

IBM SPSS Direct Marketing 22

IBM

Nota

Antes de utilizar esta información y el producto al cual proporciona soporte, lea la información de "Avisos" en la página 27.

Información de producto

Esta edición se aplica a la versión 22, release 0, modificación 0 de IBM SPSS Statistics y a todos los releases y modificaciones posteriores hasta que se indique lo contrario en ediciones nuevas.

Contenido

Capítulo 1. Marketing directo. 1

Capítulo 2. Análisis de RFM 3

Puntuaciones de RFM desde datos de transacciones . 4

Puntuaciones de RFM desde datos de cliente . . . 4

Agrupación de RFM. 5

Almacenamiento de puntuaciones de RFM desde
datos de transacciones 6

Almacenamiento de puntuaciones de RFM desde
datos de cliente 7

Resultados de RFM 8

Capítulo 3. Análisis de clústeres. 11

Clústeres 12

Capítulo 4. Perfiles de posibles clientes 13

Configuración 14

Creación de un campo de respuesta categórica . . 15

Capítulo 5. Índices de respuesta de código postal 17

Configuración 18

Creación de un campo de respuesta categórica . . 19

Capítulo 6. Propensión a la compra . . 21

Configuración 23

Creación de un campo de respuesta categórica . . 24

Capítulo 7. Prueba de paquete de control. 25

Avisos 27

Marcas comerciales 29

Índice 31

Capítulo 1. Marketing directo

La opción Marketing directo ofrece un conjunto de herramientas diseñadas para mejorar el resultado de campañas de marketing directo identificando y adquiriendo características y otras características que definen a diferentes grupos de consumidores y dirigiéndose a grupos concretos para aumentar al máximo los índices de respuesta positivos.

Análisis de RFM. Esta técnica identifica a clientes actuales con más posibilidades de responder a una nueva oferta.

Análisis de clústeres. Es una herramienta de exploración diseñada para revelar agrupaciones (o clústeres) naturales dentro de sus datos. Por ejemplo, puede identificar diferentes grupos de clientes basándose en diversas características demográficas y de adquisición .

Perfiles de posibles clientes. Esta técnica utiliza resultados de una prueba o campaña anterior para crear perfiles descriptivos. Puede utilizar los perfiles para dirigirse a grupos concretos de contactos en campañas futuras. Para obtener más información, consulte el tema Capítulo 4, “Perfiles de posibles clientes”, en la página 13.

Índices de respuesta de código postal. Esta técnica utiliza resultados de una campaña anterior para calcular índices de respuesta de código postal. Los índices de respuesta se pueden utilizar para dirigirse a códigos postales concretos en campañas futuras. Para obtener más información, consulte el tema Capítulo 5, “Índices de respuesta de código postal”, en la página 17.

Propensión a la compra. Esta técnica utiliza los resultados de un envío de correos de prueba o de campañas anteriores para generar puntuaciones de propensión. Los resultados indican los contactos que tienen más posibilidades de responder. Para obtener más información, consulte el tema Capítulo 6, “Propensión a la compra”, en la página 21.

Prueba de paquete de control. Esta técnica compara campañas de marketing para ver si hay diferencias significativas en la eficacia entre los distintos paquetes u ofertas. Para obtener más información, consulte el tema Capítulo 7, “Prueba de paquete de control”, en la página 25.

Capítulo 2. Análisis de RFM

Análisis de RFM es una técnica utilizada para identificar a clientes actuales que tienen más posibilidades de responder a una nueva oferta. Esta técnica es muy común en el marketing directo. El análisis de RFM se basa en la siguiente teoría simple:

- El factor más importante para identificar a los clientes con más posibilidades de responder a una nueva oferta es la **actualidad**. Los clientes que han realizado adquisiciones recientemente tienen más posibilidades de volver a adquirir nuevos productos que aquellos clientes que han adquirido productos en el pasado.
- El segundo factor más importante es la **frecuencia**. Los clientes que han adquirido más productos tienen más posibilidades de responder que aquellos que han adquirido menos productos.
- El tercer factor más importante es la cantidad total invertida, a la que se le denomina **valor monetario**. Los clientes que han invertido más cantidad (en total en todas las compras) en el pasado tienen más posibilidades de responder que aquellos que han invertido menos.

Funcionamiento del análisis RFM

- A los clientes se les asigna una puntuación de actualidad en función de la fecha de compra más reciente o el intervalo de tiempo transcurrido desde la compra más reciente. Esta puntuación se basa en una clasificación simple de puntuación de actualidad en un pequeño número de categorías. Por ejemplo, si utiliza cinco categorías, los clientes con las fechas de compra más recientes reciben una puntuación de 5, y los clientes con las fechas de compra menos recientes, recibe una puntuación de 1.
- De la misma forma, a los clientes se les asigna una clasificación de frecuencia, con los valores más altos para las clientes con compras más recientes. Por ejemplo, en una escala del uno al cinco, los clientes que adquieran productos con más frecuencia reciben una clasificación de frecuencia de 5.
- Finalmente, los clientes se clasifican según el valor monetario, donde los mayores valores monetarios reciben una clasificación más alta. Continuando con el ejemplo anterior, los clientes que han invertido mayores cantidades reciben un valor monetario de 5.

Al final, cada cliente recibe cuatro puntuaciones: actualidad, frecuencia, valor monetario y puntuación de RFM combinada, que combina simplemente las tres puntuaciones individuales en una única puntuación. Los "mejores" clientes (aquellos que tienen más posibilidades de responder a una oferta) son los que tengan una mayor puntuación de RFM combinada. Por ejemplo, en una escala del uno al cinco, existen un total de 125 posibles puntuaciones RFM combinadas y la mayor puntuación de RFM combinada es de 555.

Consideraciones de los datos

- Si las filas de datos representan transacciones (cada fila representa una transacción simple y cada cliente puede tener múltiples transacciones), utilice RFM desde datos de transacciones. Para obtener más información, consulte el tema "Puntuaciones de RFM desde datos de transacciones" en la página 4.
- Si las filas de datos representan clientes con información de resumen de todas las transacciones (con columnas que contienen puntuaciones de la cantidad total invertida, número total de transacciones y fecha de transacción más reciente), utilice RFM desde datos de cliente. Para obtener más información, consulte el tema "Puntuaciones de RFM desde datos de cliente" en la página 4.

Puntuaciones de RFM desde datos de transacciones

Consideraciones de los datos

El conjunto de datos debe contener variables con la siguiente información:

- Una variable o combinación de variables que identifiquen a cada caso (cliente).
- Una variable con la fecha de cada transacción.
- Una variable con el valor monetario de cada transacción.

Creación de puntuaciones de RFM desde datos de transacciones

1. En los menús, elija:

Marketing directo > Seleccionar técnica

2. Seleccione **Ayudarme a identificar los mejores contactos (Análisis RFM)** y pulse en **Continuar**.
3. Seleccione **Datos de transacciones y**, a continuación, pulse en **Continuar**.
4. Seleccione la variable que contiene las fechas de transacción.
5. Seleccione la variable que contiene la cantidad monetaria de cada transacción.
6. Seleccione el método para resumir las cantidades de transacción de cada cliente: total (suma de las transacciones), media, mediana o máxima (cantidad máxima de transacciones).
7. Seleccione la variable o combinación de variables que identifiquen de forma exclusiva cada caso del cliente. Por ejemplo, los casos se pueden identificar por un código de identificación exclusivo o una combinación del nombre y el apellido.

Puntuaciones de RFM desde datos de cliente

Consideraciones de los datos

El conjunto de datos debe contener variables con la siguiente información:

- Fecha de compra más reciente o intervalo de tiempo desde la fecha de compra más reciente. Se utilizará para calcular las puntuaciones de actualidad.
- Número total de adquisiciones. Se utilizará para calcular la puntuación de frecuencia.
- Valor monetario de resumen para todas las compras. Se utilizará para calcular valores monetarios. Normalmente, es la suma (total) de todas las compras, pero puede ser la media (promedio), valor máximo (mayor cantidad) u otras medidas de resumen.

Si desea escribir puntuaciones RFM en un nuevo conjunto de datos, el conjunto de datos activo también debe contener una variable o una combinación de variables que identifiquen a cada caso (cliente).

Creación de puntuaciones de RFM desde datos de cliente

1. En los menús, elija:

Marketing directo > Seleccionar técnica

2. Seleccione **Ayudarme a identificar los mejores contactos (Análisis RFM)** y pulse en **Continuar**.
3. Seleccione **Datos de cliente y**, a continuación, pulse en **Continuar**.
4. Seleccione la variable que contiene la fecha de transacción más reciente o un número que represente un intervalo de tiempo desde la transacción más reciente.
5. Seleccione la variable que contienen el número total de transacciones de cada cliente.
6. Seleccione la variable que contiene la cantidad monetaria de resumen de cada cliente.
7. Si desea escribir las puntuaciones RFM en un nuevo conjunto de datos, seleccione la variable o combinación de variables que identifican de forma exclusiva a cada cliente. Por ejemplo, los casos se pueden identificar por un código de identificación exclusivo o una combinación del nombre y el apellido.

Agrupación de RFM

El proceso de agrupar un gran número de valores numéricos en un pequeño número de categorías se denomina **agrupación**. En análisis de RFM, los grupos se clasifican en categorías. Puede utilizar la pestaña Agrupación para modificar el método utilizado para asignar los valores de actualidad, frecuencia y monetarios a esos intervalos.

Método para intervalo

Anidadas. En la agrupación anidada, se asigna una puntuación simple a los valores de actualidad. En cada puntuación de actualidad, a los clientes se les asigna una puntuación de frecuencia y en cada puntuación de frecuencia a los clientes se les asigna un valor monetario. Este método tiende a proporcionar una distribución más uniforme de puntuaciones de RFM combinadas, pero tiene la desventaja de hacer que las puntuaciones de frecuencia y monetarias son más difíciles de interpretar. Por ejemplo, es posible que una puntuación de frecuencia de 5 para un cliente con una puntuación de actualidad de 5 no signifique lo mismo que una puntuación de frecuencia de 5 para un cliente con una puntuación de actualidad de 4, ya que la puntuación de frecuencia depende de la puntuación de actualidad.

Independiente. Las puntuaciones simples se asignan a actualidad, frecuencia y valores monetarios. Los tres rangos se asignan de forma independiente. La interpretación de cada uno de los tres componentes de RFM no es ambigua; una puntuación de frecuencia de 5 para un cliente significa lo mismo que una puntuación de frecuencia de 5 para otro cliente, con independencia de sus puntuaciones de actualidad. Para muestras menores tiene la desventaja de producir una distribución menos uniforme de las puntuaciones de RFM combinadas.

Número de intervalos

El número de categorías (intervalos) que utiliza cada componente para crear puntuaciones RFM. El número total de puntuaciones de RFM combinadas es el producto de los tres valores. Por ejemplo, 5 grupos de actualidad, 4 grupos de frecuencia y 3 grupos monetarios crean un total de 60 puntuaciones de RFM combinadas posibles, entre 111 y 543.

- El valor predeterminado es de 5 para cada componente, que crean un total de 125 puntuaciones de RFM combinadas posibles, entre 111 y 555.
- El número máximo de grupos permitidos para cada componente es nueve.

Empates

Un "empate" son simplemente dos o más valores de actualidad, frecuencia o monetarios. Lo ideal es que desee tener aproximadamente el mismo número de clientes en cada grupo, pero un gran número de valores empatados que pueden afectar a la distribución de los grupos. Existen alternativas para gestionar los empates:

- **Asignación de empates al mismo grupo.** Este método asigna siempre valores empatados al mismo grupo, con independencia de cómo afecta a la distribución de grupos. Proporciona un método de agrupación coherente: Si dos clientes tienen el mismo valor de actualidad, siempre se les asignará la misma puntuación de actualidad. En un ejemplo extremo, sin embargo, puede tener 1.000 clientes y 500 de ellos realizando su adquisición más reciente en la misma fecha. En una clasificación del 1 al 5, el 50% de los clientes recibirían una puntuación de actualidad de 5, en lugar del valor ideal del 20%.

Tenga en cuenta que con el método de agrupación anidado, "coherencia" es algo más complicado para las puntuaciones de frecuencia y de valor monetario, ya que las puntuaciones de frecuencia se asignan en grupos de puntuación de actualidad y los valores monetarios se asignan en grupos de puntuaciones de frecuencia. Es posible que dos clientes con el mismo valor de frecuencia no tengan la misma puntuación de frecuencia si tampoco tienen la misma puntuación de actualidad, con independencia de la forma en la que se gestionan los valores empatados.

- **Asignar aleatoriamente los empates.** De esta forma se asegura una distribución uniforme de grupos al asignar un factor de varianza muy pequeño a los empates antes de la clasificación; para que al asignar los valores a los grupos de puntuaciones, no existan valores empatados. Este proceso no afecta a los valores originales. Sólo se utiliza para deshacer los empates. Mientras que este método produce una distribución uniforme de los grupos (aproximadamente el mismo número de clientes en cada grupo), puede resultar en puntuaciones muy diferentes para clientes que parecen tener valores de actualidad, frecuencia y/o monetarios iguales: especialmente si el número total de clientes es relativamente pequeño y/o el número de empates es relativamente alto.

Tabla 1. Asignación de empates al mismo grupo vs. Asignación de valores de forma aleatoria.

ID	Compra más reciente (Actualidad)	Asignación de empates al mismo grupo	Asignación de empates aleatoriamente
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

- En este ejemplo, la asignación de empates al mismo grupo resulta en una distribución de grupos no uniforme: 5 (10%), 4 (40%), 3 (20%), 2 (20%), 1 (10%).
- La asignación de empates de forma aleatoria da como resultado el 20% en cada grupo, pero para lograr este resultado, los cuatro casos con un valor de fecha de 10/28/2006 se asignan a 3 grupos diferentes y los 2 casos con un valor de fecha de 8/13/2006 también se asignan a grupos diferentes.

Tenga en cuenta que la manera en que se asignan los empates a grupos diferentes es totalmente aleatorio (con la limitación de que el resultado final es igual al número de casos de cada grupo). Si ha calculado un segundo conjunto de puntuaciones con el mismo método, la clasificación de un caso particular con un valor empatado puede cambiar. Por ejemplo, las clasificaciones de actualidad de 5 y 3 para los casos 4 y 5 respectivamente se pueden cambiar la segunda vez.

Almacenamiento de puntuaciones de RFM desde datos de transacciones

RFM desde datos de transacciones siempre crea un nuevo conjunto de datos agregado con una fila para cada cliente. Utilice la pestaña Guardar para especificar las puntuaciones y otras variables que desee guardar y la ubicación en la que desea guardarlas.

Variables

Las variables de ID que identifican de forma exclusiva a cada cliente se guardan automáticamente en el nuevo conjunto de datos. Las siguientes variables adicionales se pueden guardar en el nuevo conjunto de datos:

- **Fecha de la transacción más reciente de cada cliente.**
- **Número de transacciones.** El número total de filas de transacciones de cada cliente.
- **Cantidad.** La cantidad de resumen de cada cliente basada en el método de resumen que seleccione en la pestaña Variables.

- **Puntuación de actualidad.** La puntuación asignada a cada cliente basada en la fecha de transacción más reciente. Las puntuaciones más altas indican las fechas de transacción más recientes.
- **Puntuación de frecuencia.** La puntuación asignada a cada cliente basada en el número total de transacciones. Las puntuaciones más altas indican mayor cantidad de transacciones.
- **Valor monetario.** La puntuación asignada a cada cliente basada en la medida de resumen monetario seleccionada. Las puntuaciones más altas indican un valor monetario mayor para la medida de resumen.
- **Puntuación de RFM.** Las tres puntuaciones individuales combinadas en un único valor: $(actualidad \times 100) + (frecuencia \times 10) + valor\ monetario$.

De forma predeterminada, todas las variables disponibles se incluyen en el nuevo conjunto de datos; por lo tanto, cancele la selección de las variables que no desee incluir. También puede especificar sus propios nombres de variable. Los nombres de las variables deben cumplir las normas de denominación de variables.

Ubicación

RFM desde datos de transacciones siempre crea un nuevo conjunto de datos agregado con una fila para cada cliente. Puede crear un nuevo conjunto de datos en la sesión actual o guardar los datos de puntuación de RFM en un archivo de datos externo. Los nombres del conjunto de datos deben cumplir las normas de denominación de variables estándar. (Esta restricción no es aplicable a los nombres de archivos de datos externos.)

Almacenamiento de puntuaciones de RFM desde datos de cliente

En datos del cliente puede añadir las variables de puntuación de RFM al conjunto de datos activo o crear un nuevo conjunto de datos que contiene las variables de puntuación seleccionadas. Utilice la pestaña Guardar para especificar las variables de puntuación y otras variables que desee guardar y la ubicación en la que desea guardarlas.

Nombres de las variables guardadas

- **Generar automáticamente nombres exclusivos.** Si añade variables de puntuación al conjunto de datos activo, se asegurará de que los nuevos nombres de variable son exclusivos. Es muy útil si desea añadir varios conjuntos diferentes de puntuaciones de RFM (basado en criterios diferentes) al conjunto de datos activo.
- **Nombres personalizados.** De esta forma puede asignar sus propios nombres de variable a las variables de puntuación. Los nombres de las variables deben cumplir las normas de denominación de variables.

Variables

Seleccione las variables de puntuación que desee almacenar:

- **Puntuación de actualidad.** La puntuación asignada a cada cliente basada en el valor de la variable Fecha de transacción o Intervalo seleccionada en la pestaña Variables. Las puntuaciones más altas se asignan a las fechas más recientes o a valores de intervalos menores.
- **Puntuación de frecuencia.** La puntuación asignada a cada cliente basada en la variable Número de transacciones seleccionada en la pestaña Variables. Las puntuaciones más altas se asignan a los valores mayores.
- **Valor monetario.** La puntuación asignada a cada cliente basada en la variable Cantidades seleccionada en la pestaña Variables. Las puntuaciones más altas se asignan a los valores mayores.
- **Puntuación de RFM.** Las tres puntuaciones individuales combinadas en un único valor: $(actualidad \times 100) + (frecuencia \times 10) + valor\ monetario$.

Ubicación

Para datos de cliente, existen tres alternativas para las que puede guardar nuevas puntuaciones de RFM:

- **Conjunto de datos activo.** Las variables de puntuaciones de RFM se añaden al conjunto de datos activo.
- **Nuevo conjunto de datos.** Las variables de puntuación de RFM seleccionadas y las variables de ID que identifican de forma exclusiva a cada cliente (caso) se sobrescribirán en un nuevo conjunto de datos en la sesión actual. Los nombres del conjunto de datos deben cumplir las normas de denominación de variables estándar. Esta opción sólo está disponible si selecciona una o más variables de identificación de cliente en la pestaña Variables.
- **Archivo.** Las puntuaciones de RFM seleccionadas y las variables de ID que identifican de forma exclusiva a cada cliente (caso) se guardarán en un archivo de datos externo. Esta opción sólo está disponible si selecciona una o más variables de identificación de cliente en la pestaña Variables.

Resultados de RFM

Datos de intervalo

Los gráficos y tablas de datos agrupados se basan en las puntuaciones de actualidad y frecuencia y valores monetarios calculados.

Mapa de calor de valor monetario medio según actualidad y frecuencia. El mapa de calor de la distribución monetaria media muestra el valor monetario medio de las categorías definidas según las puntuaciones de actualidad y frecuencia. Las zonas más oscuras indican un valor monetario medio más alto.

Gráfico de recuentos de intervalos. El gráfico de recuentos de intervalos muestra la distribución de intervalos para el método de agrupación seleccionado. Cada barra representa el número de casos que se asignará a cada puntuación de RFM combinada.

- Aunque normalmente desea una distribución bastante homogénea, con todas las barras (o la mayoría de ellas) aproximadamente con la misma altura, debe esperar cierta cantidad de varianza al utilizar el método de agrupación predeterminado que asigna valores empatados al mismo grupo.
- Las fluctuaciones extremas en la distribución de grupos y/o muchos grupos vacíos pueden indicar que debería probar con otro método de agrupación (menos grupos y/o la asignación aleatoria de empates) o reconsiderar la adecuación del análisis de RFM.

Tabla de recuentos de intervalos. La misma información que se encuentra en el gráfico de recuentos de intervalos, sólo que expresada en forma de tabla, con recuentos de intervalos en cada casilla.

Datos sin intervalo

Los gráficos y tablas de datos no agrupados se basan en las variables originales utilizadas para crear puntuaciones de actualidad y frecuencia y valores monetarios.

Histogramas. Los histogramas muestran la distribución relativa de valores para las tres variables utilizadas para calcular puntuaciones de actualidad y frecuencia y valores monetarios. No es raro que estos histogramas indiquen distribuciones algo asimétricas en lugar de una distribución normal o simétrica.

El eje horizontal de cada histograma siempre está ordenado desde los valores más bajos en la izquierda hasta los valores más altos a la derecha. Sin embargo, con la actualidad, la interpretación del gráfico depende del tipo de medida de actualidad: fecha o intervalo de tiempo. En el caso de fechas, las barras de la izquierda representan valores más antiguos (una fecha menos reciente tiene un valor inferior que una fecha más reciente). En el caso de intervalos de tiempo, las barras de la izquierda representan los valores más recientes (cuanto menor sea el intervalo de tiempo, más reciente será la transacción).

Diagramas de dispersión de parejas de variables. Estos diagramas de dispersión muestran las relaciones entre las tres variables utilizadas para calcular puntuaciones de actualidad y frecuencia y valores monetarios.

También es normal ver agrupaciones lineales apreciables de puntos sobre la escala de frecuencia, ya que a menudo la frecuencia representa un rango relativamente pequeño de valores discretos. Por ejemplo, si el número total de transacciones no supera 15, entonces sólo hay 15 valores de frecuencia posibles (a menos que cuente las transacciones fraccionarias), mientras que podría haber cientos de valores de actualidad y miles de valores monetarios posibles.

La interpretación del eje de actualidad depende del tipo de medida de actualidad: fecha o intervalo de tiempo. En el caso de fechas, los puntos más cercanos al origen representan fechas más antiguas. En el caso de intervalos de tiempo, los puntos más cercanos al origen representan los valores más recientes.

Capítulo 3. Análisis de clústeres

El análisis de clústeres es una herramienta de exploración diseñada para revelar agrupaciones (o clústeres) naturales dentro de sus datos. Por ejemplo, puede identificar diferentes grupos de clientes basándose en diversas características demográficas y de adquisición .

Ejemplo. Las empresas minoristas y de venta de productos para el consumidor suelen aplicar técnicas de agrupación en clústeres a los datos que describen los hábitos de consumo, sexo, edad, nivel de ingresos, etc. de los clientes. Estas empresas adaptan sus estrategias de desarrollo de productos y de marketing en función de cada grupo de consumidores para aumentar las ventas y el nivel de fidelidad a la marca.

Análisis de clústeres: Consideraciones sobre los datos

Datos. Este procedimiento trabaja tanto con campos variables continuos como categóricos. Cada registro (fila) representa un cliente a agrupar en clústeres y los campos (variables) representan los atributos en los que se va a basar la agrupación en clústeres.

Orden de registro. Tenga en cuenta que los resultados pueden depender del orden de los registros. Para minimizar los efectos del orden, es posible que desee considerar ordenar los registros de forma aleatoria. Puede que desee ejecutar el análisis varias veces, con los registros ordenados en distintos órdenes aleatorios para comprobar la estabilidad de una solución determinada.

Nivel de medición. La asignación del nivel correcto de medición es importante porque afecta al cálculo de los resultados.

- *Nominal.* Una variable puede ser tratada como nominal cuando sus valores representan categorías que no obedecen a una clasificación intrínseca. Por ejemplo, el departamento de la compañía en el que trabaja un empleado. Algunos ejemplos de variables nominales son: región, código postal o confesión religiosa.
- *Ordinal.* Una variable puede ser tratada como ordinal cuando sus valores representan categorías con alguna clasificación intrínseca. Por ejemplo, los niveles de satisfacción con un servicio, que abarquen desde muy insatisfecho hasta muy satisfecho. Entre los ejemplos de variables ordinales se incluyen escalas de actitud que representan el grado de satisfacción o confianza y las puntuaciones de evaluación de las preferencias.
- *Continua.* Una variable puede tratarse como escala (continua) cuando sus valores representan categorías ordenadas con una métrica con significado, por lo que son adecuadas las comparaciones de distancia entre valores. Son ejemplos de variables de escala: la edad en años y los ingresos en dólares.

Un icono junto a cada campo indica el nivel de medición actual.

Tabla 2. Iconos de nivel de medición

	Numérico	Cadena	Fecha	Hora
Escala (Continuo)		n/a		
Ordinal				
Nominal				

Puede cambiar el nivel de medición en una vista de variables del Editor de datos o puede usar el cuadro de diálogo Definir propiedades de variables para sugerir un nivel de medición adecuado para cada campo .

Campos con un nivel de medición desconocido

La alerta de nivel de medición se muestra si el nivel de medición de una o más variables (campos) del conjunto de datos es desconocido. Como el nivel de medición afecta al cálculo de los resultados de este procedimiento, todas las variables deben tener un nivel de medición definido.

Explorar datos. Lee los datos del conjunto de datos activo y asigna el nivel de medición predefinido en cualquier campo con un nivel de medición desconocido. Si el conjunto de datos es grande, puede llevar algún tiempo.

Asignar manualmente. Abre un cuadro de diálogo que contiene todos los campos con un nivel de medición desconocido. Puede utilizar este cuadro de diálogo para asignar el nivel de medición a esos campos. También puede asignar un nivel de medición en la Vista de variables del Editor de datos.

Como el nivel de medición es importante para este procedimiento, no puede acceder al cuadro de diálogo para ejecutar este procedimiento hasta que se hayan definido todos los campos en el nivel de medición.

Para obtener un análisis de clústeres

En los menús, elija:

Marketing directo > Seleccionar técnica

1. Seleccione **Dividir mis contactos en clústeres**.
2. Seleccione los campos categóricos (nominal, ordinal) y continuo (escala) que desea utilizar para crear los segmentos.
3. Pulse en **Ejecutar** para ejecutar el procedimiento.

Clústeres

La pestaña Configuración permite mostrar o eliminar la visualización de gráficos y tablas que describen los segmentos, guardar un nuevo campo en el conjunto de datos que identifica el segmento (clúster) de cada registro en el conjunto de datos y especificar cuántos segmentos se incluirán en la solución de clústeres.

Visualizar gráficos y tablas. Muestra tablas y gráficos que describen los segmentos.

Pertenencia de segmento. Guarda un nuevo campo (variable) que identifica el segmento al que pertenece cada registro.

- Los nombres de los campos deben cumplir las reglas de nombres de IBM® SPSS Statistics.
- El campo de pertenencia de segmento no puede tener un nombre que ya exista en el conjunto de datos. Si ejecuta este procedimiento más de una vez en el conjunto de datos, necesitará especificar un nombre diferente cada vez.
- **Número de segmentos.** Controla cómo se determina el número de segmentos.
- **Determinar automáticamente.** El procedimiento determinará automáticamente el "mejor" número de segmentos, hasta el máximo especificado.

Especificar número fijo. El procedimiento producirá el número especificado de segmentos.

Capítulo 4. Perfiles de posibles clientes

Esta técnica utiliza resultados de una prueba o campaña anterior para crear perfiles descriptivos. Puede utilizar los perfiles para dirigirse a grupos concretos de contactos en campañas futuras. El campo Respuesta indica quién ha respondido a la campaña o prueba anterior. La lista Perfiles contiene las características que desea utilizar para crear el perfil.

Ejemplo. Basándose en los resultados de un envío de correos de prueba, el departamento de marketing directo de una empresa desea generar perfiles de los tipos de clientes con más posibilidades de responder a la oferta, en función de la información demográfica.

Resultados

El resultado incluye una tabla que proporciona una descripción de cada grupo de perfil y muestra índices de respuesta (porcentaje de respuestas positivas) e índices de respuesta acumulados y un gráfico de índices de respuesta acumulados. Si incluye un índice de respuesta de destino mínimo, la tabla tendrá una codificación de colores para mostrar los perfiles que cumplen el índice de respuesta acumulado mínimo y el gráfico incluye una línea de referencia en el valor de índice de respuesta mínimo especificado.

Consideraciones sobre los datos de perfiles de posibles clientes

Campo respuesta. El campo de respuesta puede ser nominal u ordinal. Puede ser numérico o de cadena. Si este campo contiene un valor que indica un número o cantidad de compras, necesitará crear un nuevo campo en el que un único valor represente todas las respuestas positivas. Para obtener más información, consulte el tema “Creación de un campo de respuesta categórica” en la página 15.

Valor de respuesta positiva. El valor de respuesta positiva identifica los clientes que han respondido de forma positiva (por ejemplo, los que han hecho una compra). Se asume que el resto de valores de respuesta no ausentes indican una respuesta negativa. Si hay alguna etiqueta de valor definida para el campo de respuesta, esas etiquetas se muestran en la lista desplegable.

Crear perfiles con. Estos campos pueden ser nominales, ordinales o continuos (escale). Pueden ser numéricos o de cadena.

Nivel de medición. La asignación del nivel correcto de medición es importante porque afecta al cálculo de los resultados.

- *Nominal.* Una variable puede ser tratada como nominal cuando sus valores representan categorías que no obedecen a una clasificación intrínseca. Por ejemplo, el departamento de la compañía en el que trabaja un empleado. Algunos ejemplos de variables nominales son: región, código postal o confesión religiosa.
- *Ordinal.* Una variable puede ser tratada como ordinal cuando sus valores representan categorías con alguna clasificación intrínseca. Por ejemplo, los niveles de satisfacción con un servicio, que abarquen desde muy insatisfecho hasta muy satisfecho. Entre los ejemplos de variables ordinales se incluyen escalas de actitud que representan el grado de satisfacción o confianza y las puntuaciones de evaluación de las preferencias.
- *Continua.* Una variable puede tratarse como escala (continua) cuando sus valores representan categorías ordenadas con una métrica con significado, por lo que son adecuadas las comparaciones de distancia entre valores. Son ejemplos de variables de escala: la edad en años y los ingresos en dólares.

Un icono junto a cada campo indica el nivel de medición actual.

Tabla 3. Iconos de nivel de medición

	Numérico	Cadena	Fecha	Hora
Escala (Continuo)		n/a		
Ordinal				
Nominal				

Puede cambiar el nivel de medición en una vista de variables del Editor de datos o puede usar el cuadro de diálogo Definir propiedades de variables para sugerir un nivel de medición adecuado para cada campo .

Campos con un nivel de medición desconocido

La alerta de nivel de medición se muestra si el nivel de medición de una o más variables (campos) del conjunto de datos es desconocido. Como el nivel de medición afecta al cálculo de los resultados de este procedimiento, todas las variables deben tener un nivel de medición definido.

Explorar datos. Lee los datos del conjunto de datos activo y asigna el nivel de medición predefinido en cualquier campo con un nivel de medición desconocido. Si el conjunto de datos es grande, puede llevar algún tiempo.

Asignar manualmente. Abre un cuadro de diálogo que contiene todos los campos con un nivel de medición desconocido. Puede utilizar este cuadro de diálogo para asignar el nivel de medición a esos campos. También puede asignar un nivel de medición en la Vista de variables del Editor de datos.

Como el nivel de medición es importante para este procedimiento, no puede acceder al cuadro de diálogo para ejecutar este procedimiento hasta que se hayan definido todos los campos en el nivel de medición.

Para obtener perfiles de posibles clientes

En los menús, elija:

Marketing directo > Seleccionar técnica

1. Seleccione **Generar perfiles de mis contactos que hayan respondido a una oferta**.
2. Seleccione el campo que identifica los contactos que responderán a la oferta. El campo puede ser nominal u ordinal.
3. Introduzca el valor que indica una respuesta positiva. Si algún valor ha definido etiquetas de valores, puede seleccionar la etiqueta de valor de la lista desplegable y se mostrará el valor correspondiente.
4. Seleccione los campos que desea utilizar para crear los perfiles.
5. Pulse en **Ejecutar** para ejecutar el procedimiento.

Configuración

La pestaña Configuración permite controlar el tamaño del grupo de perfil mínimo e incluir un umbral mínimo de índice de respuesta en el resultado.

Tamaño mínimo de grupo de perfil. Cada perfil representa las características compartidas de un grupo de contactos en el conjunto de datos (por ejemplo, mujeres menores de 40 años que viven en la costa

oeste). De forma predeterminada, el menor tamaño de grupo de perfil es 100. Los tamaños de grupo menores pueden revelar más grupos, pero los tamaños de grupos más grandes proporcionan resultados más fiables. El valor debe ser un entero positivo.

Incluir un mínimo de información de umbral del índice de respuesta en los resultados. Los resultados incluyen una tabla que muestra índices de respuesta (porcentaje de respuestas positivas) e índices de respuesta acumulados y un gráfico de índices de respuesta acumulados. Si introduce un índice de respuesta de destino mínimo, la tabla tendrá una codificación de colores para mostrar los perfiles que cumplen el índice de respuesta acumulado mínimo y el gráfico incluye una línea de referencia en el valor de índice de respuesta mínimo especificado. El valor debe ser mayor que 0 y menor que 100.

Creación de un campo de respuesta categórica

El campo de respuesta debe ser categórico, con un valor que representa todas las respuestas positivas. Se asume que cualquier otro valor no perdido es una respuesta negativa. Si el campo de respuesta representa un valor continuo (escala), como el número de compras o las cantidades de compras, necesita crear un nuevo campo que asigna un valor de respuesta simple a todos los valores de respuesta distintos a cero.

- Si las respuestas negativas se registran como 0 (espacio cumplimentado, que se tratan como valores perdidos), se pueden calcular con la siguiente fórmula:

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

donde *NewName* es el nombre del nuevo campo y *OldName* es el nombre del campo original. Es una expresión lógica que asigna un valor de 1 a todos los valores no perdidos mayores que 0 y 0 a todos los valores no perdidos menores o iguales que 0.

- Si no se registra ningún valor para respuestas negativas, estos valores se consideran perdidos y la fórmula se complica un poco más:

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

En esta expresión lógica, a todos los valores de respuesta no perdidos se les asigna un valor de 1 y a todos los valores de respuesta no perdidos se les asigna un valor de 0.

- Si no puede distinguir entre valores de respuesta negativos (0) y valores perdidos, no se puede calcular un valor de respuesta preciso. Si hay relativamente pocos valores perdidos, es posible que no tenga un efecto significativo en los índices de respuesta calculados. Sin embargo, si hay demasiados valores perdidos, como cuando se registra información de respuesta para una pequeña prueba de muestra del conjunto de datos total, los índices de respuesta calculados no serán significativos, ya que son significativamente menores que los índices de respuesta verdaderos.

Creación de un campo de respuesta categórica

1. En los menús, elija:

Transformar > Calcular variable

2. Para una variable de destino, introduzca el nuevo nombre de campo (variable).
3. Si se registran respuestas negativas como 0, introduzca $\text{OldName} > 0$ para Expresión numérica, donde *OldName* es el nombre del campo original.
4. Si se registran respuestas negativas como perdidas (en blanco), introduzca $\text{NOT}(\text{MISSING}(\text{OldName}))$ para Expresión numérica, donde *OldName* es el nombre del campo original.

Capítulo 5. Índices de respuesta de código postal

Esta técnica utiliza resultados de una campaña anterior para calcular índices de respuesta de código postal. Los índices de respuesta se pueden utilizar para dirigirse a códigos postales concretos en campañas futuras. El campo Respuesta indica quién ha respondido a la campaña anterior. El campo Respuesta indica quién ha respondido a la campaña anterior.

Ejemplo. Basándose en los resultados de un envío anterior, el departamento de marketing directo de una empresa genera índices de respuesta por códigos postales. Basándose en diferentes criterios como un índice de respuesta mínimo aceptable y/o el número máximo de contactos que incluirán en el envío, podrán dirigirse a códigos postales concretos.

Resultados

El resultado de este procedimiento incluye un nuevo conjunto de datos que contiene índices de respuesta por código postal y una tabla y un gráfico que resumen los resultados por deciles (10% superior, 20% superior, etc.). La tabla puede tener una codificación de colores basada en un índice de respuesta acumulado mínimo de un usuario o un número máximo de contactos.

El nuevo conjunto de datos contiene los siguientes campos:

- **Código postal.** Si los grupos de códigos postales se basan sólo en una parte del valor completo, será el valor de esa parte del código postal. La etiqueta de la fila encabezado de esta columna en el archivo Excel es el nombre del campo de código postal en el conjunto de datos original.
- **Índice de respuesta.** El porcentaje de respuestas positivas en cada código postal.
- **Respuestas.** El número de respuestas positivas en cada código postal.
- **Contactos.** El número total de contactos en cada código postal que contiene un valor no perdido en el campo de respuesta.
- **Índice.** La respuesta "ponderada" basada en la fórmula $N \times P \times (1-P)$, donde N es el número de contactos y P es el índice de respuesta expresado como proporción.
- **Rango.** Rango de decil (10% superior, 20% superior, etc.) de los índices de respuesta por código postal en orden descendente.

Consideraciones sobre los datos de índices de respuesta de código postal

Campo respuesta. El campo de respuesta puede ser de cadena o numérico. Si este campo contiene un valor que indica un número o valor monetario de compras, necesitará crear un nuevo campo en el que un único valor represente todas las respuestas positivas. Para obtener más información, consulte el tema "Creación de un campo de respuesta categórica" en la página 19.

Valor de respuesta positiva. El valor de respuesta positiva identifica los clientes que han respondido de forma positiva (por ejemplo, los que han hecho una compra). Se asume que el resto de valores de respuesta no ausentes indican una respuesta negativa. Si hay alguna etiqueta de valor definida para el campo de respuesta, esas etiquetas se muestran en la lista desplegable.

Campo Código postal. El campo código postal puede ser de cadena o numérico.

Para obtener índices de respuesta de código postal

En los menús, elija:

Marketing directo > Seleccionar técnica

1. Seleccione **Identificar principales códigos postales de respuesta**.
2. Seleccione el campo que identifica los contactos que responderán a la oferta.
3. Introduzca el valor que indica una respuesta positiva. Si algún valor ha definido etiquetas de valores, puede seleccionar la etiqueta de valor de la lista desplegable y se mostrará el valor correspondiente.
4. Seleccione el campo que contiene el código postal.
5. Pulse en **Ejecutar** para ejecutar el procedimiento.

Si lo desea, tiene la posibilidad de:

- Generar índices de respuesta basados en los primeros n caracteres o dígitos del código postal en lugar del valor completo
- Guardar automáticamente los resultados en un archivo Excel
- Controlar las opciones de visualización de los resultados

Configuración

Agrupar códigos postales basados en

Determina cómo se agrupan los registros para calcular los índices de respuesta. De forma predeterminada, se utiliza el código postal completo y todos los registros con el mismo código postal se agrupan para calcular el índice de respuesta del grupo. De forma alternativa, puede agrupar los registros en función únicamente de una parte del código postal, utilizando los primeros n dígitos o caracteres. Por ejemplo, es posible que desee agrupar los registros por los 5 primeros caracteres de un código postal de 10 caracteres o por los 3 primeros dígitos de un código postal de 5 dígitos. El conjunto de datos de resultados contendrá un registro para cada grupo de código postal. Si introduce un valor, debe ser un entero positivo.

Formato numérico de código postal

Si el campo de código postal es numérico y desea agrupar los códigos postales basados en los primeros n dígitos en lugar del valor completo, necesita especificar el número de dígitos en el valor original. El número de dígitos es el *máximo* número posible de dígitos en el código postal. Por ejemplo, si el campo del código postal contiene una mezcla de códigos postales de 5 y 9 dígitos, deberá especificar 9 como el número de dígitos.

Nota: dependiendo del formato de visualización, puede parecer que algunos códigos postales de 5 dígitos sólo contienen 4 dígitos, aunque se trata de un cero inicial.

Resultados

Además del nuevo conjunto de datos que contiene índices de respuesta por código postal, puede mostrar una tabla y un gráfico que resumen los resultados por deciles (10% superior, 20% superior, etc.). La tabla muestra índices de respuesta, índices de respuesta acumulados, número de registros y el número acumulado de registros en cada decil. El gráfico muestra los índices de respuesta acumulados y número acumulado de registros de cada decil.

Índice de respuesta mínimo aceptable. Si introduce un índice de respuesta de destino mínimo o una fórmula de equilibrio, la tabla tendrá una codificación de colores para mostrar los deciles que cumplen el índice de respuesta acumulado mínimo y el gráfico incluye una línea de referencia en el valor de índice de respuesta mínimo especificado.

- **Índice de respuesta de destino.** Índice de respuesta expresado como un porcentaje (porcentaje de respuestas positivas en cada grupo de código postal). El valor debe ser mayor que 0 y menor que 100.
- **Calcular índice de equilibrio a partir de fórmula.** Calcula el índice de respuesta acumulado mínimo basado en la fórmula: $(\text{Coste de envío de un paquete} / \text{Ingreso neto por respuesta}) \times 100$. Ambos valores deben ser números enteros positivos. El resultado debe ser un valor mayor que 0 y menor que 100. Por

ejemplo, si el coste de enviar un paquete es 0,75 dólares y el ingreso neto por respuesta es 56 dólares, el índice de respuesta mínimo es: $(0,75/56) \times 100 = 1,34\%$.

Número máximo de contactos. Si especifica un número máximo de contactos, la tabla tendrá una codificación de colores para mostrar los deciles que no superan el número máximo acumulado de contactos (registros) y el gráfico incluirá una línea de referencia en ese valor.

- **Porcentaje de contactos.** Máximo expresado como porcentaje. Por ejemplo, desea conocer los deciles con el mayor índice de respuesta que no contengan más del 50% de todos los contactos. El valor debe ser mayor que 0 y menor que 100.
- **Número de contactos.** Máximo expresado como número de contactos. Por ejemplo, si no va a enviar más de 10.000 paquetes, podrá definir el valor a 10000. El valor debe ser un entero positivo (sin símbolos de agrupamiento).

Si especifica un índice de respuesta mínimo aceptable y un número máximo de contactos, la codificación de colores de la tabla se basa en la condición que se cumpla primero.

Exportar a Excel

Este procedimiento crea automáticamente un nuevo conjunto de datos que contiene índices de respuesta por código postal. Cada registro (fila) del conjunto de datos representa un código postal. Puede guardar automáticamente la misma información en un archivo Excel. Este archivo se guarda en formato Excel 97-2003.

Creación de un campo de respuesta categórica

El campo de respuesta debe ser categórico, con un valor que representa todas las respuestas positivas. Se asume que cualquier otro valor no perdido es una respuesta negativa. Si el campo de respuesta representa un valor continuo (escala), como el número de compras o las cantidades de compras, necesita crear un nuevo campo que asigna un valor de respuesta simple a todos los valores de respuesta distintos a cero.

- Si las respuestas negativas se registran como 0 (espacio cumplimentado, que se tratan como valores perdidos), se pueden calcular con la siguiente fórmula:

$newName = OldName > 0$

donde *NewName* es el nombre del nuevo campo y *OldName* es el nombre del campo original. Es una expresión lógica que asigna un valor de 1 a todos los valores no perdidos mayores que 0 y 0 a todos los valores no perdidos menores o iguales que 0.

- Si no se registra ningún valor para respuestas negativas, estos valores se consideran perdidos y la fórmula se complica un poco más:

$newName = NOT(MISSING(OldName))$

En esta expresión lógica, a todos los valores de respuesta no perdidos se les asigna un valor de 1 y a todos los valores de respuesta no perdidos se les asigna un valor de 0.

- Si no puede distinguir entre valores de respuesta negativos (0) y valores perdidos, no se puede calcular un valor de respuesta preciso. Si hay relativamente pocos valores perdidos, es posible que no tenga un efecto significativo en los índices de respuesta calculados. Sin embargo, si hay demasiados valores perdidos, como cuando se registra información de respuesta para una pequeña prueba de muestra del conjunto de datos total, los índices de respuesta calculados no serán significativos, ya que son significativamente menores que los índices de respuesta verdaderos.

Creación de un campo de respuesta categórica

1. En los menús, elija:

Transformar > Calcular variable

2. Para una variable de destino, introduzca el nuevo nombre de campo (variable).

3. Si se registran respuestas negativas como 0, introduzca `OldName>0` para Expresión numérica, donde *OldName* es el nombre del campo original.
4. Si se registran respuestas negativas como perdidas (en blanco), introduzca `NOT(MISSING(OldName))` para Expresión numérica, donde *OldName* es el nombre del campo original.

Capítulo 6. Propensión a la compra

La propensión a la compra utiliza los resultados de un envío de correos de prueba o de campañas anteriores para generar puntuaciones. Los resultados indican los contactos que tienen más posibilidades de responder. El campo Respuesta indica quién ha respondido al envío de correos de prueba o a la campaña anterior. Los campos de propensión son las características que quiere utilizar para pronosticar las probabilidades de que respondan contactos con características similares.

Esta técnica utiliza la regresión logística binaria para generar un modelo predictivo. El proceso de generar y aplicar un modelo predictivo tiene dos pasos básicos:

1. Crear el modelo y guardar el archivo de modelo. Crea el modelo usando un conjunto de datos cuyo resultado deseado (a menudo conocido como **destino**) ya se conoce. Por ejemplo, si desea crear un modelo que prediga quién tiene más posibilidades de responder a una campaña de correo directo, deberá comenzar con un conjunto de datos que ya tenga información sobre quién respondió y quién no. Por ejemplo, esto puede tratarse de los resultados de un envío de correos de prueba a un pequeño grupo de clientes o información sobre respuestas a una campaña similar en el pasado.
2. Aplicar ese modelo a un conjunto de datos distinto (para el que no se conocen los resultados deseados) para obtener los resultados pronosticados.

Ejemplo. El departamento de marketing directo de una empresa utiliza los resultados de un envío de correos de prueba para asignar puntuaciones de propensión al resto de su base de datos de contactos, utilizando diversas características demográficas para identificar los contactos con más posibilidades de responder y realizar una compra.

Resultados

Este procedimiento crea automáticamente un nuevo campo en el conjunto de datos que contienen puntuaciones de propensión para los datos de prueba y un archivo XML de modelo que se puede utilizar para puntuar otros conjuntos de datos. Los resultados estadísticos opcionales incluyen un gráfico de calidad global de modelo y una tabla de clasificación que compara las respuestas pronosticadas con respuestas actuales.

Consideraciones de los datos de propensión a la compra

Campo respuesta. El campo de respuesta puede ser de cadena o numérico. Si este campo contiene un valor que indica un número o valor monetario de compras, necesitará crear un nuevo campo en el que un único valor represente todas las respuestas positivas. Para obtener más información, consulte el tema “Creación de un campo de respuesta categórica” en la página 24.

Valor de respuesta positiva. El valor de respuesta positiva identifica los clientes que han respondido de forma positiva (por ejemplo, los que han hecho una compra). Se asume que el resto de valores de respuesta no ausentes indican una respuesta negativa. Si hay alguna etiqueta de valor definida para el campo de respuesta, esas etiquetas se muestran en la lista desplegable.

Pronosticar propensión con. Los campos utilizados para pronosticar pueden ser numéricos o de cadena y pueden ser nominales, ordinales o continuos (escala), pero es importante asignar el nivel de medición adecuado a todos los campos predictores.

Nivel de medición. La asignación del nivel correcto de medición es importante porque afecta al cálculo de los resultados.

- *Nominal*. Una variable puede ser tratada como nominal cuando sus valores representan categorías que no obedecen a una clasificación intrínseca. Por ejemplo, el departamento de la compañía en el que trabaja un empleado. Algunos ejemplos de variables nominales son: región, código postal o confesión religiosa.
- *Ordinal*. Una variable puede ser tratada como ordinal cuando sus valores representan categorías con alguna clasificación intrínseca. Por ejemplo, los niveles de satisfacción con un servicio, que abarquen desde muy insatisfecho hasta muy satisfecho. Entre los ejemplos de variables ordinales se incluyen escalas de actitud que representan el grado de satisfacción o confianza y las puntuaciones de evaluación de las preferencias.
- *Continua*. Una variable puede tratarse como escala (continua) cuando sus valores representan categorías ordenadas con una métrica con significado, por lo que son adecuadas las comparaciones de distancia entre valores. Son ejemplos de variables de escala: la edad en años y los ingresos en dólares.

Un icono junto a cada campo indica el nivel de medición actual.

Tabla 4. Iconos de nivel de medición

	Numérico	Cadena	Fecha	Hora
Escala (Continuo)		n/a		
Ordinal				
Nominal				

Puede cambiar el nivel de medición en una vista de variables del Editor de datos o puede usar el cuadro de diálogo Definir propiedades de variables para sugerir un nivel de medición adecuado para cada campo .

Campos con un nivel de medición desconocido

La alerta de nivel de medición se muestra si el nivel de medición de una o más variables (campos) del conjunto de datos es desconocido. Como el nivel de medición afecta al cálculo de los resultados de este procedimiento, todas las variables deben tener un nivel de medición definido.

Explorar datos. Lee los datos del conjunto de datos activo y asigna el nivel de medición predefinido en cualquier campo con un nivel de medición desconocido. Si el conjunto de datos es grande, puede llevar algún tiempo.

Asignar manualmente. Abre un cuadro de diálogo que contiene todos los campos con un nivel de medición desconocido. Puede utilizar este cuadro de diálogo para asignar el nivel de medición a esos campos. También puede asignar un nivel de medición en la Vista de variables del Editor de datos.

Como el nivel de medición es importante para este procedimiento, no puede acceder al cuadro de diálogo para ejecutar este procedimiento hasta que se hayan definido todos los campos en el nivel de medición.

Para obtener puntuaciones de propensión a la compra

En los menús, elija:

Marketing directo > Seleccionar técnica

1. Seleccione **Seleccionar contactos con más probabilidades de realizar la compra**.

2. Seleccione el campo que identifica los contactos que responderán a la oferta.
3. Introduzca el valor que indica una respuesta positiva. Si algún valor ha definido etiquetas de valores, puede seleccionar la etiqueta de valor de la lista desplegable y se mostrará el valor correspondiente.
4. Seleccione los campos que desea utilizar para pronosticar la propensión.
Para guardar un archivo XML modelo para puntuar otros archivos de datos:
5. Seleccione (active) **Exportar información del modelo a un archivo XML**.
6. Introduzca una ruta de directorio y el nombre del archivo o pulse en **Examinar** para desplazarse hasta la ubicación en la que desea guardar el archivo XML de modelo.
7. Pulse en **Ejecutar** para ejecutar el procedimiento.
Para utilizar el archivo modelo para puntuar otros conjuntos de datos:
8. Abra el conjunto de datos que desee puntuar.
9. Utilice el Asistente para puntuación para aplicar el modelo al conjunto de datos. En los menús, elija: **Utilidades > Asistente para puntuación**.

Configuración

Validación de modelos

Validación de modelos crea grupos de entrenamiento y prueba para tareas de diagnóstico. Si selecciona la tabla de clasificación en la sección Resultados diagnósticos, la tabla se dividirá en entrenamiento (seleccionada) y prueba (sin seleccionar) para su comparación. No seleccione la validación de modelos salvo que también seleccione la tabla de clasificación. Los resultados se basan en el modelo generado a partir de la muestra de entrenamiento, que siempre contendrá menos resultados que el número total de registros disponibles. Por ejemplo, el tamaño predeterminado de la muestra de entrenamiento es 50%, y es posible que un modelo creado con sólo la mitad de los registros disponibles no sea tan fiable como el modelo basado en todos los registros disponibles.

- **Tamaño de partición de muestra de entrenamiento (%)**. Especifique el porcentaje de registros que se asignarán a la muestra de entrenamiento. El resto de registros con valores no perdidos para el campo de respuesta se asignan a la muestra de entrenamiento. El valor debe ser mayor que 0 y menor que 100.
- **Establecer semilla para replicar resultados**. Como los registros se asignan de forma aleatoria a las muestras de entrenamiento y prueba, cada vez que ejecute el procedimiento podrá obtener resultados diferentes, salvo que especifique siempre el mismo valor de semilla de aleatorización de inicio.

Resultados diagnósticos

Calidad global de modelo. Muestra un gráfico de la calidad global de modelo, expresada como un valor entre 0 y 1. Un buen modelo debe tener un valor superior a 0,5.

Tabla de clasificación. Muestra una tabla que compara las respuestas positivas y negativas pronosticadas con las respuestas positivas y negativas reales. El porcentaje de precisión global puede proporcionar algún tipo de indicación del funcionamiento del modelo, pero es posible que esté más interesado en el porcentaje de respuestas positivas correctas pronosticadas.

- **Probabilidad mínima**. Asigna registros con una puntuación mayor que el valor especificado en la categoría de respuesta positiva pronosticada en la tabla de clasificación. Las puntuaciones generadas por el procedimiento representan la probabilidad de que el contacto responderá de forma positiva (por ejemplo, realizar una compra). Como norma general, debe especificar un valor cercano a su índice de respuesta de objetivo mínimo, expresado como una proporción. Por ejemplo, si está interesado en un índice de respuesta mínimo del 5%, especifique 0,05. El valor debe ser mayor que 0 y menor que 1.

Nombre y etiqueta del campo de respuesta recodificado

Este procedimiento recodifica automáticamente el campo de respuesta como un nuevo campo en el que 1 representa respuestas positivas y 0 representa respuestas negativas; y el análisis se realiza en el campo recodificado. Puede sustituir el nombre predeterminado y asignar uno de su elección. Los nombres deben cumplir las reglas de nombres de IBM SPSS Statistics .

Guardar puntuaciones

Un nuevo campo con puntuaciones de propensión se guarda automáticamente en el conjunto de datos original. Las puntuaciones representan la probabilidad de una respuesta positiva, expresada como una proporción.

- Los nombres de los campos deben cumplir las reglas de nombres de IBM SPSS Statistics.
- El campo no puede tener un nombre que ya exista en el conjunto de datos. Si ejecuta este procedimiento más de una vez en el conjunto de datos, necesitará especificar un nombre diferente cada vez.

Creación de un campo de respuesta categórica

El campo de respuesta debe ser categórico, con un valor que representa todas las respuestas positivas. Se asume que cualquier otro valor no perdido es una respuesta negativa. Si el campo de respuesta representa un valor continuo (escala), como el número de compras o las cantidades de compras, necesita crear un nuevo campo que asigna un valor de respuesta simple a todos los valores de respuesta distintos a cero.

- Si las respuestas negativas se registran como 0 (espacio cumplimentado, que se tratan como valores perdidos), se pueden calcular con la siguiente fórmula:

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

donde *NewName* es el nombre del nuevo campo y *OldName* es el nombre del campo original. Es una expresión lógica que asigna un valor de 1 a todos los valores no perdidos mayores que 0 y 0 a todos los valores no perdidos menores o iguales que 0.

- Si no se registra ningún valor para respuestas negativas, estos valores se consideran perdidos y la fórmula se complica un poco más:

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

En esta expresión lógica, a todos los valores de respuesta no perdidos se les asigna un valor de 1 y a todos los valores de respuesta no perdidos se les asigna un valor de 0.

- Si no puede distinguir entre valores de respuesta negativos (0) y valores perdidos, no se puede calcular un valor de respuesta preciso. Si hay relativamente pocos valores perdidos, es posible que no tenga un efecto significativo en los índices de respuesta calculados. Sin embargo, si hay demasiados valores perdidos, como cuando se registra información de respuesta para una pequeña prueba de muestra del conjunto de datos total, los índices de respuesta calculados no serán significativos, ya que son significativamente menores que los índices de respuesta verdaderos.

Creación de un campo de respuesta categórica

1. En los menús, elija:

Transformar > Calcular variable

2. Para una variable de destino, introduzca el nuevo nombre de campo (variable).
3. Si se registran respuestas negativas como 0, introduzca $\text{OldName}>0$ para Expresión numérica, donde *OldName* es el nombre del campo original.
4. Si se registran respuestas negativas como perdidas (en blanco), introduzca $\text{NOT}(\text{MISSING}(\text{OldName}))$ para Expresión numérica, donde *OldName* es el nombre del campo original.

Capítulo 7. Prueba de paquete de control

Esta técnica compara campañas de marketing para ver si hay diferencias significativas en la eficacia entre los distintos paquetes u ofertas. La eficacia de una campaña se mide por las respuestas. El campo Campaña identifica diferentes campañas, por ejemplo, Oferta A y Oferta B. El campo Respuesta indica si un contacto ha respondido a la campaña. Seleccione Cantidad de adquisición si la respuesta se registra como una cantidad de adquisición, por ejemplo "99,99". Seleccione Respuesta si la respuesta indica simplemente si el contacto ha respondido positivamente o no, por ejemplo "Sí" o "No".

Ejemplo. La división de marketing directo de una empresa quiere ver si un nuevo diseño de paquete generará respuestas más positivas que el paquete actual. Efectúan un envío de correos de prueba para determinar si el nuevo paquete genera un índice de respuesta significativamente más positivo. El envío de correos de prueba consiste en un grupo de control que recibe el paquete actual y un grupo de prueba que recibe el nuevo diseño del paquete. Los resultados de los dos grupos se comparan para ver si hay una diferencia significativa.

Resultados

Los resultados incluyen una tabla que muestra los recuentos y porcentajes de respuestas positivas y negativas de cada grupo definidos por el campo Campaña y una tabla que identifica los grupos que pueden ser muy diferentes entre sí.

Consideraciones y supuestos de datos de prueba de paquete de control

Campo Campaña. El campo Campaña debe ser categórico (nominal u ordinal).

Campo Respuesta de eficacia. Si selecciona Cantidad de adquisición para el campo Eficacia, el campo debe ser numérico y el nivel de medición debe ser continuo (escala).

Si no puede distinguir entre valores de respuesta negativos (para la cantidad de adquisición, un valor de 0) y valores perdidos, no se puede calcular un índice de respuesta preciso. Si hay relativamente pocos valores perdidos, es posible que no tenga un efecto significativo en los índices de respuesta calculados. Sin embargo, si hay demasiados valores perdidos, como cuando se registra información de respuesta para una pequeña prueba de muestra del conjunto de datos total, los índices de respuesta calculados no serán significativos, ya que son significativamente menores que los índices de respuesta verdaderos.

Supuestos. Este procedimiento supone que los contactos se han asignado aleatoriamente a cada grupo de campaña. En otras palabras, ningún tipo de información demográfica, historial de adquisiciones u otras características afectará a la asignación de grupos y todos los contactos tienen una probabilidad igual de ser asignados a cualquier grupo.

Para obtener una prueba de paquete de control

En los menús, elija:

Marketing directo > Seleccionar técnica

1. Seleccione **Comparar la eficacia de las campañas**.
2. Seleccione el campo que identifica el grupo de campaña al que pertenece cada contacto (por ejemplo, Oferta A y Oferta B, etc.). Este campo debe ser nominal u ordinal.
3. Seleccione el campo que indica la respuesta de eficacia.

Si el campo de respuesta es una cantidad de adquisición, el campo debe ser numérico.

Si el campo de respuesta indica si el contacto han respondido positivamente o no (por ejemplo, "Sí" o "No"), seleccione **Respuesta** e introduzca el valor que representa una respuesta positiva. Si algún valor ha definido etiquetas de valores, puede seleccionar la etiqueta de valor de la lista desplegable y se mostrará el valor correspondiente.

Se crea automáticamente un nuevo campo, en el que 1 representa las respuestas positivas y 0 representa las respuestas negativas y el análisis se ejecuta en el nuevo campo. Puede sustituir el nombre predeterminado y asignar uno de su elección. Los nombres deben cumplir las reglas de nombres de IBM SPSS Statistics .

4. Pulse en **Ejecutar** para ejecutar el procedimiento.

Avisos

Esta información se ha escrito para productos y servicios ofrecidos en los EE.UU.

Es posible que IBM no ofrezca los productos, servicios o las características que se describen en este documento en otros países. Consulte al representante local de IBM para obtener información sobre los productos y servicios disponibles actualmente en su zona. Cualquier referencia a un producto, programa o servicio de IBM no pretende afirmar ni implicar que sólo se pueda utilizar dicho producto, programa o servicio de IBM. En su lugar se puede utilizar cualquier producto, programa o servicio funcionalmente equivalente que no vulnere ningún derecho de propiedad intelectual de IBM. Sin embargo, es responsabilidad del usuario evaluar y comprobar el funcionamiento de todo producto, programa o servicio que no sea de IBM.

IBM puede tener patentes o solicitudes de patentes pendientes que cubran la materia descrita en este documento. Este documento no le otorga ninguna licencia para estas patentes. Puede enviar preguntas acerca de las licencias, por escrito, a:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
Estados Unidos

Para consultas sobre licencias relacionadas con información de doble byte (DBCS), póngase en contacto con el departamento de propiedad intelectual de IBM de su país o envíe sus consultas, por escrito, a:

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
1623-14, Shimotsuruma, Yamato-shi
Kanagawa 242-8502 Japan

El párrafo siguiente no se aplica al Reino Unido ni a ningún otro país donde estas disposiciones sean incompatibles con la legislación local: INTERNATIONAL BUSINESS MACHINES CORPORATION PROPORCIONA ESTA PUBLICACIÓN "TAL CUAL" SIN GARANTÍAS DE NINGÚN TIPO, NI EXPLÍCITAS NI IMPLÍCITAS, INCLUIDAS, PERO SIN LIMITARSE A ELLAS, LAS GARANTÍAS IMPLÍCITAS DE NO VULNERACIÓN, COMERCIALIZACIÓN O IDONEIDAD PARA UN FIN CONCRETO. Algunos estados no permiten la renuncia a expresar o a garantías implícitas en determinadas transacciones, por lo tanto, esta declaración no se aplique a usted.

Esta información puede incluir imprecisiones técnicas o errores tipográficos. Periódicamente, se efectúan cambios en la información aquí y estos cambios se incorporarán en nuevas ediciones de la publicación. IBM puede efectuar mejoras y/o cambios en los productos y/o programas descritos en esta publicación en cualquier momento y sin previo aviso.

Cualquier referencia a sitios Web que no sean de IBM en esta información sólo es ofrecida por comodidad y de ningún modo sirve como aprobación de esos sitios Web. El material de esos sitios web no forma parte del material de este producto de IBM y el uso de dichos sitios web es responsabilidad del usuario.

IBM puede utilizar o distribuir cualquier información que proporcione en la forma que considere adecuada sin incurrir en ninguna obligación con el usuario.

Los propietarios de licencia de este programa que deseen tener información sobre el mismo con el fin de poder: (i) intercambiar información entre programas creados de forma independiente y otros programas (incluido éste) y (ii) utilizar de forma mutua la información que se ha intercambiado, deberán ponerse en contacto con:

Tel. 901 100 400
ATTN: Licensing
200 W. Madison St.
Chicago, IL; 60606
Estados Unidos de América

Esta información estará disponible, bajo las condiciones adecuadas, incluyendo en algunos casos el pago de una cuota.

El programa bajo licencia que se describe en este documento y todo el material bajo licencia disponible los proporciona IBM bajo los términos de las Condiciones Generales de IBM, Acuerdo Internacional de Programas Bajo Licencia de IBM o cualquier acuerdo equivalente entre las partes.

Cualquier dato de rendimiento mencionado aquí ha sido determinado en un entorno controlado. Por lo tanto, los resultados obtenidos en otros entornos operativos pueden variar de forma significativa. Es posible que algunas mediciones se hayan realizado en sistemas en desarrollo y no existe ninguna garantía de que estas mediciones sean las mismas en los sistemas comerciales. Además, es posible que algunas mediciones hayan sido estimadas a través de extrapolación. Los resultados reales pueden variar. Los usuarios de este documento deben consultar los datos que corresponden a su entorno específico.

Se ha obtenido información acerca de productos que no son de IBM de los proveedores de esos productos, de sus publicaciones anunciadas o de otros orígenes disponibles públicamente. IBM no ha comprobado estos productos y no puede confirmar la precisión de su rendimiento, compatibilidad ni contemplar ninguna otra reclamación relacionada con los productos que no son de IBM. Las preguntas acerca de las aptitudes de productos que no sean de IBM deben dirigirse a los proveedores de dichos productos.

Todas las declaraciones sobre el futuro del rumbo y la intención de IBM están sujetas a cambio o retirada sin previo aviso y representan únicamente metas y objetivos.

Esta información contiene ejemplos de datos e informes utilizados en operaciones comerciales diarias. Para ilustrarlos lo máximo posible, los ejemplos incluyen los nombres de las personas, empresas, marcas y productos. Todos esos nombres son ficticios y cualquier parecido con los nombres y direcciones utilizados por una empresa real es pura coincidencia.

Cada copia, parcial o completa, de estos programas de ejemplo, o cualquier trabajo obtenido a partir de los mismos, debe incluir el siguiente aviso de copyright:

Esta información contiene ejemplos de datos e informes utilizados en operaciones comerciales diarias. Para ilustrarlos lo máximo posible, los ejemplos incluyen los nombres de las personas, empresas, marcas y productos. Todos esos nombres son ficticios y cualquier parecido con los nombres y direcciones utilizados por una empresa real es pura coincidencia.

Cada copia, parcial o completa, de estos programas de ejemplo, o cualquier trabajo obtenido a partir de los mismos, debe incluir el siguiente aviso de copyright:

© el nombre de su empresa) (año). Algunas partes de este código proceden de IBM Corp. Sample Programs.

© Copyright IBM Corp. _especifique el año o los años_. Todos los derechos reservados.

Marcas comerciales

IBM, el logotipo de IBM e ibm.com son marcas registradas o marcas comerciales registradas de International Business Machines Corp., registrada en muchas jurisdicciones en todo el mundo. Otros nombres de servicios y productos podrían ser marcas registradas de IBM u otras compañías. Hay disponible una lista actual de marcas registradas de IBM en la web en "Información de marca registrada y copyright en "www.ibm.com/legal/copytrade.shtml.

Adobe, el logotipo Adobe, PostScript y el logotipo PostScript son marcas registradas o marcas comerciales de Adobe Systems Incorporated en Estados Unidos y/o otros países.

Intel, el logotipo de Intel, Intel Inside, el logotipo de Intel Inside, Intel Centrino, el logotipo de Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium y Pentium son marcas comerciales o marcas registradas de Intel Corporation o sus filiales en Estados Unidos y otros países.

Linux es una marca registrada de Linus Torvalds en Estados Unidos, otros países o ambos.

Microsoft, Windows, Windows NT, y el logotipo de Windows son marcas comerciales de Microsoft Corporation en Estados Unidos, otros países o ambos.

UNIX es una marca registrada de The Open Group en Estados Unidos y otros países.

Java y todas las marcas registradas y logotipos basados en Java son marcas registradas de Oracle o sus filiales.

Índice

A

análisis de clúster (opción Marketing
directo) 11
análisis de clústeres 11

C

clúster 11

I

índices de respuesta de código postal 17

P

perfiles de posibles clientes (opción
Marketing directo) 13
propensión a la compra 21
prueba de paquete de control 25

R

regresión logística (Marketing
directo) 21
RFM 3, 6, 7, 8
 agrupación 5
 datos de cliente 4
 datos de transacciones 4



Impreso en España