrato, Chebychev, City-Block, Minkowski o Personalizzato.

- **Conteggi.** Misura chi-quadrato e Misura phi-quadrato.
- **Binaria.** Distanza euclidea, Distanza euclidea al quadrato, Differenza di dimensione, Differenza di modello, Varianza o Lance e Williams.

**Crea matrice delle distanze.** Consente di scegliere l'unità di analisi. Le alternative sono Fra variabili o Fra casi.

**Trasforma valori.** In certi casi, come quando le variabili vengono misurate su scale molto diverse, si desidera standardizzare i valori prima di computare le prossimità (non applicabile ai dati binari). Selezionare un metodo di standardizzazione dall'elenco a discesa Standardize (se non è richiesta alcuna standardizzazione, selezionare **Nessuno**).

## Definire un modello di Scaling Multidimensionale

La finestra di dialogo Modello consente di specificare un modello di scaling, il suo numero minimo e massimo di dimensioni, la struttura della matrice di prossimità, la trasformazione da utilizzare sulle prossimità e se le prossimità vengono trasformate all'interno di ogni sorgente separatamente o incondizionatamente sulla sorgente.

**Modello di scala.** Scegli tra le seguenti alternative:

- **Identità.** Tutte le fonti hanno la stessa configurazione.
- **Euclideo Ponderato.** Questo modello è un modello di differenze individuale. Ogni sorgente ha uno spazio individuale in cui ogni dimensione dello spazio comune è ponderata differentemente.
- **Euclideo generalizzato.** Questo modello è un modello di differenze individuale. Ogni sorgente ha uno spazio individuale uguale a una rotazione dello spazio comune, seguito da una ponderazione differenziale delle dimensioni.
- **Rango ridotto.** Questo modello è un modello euclideo generalizzato per il quale è possibile specificare il rango dello spazio individuale. È necessario specificare una classifica maggiore o uguale a 1 e inferiore al numero massimo di dimensioni.

**Forma.** Specificare se le prossimità devono essere prese dalla parte inferiore - triangolare o dalla parte superiore - triangolare della matrice di prossimità. È possibile specificare che venga utilizzata la matrice completa, in tal caso verrà analizzato la somma ponderata della parte superiore - triangolare e la parte inferiore - triangolare. In ogni caso, deve essere specificata la matrice completa, compresa la diagonale, anche se verranno utilizzate solo le parti specificate.

**Prossimità.** Specificare se la matrice di prossimità contiene misure di similitudine o dissimilarità.

**Trasformazioni di prossimità.** Scegli tra le seguenti alternative:

- **Rapporto.** Le prossimità trasformate sono proporzionali alle prossimità originali. Questo è consentito solo per le prossimità valutate positivamente.
- **Intervallo.** Le prossimità trasformate sono proporzionali alle prossimità originali, più un termine di intercettazione. L'intercettazione assicura tutte le prossimità trasformate ad essere positive.
- **Ordinale.** Le prossimità trasformate hanno lo stesso ordine delle prossimità originali. Si specifica se le prossimità legate devono essere tenute legate o lasciate svincolate.
- **Spline.** Le prossimità trasformate sono una fluida trasformazione polinomica non decrescente delle prossimità originali. Si specifica il grado del polinomio e il numero di nodi interni.

**Applicare le trasformazioni.** Specificare se solo le prossimità all'interno di ogni origine sono confrontati tra di loro o se i confronti sono incondizionati sulla fonte.

**Dimensioni.** Per impostazione predefinita, una soluzione viene calcolata in due dimensioni (minima = 2, Massimo = 2). Si sceglie un numero intero minimo e massimo da 1 al numero di oggetti meno 1 (a patto che il minimo sia inferiore o uguale al massimo). La procedura calcola una soluzione nelle dimensioni massime e poi riduce la dimensionalità in gradini fino a raggiungere il minimo.

## Restrizioni Scaling Multidimensionali

La Finestra di dialogo Limiti consente di posizionare le restrizioni sullo spazio comune.

**Limitazioni su spazio comune.** Specificare il tipo di restrizione desiderata.

- **Nessuna limitazione.** Non sono poste restrizioni sullo spazio comune.
- **Alcune coordinate fisse.** La prima variabile selezionata contiene le coordinate degli oggetti sulla prima dimensione, la seconda variabile corrisponde a coordinate sulla seconda dimensione, e così via. Un valore mancante indica che una coordinata su una dimensione è gratuita. Il numero di variabili selezionate deve essere uguale al numero massimo di dimensioni richieste.
- **Combinazione lineare di variabili indipendenti.** Lo spazio comune è limitato ad essere una combinazione lineare delle variabili selezionate.

**Variabili di restrizione.** Selezionare le variabili che definiscono le restrizioni sullo spazio comune. Se è stata specificata una combinazione lineare, si specifica un intervallo, una trasformazione nominale, ordinale o spline per le variabili di restrizione. In entrambi i casi, il numero di casi per ciascuna variabile deve essere uguale al numero di oggetti.

## Scaling multidimensionale: Opzioni

La finestra di dialogo Opzioni consente di selezionare lo stile di configurazione iniziale, specificare i criteri di iterazione e convergenza e selezionare gli aggiornamenti standard o rilassati.

**Configurazione iniziale.** Scegli una delle seguenti alternative:

- **Simplex.** Gli oggetti sono posizionati alla stessa distanza l'uno dall'altro nella dimensione massima. Una iterazione viene presa per migliorare questa configurazione ad alta quota, seguita da un'operazione di riduzione della quota per ottenere una configurazione iniziale che abbia il numero massimo di dimensioni specificato nella finestra di dialogo Modello.
- **Torgerson.** Viene utilizzata una soluzione di scaling classica come configurazione iniziale.
- **avvio casuale singolo.** Una configurazione viene scelta a caso.
- **Inietta più casuale.** Diverse configurazioni vengono scelte a caso, e la configurazione con lo stress grezzo normalizzato più basso viene utilizzata come configurazione iniziale.
- **Personalizzato.** Si selezionano variabili che contengono le coordinate della propria configurazione iniziale. Il numero di variabili selezionate dovrebbe essere uguale al numero massimo di dimensioni specificato, con la prima variabile corrispondente a coordinate sulla dimensione 1, la seconda variabile corrispondente alle coordinate sulla dimensione 2, e così via. Il numero di casi in ogni variabile dovrebbe essere eguali al numero di oggetti.

**Criteri di iterazione.** Specificare i valori dei criteri di iterazione.

- **Convergenza di stress.** L'algoritmo smetterà di iterare quando la differenza in valori di stress raw normalizzati consecutivi è inferiore al numero che qui viene specificato, che deve sdraiarsi tra 0,0 e 1,0.
- **Lo stress minimo.** L'algoritmo si ferma quando lo stress grezzo normalizzato scende al di sotto del numero che qui viene specificato, che deve sdraiarsi tra il 0,0 e il 1,0.
- **Numero massimo di iterazioni.** L'algoritmo eseguirà il numero di iterazioni specificate, a meno che uno dei criteri di cui sopra non sia soddisfatto per primo.
- **Utilizza aggiornamenti rilassati.** Gli aggiornamenti rilassati accelereranno l'algoritmo; questi aggiornamenti non possono essere utilizzati con modelli diversi dal modello di identità o utilizzati con limitazioni.

## Grafici Scaling multidimensionali, Versione 1

La finestra di dialogo Plotti consente di specificare quali complotti verranno prodotti. In questo argomento viene descritta la finestra di dialogo Pagine se si dispone del formato dati Prossimo in Columns. Per **Pesi Spaziali Individuali, Originali vs. prossimità trasformate** e **Le prossimità Trasformate vs. distanze**, si

specificano le fonti per le quali i tracciati devono essere prodotti. L'elenco delle fonti disponibili è l'elenco delle variabili di prossimità nella finestra di dialogo principale.

**Stress.** Una trama è prodotta da stress grezzo normalizzato versus dimensioni. Questa trama viene prodotta solo se il numero massimo di dimensioni è superiore al numero minimo di dimensioni.

**Spazio comune.** Viene visualizzata una matrice di dispersione delle coordinate dello spazio comune.

**Spazi individuali.** Per ogni sorgente, le coordinate dei singoli spazi vengono visualizzate in matrici scatter. Questo è possibile solo se nella finestra di dialogo Modello viene specificato uno dei singoli modelli di differenze.

**pesi spaziali individuali.** Viene prodotta una trama di dispersione dei singoli pesi spaziali. Questo è possibile solo se nella finestra di dialogo Modello viene specificato uno dei singoli modelli di differenze. Per il modello Euclideo ponderato, i pesi sono stampati in tracciati, con una dimensione su ogni asse. Per il modello Euclideo generalizzato si produce una trama per dimensione, indicando sia la rotazione che la ponderazione di quella dimensione. Il modello di classifica ridotto produce la stessa trama del modello Euclideo generalizzato ma riduce il numero di dimensioni per i singoli spazi.

**Prossimità a prossimità trasformate.** I tracciati sono prodotti delle prossimità originali contro le prossimità trasformate.

**Prossimità trasformate vs. distanze.** Le prossimità trasformate contro le distanze sono tracciate.

**Variabili indipendenti trasformate.** I tracciati di trasformazione sono prodotti per le variabili indipendenti.

**Correlazioni variabili e di dimensione.** Viene visualizzata una trama di correlazioni tra le variabili indipendenti e le dimensioni dello spazio comune.

## Grafici Scaling multidimensionali, Versione 2

La finestra di dialogo Plotti consente di specificare quali complotti verranno prodotti. Questo argomento descrive la finestra di dialogo Pagine se il tuo formato dati è altro che Prossimo in Colonne. Per **Pesi Spaziali Individuali, Originali vs. prossimità trasformate e Le prossimità Trasformate vs. distanze**, si specificano le fonti per le quali i tracciati devono essere prodotti. I numeri di origine inseriti devono essere valori della variabile di origine che viene specificato nella finestra di dialogo principale e devono spaziare dal 1 al numero di sorgenti.

## Output di Scaling multidimensionale

La finestra di dialogo Output consente di controllare la quantità di output visualizzato e di salvarle alcune in file separati.

**Visualizza.** Selezionare uno o più dei seguenti elementi per la visualizzazione:

- **Coordinate spazio comune.** Visualizza le coordinate dello spazio comune.
- **Coordinate dello spazio individuale.** Le coordinate dei singoli spazi vengono visualizzate solo se il modello non è il modello di identità.
- **pesi spaziali individuali.** Visualizza i singoli pesi dello spazio solo se viene specificato uno dei singoli modelli di differenze. A seconda del modello, i pesi dello spazio sono decomposti in pesi di rotazione e pesi di quota, che vengono visualizzati anche.
- **Distanze.** Visualizza le distanze tra gli oggetti nella configurazione.
- **Prossimità trasformate.** Visualizza le prossimità trasformate tra gli oggetti nella configurazione.
- **Dati di input.** Include le prossimità originali e, se presenti, i pesi dei dati, la configurazione iniziale e le coordinate fisse delle variabili indipendenti.
- **Stress per gli avvisi casuali.** Visualizza il seed numero casuale e il valore di stress grezzo normalizzato di ogni inizio casuale.
- **Cronologia iterazioni.** Visualizza la cronologia delle iterazioni dell'algoritmo principale.

- **Misure di stress multiple.** Visualizza diversi valori di stress. La tabella contiene valori per lo stress grezzo normalizzato, Stress-I, Stress - II, S-Stress, Dispersione Contabilizzato (DAF) e Coefficiente di Tucker di Congruenza.
- **Decomposizione dello stress.** Visualizza un oggetto e le sorgenti decomposizione dello stress grezzo normalizzato finale, inclusa la media per oggetto e la media per fonte.
- **Variabili indipendenti trasformate.** Se è stata selezionata una restrizione di combinazione lineare, vengono visualizzate le variabili indipendenti trasformate e i pesi di regressione corrispondenti.
- **Correlazioni variabili e di dimensione.** Se è stata selezionata una restrizione di combinazione lineare, vengono visualizzate le correlazioni tra le variabili indipendenti e le dimensioni dello spazio comune.

**Salva su Nuovo file.** È possibile salvare le coordinate comuni dello spazio, i singoli pesi spaziali, le distanze, le prossimità trasformate e le variabili indipendenti trasformate in file di dati IBM SPSS Statistics separati.

## PROXSCAL Command Aggiuntive Features

Puoi personalizzare il tuo scaling multidimensionale di analisi di prossimità se incolla le selezioni in una finestra di sintassi e modifica la sintassi del comando PROXSCAL risultante. Il linguaggio della sintassi dei comandi consente inoltre di:

- Specificare elenchi variabili separati per le trasformazioni e i tracciati dei residui (con il comando PLOT).
- Specificare liste di origine separate per i singoli pesi dello spazio, le trasformazioni e i tracciati dei residui (con il comando PLOT).
- Specificare un sottoinsieme delle variabili di trasformazione delle variabili indipendenti da visualizzare (con il comando PLOT).

Vedere *Command Syntax Reference* per informazioni dettagliate sulla sintassi.

## Unfolding multidimensionale (PREFSCAL)

La procedura Unfolding Multidimensionale tenta di trovare una scala quantitativa comune che consente di esaminare visivamente le relazioni tra due insiemi di oggetti.

**Esempi.** Hai chiesto a 21 individui di classificare 15 articoli per la colazione in ordine di preferenza, 1 a 15. Utilizzando Unfolding Multidimensionale, è possibile stabilire che gli individui discriminano tra gli articoli per la colazione in due modi primari: tra i batti morbidi e duri, e tra gli articoli da ingrasso e non ingrassanti.

In alternativa, avete chiesto a un gruppo di conducenti di valutare 26 modelli di auto su 10 attributi su una scala di 6 punti che varia da 1 = "non vero per tutti" a 6 = "molto vero". Sconfinati sugli individui, i valori sono presi come similitudini. Utilizzando Unfolding Multidimensionale, trovate clusterings di modelli simili e gli attributi con cui sono maggiormente associati.

**Statistiche e grafici.** La procedura di Unfolding Multidimensionale può produrre una storia di iterazione, misure di stress, decomposizione dello stress, coordinate dello spazio comune, distanze degli oggetti all'interno della configurazione finale, pesi singoli spazi, spazi individuali, prossimità trasformate, tracciati di stress, comuni scattergrafici spaziali, singoli scatterplotti di peso spazio, singoli spazi scattergrafici, tracciati di trasformazione e grafici residui di Shepard.

Considerazioni Dati Multidimensionali Unfolding

**Dati.** I dati sono forniti sotto forma di matrici di prossimità rettangolari. Ogni colonna è considerata un oggetto di colonna separato. Ogni riga di una matrice di prossimità è considerata un oggetto di riga separato. Quando ci sono più fonti di prossimità, le matrici sono impilate.

**Ipotesi.** Devono essere specificate almeno due variabili. Il numero di dimensioni nella soluzione non può superare il numero di oggetti meno uno. Se viene specificata una sola origine, tutti i modelli sono equivalenti al modello di identità; pertanto, l'analisi viene impostata per default sul modello di identità.

Per Ottenere un Unfolding Multidimensionale

1. Dai menu, scegliere:

**Analizzare > Scala > Unfolding multidimensionale (PREFSCAL) ...**

2. Selezionare due o più variabili che identificano le colonne nella matrice di prossimità rettangolare. Ogni variabile rappresenta un oggetto colonna separato.
3. Opzionalmente, selezionare una serie di variabili di peso pari al numero di variabili oggetto colonne. L'ordine delle variabili di peso dovrebbe corrispondere all'ordine degli oggetti di colonna che pesano.
4. Opzionalmente, selezionare una variabile di righe. I valori (o le etichette di valore) di questa variabile vengono utilizzati per etichettare gli oggetti di riga nell'output.
5. Se ci sono più fonti, opzionalmente selezionare una variabile di origini. Il numero di casi nel file dati dovrebbe pari al numero di oggetti di riga che si riferiscono al numero di sorgenti.

Inoltre, è possibile definire un modello per l'unfolding multidimensionale, porre limitazioni allo spazio comune, impostare criteri di convergenza, specificare la configurazione iniziale da utilizzare e scegliere i grafici e l'output.

## Definire un modello di unidimensionazione multidimensionale

La finestra di dialogo Modello consente di specificare un modello di scaling, il suo numero minimo e massimo di dimensioni, la struttura della matrice di prossimità, la trasformazione da utilizzare sulle prossimità, e se le prossimità sono trasformate condizionate sulla riga, condizionate alla sorgente o incondizionatamente sulla sorgente.

**Modello di scala.** Scegli tra le seguenti alternative:

- **Identità.** Tutte le fonti hanno la stessa configurazione.
- **Euclideo Ponderato.** Questo modello è un modello di differenze individuale. Ogni sorgente ha uno spazio individuale in cui ogni dimensione dello spazio comune è ponderata differentemente.
- **Euclideo generalizzato.** Questo modello è un modello di differenze individuale. Ogni sorgente ha uno spazio individuale uguale a una rotazione dello spazio comune, seguito da una ponderazione differenziale delle dimensioni.

**Prossimità.** Specificare se la matrice di prossimità contiene misure di similitudine o dissimilarità.

**Dimensioni.** Per impostazione predefinita, una soluzione viene calcolata in due dimensioni (minima = 2, Massimo = 2). È possibile scegliere un numero intero minimo e massimo da 1 al numero di oggetti meno 1 finché il minimo è inferiore o uguale al massimo. La procedura calcola una soluzione nella dimensionalità massima e poi riduce la dimensionalità in step fino a raggiungere il livello più basso.

**Trasformazioni di prossimità.** Scegli tra le seguenti alternative:

- **Messuno.** Le prossimità non sono trasformate. È possibile opzionalmente selezionare **Includi intercetta**, in tal caso le prossimità possono essere spostate da un termine costante.
- **Lineare.** Le prossimità trasformate sono proporzionali alle prossimità originali; cioè la funzione di trasformazione stima una pendenza e l'intercettazione è fissata a 0. Questa è anche chiamata trasformazione di rapporto. È possibile opzionalmente selezionare **Includi intercetta**, in tal caso le prossimità possono anche essere spostate da un termine costante. Questa viene anche chiamata trasformazione dell'intervallo.
- **Spline.** Le prossimità trasformate sono una fluida trasformazione polinomica non decrescente delle prossimità originali. È possibile specificare il grado del polinomio e il numero di nodi interni. È possibile opzionalmente selezionare **Includi intercetta**, in tal caso le prossimità possono anche essere spostate da un termine costante.
- **Smooth.** Le prossimità trasformate hanno lo stesso ordine delle prossimità originali, tra cui una restrizione che prende in considerazione le differenze tra i valori successivi. Il risultato è una trasformazione "ordinata ordinale". È possibile specificare se le prossimità legate devono essere tenute legate o lasciate svincolate.
- **Ordinale.** Le prossimità trasformate hanno lo stesso ordine delle prossimità originali. È possibile specificare se le prossimità legate devono essere tenute legate o lasciate svincolate.

**Applicare le trasformazioni.** Specificare se solo le prossimità all'interno di ciascuna riga vengono confrontati tra di loro, o solo le prossimità all'interno di ogni sorgente vengono confrontati tra loro, oppure i confronti sono incondizionati sulla riga o sull'origine; cioè se le trasformazioni vengono eseguite per riga, per fonte o su tutte le prossimità in una volta.

## Limitazioni Unfolding Multidimensionali

La Finestra di dialogo Limiti consente di posizionare le restrizioni sullo spazio comune.

**Limitazioni su spazio comune.** È possibile scegliere di fissare le coordinate di oggetti di riga e / o colonna nello spazio comune.

**Variabili di restrizione della riga / colonna.** Scegliere il file contenente le limitazioni e selezionare le variabili che definiscono le limitazioni allo spazio comune. La prima variabile selezionata contiene le coordinate degli oggetti sulla prima dimensione, la seconda variabile corrisponde a coordinate sulla seconda dimensione, e così via. Un valore mancante indica che una coordinata su una dimensione è gratuita. Il numero di variabili selezionate deve essere uguale al numero massimo di dimensioni richieste. Il numero di casi per ogni variabile deve essere uguale al numero di oggetti.

## Opzioni di sfolding multidimensionali

La finestra di dialogo Opzioni consente di selezionare lo stile di configurazione iniziale, specificare l'iterazione e i criteri di convergenza e impostare il termine di penalità per lo stress.

**Configurazione iniziale.** Scegli una delle seguenti alternative:

- **Classico.** La matrice di prossimità rettangolare viene utilizzata per integrare gli intralocchi (valori tra righe e tra colonne) della matrice MDS simmetrica completa. Una volta formata la matrice completa, viene utilizzata una soluzione di scaling classica come configurazione iniziale. Gli intralocchi possono essere compilati tramite imputazione utilizzando la disuguaglianza triangolare o le distanze Spearman.
- **Ross - Cliff.** La partenza di Ross - Cliff utilizza i risultati di una singolare decomposizione del valore sulla doppia matrice di prossimità centrata e quadrata come valori iniziali per gli oggetti di riga e colonna.
- **Corrispondenza.** La corrispondenza inizia utilizza i risultati di un'analisi della corrispondenza sui dati invertiti (similitudini invece di dissimilarità), con la normalizzazione simmetrica dei punteggi di riga e colonna.
- **Centroidi.** La procedura inizia posizionando gli oggetti di riga nella configurazione utilizzando una decomposizione eigena. Poi gli oggetti della colonna sono posizionati al centroide delle scelte specificate. Per il numero di scelte, specificare un intero positivo compreso tra 1 e il numero di variabili di prossimità.
- **Inietta più casuale.** Le soluzioni sono calcolate per diverse configurazioni iniziali scelte a caso, e quella con lo stress penalizzato più basso viene mostrata come la soluzione migliore.
- **Personalizzato.** È possibile selezionare le variabili che contengono le coordinate della propria configurazione iniziale. Il numero di variabili selezionate dovrebbe essere uguale al numero massimo di dimensioni specificato, con la prima variabile corrispondente a coordinate sulla dimensione 1, la seconda variabile corrispondente alle coordinate sulla dimensione 2, e così via. Il numero di casi in ogni variabile dovrebbe essere pari al numero combinato di oggetti di righe e colonne. Le coordinate di riga e colonna devono essere impilate, con le coordinate della colonna seguendo le coordinate di riga.

**Criteri di iterazione.** Specificare i valori dei criteri di iterazione.

- **Convergenza di stress.** L'algoritmo smetterà di iterare quando la differenza relativa in valori di stress penalizzati consecutivi è inferiore al numero specificato qui, che deve essere non negativo.
- **Lo stress minimo.** L'algoritmo si ferma quando lo stress penalizzato scende al di sotto del numero specificato qui, che deve essere non negativo.
- **Numero massimo di iterazioni.** L'algoritmo eseguirà il numero di iterazioni qui specificate a meno che uno dei criteri di cui sopra non sia soddisfatto per primo.

**Termine di penalità.** L'algoritmo tenta di minimizzare lo stress penalizzato, una misura di bontà di adattamento uguale al prodotto di Kruskal's Stress-I e un termine di penalità basato sul coefficiente di

variazione delle prossimità trasformate. Questi controlli consentono di impostare la forza e la portata del termine di penalità.

- **Forza.** Più piccolo è il valore del parametro di forza, più forte è la pena. Specificare un valore compreso tra 0,0 e 1,0.
- **Intervallo.** Questo parametro imposta il momento in cui la penalità diventa attiva. Se impostato su 0,0, la penalità è inattiva. Aumentare il valore provoca l'algoritmo alla ricerca di una soluzione con maggiore variazione tra le prossimità trasformate. Specificare un valore non negativo.

## Grafici di sfolding multidimensionali

La finestra di dialogo Plotti consente di specificare quali complotti verranno prodotti.

**Grafici.** Sono disponibili i seguenti grafici:

- **Più inizia.** Visualizza un istogramma impilato di stress penalizzato che visualizza sia stress che penalità.
- **Spazio comune iniziale.** Visualizza una matrice di dispersione delle coordinate dello spazio comune iniziale.
- **Stress per dimensione.** Produce un lineplot di stress penalizzato contro dimensionalità. Questa trama viene prodotta solo se il numero massimo di dimensioni è superiore al numero minimo di dimensioni.
- **Spazio comune finale.** Viene visualizzata una matrice di dispersione delle coordinate dello spazio comune.
- **pesi dello spazio.** Viene prodotta una trama di dispersione dei singoli pesi spaziali. Questo è possibile solo se nella finestra di dialogo Modello viene specificato uno dei singoli modelli di differenze. Per il modello Euclideo ponderato, i pesi per tutte le fonti sono visualizzati in una trama, con una dimensione su ogni asse. Per il modello Euclideo generalizzato viene prodotta una trama per dimensione, indicando sia la rotazione che la ponderazione di quella dimensione per ogni sorgente.
- **Spazi individuali.** Viene visualizzata una matrice di dispersione di coordinate dello spazio individuale di ogni sorgente. Questo è possibile solo se nella finestra di dialogo Modello viene specificato uno dei singoli modelli di differenze.
- **Complotti di trasformazione.** Una trama scatter è prodotta delle prossimità originali contro le prossimità trasformate. A seconda di come vengono applicate le trasformazioni, viene assegnato un colore separato ad ogni riga o fonte. Una trasformazione incondizionata produce un solo colore.
- **grafici di Shepard.** Le prossimità originali contro entrambe le prossimità trasformate e le distanze. Le distanze sono indicate da punti, e le prossimità trasformate sono indicate da una linea. A seconda di come vengono applicate le trasformazioni, viene prodotta una linea separata per ogni riga o fonte. Una trasformazione incondizionata produce una sola linea.
- **Scatto di adattamento.** Viene visualizzata una trama di dispersione delle prossimità trasformate contro le distanze. Un colore separato viene assegnato ad ogni origine se vengono specificate più fonti.
- **Grafici dei residui.** Viene visualizzata una trama di dispersione delle prossimità trasformate contro i residui (prossimità trasformate meno le distanze). Un colore separato viene assegnato ad ogni origine se vengono specificate più fonti.

**Stili oggetto di riga.** Questi ti danno un ulteriore controllo della visualizzazione degli oggetti di riga nei tracciati. I valori della variabile colori opzionali sono utilizzati per ciclo attraverso tutti i colori. I valori della variabile di marcatori opzionali sono utilizzati per ciclo attraverso tutti i possibili marcatori.

**Pagine di origine.** Per **Spazi individuali**, **Scatterplot di fite Pitti dei residui**- e se le trasformazioni vengono applicate per fonte, per **lotti di trasformazione** e **Scherni di Shepard**- è possibile specificare le fonti per le quali i tracciati devono essere prodotti. I numeri di origine inseriti devono essere valori della variabile di origine specificati nella finestra di dialogo principale e spaziano dal 1 al numero di sorgenti.

**Pagine riga.** Se le trasformazioni vengono applicate per riga, per **grafici di trasformazione** e **Scherni di Shepard** è possibile specificare la riga per la quale devono essere prodotti i tracciati. I numeri di riga inseriti devono spaziare da 1 al numero di righe.

## Output unificato multidimensionale

La finestra di dialogo Output consente di controllare la quantità di output visualizzato e di salvarle alcune in file separati.

**Visualizza.** Seleziona una o più delle seguenti per la visualizzazione:

- **Dati di input.** Include le prossimità originali e, se presenti, i pesi dei dati, la configurazione iniziale e le coordinate fisse.
- **Più inizia.** Visualizza il seed numero casuale e il valore di stress penalizzato di ogni inizio casuale.
- **Dati iniziali.** Visualizza le coordinate dello spazio comune iniziale.
- **Cronologia iterazioni.** Visualizza la cronologia delle iterazioni dell'algoritmo principale.
- **Misure di adattamento.** Visualizza diverse misure. La tabella contiene diverse misure di bontà, badante - di - fit, correlazione, variazione e nondegeneracy.
- **Decomposizione dello stress.** Visualizza un oggetto, righe e sorgenti decomposizione di stress penalizzato, tra cui riga, colonna e mezzi di origine e deviazioni standard.
- **Prossimità trasformate.** Visualizza le prossimità trasformate.
- **Spazio comune finale.** Visualizza le coordinate dello spazio comune.
- **pesi dello spazio.** Visualizza i singoli pesi dello spazio. Questa opzione è disponibile solo se viene specificato uno dei singoli modelli di differenze. A seconda del modello, i pesi dello spazio sono decomposti in pesi di rotazione e pesi di quota, che vengono visualizzati anche.
- **Spazi individuali.** Vengono visualizzate le coordinate dei singoli spazi. Questa opzione è disponibile solo se viene specificato uno dei singoli modelli di differenze.
- **Distanze installate.** Visualizza le distanze tra gli oggetti nella configurazione.

**Salva su Nuovo file.** È possibile salvare le coordinate dello spazio comune, i singoli pesi spaziali, le distanze e le prossimità trasformate in file di dati IBM SPSS Statistics separati.

## Funzioni Aggiuntive di comando PREFSCAL

È possibile personalizzare l'analisi Multidimensionale di prossimità di prossimità se incolla le selezioni in una finestra di sintassi e modificare la sintassi del comando PREFSCAL risultante. Il linguaggio della sintassi dei comandi consente inoltre di:

- Specificare più liste di origine per gli spazi individuali, Scatterlotti di fit, e i tracciati dei residui - e in caso di trasformazioni condizionate a matrice, per i complotti di Trasformazione e i tracciati di Shepard - quando sono disponibili più fonti (con il comando PLOT).
- Specificare più liste di righe per i complotti di Trasformazione e i tracciati di Shepard in caso di trasformazioni condizionate a righe (con il comando PLOT).
- Specificare un numero di righe invece di una variabile ID riga (con il comando INPUT).
- Specificare un numero di fonti invece di una variabile ID di origine (con il comando INPUT).

Vedere *Command Syntax Reference* per informazioni dettagliate sulla sintassi.



## Informazioni particolari

---

Queste informazioni sono state sviluppate per prodotti e servizi offerti negli Stati Uniti. Questo materiale potrebbe essere disponibile da IBM in altre lingue. Tuttavia, all'utente potrebbe essere richiesto di possedere una copia del prodotto o una versione del prodotto in tale lingua per accedervi.

IBM può non offrire i prodotti, i servizi o le funzioni presentati in questo documento in altri paesi. Consultare il proprio rappresentante locale IBM per informazioni sui prodotti ed i servizi attualmente disponibili nella propria zona. Qualsiasi riferimento ad un prodotto, programma o servizio IBM non implica o intende dichiarare che solo quel prodotto, programma o servizio IBM può essere utilizzato. In sostituzione a quelli forniti da IBM, è possibile usare prodotti, programmi o servizi funzionalmente equivalenti che non comportino violazione dei diritti di proprietà intellettuale o di altri diritti di IBM. Tuttavia, è responsabilità dell'utente valutare e verificare il funzionamento di qualsiasi prodotto, programma o servizio non IBM.

IBM può avere applicazioni di brevetti o brevetti in corso relativi all'argomento descritto in questo documento. La fornitura di questa documentazione non concede alcuna licenza su questi brevetti. È possibile inviare per iscritto richieste di licenze a:

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

Per richieste di licenze relative ad informazioni double-byte (DBCS), contattare il Dipartimento di Proprietà Intellettuale IBM nel proprio paese o inviare richieste per iscritto a:

*Intellectual Property Licensing*  
*Legal and Intellectual Property Law*  
*IBM Japan Ltd.*  
*19-21, Nihonbashi-Hakozakicho, Chuo-ku*  
*Tokyo 103-8510, Japan*

IBM (INTERNATIONAL BUSINESS MACHINES CORPORATION) FORNISCE LA PRESENTE PUBBLICAZIONE "NELLO STATO IN CUI SI TROVA" SENZA GARANZIE DI ALCUN TIPO, ESPRESSE O IMPLICITE, IVI INCLUSE, A TITOLO DI ESEMPIO, GARANZIE IMPLICITE DI NON VIOLAZIONE, DI COMMERCIALIZZABILITÀ E DI IDONEITÀ PER UNO SCOPO PARTICOLARE. Alcune giurisdizioni non consentono la rinuncia ad alcune garanzie espresse o implicite in determinate transazioni, pertanto, la presente dichiarazione può non essere applicabile.

Questa pubblicazione potrebbe contenere imprecisioni tecniche o errori tipografici. Le modifiche vengono apportate periodicamente alle informazioni qui contenute; questi cambiamenti saranno incorporati nelle nuove edizioni della pubblicazione. IBM si riserva il diritto di apportare miglioramenti e/o modifiche al prodotto o al programma descritto nel manuale in qualsiasi momento e senza preavviso.

I riferimenti in queste informazioni a siti Web non IBM vengono forniti solo per comodità e non implicano in alcun modo l'approvazione di tali siti web. I materiali disponibili su tali siti Web non fanno parte del materiale relativo a questo prodotto IBM e l'utilizzo di questi è a discrezione dell'utente.

IBM può utilizzare o distribuire qualsiasi informazione fornita in qualsiasi modo ritenga appropriato senza incorrere in alcun obbligo verso l'utente.

Coloro che detengano la licenza su questo programma e desiderano avere informazioni su di esso allo scopo di consentire: (i) uno scambio di informazioni tra programmi indipendenti ed altri (compreso questo) e (ii) l'utilizzo reciproco di tali informazioni, dovrebbe rivolgersi a:

*IBM Director of Licensing*  
*IBM Corporation*

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

Tali informazioni potrebbero essere disponibili secondo termini e condizioni appropriati compreso, in alcuni casi, il pagamento di un corrispettivo.

Il programma concesso in licenza descritto nel presente documento e tutto il materiale concesso in licenza disponibile sono forniti da IBM in base alle clausole dell'Accordo per Clienti IBM (IBM Customer Agreement), dell'IBM IPLA (IBM International Program License Agreement) o qualsiasi altro accordo equivalente tra le parti.

I dati delle prestazioni e gli esempi client citati vengono presentati solo a scopo illustrativo. Gli effettivi risultati delle prestazioni possono variare in base alle configurazioni e alle condizioni operative specifiche.

Le informazioni relative a prodotti non IBM sono ottenute dai fornitori di quei prodotti, dagli annunci pubblicati e da altre fonti disponibili al pubblico. IBM non ha testato quei prodotti e non può confermarne la precisione della prestazione, la compatibilità o qualsiasi altro reclamo relativo ai prodotti non IBM. Le domande sulle funzionalità dei prodotti non IBM devono essere indirizzate ai fornitori di tali prodotti.

Qualsiasi affermazione relativa agli obiettivi e alla direzione futura di IBM è soggetta a modifica o revoca senza preavviso e concerne esclusivamente gli scopi dell'azienda.

Queste informazioni contengono esempi di dati e report utilizzati nelle operazioni aziendali quotidiane. Pertanto, per maggiore completezza, gli esempi includono nomi di persone, società, marchi e prodotti. Tutti i nomi contenuti nel manuale sono fittizi e ogni riferimento a persone o aziende reali è puramente casuale.

#### LICENZA DI COPYRIGHT:

Queste informazioni contengono programmi campione di applicazione nella lingua di origine, i quali illustrano le tecniche di programmazione su varie piattaforme operative. È possibile copiare, modificare e distribuire questi programmi di esempio sotto qualsiasi forma senza alcun pagamento a IBM, allo scopo di sviluppare, utilizzare, commercializzare o distribuire i programmi applicativi in conformità alle API (application programming interface) a seconda della piattaforma operativa per cui i programmi di esempio sono stati scritti. Questi esempi non sono stati testati approfonditamente tenendo conto di tutte le condizioni possibili. IBM, quindi, non può garantire o sottintendere l'affidabilità, l'utilità o il funzionamento di questi programmi. I programmi di esempio sono forniti "COSÌ COME SONO", senza garanzie di alcun tipo. IBM non intende essere responsabile per alcun danno derivante dall'uso dei programmi di esempio.

Ogni copia o qualsiasi parte di questi programmi di esempio o qualsiasi lavoro derivato, devono contenere le seguenti informazioni relative alle leggi sul diritto d'autore:

© Copyright IBM Corp. 2021. Le porzioni di questo codice derivano da IBM Corp. Programmi Di Esempio.

© Copyright IBM Corp. 1989 - 2021. Tutti i diritti riservati.

## Marchi

---

IBM, il logo IBM e [ibm.com](http://ibm.com) sono marchi o marchi registrati di International Business Machines Corp., registrati in molte giurisdizioni in tutto il mondo. Altri nomi di prodotti e servizi possono essere marchi di IBM o di altre società. Un elenco corrente dei marchi IBM è disponibile sul web in "Copyright and trademark information" all'indirizzo [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml).

Adobe, il logo Adobe, PostScript e il logo PostScript sono marchi o marchi registrati di Adobe Systems Incorporated negli Stati Uniti e/o in altri paesi.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium e Pentium sono marchi o marchi registrati di Intel Corporation o relative controllate negli Stati Uniti e altri paesi.

Linux è un marchio registrato di Linus Torvalds negli Stati Uniti e/o in altri paesi.

Microsoft, Windows, Windows NT e il logo Windows sono marchi di Microsoft Corporation negli Stati Uniti e/o in altri paesi.

UNIX è un marchio della The Open Group negli Stati Uniti e/o negli altri paesi.

Java e tutti i marchi e i logo basati su Java sono marchi o marchi registrati di Oracle e/o associate.



---

# Indice analitico

## A

- Adatta
  - in Analisi Correlazione Canonica non lineare [23](#)
- aggiornamenti rilassati
  - in scaling multidimensionale [37](#)
- Analisi Componenti Principali Categoriali
  - bootstrap [20](#)
  - funzioni aggiuntive del comando [21](#)
  - Livello di scala ottimale [16](#)
  - Salva variabili [19](#)
- Analisi della correlazione canonica non lineare (OVERALS)
  - funzioni aggiuntive del comando [24](#)
  - grafici [21](#)
  - Statistiche [21](#)
- Analisi delle corrispondenze
  - funzioni aggiuntive del comando [28](#)
  - grafici [24](#)
  - Statistiche [24](#)
- Analisi delle corrispondenze multiple
  - funzioni aggiuntive del comando [33](#)
  - Livello di scala ottimale [29](#)
  - Salva variabili [32](#)
- ANOVA
  - in Regressione categoriale [13](#)

## B

- Baricentri
  - in Analisi Correlazione Canonica non lineare [23](#)
- bigrafici
  - in Analisi Categorie Principali di categoria [19](#)
  - in Analisi Corrispondenza Multipla [32](#)
  - in Analisi della corrispondenza [28](#)
- bootstrap
  - Analisi Componenti Principali Categoriali [20](#)

## C

- Caricamenti di componenti
  - in Analisi Categorie Principali di categoria [19](#)
  - in Analisi Correlazione Canonica non lineare [23](#)
- coefficienti di regressione.
  - in Regressione categoriale [13](#)
- complotti di avvio multipli
  - in Unfolding Multidimensionale [42](#)
- complotti di caricamenti dei componenti
  - in Analisi Categorie Principali di categoria [20](#)
- complotti di pesi singoli spazi
  - in scaling multidimensionale [37](#)
  - in Unfolding Multidimensionale [42](#)
- complotti punti oggetto
  - in Analisi Categorie Principali di categoria [19](#)
  - in Analisi Corrispondenza Multipla [32](#)
- complotti singoli spazi
  - in scaling multidimensionale [37](#)
  - in Unfolding Multidimensionale [42](#)

- complotti spazi comuni
  - in scaling multidimensionale [37](#)
  - in Unfolding Multidimensionale [42](#)
- complotti spazi comuni finali
  - in Unfolding Multidimensionale [42](#)
- Configurazione iniziale
  - in Analisi Correlazione Canonica non lineare [23](#)
  - in Regressione categoriale [12](#)
  - in scaling multidimensionale [37](#)
  - in Unfolding Multidimensionale [41](#)
- Coordinate spazio comune
  - in scaling multidimensionale [38](#)
  - in Unfolding Multidimensionale [43](#)
- Coordinate spazio individuale
  - in Unfolding Multidimensionale [43](#)
- correlazioni
  - in scaling multidimensionale [38](#)
- Criteri di iterazione
  - in scaling multidimensionale [37](#)
  - in Unfolding Multidimensionale [41](#)
- cronologia delle iterazioni
  - in Analisi Categorie Principali di categoria [19](#)
  - in Analisi Corrispondenza Multipla [31](#)
  - in scaling multidimensionale [38](#)
  - in Unfolding Multidimensionale [43](#)

## D

- dimensioni
  - in Analisi della corrispondenza [26](#)
- discretizzazione
  - in Analisi Categorie Principali di categoria [16](#)
  - in Analisi Corrispondenza Multipla [30](#)
  - in Regressione categoriale [11](#)
- distanze
  - in scaling multidimensionale [38](#)
  - in Unfolding Multidimensionale [43](#)

## G

- grafici
  - in Analisi Correlazione Canonica non lineare [23](#)
  - in Analisi della corrispondenza [28](#)
  - in Regressione categoriale [14](#)
  - in scaling multidimensionale [37](#), [38](#)
- grafici a centroidi proiettati
  - in Analisi Categorie Principali di categoria [20](#)
- Grafici correlazioni
  - in scaling multidimensionale [37](#)
- Grafici dei residui
  - in Unfolding Multidimensionale [42](#)
- Grafici di categoria
  - in Analisi Categorie Principali di categoria [20](#)
  - in Analisi Corrispondenza Multipla [32](#)
- Grafici di categoria congiunti
  - in Analisi Categorie Principali di categoria [20](#)
  - in Analisi Corrispondenza Multipla [32](#)

Grafici di Shepard  
in [Unfolding Multidimensionale 42](#)  
Grafici di trasformazione  
in [Analisi Categorie Principali di categoria 20](#)  
in [Analisi Corrispondenza Multipla 32](#)  
in [scaling multidimensionale 37](#)  
in [Unfolding Multidimensionale 42](#)  
Grafici pesi dello spazio  
in [Unfolding Multidimensionale 42](#)  
Grafico a dispersione dell'adattamento  
in [Unfolding Multidimensionale 42](#)

## I

Inerzia  
in [Analisi della corrispondenza 27](#)

## L

Lasso  
in [Regressione categoriale 13](#)  
Limitazioni allo spazio comune  
in [Unfolding Multidimensionale 41](#)  
Livello di scala ottimale  
in [Analisi Categorie Principali di categoria 16](#)  
in [Analisi Corrispondenza Multipla 29](#)

## M

matrice di correlazione  
in [Analisi Categorie Principali di categoria 19](#)  
in [Analisi Corrispondenza Multipla 31](#)  
misure della distanza  
in [Analisi della corrispondenza 26](#)  
Misure di discriminazione  
in [Analisi Corrispondenza Multipla 31](#)  
misure di discriminazione complotti  
in [Analisi Corrispondenza Multipla 32](#)  
misure di stress  
in [scaling multidimensionale 38](#)  
in [Unfolding Multidimensionale 43](#)  
modello di identità  
in [Unfolding Multidimensionale 40](#)  
Modello di scala  
in [Unfolding Multidimensionale 40](#)  
modello euclideo generalizzato  
in [Unfolding Multidimensionale 40](#)  
modello euclideo ponderato  
in [Unfolding Multidimensionale 40](#)

## N

normalizzazione  
in [Analisi della corrispondenza 26](#)

## O

Oggetti supplementari  
in [Regressione categoriale 12](#)

## P

pesi  
in [Analisi Correlazione Canonica non lineare 23](#)  
Pesi dello spazio individuale  
in [scaling multidimensionale 38](#)  
in [Unfolding Multidimensionale 43](#)  
Peso variabile  
in [Analisi Categorie Principali di categoria 16](#)  
in [Analisi Corrispondenza Multipla 29](#)  
Prefsal [39](#)  
Prossimità trasformate  
in [scaling multidimensionale 38](#)  
in [Unfolding Multidimensionale 43](#)  
Punteggi oggetto  
in [Analisi Categorie Principali di categoria 19](#)  
in [Analisi Correlazione Canonica non lineare 23](#)  
in [Analisi Corrispondenza Multipla 31](#)

## Q

Quantificazioni di categoria  
in [Analisi Categorie Principali di categoria 19](#)  
in [Analisi Correlazione Canonica non lineare 23](#)  
in [Analisi Corrispondenza Multipla 31](#)  
in [Regressione categoriale 13](#)

## R

R multiplo  
in [Regressione categoriale 13](#)  
Regressione categoriale  
funzioni aggiuntive del comando [14](#)  
grafici [10](#)  
Livello di scala ottimale [11](#)  
Regolarizzazione [13](#)  
Salva [14](#)  
Statistiche [10](#)  
Regressione Ridge  
in [Regressione categoriale 13](#)  
restrizioni  
in [scaling multidimensionale 37](#)  
Rete elastica  
in [Regressione categoriale 13](#)

## S

Scaling multidimensionale  
funzioni aggiuntive del comando [39](#)  
grafici [33](#), [37](#), [38](#)  
modello [36](#)  
Opzioni [37](#)  
output [38](#)  
restrizioni [37](#)  
Statistiche [33](#)  
standardizzazione  
in [Analisi della corrispondenza 26](#)  
statistica descrittiva  
in [Regressione categoriale 13](#)  
statistiche di fiducia  
in [Analisi della corrispondenza 27](#)

## T

- Termine di penalità
  - in Unfolding Multidimensionale [41](#)
- tracciati di stress
  - in scaling multidimensionale [37](#)
  - in Unfolding Multidimensionale [42](#)
- tracciati spazi comuni iniziali
  - in Unfolding Multidimensionale [42](#)
- Trasformazioni di prossimità
  - in Unfolding Multidimensionale [40](#)
- triplotti
  - in Analisi Categorie Principali di categoria [19](#)

## U

- Unfolding multidimensionale
  - funzioni aggiuntive del comando [43](#)
  - grafici [39](#), [42](#)
  - Limitazioni allo spazio comune [41](#)
  - modello [40](#)
  - Opzioni [41](#)
  - output [43](#)
  - Statistiche [39](#)

## V

- valori mancanti
  - in Analisi Categorie Principali di categoria [17](#)
  - in Analisi Corrispondenza Multipla [30](#)
  - in Regressione categoriale [12](#)
- Variabili indipendenti trasformate
  - in scaling multidimensionale [38](#)
- Varianza spiegata
  - in Analisi Categorie Principali di categoria [19](#)





