

IBM SPSS Direct Marketing 30

IBM

Nota

Prima di utilizzare queste informazioni e il prodotto che supportano, leggere le informazioni in [“Informazioni particolari” a pagina 19](#).

Informazioni sul prodotto

Questa edizione si applica alla versione 30, release 0, modifica 0 di IBM® SPSS Statistiche e a tutte le release e modifiche successive se non diversamente indicato in nuove edizioni.

© **Copyright International Business Machines Corporation .**

Indice

Capitolo 1. Marketing diretto.....	1
Direct marketing.....	1
Analisi RFM.....	1
Punteggi RFM dai dati delle transazioni.....	2
Punteggi RFM da Customer Data.....	2
Binning RFM.....	3
Risparmio RFM Punteggio da Dati di transazione.....	5
Salvataggio Dei Punteggi RFM dai Dati del Cliente.....	5
Output RFM.....	6
Analisi del cluster.....	7
Impostazioni.....	8
Profili potenziali clienti.....	8
Impostazioni.....	10
Creazione di un campo di risposta categoriale.....	10
Tassi Di Risposta Del Codice Postale.....	11
Impostazioni.....	12
Creazione di un campo di risposta categoriale.....	13
Propensione all'acquisto.....	14
Impostazioni.....	16
Creazione di un campo di risposta categoriale.....	17
Test package di controllo.....	17
Informazioni particolari.....	19
Marchi.....	20
Indice analitico.....	23

Capitolo 1. Marketing diretto

Le seguenti funzioni di marketing diretto sono incluse in SPSS Statistiche Premium Edition o l'opzione Direct Marketing.

Direct marketing

L'opzione Direct Marketing include una serie di strumenti pensati per migliorare i risultati delle campagne di direct marketing identificando le caratteristiche demografiche, di acquisto e altre caratteristiche che definiscono vari gruppi di consumatori e puntando su gruppi specifici per massimizzare i tassi di risposta positivi.

Analisi RFM. Questa tecnica identifica i clienti esistenti che hanno più probabilità di rispondere ad una nuova offerta.

Analisi del cluster. Questo è uno strumento esplorativo progettato per rivelare raggruppamenti naturali (o cluster) all'interno dei tuoi dati. Ad esempio, è in grado di identificare gruppi diversi di clienti in base a varie caratteristiche demografiche e di acquisto.

Profili di prospettiva. Questa tecnica utilizza i risultati di una campagna precedente o di test per creare profili descrittivi. È possibile utilizzare i profili per definire gruppi specifici di contatti nelle campagne successive. Per ulteriori informazioni, consultare l'argomento [“Profili potenziali clienti”](#) a pagina 8.

Tassi di risposta del codice postale. Questa tecnica utilizza i risultati di una campagna precedente per calcolare i tassi di risposta relativi al codice postale. Tali tassi possono essere utilizzati per identificare codici postali specifici nelle campagne successive. Per ulteriori informazioni, consultare l'argomento [“Tassi Di Risposta Del Codice Postale”](#) a pagina 11.

Propensione all'acquisto. Questa tecnica utilizza i risultati di una mailing test o di una campagna precedente per generare punteggi di propensione. I punteggi indicano i contatti che risponderanno con maggiore probabilità. Per ulteriori informazioni, consultare l'argomento [“Propensione all'acquisto”](#) a pagina 14.

Test del pacchetto di controllo. Questa tecnica confronta le campagne di marketing per rilevare eventuali differenze significative in termini di efficacia tra i diversi package o offerte. Per ulteriori informazioni, consultare l'argomento [“Test package di controllo”](#) a pagina 17.

Analisi RFM

L'analisi RFM è una tecnica utilizzata per identificare i clienti esistenti che hanno più probabilità di rispondere ad una nuova offerta. Questa tecnica è comunemente utilizzata nel marketing diretto. L'analisi RFM si basa sulla seguente semplice teoria:

- Il fattore più importante per identificare i clienti che probabilmente risponderanno a una nuova offerta è **recenza**. I clienti che hanno acquistato più recentemente hanno più probabilità di acquistare di nuovo rispetto ai clienti che hanno acquistato ulteriormente in passato.
- Il secondo fattore più importante è **frequenza**. I clienti che hanno fatto più acquisti in passato hanno più probabilità di rispondere di quelli che hanno fatto meno acquisti.
- Il terzo fattore più importante è l'importo totale speso, che viene indicato come **monetario**. I clienti che hanno speso di più (in totale per tutti gli acquisti) in passato hanno più probabilità di rispondere di quelli che hanno speso di meno.

Come funziona l'analisi RFM

- Ai clienti viene assegnato un punteggio di recenza basato sulla data dell'acquisto più recente o dell'intervallo di tempo dall'acquisto più recente. Questo punteggio si basa su una semplice graduatoria dei valori di recenza in un numero ridotto di categorie. Ad esempio, se si utilizzano cinque categorie, i

clienti con le date di acquisto più recenti ricevono una graduatoria di recenza del 5, e quelle con date di acquisto in passato ricevono una graduatoria di recenza del 1.

- In una moda simile, ai clienti viene quindi assegnata una graduatoria di frequenza, con valori più elevati che rappresentano una maggiore frequenza degli acquisti. Ad esempio, in un regime di graduatoria di cinque categorie, i clienti che acquistano più spesso ricevono una graduatoria di frequenza di 5.
- Infine, i clienti sono classificati per valore monetario, con i massimi valori monetari che ricevono la classifica più alta. Continuando l'esempio a cinque categorie, i clienti che hanno speso di più riceverebbero una graduatoria monetaria del 5.

Il risultato è di quattro punteggi per ogni cliente: recenza, frequenza, monetario e punteggio combinato RFM, ovvero semplicemente i tre punteggi singoli concatenati in un unico valore. I clienti "migliori" (quelli che più probabilmente rispondono a un'offerta) sono quelli con i più alti punteggi RFM combinati. Ad esempio, in una graduatoria di cinque categorie, c'è un totale di 125 possibili punteggi RFM combinati e il punteggio RFM combinato più alto è di 555.

Considerazioni sui dati

- Se le righe di dati rappresentano le transazioni (ogni riga rappresenta una singola transazione e ci possono essere più transazioni per ogni cliente), utilizzare RFM da Transactions. Per ulteriori informazioni, consultare l'argomento [“Punteggi RFM dai dati delle transazioni”](#) a pagina 2.
- Se le righe di dati rappresentano i clienti con informazioni di riepilogo per tutte le transazioni (con colonne che contengono valori per l'importo totale speso, numero totale di transazioni e data di transazione più recente), utilizzare RFM da Customer Data. Per ulteriori informazioni, consultare l'argomento [“Punteggi RFM da Customer Data”](#) a pagina 2.

Punteggi RFM dai dati delle transazioni

Considerazioni sui dati

Il dataset deve contenere variabili che contengono le seguenti informazioni:

- Una variabile o combinazione di variabili che identificano ogni caso (cliente).
- Una variabile con la data di ogni transazione.
- Una variabile con il valore monetario di ogni transazione.

Creazione Dei Punteggi RFM dai dati delle transazioni

1. Dai menu, scegliere:

Marketing Diretto > Scegliere Tecnica

2. Selezionare **Guida identificare i miei migliori contatti (Analisi RFM)** e fare clic su **Continua**.

3. Selezionare **Dati sulle transazioni** e fare clic su **Continua**.

4. Selezionare la variabile che contiene date di transazione.

5. Selezionare la variabile che contiene l'importo monetario per ogni transazione.

6. Selezionare il metodo per riepilogare gli importi delle transazioni per ciascun cliente: Totale (somma di tutte le transazioni), media, mediana o massima (importo massimo della transazione).

7. Selezionare la variabile o la combinazione di variabili che identifica univocamente ogni cliente. Ad esempio, i casi potrebbero essere identificati da un codice ID univoco o da una combinazione di cognome e nome.

Punteggi RFM da Customer Data

Considerazioni sui dati

Il dataset deve contenere variabili che contengono le seguenti informazioni:

- Data di acquisto più recente o un intervallo di tempo dalla data di acquisto più recente. Questo verrà utilizzato per calcolare i punteggi di recenza.
- Numero totale di acquisti. Questo verrà utilizzato per calcolare i punteggi di frequenza.

- Valore monetario riassunto per tutti gli acquisti. Questo verrà utilizzato per calcolare i punteggi monetari. Tipicamente, questa è la somma (totale) di tutti gli acquisti, ma potrebbe essere la media (media), massima (importo massimo), o altra misura di riepilogo.

Se si desidera scrivere punteggi RFM in un nuovo dataset, il dataset attivo deve contenere anche una variabile o una combinazione di variabili che identificano ogni caso (cliente).

Creazione Punteggi RFM da Data Customer

1. Dai menu, scegliere:

Marketing Diretto > Scegliere Tecnica

2. Selezionare **Guida identificare i miei migliori contatti (Analisi RFM)** e fare clic su **Continua**.
3. Selezionare **Dati del cliente** e fare clic su **Continua**.
4. Selezionare la variabile che contiene la data di transazione più recente o un numero che rappresenta un intervallo di tempo dalla transazione più recente.
5. Selezionare la variabile che contiene il numero totale di transazioni per ogni cliente.
6. Selezionare la variabile che contiene l'importo monetario di riepilogo per ciascun cliente.
7. Se si desidera scrivere punteggi RFM in un nuovo dataset, selezionare la variabile o la combinazione di variabili che identifica univocamente ogni cliente. Ad esempio, i casi potrebbero essere identificati da un codice ID univoco o da una combinazione di cognome e nome.

Binning RFM

Il processo di raggruppamento di un gran numero di valori numerici in un numero ridotto di categorie è a volte indicato come **binning**. Nell'analisi RFM, i bidoni sono le categorie classificati. È possibile utilizzare la scheda di Binning per modificare il metodo utilizzato per assegnare la recenza, la frequenza e i valori monetari a quei bidoni.

Metodo di raccolta

Nidificato. Nel binning nidificato, una semplice classifica viene assegnata ai valori di recenza. All'interno di ogni rango di recenza i clienti vengono quindi assegnati una classifica di frequenza, e all'interno di ogni rango di frequenza, al cliente viene assegnato un rango monetario. Questo tende a fornire una distribuzione più uniforme dei punteggi combinati RFM, ma ha lo svantaggio di rendere più difficili i punteggi di frequenza e di classifica monetaria. Ad esempio, una classifica di frequenza di 5 per un cliente con rango di recenza del 5 maggio non significa la stessa cosa di una classifica di frequenza di 5 per un cliente con rango di recenza di 4, dato che la classifica di frequenza dipende dal grado di recenza.

Independent. I ranghi semplici sono assegnati alla recenza, alla frequenza e ai valori monetari. I tre ranghi sono assegnati in modo indipendente. L'interpretazione di ciascuno dei tre componenti RFM è quindi univoca; un punteggio di frequenza di 5 per un cliente significa lo stesso di un punteggio di frequenza di 5 per un altro cliente, indipendentemente dai loro punteggi di recenza. Per i campioni più piccoli, questo ha lo svantaggio di ottenere una minore distribuzione dei punteggi combinati RFM.

Numero di bin

Il numero di categorie (bins) da utilizzare per ogni componente per creare punteggi RFM. Il numero totale di possibili punteggi RFM combinati è il prodotto dei tre valori. Ad esempio, 5 bidoni di recenza, 4 bidoni di frequenza e 3 bins monetari creerebbero un totale di 60 possibili punteggi combinati RFM, che vanno dal 111 al 543.

- Il default è di 5 per ogni componente, che creerà 125 possibili punteggi combinati RFM, che variano da 111 a 555.
- Il numero massimo di bidoni consentito per ogni componente di punteggio è nove.

Correlazioni

Una "cravatta" è semplicemente due o più uguali retaggio, frequenza o valori monetari. Idealmente, si desidera avere approssimativamente lo stesso numero di clienti in ogni vassoio, ma un gran numero

di valori legati può influenzare la distribuzione del vassoio. Ci sono due alternative per la gestione dei legami:

- **Assegnare i collegamenti allo stesso bin.** Questo metodo assegna sempre valori legati allo stesso bin, indipendentemente da come questo influisce sulla distribuzione del bin. Questo fornisce un metodo di binning coerente: Se due clienti hanno lo stesso valore di recenza, allora verranno sempre assegnati lo stesso punteggio di recenza. In un esempio estremo, però, potreste avere 1.000 clienti, con 500 di loro che effettuano il loro acquisto più recente nella stessa data. In una graduatoria di 5 bin, il 50% dei clienti riceverà quindi un punteggio di recenza di 5, invece del valore ideale di 20%.

Si noti che con il metodo di binning nidificato "consistenza" è alquanto più complicato per i punteggi di frequenza e monetari, dato che i punteggi di frequenza sono assegnati all'interno di bins di punteggio di recenza, e i punteggi monetari vengono assegnati all'interno dei pin di punteggio di frequenza. Quindi due clienti con lo stesso valore di frequenza potrebbero non avere lo stesso punteggio di frequenza se non hanno anche lo stesso punteggio di recenza, indipendentemente da come vengono gestiti i valori legati.

- **Assegni di assegnazione casuali.** Questo garantisce una distribuzione anche bin, assegnando un piccolissimo fattore di varianza casuale ai legami precedenti alla graduatoria; quindi allo scopo di assegnare valori ai bidoni classificati, non ci sono valori legati. Questo processo non ha alcun effetto sui valori originali. Si usa solo per disambiguare i legami. Mentre questo produce una distribuzione anche bin (approssimativamente lo stesso numero di clienti in ogni vassoio), può comportare risultati di punteggio completamente diversi per i clienti che sembrano avere valori di recenza, frequenza e / o valori monetari simili o identici - in particolare se il numero totale dei clienti è relativamente piccolo e / o il numero di legami è relativamente alto.

Tabella 1. Assegna Ties a Stesso Bin vs. Assegni Casuali

ID	Più Recente Acquisto (Recenza)	Assegna Legami a Stesso Bin	Assegni Assegni Casuali
1	10/29/2006	5	5
2	10/28/2006	4	4
3	10/28/2006	4	4
4	10/28/2006	4	5
5	10/28/2006	4	3
6	9/21/2006	3	3
7	9/21/2006	3	2
8	8/13/2006	2	2
9	8/13/2006	2	1
10	6/20/2006	1	1

- In questo esempio, assegnando i legami con lo stesso bin si ottiene una distribuzione senza pari: 5 (10%), 4 (40%), 3 (20%), 2 (20%), 1 (10%).
- Assegnando casualmente i rapporti risultati in 20% minuti in ogni vassoio, ma per ottenere questo risultato i quattro casi con un valore di data 28.10.2006 sono assegnati a 3 differenti bidoni, e i 2 casi con un valore di data del 13.8.2006 sono assegnati anche a diversi bidoni.

Da notare che il modo in cui i legami vengono assegnati a diversi bidoni è interamente casuale (all'interno dei vincoli del risultato finale si intende un numero uguale di casi in ogni vassoio). Se si calcola una seconda serie di punteggi utilizzando lo stesso metodo, la graduatoria per qualsiasi caso particolare con un valore legato potrebbe cambiare. Ad esempio, le classifiche di recensione rispettivamente del 5 e del 3 per i casi 4 e 5 potrebbero essere commutate la seconda volta.

Risparmio RFM Punteggio da Dati di transazione

RFM da Transaction Data crea sempre un nuovo dataset aggregato con una sola riga per ogni cliente. Utilizzare la scheda Salva per specificare quali punteggi e altre variabili si desidera salvare e dove si desidera salvarle.

Variabili

Le variabili ID che identificano in modo univoco ogni cliente vengono salvate automaticamente nel nuovo dataset. Le seguenti variabili aggiuntive possono essere salvate nel nuovo dataset:

- **Data della transazione più recente per ogni cliente.**
- **Numero di transazioni.** Il numero totale di righe di transazione per ogni cliente.
- **Importo.** L'importo di riepilogo per ogni cliente in base al metodo di riepilogo selezionato nella scheda Variabili.
- **Punteggio attualità.** Il punteggio assegnato ad ogni cliente in base alla data di transazione più recente. I punteggi superiori indicano date di transazione più recenti.
- **Punteggio frequenza.** Il punteggio assegnato ad ogni cliente in base al numero totale di transazioni. I punteggi superiori indicano più transazioni.
- **Punteggio valore monetario.** Il punteggio assegnato ad ogni cliente in base alla misura di riepilogo monetaria selezionata. I punteggi più alti indicano un valore superiore per la misura di riepilogo monetario.
- **Punteggio RFM.** I tre punteggi individuali combinati in un unico valore: $(recenza \times 100) + (frequenza \times 10) + monetario$.

Per impostazione predefinita tutte le variabili disponibili sono incluse nel nuovo dataset; così deselezionare quelli che non si desidera includere. Facoltativamente, è possibile specificare i propri nomi variabili. I nomi variabili devono essere conformi alle regole di denominazione delle variabili standard.

Ubicazione

RFM da Transaction Data crea sempre un nuovo dataset aggregato con una sola riga per ogni cliente. È possibile creare un nuovo dataset nella sessione corrente o salvare i dati di punteggio RFM in un file di dati esterno. I nomi di dataset devono essere conformi alle regole di denominazione delle variabili standard. (Questa limitazione non si applica ai nomi dei file di dati esterni.)

Salvataggio Dei Punteggi RFM dai Dati del Cliente

Per i dati dei clienti è possibile aggiungere le variabili di punteggio RFM al dataset attivo o creare un nuovo dataset che contenga le variabili di punteggi selezionati. Usa il Salva Tab per specificare quali sono le variabili di punteggio che vuoi salvare e dove vuoi salvarle.

Nomi delle variabili salvate

- **Automaticamente genera nomi univoli.** Quando si aggiungono le variabili di punteggio al dataset attivo, questo garantisce che i nuovi nomi delle variabili siano unici. Questo è particolarmente utile se si desidera aggiungere più serie di punteggi RFM (in base a criteri diversi) al dataset attivo.
- **Nomi personalizzati.** Ciò consente di assegnare i propri nomi variabili alle variabili di punteggio. I nomi variabili devono essere conformi alle regole di denominazione delle variabili standard.

Variabili

Selezionare (controllare) le variabili di punteggio che si desidera salvare:

- **Punteggio attualità.** Il punteggio assegnato ad ogni cliente in base al valore della variabile Data o variabile Interval selezionata nella scheda Variabili. I punteggi più alti vengono assegnati a date più recenti o a valori di intervallo più bassi.
- **Punteggio frequenza.** Il punteggio assegnato ad ogni cliente in base alla variabile Numero di transazioni selezionate nella scheda Variabili. I punteggi più alti vengono assegnati a valori superiori.

- **Punteggio valore monetario.** Il punteggio assegnato ad ogni cliente in base alla variabile Importo selezionata nella scheda Variabili. I punteggi più alti vengono assegnati a valori superiori.
- **Punteggio RFM.** I tre punteggi individuali combinati in un unico valore: $(recency*100) + (frequency*10) + monetario$.

Ubicazione

Per i dati dei clienti ci sono tre alternative per dove è possibile salvare nuovi punteggi RFM:

- **dataset attivo.** Le variabili di punteggio RFM selezionate vengono aggiunte al dataset attivo.
- **Nuovo dataset.** Le variabili di punteggio RFM selezionate e le variabili ID che identificano in modo univoco ogni cliente (caso) verranno scritte in un nuovo dataset nella sessione corrente. I nomi di dataset devono essere conformi alle regole di denominazione delle variabili standard. Questa opzione è disponibile solo se si selezionano una o più variabili identificativi del Cliente sulla scheda Variabili.
- **File.** I punteggi RFM selezionati e le variabili ID che identificano univocamente ogni cliente (caso) verranno salvati in un file di dati esterno. Questa opzione è disponibile solo se si selezionano una o più variabili identificativi del Cliente sulla scheda Variabili.

Output RFM

Dati raccolti

I grafici e le tabelle per i dati binati si basano sulla ricettività calcolata, sulla frequenza e sui punteggi monetari.

Mappa termica del valore monetario medio per ricenza e frequenza. La mappa termica della distribuzione monetaria media visualizza il valore monetario medio per le categorie definite dai punteggi di attualità e frequenza. Le aree più scure indicano un valore monetario medio più elevato.

Grafico di bin conta. Il grafico dei conteggi dei bin visualizza la distribuzione bin relativa al metodo di raccolta selezionato. Ogni barra rappresenta il numero di casi che verranno assegnati ad ogni punteggio combinato RFM.

- Anche se di norma si desidera una distribuzione piuttosto omogenea, con tutte le barre (o buona parte di esse) all'incirca della stessa altezza, si deve prevedere un certo grado di varianza quando si utilizza il metodo di raccolta predefinito che assegna i valori correlati allo stesso bin.
- Estreme fluttuazioni nella distribuzione dei bin e/o numerosi bin vuoti indica che è necessario tentare un altro metodo di raccolta (un numero inferiore di bin e/o un'assegnazione casuale delle correlazioni) o riconsiderare l'adeguatezza dell'analisi RFM.

Tabella dei conteggi bin. Le stesse informazioni che si trovano nel grafico di bin conta, tranne che espresse sotto forma di tabella, con bin conta in ogni cella.

Dati non raccolti

Grafico e tabelle per i dati non binati si basano sulle variabili originali utilizzate per creare recensioni, frequenza e punteggi monetari.

Histogrammi. Gli istogrammi mostrano la relativa distribuzione dei valori per le tre variabili utilizzate per calcolare la recenza, la frequenza e i punteggi monetari. Non è insolito per tali istogrammi indicare distribuzioni in qualche modo asimmetriche invece che una distribuzione normale o simmetrica.

L'asse orizzontale di ogni istogramma è sempre ordinato da valori bassi sulla sinistra ad alti valori a destra. Nel caso dell'attualità, tuttavia, l'interpretazione del grafico dipende dal tipo di misura dell'attualità: data o intervallo di tempo. Per le date, le barre a sinistra rappresentano i valori meno recenti (una data meno recente ha un valore meno elevato rispetto a una data più recente). Nel caso degli intervalli di tempo, le barre a sinistra rappresentano i valori più recenti (più limitato è l'intervallo di tempo, più recente è la transazione).

Scatterplot di coppie di variabili. Questi grafici a dispersione mostrano le relazioni tra le tre variabili utilizzate per calcolare i punteggi di attualità, frequenza e valore monetario.

È comune vedere raggruppamenti lineari apprezzabili di punti sulla scala di frequenza, dato che la frequenza rappresenta spesso una gamma relativamente piccola di valori discreti. Ad esempio, se il numero totale delle transazioni non supera i 15, allora ci sono solo 15 possibili valori di frequenza (a meno che non si contino le transazioni frazionali), mentre ci potrebbero essere centinaia di valori di ricettività possibili e migliaia di valori monetari.

L'interpretazione dell'asse di recenza dipende dal tipo di misura di recenza: data o intervallo di tempo. Nel caso delle date, i punti più vicini all'origine rappresentano le date meno recenti. Nel caso degli intervalli di tempo, i punti più vicini all'origine rappresentano i valori più recenti.

Analisi del cluster

L'analisi cluster è uno strumento esplorativo progettato per rivelare i raggruppamenti naturali (o cluster) all'interno dei dati. Ad esempio, è in grado di identificare gruppi diversi di clienti in base a varie caratteristiche demografiche e di acquisto.

Esempio. Le aziende di prodotti retail e di consumo applicano regolarmente le tecniche di clustering ai dati che descrivono le abitudini di acquisto dei propri clienti, il sesso, l'età, il livello di reddito, ecc. Tali aziende adattano le loro strategie di sviluppo dei prodotti e di marketing a ciascun gruppo di consumo per incrementare le vendite e creare una fidelizzazione al marchio.

Considerazioni sui dati di analisi cluster












Dati. Questa procedura funziona con campi sia continui che categoriali. Ogni record (riga) rappresenta un cliente da clustered e i campi (variabili) rappresentano attributi su cui si basa il clustering.

Ordine record. Da notare che i risultati possono dipendere dall'ordine dei record. Per ridurre al minimo gli effetti di ordine, si può voler considerare ordinatamente ordinare i record. Si potrebbe voler eseguire più volte l'analisi, con record ordinati in diversi ordini casuali per verificare la stabilità di una determinata soluzione.

Livello di misurazione. L'assegnazione del livello di misurazione corretto è importante perché influiscono sul calcolo dei risultati.

- *Nominale.* Una variabile può essere considerata nominale quando i suoi valori rappresentano categorie prive di classificazione intrinseca (ad esempio, il dipartimento della società in cui lavora un dipendente). Degli esempi di variabili nominali includono la regione, il codice postale e l'affiliazione religiosa.
- *Ordinale.* Una variabile può essere trattata come ordinale quando i suoi valori rappresentano categorie con una classificazione intrinseca (ad esempio, livelli di soddisfazione del servizio da molto insoddisfatti a molto soddisfatti). Degli esempi di variabili ordinali includono i punteggi di atteggiamento che rappresentano i gradi di soddisfazione o fiducia e i punteggi di classificazione delle preferenze.
- *Continua.* Una variabile può essere trattata come una scala (continua) quando i suoi valori rappresentano categorie ordinate con una metrica significativa, in modo che i confronti di distanza tra i valori siano appropriati. Esempi di variabili di scala sono l'età espressa in anni o il reddito espresso in migliaia di Euro.

Un'icona accanto a ciascun campo indica il livello di misurazione corrente.

	Numerico	Stringa	Dati	Ora
Scala (continuo)		n/d		
Ordinale				
Nominale				

È possibile modificare il livello di misurazione in Variabile Vista dell'editor dei dati oppure è possibile utilizzare la finestra di dialogo Definisci proprietà variabile per suggerire un livello di misurazione appropriato per ogni campo.

Campi con livello di misurazione sconosciuto

L'avviso Livello di misurazione viene visualizzato quando il livello di misurazione di una o più variabili (campi) del dataset è sconosciuto. Poiché influisce sul calcolo dei risultati di questa procedura, il livello di misurazione deve essere definito per tutte le variabili.

Esegui scansione dati. Legge i dati del dataset attivo e assegna un livello di misurazione predefinito a tutti i campi con livello di misurazione sconosciuto. Con dataset di grandi dimensioni, questa operazione può richiedere del tempo.

Assegna manualmente. Apre una finestra di dialogo che elenca tutti i campi con livello di misurazione sconosciuto, mediante la quale è possibile assegnare un livello di misurazione a questi campi. Il livello di misurazione si può assegnare anche nella Vista variabile dell'Editor dei dati.

Dal momento che il livello di misurazione è importante per questa procedura, è possibile accedere alla finestra di dialogo per la sua esecuzione solo quando per tutti i campi è stato definito un livello di misurazione.

Per ottenere l'analisi del Cluster

Dai menu, scegliere:

Direct marketing > Scegli tecnica

1. Selezionare **Segmento i miei contatti in cluster**.
2. Selezionare i campi categoriali (nominali, ordinali) e continui (scala) che si desidera utilizzare per creare segmenti.
3. Fare clic su **Esegui** per eseguire la procedura.

Impostazioni

La scheda Impostazioni consente di mostrare o sopprimere la visualizzazione di grafici e tabelle che descrivono i segmenti, salvare un nuovo campo nel dataset che identifica il segmento (cluster) per ogni record nel dataset e specificare quanti segmenti da includere nella soluzione cluster.

Visualizza grafici e tabelle. Visualizza tabelle e grafici che descrivono i segmenti.

Membership di segmento. Salva un nuovo campo (variabile) che identifica il segmento a cui ogni record appartiene.

- I nomi di campo devono essere conformi alle regole di denominazione IBM SPSS Statistiche .
- Il nome campo di appartenenza del segmento non può duplicare un nome di campo che già esiste nel dataset. Se si esegue questa procedura più di una volta sullo stesso dataset, sarà necessario specificare ogni volta un nome diverso.
- **Numero di Segmenti.** Controlla come viene determinato il numero di segmenti.
- **Determina automaticamente.** La procedura determinerà automaticamente il "miglior" numero di segmenti, fino al massimo specificato.

Specifica fisso. La procedura produrrà il numero specificato di segmenti.

Profili potenziali clienti

Questa tecnica utilizza i risultati di una campagna precedente o di test per creare profili descrittivi. È possibile utilizzare i profili per definire gruppi specifici di contatti nelle campagne successive. Il campo Response indica chi ha risposto alla campagna precedente o di test. L'elenco Profili contiene le caratteristiche che si desidera utilizzare per creare il profilo.

Esempio. In base ai risultati di un mailing test, la divisione diretta di marketing di un'azienda vuole generare profili dei tipi di clienti più suscettibili di rispondere ad un'offerta, basata sull'informazione demografica.

Output

L'output include una tabella che fornisce una descrizione di ciascun gruppo di profili e visualizza i tassi di risposta (percentuale di risposte positive) e i tassi di risposta cumulativi e un grafico dei tassi di risposta cumulativi. Se si include una velocità di risposta minima di destinazione, la tabella sarà colorata per mostrare quali profili soddisfano il tasso di risposta cumulativo minimo e il grafico includerà una linea di riferimento al valore minimo di risposta minimo specificato.

considerazioni sui dati Profili di prospettiva

Campo di risposta. Il campo di risposta deve essere nominale o ordinale. Può essere stringa o numerica. Se questo campo contiene un valore che indica il numero o l'importo degli acquisti, sarà necessario creare un nuovo campo in cui un singolo valore rappresenta tutte le risposte positive. Per ulteriori informazioni, consultare l'argomento [“Creazione di un campo di risposta categoriale”](#) a pagina 10.

Valore di risposta positivo. Il valore di risposta positivo identifica i clienti che hanno risposto positivamente (ad esempio, effettuato un acquisto). Tutti gli altri valori di risposta non mancanti sono assunti per indicare una risposta negativa. Se ci sono etichette di valore definite per il campo di risposta, tali etichette vengono visualizzate nell'elenco a discesa.

Crea profili con. Questi campi possono essere nominali, ordinali o continui (scala). Possono essere stringa o numerica.

Livello di misurazione. L'assegnazione del livello di misurazione corretto è importante perché influiscono sul calcolo dei risultati.

- *Nominale.* Una variabile può essere considerata nominale quando i suoi valori rappresentano categorie prive di classificazione intrinseca (ad esempio, il dipartimento della società in cui lavora un dipendente). Degli esempi di variabili nominali includono la regione, il codice postale e l'affiliazione religiosa.
- *Ordinale.* Una variabile può essere trattata come ordinale quando i suoi valori rappresentano categorie con una classificazione intrinseca (ad esempio, livelli di soddisfazione del servizio da molto insoddisfatti a molto soddisfatti). Degli esempi di variabili ordinali includono i punteggi di atteggiamento che rappresentano i gradi di soddisfazione o fiducia e i punteggi di classificazione delle preferenze.
- *Continua.* Una variabile può essere trattata come una scala (continua) quando i suoi valori rappresentano categorie ordinate con una metrica significativa, in modo che i confronti di distanza tra i valori siano appropriati. Esempi di variabili di scala sono l'età espressa in anni o il reddito espresso in migliaia di Euro.

Un'icona accanto a ciascun campo indica il livello di misurazione corrente.

Tabella 3. Icone del livello di misurazione				
	Numerico	Stringa	Dati	Ora
Scala (continuo)		n/d		
Ordinale				
Nominale				

È possibile modificare il livello di misurazione in Variabile Vista dell'editor dei dati oppure è possibile utilizzare la finestra di dialogo Definisci proprietà variabile per suggerire un livello di misurazione appropriato per ogni campo.

Campi con livello di misurazione sconosciuto

L'avviso Livello di misurazione viene visualizzato quando il livello di misurazione di una o più variabili (campi) del dataset è sconosciuto. Poiché influisce sul calcolo dei risultati di questa procedura, il livello di misurazione deve essere definito per tutte le variabili.

Esegui scansione dati. Legge i dati del dataset attivo e assegna un livello di misurazione predefinito a tutti i campi con livello di misurazione sconosciuto. Con dataset di grandi dimensioni, questa operazione può richiedere del tempo.

Assegna manualmente. Apre una finestra di dialogo che elenca tutti i campi con livello di misurazione sconosciuto, mediante la quale è possibile assegnare un livello di misurazione a questi campi. Il livello di misurazione si può assegnare anche nella Vista variabile dell'Editor dei dati.

Dal momento che il livello di misurazione è importante per questa procedura, è possibile accedere alla finestra di dialogo per la sua esecuzione solo quando per tutti i campi è stato definito un livello di misurazione.

Per ottenere profili di prospettiva

Dai menu, scegliere:

Direct marketing > Scegli tecnica

1. Selezionare **Generare i profili dei miei contatti che hanno risposto ad un'offerta.**
2. Selezionare il campo che identifica quali contatti hanno risposto all'offerta. Questo campo deve essere nominale o ordinale.
3. Inserire il valore che indica una risposta positiva. Se eventuali valori hanno definito etichette di valore, è possibile selezionare l'etichetta del valore dall'elenco a discesa e verrà visualizzato il valore corrispondente.
4. Selezionare i campi che si desidera utilizzare per creare i profili.
5. Fare clic su **Esegui** per eseguire la procedura.

Impostazioni

La scheda Impostazioni consente di controllare la dimensione del gruppo di profili minimo e includere una soglia minima di velocità di risposta nell'output.

Dimensione del gruppo di profilo minimo. Ogni profilo rappresenta le caratteristiche condivise di un gruppo di contatti nel dataset (ad esempio le femmine sotto i 40 che vivono a ovest). Per impostazione predefinita, la dimensione del gruppo di profili più piccolo è di 100. Dimensioni di gruppo minori possono rivelare più gruppi, ma dimensioni di gruppo più grandi possono fornire risultati più affidabili. Il valore deve essere un intero positivo.

Include le informazioni minime sulla soglia di risposta nei risultati. I risultati includono una tabella che visualizza i tassi di risposta (percentuale di risposte positive) e i tassi di risposta cumulativi e un grafico dei tassi di risposta cumulativi. Se si immette un tasso di risposta minimo di destinazione, la tabella sarà colorata per mostrare quali profili soddisfano il tasso di risposta cumulativo minimo e il grafico includerà una linea di riferimento al valore minimo di risposta minimo specificato. Il valore deve essere maggiore di 0 e minore di 100.

Creazione di un campo di risposta categoriale

Il campo di risposta dovrebbe essere categoriale, con un valore che rappresenta tutte le risposte positive. Qualsiasi altro valore non mancante si assume una risposta negativa. Se il campo di risposta rappresenta un valore continuo (scala), come il numero di acquisti o l'importo monetario degli acquisti, è necessario creare un nuovo campo che assegna un valore di risposta positivo unico a tutti i valori di risposta non zero.

- Se le risposte negative sono registrate come 0 (non in bianco, che viene trattata come mancante), questo può essere calcolato con la seguente formula:

$$\text{NewName} = \text{OldName} > 0$$

dove *NewName* è il nome del nuovo campo e *OldName* è il nome del campo originale. Si tratta di un'espressione logica che assegna un valore di 1 a tutti i valori non mancanti superiori a 0 e 0 a tutti i valori non mancanti inferiori o uguali a 0.

- Se non viene registrato alcun valore per le risposte negative, allora questi valori sono trattati come mancanti e la formula è un po' più complicata:

`NewName= NOT (MISSING (OldName))`

In questa espressione logica, tutti i valori di risposta non mancanti vengono assegnati un valore di 1 e tutti i valori di risposta mancanti vengono assegnati un valore di 0.

- Se non è possibile distinguere tra valori di risposta negativi (0) e valori mancanti, non è possibile calcolare un valore di risposta accurato. Se ci sono relativamente pochi valori realmente mancanti, questo potrebbe non avere un effetto significativo sui tassi di risposta calcolati. Se però ci sono molti valori mancanti - come ad esempio quando le informazioni di risposta vengono registrate per solo un piccolo campione di test del dataset totale - allora i tassi di risposta calcolati saranno privi di significato, dato che saranno significativamente inferiori ai tassi di risposta veri.

Per creare un campo di risposta categoriale

1. Dai menu, scegliere:

Trasforma > Variabile variabile

2. Per variabile di destinazione, inserire il nuovo nome di campo (variabile).
3. Se le risposte negative vengono registrate come 0, per l'espressione numerica immettere `OldName>0`, dove *OldName* è il nome del campo originale.
4. Se le risposte negative vengono registrate come mancanti (vuoto), per l'espressione numerica immettere `NOT (MISSING (OldName))`, dove *OldName* è il nome del campo originale.

Tassi Di Risposta Del Codice Postale

Questa tecnica utilizza i risultati di una campagna precedente per calcolare i tassi di risposta relativi al codice postale. Tali tassi possono essere utilizzati per identificare codici postali specifici nelle campagne successive. Il campo Risposta indica chi ha risposto alla campagna precedente. Il campo Codice postale identifica il campo contenente i codici postali.

Esempio. In base ai risultati di una precedente mailing, la divisione diretta di marketing di un'azienda genera tassi di risposta tramite codici postali. In base a vari criteri, come ad esempio un tasso di risposta accettabile minimo e / o il numero massimo di contatti da includere nella mailing, possono quindi mirare codici postali specifici.

Output

L'output da questa procedura include un nuovo dataset che contiene tassi di risposta per codice postale, e un tavolo e grafico che sintetizzano i risultati per rango decile (top 10%, top 20%, ecc.). La tabella può essere colorata in base a un tasso di risposta cumulativo minimo specificato dall'utente o al numero massimo di contatti.

Il nuovo dataset contiene i seguenti campi:

- **Codice postale.** Se i gruppi di codici postali si basano solo su una parte del valore completo, allora questo è il valore di quella porzione di codice postale. L'etichetta di riga di intestazione per questa colonna nel file Excel è il nome del campo del codice postale nel dataset originale.
- **ResponseRate.** La percentuale di risposte positive in ogni codice postale.
- **Risposte.** Il numero di risposte positive in ogni codice postale.
- **Contatti.** Il numero totale di contatti in ogni codice postale che contengono un valore non mancante per il campo di risposta.
- **Indice.** La risposta "ponderata" in base alla formula $N \times P \times (1-P)$, dove N è il numero di contatti, e P è il tasso di risposta espresso in proporzione.

- **Rango.** Rango decile (top 10%, top 20%, ecc.) dei tassi di risposta del codice postale cumulativo in ordine decrescente.

Considerazioni Sui Tassi Di Risposta Del Codice Postale Considerazioni

Campo di risposta. Il campo di risposta può essere stringa o numerica. Se questo campo contiene un valore che indica il numero o il valore monetario degli acquisti, sarà necessario creare un nuovo campo in cui un singolo valore rappresenta tutte le risposte positive. Per ulteriori informazioni, consultare l'argomento [“Creazione di un campo di risposta categoriale”](#) a pagina 13.

Valore di risposta positivo. Il valore di risposta positivo identifica i clienti che hanno risposto positivamente (ad esempio, effettuato un acquisto). Tutti gli altri valori di risposta non mancanti sono assunti per indicare una risposta negativa. Se ci sono etichette di valore definite per il campo di risposta, tali etichette vengono visualizzate nell'elenco a discesa.

Campo codice postale. Il campo del codice postale può essere stringa o numerica.

Per Ottenere Tassi Di Risposta Del Codice Postale

Dai menu, scegliere:

Direct marketing > Scegli tecnica

1. Selezionare **Identificare i codici postali di risposta principali**.
2. Selezionare il campo che identifica quali contatti hanno risposto all'offerta.
3. Inserire il valore che indica una risposta positiva. Se eventuali valori hanno definito etichette di valore, è possibile selezionare l'etichetta del valore dall'elenco a discesa e verrà visualizzato il valore corrispondente.
4. Selezionare il campo che contiene il codice postale.
5. Fare clic su **Esegui** per eseguire la procedura.

In alternativa, è possibile anche:

- Generare tassi di risposta in base ai primi n caratteri o cifre del codice postale invece del valore completo
- Salvare automaticamente i risultati in un file Excel
- Opzioni di visualizzazione di uscita di

Impostazioni

Raggruppa codici postali in base a

Questo determina come i record sono raggruppati per calcolare i tassi di risposta. Per impostazione predefinita viene utilizzato l'intero codice postale e tutti i record con lo stesso codice postale sono raggruppati per calcolare il tasso di risposta del gruppo. In alternativa, è possibile raggruppare record in base a solo una parte del codice postale completo, composto dalle prime n cifre o caratteri. Ad esempio, si potrebbe voler raggruppare record in base ai soli primi 5 caratteri di un codice postale di 10 caratteri o alle prime tre cifre di un codice postale a 5 cifre. Il dataset di output conterrà un solo record per ogni gruppo di codici postali. Se si immette un valore, deve essere un intero positivo.

Formato codice postale numerico

Se il campo del codice postale è numerico e si desidera raggruppare i codici postali in base alla prima cifra n invece dell'intero valore, è necessario specificare il numero di cifre nel valore originale. Il numero di cifre è il numero *massimo* possibile di cifre nel codice postale. Ad esempio, se il campo del codice postale contiene un mix di codici zip a 5 cifre e 9 cifre, è necessario specificare 9 come numero di cifre.

Nota: A seconda del formato di visualizzazione, alcuni codici zip a 5 cifre possono sembrare contenere solo 4 cifre, ma c'è uno zero implicito.

Output

Oltre al nuovo dataset che contiene tassi di risposta per codice postale, è possibile visualizzare una tabella e un grafico che riepilogano i risultati per rango decile (top 10%, top 20%, ecc.). La tabella

visualizza i tassi di risposta, i tassi di risposta cumulativi, il numero di record e il numero cumulativo di record in ogni decile. Il grafico visualizza i tassi di risposta cumulativi e il numero cumulativo di record in ogni decile.

Tasso di risposta minimo accettabile. Se si immette una velocità di risposta minima o una formula di break-even, la tabella sarà colorata per mostrare quali decili soddisfano il tasso di risposta cumulativo minimo e il grafico includerà una linea di riferimento al valore minimo di risposta specificato.

- **Velocità di risposta di destinazione.** Tasso di risposta espresso come pernottamento (percentuale di risposte positive in ogni gruppo di codici postali). Il valore deve essere maggiore di 0 e minore di 100.
- **Calcolare la rottura - anche la velocità dalla formula.** Calcolare il tasso di risposta cumulativo minimo in base alla formula: $(Costo\ di\ mailing\ un\ pacchetto / Entrate\ Net\ per\ risposta) \times 100$. Entrambi i valori devono essere numeri positivi. Il risultato dovrebbe essere un valore superiore a 0 e inferiore a 100. Ad esempio, se il costo del servizio di mailing di un pacchetto è \$0.75 e il ricavo netto per risposta è \$56, il tasso di risposta minimo è: $(0.75/56) \times 100 = 1.34\%$.

Numero massimo di contatti. Se si specifica un numero massimo di contatti, la tabella sarà colorata per mostrare quali decili non superano il numero massimo cumulativo di contatti (record) e il grafico includerà una linea di riferimento a quel valore.

- **Percentuale di contatti.** Massimo espresso in percentuale. Ad esempio, si potrebbe voler conoscere i decili con i tassi di risposta più alti che contengono non più del 50% di tutti i contatti. Il valore deve essere maggiore di 0 e minore di 100.
- **Numero di contatti.** Massimo espresso come numero di contatti. Ad esempio, se non si intende spedire più di 10.000 pacchetti, si potrebbe impostare il valore a 10000. Il valore deve essere un intero positivo (senza simboli di raggruppamento).

Se si specifica sia un tasso di risposta accettabile minimo che un numero massimo di contatti, la codifica colorata della tabella sarà basata su qualunque condizione venga soddisfatta prima.

Esporta in Excel

Questa procedura crea automaticamente un nuovo dataset che contiene tassi di risposta per codice postale. Ogni record (riga) nel dataset rappresenta un codice postale. È possibile salvare automaticamente le stesse informazioni in un file Excel. Questo file viene salvato in formato Excel 97 - 2003.

Creazione di un campo di risposta categoriale

Il campo di risposta dovrebbe essere categoriale, con un valore che rappresenta tutte le risposte positive. Qualsiasi altro valore non mancante si assume una risposta negativa. Se il campo di risposta rappresenta un valore continuo (scala), come il numero di acquisti o l'importo monetario degli acquisti, è necessario creare un nuovo campo che assegna un valore di risposta positivo unico a tutti i valori di risposta non zero.

- Se le risposte negative sono registrate come 0 (non in bianco, che viene trattata come mancante), questo può essere calcolato con la seguente formula:

$NewName = OldName > 0$

dove *NewName* è il nome del nuovo campo e *OldName* è il nome del campo originale. Si tratta di un'espressione logica che assegna un valore di 1 a tutti i valori non mancanti superiori a 0 e 0 a tutti i valori non mancanti inferiori o uguali a 0.

- Se non viene registrato alcun valore per le risposte negative, allora questi valori sono trattati come mancanti e la formula è un po' più complicata:

$NewName = NOT (MISSING (OldName))$

In questa espressione logica, tutti i valori di risposta non mancanti vengono assegnati un valore di 1 e tutti i valori di risposta mancanti vengono assegnati un valore di 0.

- Se non è possibile distinguere tra valori di risposta negativi (0) e valori mancanti, non è possibile calcolare un valore di risposta accurato. Se ci sono relativamente pochi valori realmente mancanti, questo potrebbe non avere un effetto significativo sui tassi di risposta calcolati. Se però ci sono molti

valori mancanti - come ad esempio quando le informazioni di risposta vengono registrate per solo un piccolo campione di test del dataset totale - allora i tassi di risposta calcolati saranno privi di significato, dato che saranno significativamente inferiori ai tassi di risposta veri.

Per creare un campo di risposta categoriale

1. Dai menu, scegliere:

Trasforma > Variabile variabile

2. Per variabile di destinazione, inserire il nuovo nome di campo (variabile).
3. Se le risposte negative vengono registrate come 0, per l'espressione numerica immettere $OldName > 0$, dove *OldName* è il nome del campo originale.
4. Se le risposte negative vengono registrate come mancanti (vuoto), per l'espressione numerica immettere $NOT (MISSING (OldName))$, dove *OldName* è il nome del campo originale.

Propensione all'acquisto

Propensione all'acquisto utilizza risultati da una mailing testata o da una campagna precedente per generare punteggi. I punteggi indicano i contatti che risponderanno con maggiore probabilità. Il campo Risposta indica chi ha risposto alla mailing di test o alla campagna precedente. I campi Propensione contengono le caratteristiche che si desidera utilizzare per prevedere la probabilità di risposta di contatti con caratteristiche simili.

Questa tecnica utilizza la regressione logistica binaria per costruire un modello predittivo. Il processo di costruzione e applicazione di un modello predittivo ha due passaggi fondamentali:

1. Creazione del modello e salvataggio del file del modello. La creazione del modello avviene utilizzando un dataset per il quale il risultato di interesse (spesso indicato come **obiettivo**) è noto. Ad esempio, se si desidera creare un modello che prevede chi è probabile che risponda ad una campagna tramite direct mailing, è necessario iniziare con un dataset che contiene già le informazioni su chi ha risposto e chi non ha risposto. Ad esempio, potrebbero essere i risultati di un mailing di test ad un piccolo gruppo di clienti o le informazioni sulle risposte ad una campagna simile in passato.
2. Applicare quel modello ad un dataset diverso (per il quale il risultato di interesse non è noto) per ottenere i risultati previsti.

Esempio. La divisione Direct marketing di una società utilizza i risultati di un mailing del test per assegnare dei punteggi di propensione al resto dei contatti presenti nel database, identificando i contatti maggiormente propensi a rispondere e a effettuare un acquisto secondo varie caratteristiche demografiche.

Output

Questa procedura crea automaticamente un nuovo campo nel dataset che contiene punteggi di propensione per i dati di prova e un file di modello XML utilizzabile per segnare altri datasets. L'output diagnostico opzionale include un grafico di qualità del modello complessivo e una tabella di classificazione che confronta le risposte previste a risposte reali.

Propensione alle considerazioni sui dati di acquisto

Campo di risposta. Il campo di risposta può essere stringa o numerica. Se questo campo contiene un valore che indica il numero o il valore monetario degli acquisti, sarà necessario creare un nuovo campo in cui un singolo valore rappresenta tutte le risposte positive. Per ulteriori informazioni, consultare l'argomento ["Creazione di un campo di risposta categoriale"](#) a pagina 17.

Valore di risposta positivo. Il valore di risposta positivo identifica i clienti che hanno risposto positivamente (ad esempio, effettuato un acquisto). Tutti gli altri valori di risposta non mancanti sono assunti per indicare una risposta negativa. Se ci sono etichette di valore definite per il campo di risposta, tali etichette vengono visualizzate nell'elenco a discesa.

Prevedere Propensione con. I campi utilizzati per prevedere la propensione possono essere stringa o numerica, e possono essere nominali, ordinali o continui (scala) - ma è importante assegnare il corretto livello di misurazione a tutti i campi predittivi.

Livello di misurazione. L'assegnazione del livello di misurazione corretto è importante perché influiscono sul calcolo dei risultati.

- **Nominale.** Una variabile può essere considerata nominale quando i suoi valori rappresentano categorie prive di classificazione intrinseca (ad esempio, il dipartimento della società in cui lavora un dipendente). Degli esempi di variabili nominali includono la regione, il codice postale e l'affiliazione religiosa.
- **Ordinale.** Una variabile può essere trattata come ordinale quando i suoi valori rappresentano categorie con una classificazione intrinseca (ad esempio, livelli di soddisfazione del servizio da molto insoddisfatti a molto soddisfatti). Degli esempi di variabili ordinali includono i punteggi di atteggiamento che rappresentano i gradi di soddisfazione o fiducia e i punteggi di classificazione delle preferenze.
- **Continua.** Una variabile può essere trattata come una scala (continua) quando i suoi valori rappresentano categorie ordinate con una metrica significativa, in modo che i confronti di distanza tra i valori siano appropriati. Esempi di variabili di scala sono l'età espressa in anni o il reddito espresso in migliaia di Euro.

Un'icona accanto a ciascun campo indica il livello di misurazione corrente.

Tabella 4. Icone del livello di misurazione

	Numerico	Stringa	Dati	Ora
Scala (continuo)		n/d		
Ordinale				
Nominale				

È possibile modificare il livello di misurazione in Variabile Vista dell'editor dei dati oppure è possibile utilizzare la finestra di dialogo Definisci proprietà variabile per suggerire un livello di misurazione appropriato per ogni campo.

Campi con livello di misurazione sconosciuto

L'avviso Livello di misurazione viene visualizzato quando il livello di misurazione di una o più variabili (campi) del dataset è sconosciuto. Poiché influisce sul calcolo dei risultati di questa procedura, il livello di misurazione deve essere definito per tutte le variabili.

Esegui scansione dati. Legge i dati del dataset attivo e assegna un livello di misurazione predefinito a tutti i campi con livello di misurazione sconosciuto. Con dataset di grandi dimensioni, questa operazione può richiedere del tempo.

Assegna manualmente. Apre una finestra di dialogo che elenca tutti i campi con livello di misurazione sconosciuto, mediante la quale è possibile assegnare un livello di misurazione a questi campi. Il livello di misurazione si può assegnare anche nella Vista variabile dell'Editor dei dati.

Dal momento che il livello di misurazione è importante per questa procedura, è possibile accedere alla finestra di dialogo per la sua esecuzione solo quando per tutti i campi è stato definito un livello di misurazione.

Per ottenere propensione ai punteggi di acquisto

Dai menu, scegliere:

Direct marketing > Scegli tecnica

1. Selezionare **Seleziona i contatti più probabili da acquistare.**
2. Selezionare il campo che identifica quali contatti hanno risposto all'offerta.

3. Inserire il valore che indica una risposta positiva. Se eventuali valori hanno definito etichette di valore, è possibile selezionare l'etichetta del valore dall'elenco a discesa e verrà visualizzato il valore corrispondente.
4. Selezionare i campi che si desidera utilizzare per prevedere la propensione.
Per salvare un file XML di modello per segnare altri file di dati:
5. Selezionare **Esporta informazioni modello in file XML**.
6. Inserire un percorso di directory e un nome file o fare clic su **Sfoggia** per navigare nella località in cui si desidera salvare il file XML del modello.
7. Fare clic su **Esegui** per eseguire la procedura.
Per utilizzare il file modello per segnare altri datasets:
8. Aprire il dataset di cui si desidera calcolare il punteggio.
9. Utilizzare la procedura guidata di Scoring per applicare il modello al dataset. Dai menu, scegliere:
Programmi di utilità > Scoring Wizard.

Impostazioni

Convalida modello

La validazione del modello crea gruppi di formazione e verifica per scopi diagnostici. Se si seleziona la tabella di classificazione nella sezione Diagnostica Output, la tabella sarà suddivisa in sezioni di formazione (selezionate) e di verifica (non selezionate) a fini di confronto. Non selezionare la validazione del modello a meno che non venga selezionata anche la tabella di classificazione. I punteggi si basano sul modello generato dal campione di allenamento, che conterrà sempre meno record rispetto al numero totale di record disponibili. Ad esempio, la dimensione del campione di allenamento predefinito è di 50% e un modello costruito su solo la metà dei record disponibili potrebbe non essere affidabile come un modello costruito su tutti i record disponibili.

- **Dimensione partizione del campione di formazione (%).** Specificare la percentuale di record da assegnare al campione di formazione. Il resto dei record con valori non mancanti per il campo di risposta sono assegnati al campione di prova. Il valore deve essere maggiore di 0 e minore di 100.
- **Imposta seme per replicare i risultati.** Dal momento che i record vengono assegnati casualmente ai campioni di formazione e verifica, ogni volta che si esegue la procedura si possono ottenere risultati diversi, a meno che non si specificasse sempre lo stesso valore di partenza casuale del numero di partenza.

Output diagnostico

Qualità modello globale. Visualizza un grafico a barre di qualità del modello complessivo, espresso come valore compreso tra 0 e 1. Un modello valido deve avere un valore maggiore di 0.5.

Tabella di classificazione. Visualizza una tabella che confronta le risposte positive e negative a risposte positive e negative reali. La velocità di precisione complessiva può fornire qualche indicazione su quanto funziona bene il modello, ma si può essere più interessati alla percentuale di risposte positive previste corrette.

- **probabilità minima.** Assegna record con un valore di punteggio superiore al valore specificato alla categoria di risposta positiva prevista nella tabella di classificazione. I punteggi generati dalla procedura rappresentano la probabilità che il contatto risponda positivamente (ad esempio, effettuare un acquisto). Come regola generale è bene specificare un valore vicino a quello del tasso di risposta obiettivo minimo, espresso come proporzione. Ad esempio, se si è interessati a una percentuale di risposta pari almeno al 5%, specificare 0.05. Il valore deve essere superiore a 0 e inferiore a 1.

Nome ed etichetta del campo risposta ricodificata

Questa procedura recupera automaticamente il campo di risposta in un nuovo campo in cui il 1 rappresenta risposte positive e 0 rappresenta risposte negative, e l'analisi viene eseguita sul campo recuperato. Puoi sovrascrivere il nome e l'etichetta di default e fornire il tuo. I nomi devono essere conformi alle regole di denominazione IBM SPSS Statistiche .

Salva punteggi

Un nuovo campo contenente punteggi di propensione viene salvato automaticamente al dataset originale. I punteggi rappresentano la probabilità di una risposta positiva, espressa in proporzione.

- I nomi di campo devono essere conformi alle regole di denominazione IBM SPSS Statistiche .
- Il nome del campo non può duplicare un nome di campo già esistente nel dataset. Se si esegue questa procedura più di una volta sullo stesso dataset, sarà necessario specificare ogni volta un nome diverso.

Creazione di un campo di risposta categoriale

Il campo di risposta dovrebbe essere categoriale, con un valore che rappresenta tutte le risposte positive. Qualsiasi altro valore non mancante si assume una risposta negativa. Se il campo di risposta rappresenta un valore continuo (scala), come il numero di acquisti o l'importo monetario degli acquisti, è necessario creare un nuovo campo che assegna un valore di risposta positivo unico a tutti i valori di risposta non zero.

- Se le risposte negative sono registrate come 0 (non in bianco, che viene trattata come mancante), questo può essere calcolato con la seguente formula:

$$\text{NewName} = \text{OldName} > 0$$

dove *NewName* è il nome del nuovo campo e *OldName* è il nome del campo originale. Si tratta di un'espressione logica che assegna un valore di 1 a tutti i valori non mancanti superiori a 0 e 0 a tutti i valori non mancanti inferiori o uguali a 0.

- Se non viene registrato alcun valore per le risposte negative, allora questi valori sono trattati come mancanti e la formula è un po' più complicata:

$$\text{NewName} = \text{NOT} (\text{MISSING} (\text{OldName}))$$

In questa espressione logica, tutti i valori di risposta non mancanti vengono assegnati un valore di 1 e tutti i valori di risposta mancanti vengono assegnati un valore di 0.

- Se non è possibile distinguere tra valori di risposta negativi (0) e valori mancanti, non è possibile calcolare un valore di risposta accurato. Se ci sono relativamente pochi valori realmente mancanti, questo potrebbe non avere un effetto significativo sui tassi di risposta calcolati. Se però ci sono molti valori mancanti - come ad esempio quando le informazioni di risposta vengono registrate per solo un piccolo campione di test del dataset totale - allora i tassi di risposta calcolati saranno privi di significato, dato che saranno significativamente inferiori ai tassi di risposta veri.

Per creare un campo di risposta categoriale

1. Dai menu, scegliere:

Trasforma > Variabile variabile

2. Per variabile di destinazione, inserire il nuovo nome di campo (variabile).
3. Se le risposte negative vengono registrate come 0, per l'espressione numerica immettere $\text{OldName} > 0$, dove *OldName* è il nome del campo originale.
4. Se le risposte negative vengono registrate come mancanti (vuoto), per l'espressione numerica immettere $\text{NOT} (\text{MISSING} (\text{OldName}))$, dove *OldName* è il nome del campo originale.

Test package di controllo

Questa tecnica confronta le campagne di marketing per rilevare eventuali differenze significative in termini di efficacia tra i diversi package o offerte. L'efficacia di una campagna viene misurata in base alle risposte. Il Campo della Campagna identifica diverse campagne, ad esempio Offer A e Offer B. Il Campo di risposta indica se un contatto ha risposto alla campagna. Selezionare Importo acquisto quando la risposta viene registrata come importo di acquisto, ad esempio "99.99". Selezionare Rispondi quando la risposta indica semplicemente se il contatto ha risposto positivamente o meno, ad esempio "Sì" o "No".

Esempio. La divisione marketing diretto di un'azienda vuole vedere se un nuovo design del pacchetto genererà risposte più positive rispetto al pacchetto esistente. Così mandano un mailing test per determinare se il nuovo pacchetto genera un tasso di risposta positivo significativamente più alto. Il

mailing test consiste in un gruppo di controllo che riceve il pacchetto esistente e un gruppo di prova che riceve il nuovo design del pacchetto. I risultati per i due gruppi sono poi confrontati per vedere se c'è una differenza significativa.

Output

L'output include una tabella che visualizza i conteggi e le percentuali di risposte positive e negative per ogni gruppo definito dal Campo della Campagna e da una tabella che identifica quali gruppi differiscono in modo significativo l'uno dall'altro.

Control Package Test Data Considerazioni e Presupposti

Campo della campagna. Il Campo della Campagna dovrebbe essere categorico (nominale o ordinale).

Campo di risposta dell'efficacia. Se si seleziona l'importo di acquisto per il Campo di efficacia, il campo deve essere numerico e il livello di misurazione deve essere continuo (scala).

Se non è possibile distinguere tra negativi (per l'importo di acquisto, un valore di 0) valori di risposta e valori mancanti, allora non è possibile calcolare un tasso di risposta accurato. Se ci sono relativamente pochi valori realmente mancanti, questo potrebbe non avere un effetto significativo sui tassi di risposta calcolati. Se però ci sono molti valori mancanti - come ad esempio quando le informazioni di risposta vengono registrate per solo un piccolo campione di test del dataset totale - allora i tassi di risposta calcolati saranno privi di significato, dato che saranno significativamente inferiori ai tassi di risposta veri.

Ipotesi. Questa procedura presuppone che i contatti siano stati assegnati casualmente ad ogni gruppo di campagna. In altre parole, nessun particolare demografico, la storia dell'acquisto o altre caratteristiche influenzano l'assegnazione del gruppo e tutti i contatti hanno una probabilità uguale di essere assegnati a qualsiasi gruppo.

Per Ottenere un Test del pacchetto di controllo

Dai menu, scegliere:

Direct marketing > Scegli tecnica

1. Selezionare **Confronta l'efficacia delle campagne**.
2. Selezionare il campo che identifica quale gruppo di campagna ogni contatto appartiene (ad esempio, offrire A, offrire B, ecc.) Questo campo deve essere nominale o ordinale.
3. Selezionare il campo che indica efficacia di risposta.

Se il campo di risposta è un importo di acquisto, il campo deve essere numerico.

Se il campo di risposta indica semplicemente se il contatto ha risposto positivamente o meno (ad esempio "Sì" o "No"), selezionare **Rispondi** ed inserire il valore che rappresenta una risposta positiva. Se eventuali valori hanno definito etichette di valore, è possibile selezionare l'etichetta del valore dall'elenco a discesa e verrà visualizzato il valore corrispondente.

Viene creato automaticamente un nuovo campo, in cui 1 rappresenta risposte positive e 0 rappresenta risposte negative e l'analisi viene eseguita sul nuovo campo. Puoi sovrascrivere il nome e l'etichetta di default e fornire il tuo. I nomi devono essere conformi alle regole di denominazione IBM SPSS Statistiche .

4. Fare clic su **Esegui** per eseguire la procedura.

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, può essere necessario possedere una copia del prodotto o versione del prodotto in tale lingua per potervi accedere.

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. Qualsiasi prodotto funzionalmente equivalente al prodotto, programma o servizio che non violi alcun diritto di proprietà intellettuale IBM può essere utilizzato. È comunque responsabilità dell'utente valutare e verificare la possibilità di utilizzare altri prodotti, programmi o servizi non IBM.

IBM può avere applicazioni di brevetti o brevetti in corso relativi all'argomento descritto in questo documento. La fornitura del presente documento non concede alcuna licenza a tali 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
US*

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.

Queste informazioni potrebbero contenere imprecisioni tecniche o errori tipografici. Le modifiche alle presenti informazioni vengono effettuate periodicamente; tali modifiche saranno incorporate nelle nuove pubblicazioni 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 presenti in tali siti web non sono parte dei materiali per questo prodotto IBM e l'utilizzo di tali siti è a proprio rischio.

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

US

Tali informazioni possono essere disponibili, in base ad appropriate clausole e condizioni, includendo in alcuni casi, il pagamento di una tassa.

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.

Gli esempi client e dati delle prestazioni citati sono presentati esclusivamente 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 o 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.

Le dichiarazioni relative all'orientamento o alle intenzioni future di IBM sono soggette a modifica o a ritiro senza preavviso e rappresentano solo mete e obiettivi.

Questa pubblicazione contiene esempi di dati e prospetti utilizzati quotidianamente nelle operazioni aziendali. 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 applicativi di esempio in linguaggio sorgente, che illustrano 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 "NELLO STATO IN CUI SI TROVANO", senza alcun tipo di garanzia. 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 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.

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, il logo Intel, Intel Inside, il logo Intel Inside, Intel Centrino, il logo Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium e Pentium sono marchi o marchi registrati di Intel Corporation o delle sue consociate negli Stati Uniti e in altri paesi.

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

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

analisi cluster [7](#)
analisi cluster (opzione Marketing Direct) [7](#)

C

cluster [7](#)

P

profili di prospettiva (opzione Marketing Direct) [8](#)
Propensione all'acquisto [14](#)

R

regressione logistica (Direct Marketing) [14](#)
RFM
 dati del cliente [2](#)
 dati sulle transazioni [2](#)
 raccolta [3](#)

T

tariffe di risposta del codice postale [11](#)
tariffe di risposta del codice zip [11](#)
Test package di controllo [17](#)

