

IBM® Analytical Decision Management
Versión 18 Release 0

IBM® SPSS® Modeler Advantage

IBM

Nota

Antes de utilizar esta información y el producto al que da soporte, lea la información contenida en "Avisos" en la página 97.

Información sobre el producto

Esta edición se aplica a la versión 18, release 0, modificación 0 de IBM Analytical Decision Management y a todos los releases y las modificaciones posteriores, hasta que se indique lo contrario en nuevas ediciones.

© Copyright IBM Corporation 2010, 2015.

Contenido

Prefacio	v	Selección de campos y valores	48
		Comprobación de expresiones	49
		Uso de expresiones que devuelven un booleano	49
Capítulo 1. Acerca de IBM Analytical Decision Management	1	Capítulo 5. Trabajar con reglas	51
Visión general del producto	1	Creación de reglas	51
Novedades en la versión 18	1	Definición de reglas de segmentación.	51
Versiones anteriores	2	Definición de reglas de selección	53
Capítulo 2. Ejemplo de modelado de respuestas de cliente	9	Exportación y reutilización de reglas	54
IBM SPSS Modeler Advantage	9	Añadir anotaciones	55
Maximización del beneficio en una campaña de correo directo	9	Utilización de reglas de IBM Operational Decision Management	55
Archivos de muestra.	9	Capítulo 6. Creación de modelos predictivos	57
Lanzamiento de la aplicación	11	Acerca de modelos predictivos	57
Configuración de un origen de datos	11	Generación de modelos	57
Construcción del modelo	13	Generación de un modelo predictivo	57
Evaluación del modelo	15	Resultados de modelado predictivo	58
Modelado de reglas predictivo	18	Creación de un modelo de regla predictiva	59
Prueba del modelo	19	Generación de un modelo de agrupación en clúster	61
Puntuación del modelo	19	Creación de un modelo de asociación.	62
Ejemplo de modelado de clúster	20	Valores de modelo opcionales	63
Ejemplo de modelado de asociaciones	21	Evaluación de modelos	67
		Simulación de beneficio máximo	69
Capítulo 3. Navegación por el espacio de trabajo	25	Probar un modelo	70
Cambiar el idioma visualizado	25	Utilización de puntuaciones de modelo en aplicaciones	70
Inicio de aplicaciones	25	Capítulo 7. Modelos de puntuación ..	73
La página de inicio	26	Puntuación de modelos en una tabla de base de datos, un archivo, Analytic Server o un servidor IBM Cognos BI	73
Galería	26	Selección de datos y subconjuntos	75
Establecimiento de preferencias.	27	Selección de registros para puntuar	75
Abrir y guardar proyectos, modelos y reglas	28	Selección del destino de puntuación	75
Abrir y guardar en el repositorio	29	Selección de campos de salida	78
Trabajar con etiquetas	30	Modelos de puntuación	79
Bloqueo de proyectos	30	Informe de resumen de puntuación	80
Propiedades	31	Capítulo 8. Administración de aplicaciones	81
Capítulo 4. Gestión de orígenes de datos	33	Opciones generales	82
Creación de un nuevo origen de datos	35	Control de atajos de aplicación	82
Origen de archivo	36	Ocultar la página de inicio	83
Origen de base de datos	39	Bloqueo de opciones de datos	83
Origen de la vista de datos	39	Bloqueo de opciones de destino de puntuación	84
Origen de IBM Cognos BI	40	Descarga de metadatos de proyecto para su uso en la creación de reglas externas	84
Origen de Analytic Server	42	Edición de la definición de esquema XML	85
Selección de campos de entrada	43	Capítulo 9. Obtener ayuda	91
Vista previa de datos	43	Documentación disponible	91
Visión general de datos	43		
Niveles de medición	44		
Edición de valores de campo de datos	45		
Campos de correlaciones	45		
Derivación de campos	46		
Creador de expresiones	47		
Selección de funciones.	47		

Apéndice A. Compartir rutas	93
Apéndice B. Accesibilidad	95
Avisos	97
Marcas registradas	98

Términos y condiciones de la documentación del producto	99
Índice.	101

Prefacio

Acerca de IBM Business Analytics

El software de IBM Business Analytics proporciona información completa, coherente y precisa en la que confían los responsables de tomas de decisiones para mejorar el rendimiento de la empresa. Una cartera integral de inteligencia empresarial, análisis predictivo, gestión de rendimiento y estrategia financieros y aplicaciones de análisis que ofrece una perspectiva clara, inmediata e interactiva del rendimiento actual y la capacidad de predecir resultados futuros. Combinado con soluciones sectoriales avanzadas, prácticas probadas y servicios profesionales, las organizaciones de cualquier tamaño puede impulsar la mayor productividad, automatizar decisiones con confianza y ofrecer los mejores resultados.

Como parte de esta cartera de productos, el software de IBM SPSS Predictive Analytics ayuda a las organizaciones a predecir eventos futuros y actuar de forma proactiva con respecto a las informaciones que generan unos mejores resultados empresariales. Clientes comerciales, gubernamentales y académicos de todo el mundo confían en la tecnología de IBM® SPSS como una ventaja competitiva para atraer, retener y aumentar los clientes y, a la vez, reducir el fraude y mitigar el riesgo. Mediante la incorporación del software de IBM SPSS en sus operaciones diarias, las organizaciones se convierten en empresas predictivas, que pueden dirigir y automatizar decisiones para cumplir objetivos de negocio y conseguir una ventaja competitiva mensurable. Si desea más información o contactar con un representante, visite <http://www.ibm.com/spss>.

Soporte técnico

El soporte técnico está disponible para los clientes del servicio de mantenimiento. Los clientes pueden contactar con el soporte técnico para recibir ayuda para utilizar productos IBM Corp. o para la instalación de uno de los entornos de hardware soportados. Para acceder al soporte técnico, consulte el sitio web de IBM Corp. en <http://www.ibm.com/support>. Tenga a mano su identificación, la de su organización y su acuerdo de soporte cuando solicite ayuda.

Capítulo 1. Acerca de IBM Analytical Decision Management

Visión general del producto

IBM Analytical Decision Management aporta los beneficios del análisis predictivo a los problemas empresariales reales, lo que le permite crear aplicaciones personalizadas adaptadas a sus clientes o sector. Aunque las aplicaciones se suelen configurar para resolver problemas muy específicos, todas se basan en un conjunto común de funciones:

- Automatizar decisiones utilizando reglas empresariales.
- Añadir información utilizando modelos predictivos.
- Utilizar la priorización, optimización o simulación para tomar la mejor decisión basándose en lo anterior.

Está disponible una serie de aplicaciones empaquetadas, adaptadas para resolver problemas empresariales específicos. Póngase en contacto con su representante de ventas si desea más información.

Novedades en la versión 18

IBM Analytical Decision Management 18 (18.0.0.0) incorpora las siguientes funcionalidades nuevas:

A continuación se proporciona una descripción de las principales funciones nuevas:

Entorno de ejecución de escenarios. Hay un nuevo icono de propiedades en el diálogo Gestionar escenarios que permite visualizar detalles de un escenario como, por ejemplo, las horas de inicio y finalización, y el tiempo de ejecución.

Registro de CPLEX. Si desea guardar el registro de CPLEX en un archivo del sistema de archivos de IBM SPSS Modeler Server a efectos de resolución de problemas, está disponible una nueva opción **Redirigir el motor de anotaciones cronológicas de CPLEX a archivo** en el diálogo Opciones avanzadas de las aplicaciones que tienen una pestaña Optimizar. Tenga en cuenta que, en el improbable caso de que una aplicación CPLEX esté desplegada en el servicio de puntuaciones, deberá desactivar el registro para mejorar el rendimiento.

Carga masiva y tamaño de confirmación de lote. En la pestaña Desplegar, al generar salida para una base de datos, puede ahora establecer el tamaño del lote y habilitar la carga masiva. Esto mejora el rendimiento al puntuar tablas de gran tamaño y grabar el resultado en una base de datos.

Visión general de datos en contexto. La Visión general de datos se muestra ahora en el contexto donde se está utilizando el campo dentro del proyecto. Por ejemplo, se tiene en cuenta cualquier regla creada anteriormente en el proyecto.

Nueva restricción de optimización. Se ha añadido una nueva restricción a la aplicación IBM Analytical Decision Management for Campaign Optimization para el *Número mínimo de ofertas que se deben aceptar para este tipo de oferta*.

Mejoras en la memoria. Se han efectuado mejoras en la memoria para la aplicación IBM Analytical Decision Management for Campaign Optimization. Para obtener más información, consulte **Mejores prácticas de rendimiento > Optimización CPLEX > Requisitos mínimos de memoria** en la *Guía de resolución de problemas y rendimiento*.

Versiones anteriores

Versión 17

IBM Analytical Decision Management 17 (17.0.0.0) incorpora las siguientes funcionalidades nuevas:

A continuación se muestra una descripción de las principales funcionalidades nuevas que se han añadido en la versión 17:

Creador de expresiones. El creador de expresiones se ha mejorado para permitir expresiones más complejas y potentes. Ahora el creador de expresiones en IBM Analytical Decision Management se parece más al creador de expresiones en el cliente IBM SPSS Modeler. Consulte http://www-01.ibm.com/support/knowledgecenter/SS6A3P_17.0.0/adm_user_guide_ddita/configurableapps/dm_expressions.dita.

Pestaña Derivado. En la mayoría de las aplicaciones de IBM Analytical Decision Management, se pueden añadir campos adicionales al modelo de datos de proyecto como, por ejemplo, expresiones, salida de modelo, campos creados a partir de reglas de segmentación o campos de tablas secundarias. La nueva subpestaña **Derivado** de la pestaña Datos muestra los campos que se van a incluir como parte de los datos. Consulte http://www-01.ibm.com/support/knowledgecenter/SS6A3P_17.0.0/adm_user_guide_ddita/configurableapps/dm_app_derived.dita.

Sección Tablas secundarias. Se ha añadido una nueva sección de tablas secundarias a la mayoría de las aplicaciones IBM Analytical Decision Management. Esto sustituye los diversos enlaces **Añadir/Editar tablas adicionales** de los releases anteriores. Consulte http://www-01.ibm.com/support/knowledgecenter/SS6A3P_17.0.0/adm_user_guide_ddita/configurableapps/dm_app_additional_fields.dita.

Autenticación multifactorial La funcionalidad que soporta la autenticación multifactorial (MFA) se ha añadido a IBM SPSS Collaboration and Deployment Services 7, IBM SPSS Modeler 17 e IBM Analytical Decision Management 17. Estas mejoras pueden mejorar la seguridad en entornos empresariales de gran tamaño. Por ejemplo, se pueden integrar productos SPSS con la infraestructura de inicio de sesión único de Kerberos existente en la empresa y con infraestructuras de tarjeta inteligente (cuando los empleados de una empresa utilicen identificadores de seguridad y lectores de tarjetas para la autenticación por máquina y por aplicación, por ejemplo). Consulte http://www-01.ibm.com/support/knowledgecenter/SS6A3P_17.0.0/adm_install_guide_ddita/configurableapps/dms_install_security.dita.

Correlación de origen de datos. En la medida de lo posible, ahora IBM Analytical Decision Management correlaciona automáticamente tipos de almacenamiento compatibles cuando un usuario añade un origen de datos adicional. Si el tipo no es compatible, se autorizará al usuario que convierta los tipos.

Tipos de almacenamiento de archivo. Ahora, los usuarios pueden definir el tipo de almacenamiento y el formato al leer archivos de datos de formato fijo y variable para garantizar que los valores se lean correctamente. Por ejemplo, si el tipo del origen de datos es **Archivo** y el formato del archivo es **Campos de longitud variable en texto plano**, al añadir un nuevo origen de datos en IBM Analytical Decision Management, el diálogo de origen de datos tendrá una nueva sección **Especificar formato de datos** que permite al usuario especificar el tipo de entrada y el formato de cada campo. Esto es similar a la funcionalidad disponible en el cliente IBM SPSS Modeler.

Resumen de proyecto. En las aplicaciones con una pestaña Desplegar se ha añadido una nueva sección **Resumen de proyecto**. Esta sección se utiliza para validar el proyecto y mostrar información importante de resumen de proyecto, incluyendo los errores que puedan afectar al despliegue. El resumen del proyecto muestra información como, por ejemplo, todas las referencias de atributo y todas las referencias de repositorio, y se puede exportar como un archivo *.csv. También se puede ir a otras pestañas de la aplicación para corregir errores y luego volver al resumen del proyecto en la pestaña Desplegar.

Referencias de modelo. Cuando se referencian modelos desde dentro de una aplicación IBM Analytical Decision Management, el comportamiento del enlace del modelo ha cambiado. Cuando se pulsa un

enlace de referencia de modelo, se abre el diálogo del visor del modelo, que permite ver detalles del modelo. O bien, ahora puede pulsar con el botón derecho del ratón en la referencia del modelo para seleccionar la opción **Referenciar** o **Generar**. Cuando se pulsa en una referencia de regla, se abre el visor de reglas (como en releases anteriores). O bien, se puede pulsar con el botón derecho del ratón en una referencia de regla y seleccionar **Referenciar**.

Sufijo de credencial de Analytic Server. Los administradores pueden utilizar la IBM SPSS Collaboration and Deployment Services Deployment Manager basada en navegador para configurar un sufijo que se vaya a añadir a las credenciales de Analytic Server. Esto derivará el nombre de la credencial que se va a utilizar en las conexiones de Analytic Server en IBM Analytical Decision Management.

Credenciales de IBM Cognos. Al importar datos de la base de datos de IBM Cognos BI o informes de lista única para su uso en IBM Analytical Decision Management, ahora se puede seleccionar una credencial de IBM Cognos almacenada en IBM SPSS Collaboration and Deployment Services Repository en lugar de especificar la información de nombre de usuario y contraseña o utilizar una conexión anónima.

Versión 8

IBM SPSS Analytical Decision Management 8 ha ampliado la capacidad de los usuarios para desplegar la analítica en los procesos empresariales cotidianos, que soporta una amplia gama de decisiones con mayor información y control que nunca.

- Ajustar de forma precisa y visualizar el impacto combinado de reglas, modelos y optimización en los resultados empresariales, como mejores prestaciones de simulación.
- Resolver una amplia gama de problemas con nuevas aplicaciones creadas previamente y técnicas de optimización nuevas como, por ejemplo, la optimización del precio y la optimización de abastecimiento.
- Mejor integración con IBM Operational Decision Management, y los grandes volúmenes de datos a través de la integración de IBM SPSS Analytic Server.

A continuación, aparece una descripción de las principales características nuevas que se han añadido en la versión 8:

Nueva tecnología de instalación. La instalación del producto se ha simplificado y mejorado. Ahora utiliza la tecnología Installation Manager. Si desea más información, consulte la *Guía de instalación de IBM SPSS Analytical Decision Management*.

Mejoras de la interfaz de usuario. Se han realizado muchos cambios en la interfaz de usuario para mejorar la capacidad de uso.

IBM Analytical Decision Management for Predictive Maintenance. IBM SPSS Analytical Decision Management for Predictive Maintenance es una nueva aplicación creada previamente que utiliza la optimización para anticipar y evitar anomalías del equipo antes de que se produzcan. Puede minimizar el tiempo de inactividad costoso y optimizar el uso de infraestructuras y recursos escasos prediciendo anomalías y planificando las acciones de corrección correctas en el momento oportuno. Podría anticipar qué partes es más probable que fallen primero y sustituirlas en el momento adecuado. Predecir cuándo fallarán las máquinas ayuda a maximizar los beneficios y evitar pérdidas de máquinas que fallan y que dejan de estar operativas.

IBM Analytical Decision Management for Demand Optimization. IBM SPSS Analytical Decision Management for Demand Optimization es una nueva aplicación construida previamente que utiliza la optimización para evitar los costosos exceso y escasez de existencias asignando la cantidad adecuada de cada producto a cada tienda. Puede maximizar los beneficios y evitar pérdidas debido a la escasez de existencias y las comisiones por reabastecimiento asignando la cantidad correcta de inventario en cada establecimiento.

Mejoras de simulación. Las nuevas funciones de simulación combinan las funciones del botón anterior **Simular** (anteriormente en la pestaña Definir) con las funciones hipotéticas anteriores **¿Y si?** (anteriormente en la pestaña Optimizar, Priorizar o Combinar) en una sola característica. La nueva y mejorada función de simulación tiene su propia barra de herramientas en la esquina superior derecha de la interfaz de usuario, y permite que se guarden escenarios con el proyecto y persisten entre sesiones.

Despliegue y puntuación. Se han realizado varias mejoras en la pestaña Desplegar/Puntuar.

URL de aplicación. IBM SPSS Collaboration and Deployment Services ha añadido la capacidad de personalizar el URL para aplicaciones basadas en web, incluido el URL para iniciar IBM SPSS Analytical Decision Management. Póngase en contacto con el administrador para obtener detalles. Está disponible más información en el capítulo *Raíces de contexto de aplicación* de la *Guía de instalación y configuración de IBM SPSS Collaboration and Deployment Services Repository*.

Mejoras del archivo de estadística. Los archivos de datos IBM SPSS Statistics comprimidos protegidos con contraseña (.sav y .zsav) ahora están soportados en IBM SPSS Analytical Decision Management.

Mejoras del visor de árbol Al visualizar resultados de modelos individuales en IBM SPSS Modeler Advantage, el nombre de cada modelo es un enlace que abre un nuevo diálogo que contiene información sobre el modelo como, por ejemplo, el campo objetivo y campos de entrada. Si el modelo seleccionado es un *modelo de árbol*, también se visualizará un visor de árbol. El visor de árbol se ha mejorado. Ahora puede realizar acciones como, por ejemplo, alejar y acercar el árbol, cambiar la orientación del árbol, navegar por árboles grandes, y expandir o contraer nodos individuales dentro del árbol.

Opciones avanzadas en la pestaña Optimizar. Para aplicaciones que utilizan la optimización matemática compleja (CPLEX) como, por ejemplo, IBM SPSS Analytical Decision Management for Campaign Optimization, IBM SPSS Analytical Decision Management for Predictive Maintenance e IBM SPSS Analytical Decision Management for Demand Optimization, ahora las opciones avanzadas están disponibles en relación con el algoritmo de optimización CPLEX. Ahora puede exportar un archivo de modelo de lenguaje de programación de optimización (OPL) y añadir valores personalizados.

Origen de vista de datos. Se ha añadido un nuevo nodo de *vista de datos* a IBM SPSS Modeler 16 para incluir datos en rutas que se han definido en una **vista de datos analíticos** de IBM SPSS Collaboration and Deployment Services. El nuevo origen de datos de vista de datos también se puede utilizar en la pestaña Datos en IBM SPSS Analytical Decision Management. Puede crear una conexión para utilizar con reglas de IBM Operational Decision Management. Para reglas complejas que utilizan un modelo de datos jerárquico, tenga en cuenta que los administradores pueden restringir el uso de dichas reglas porque pueden degradar el rendimiento.

Antes de la versión 8, las reglas externas necesitaban parámetros de entrada sin formato que coincidían con el modelo de datos de proyecto de IBM SPSS Analytical Decision Management. Pero con este nuevo soporte de vista de datos analíticos, ahora las reglas externas pueden tener parámetros de entrada jerárquica. Consulte los temas “Origen de la vista de datos” en la página 39 y “Bloqueo de opciones de datos” en la página 83 si desea más información, y el *Capítulo 5* de la *Guía del diseñador de aplicaciones*.

Integración de IBM SPSS Analytic Server. En IBM SPSS Modeler 16, se ha introducido IBM SPSS Analytic Server, que permite el uso del sistema de archivos distribuido de Hadoop (HDFS) como origen de datos. Los usuarios de IBM SPSS Modeler pueden crear rutas utilizando el nodo de origen de Analytic Server del HDFS y generar modelos basándose en el sistema de datos grandes de Hadoop. Los usuarios de IBM SPSS Analytical Decision Management pueden abrir estas rutas de modelo en IBM SPSS Modeler Advantage para la visualización (no pueden crear proyectos nuevos utilizando orígenes de datos de Analytic Server).

Nuevo capítulo en la guía del diseñador de aplicaciones. Se ha añadido un nuevo capítulo a la *Guía del diseñador de aplicaciones de IBM SPSS Analytical Decision Management*. Proporciona una configuración de ejemplo de una aplicación de optimización personalizada. La aplicación estándar de IBM SPSS Analytical

Decision Management for Campaign Optimization utiliza un archivo de modelo de lenguaje de programación de optimización (OPL) llamado *CampaignOptimization.mod*. En algunos casos, la aplicación Campaign Optimization estándar y su modelo OPL podrían no adaptarse a las finalidades del problema empresarial en cuestión. El nuevo capítulo le guía a través de un ejemplo de un modelo OPL personalizado creado utilizando IBM ILOG CPLEX Optimization Studio, y una aplicación de optimización personalizada de IBM SPSS Analytical Decision Management, para una empresa ficticia.

Resultados de la evaluación en forma de tabla. Está disponible una nueva opción **Incluir resultados en forma de tabla** al ejecutar evaluaciones. Si se deselecciona la opción, se mejora el rendimiento porque se tarda más tiempo en ejecutar evaluaciones que en incluir resultados en forma de tabla.

Nuevo valor para número máximo de campos nominales. Se ha añadido un valor nuevo al diálogo Propiedades para definir el número máximo de miembros permitidos para campos de datos *nominales*. Este se utiliza en la pestaña Datos. Si desea más información, consulte "Propiedades" en la página 31.

Inicio de sesión único. Las conexiones ODBC ahora pueden utilizar el inicio de sesión único (SSO) para IBM SPSS Modeler Server y bases de datos ODBC con IBM SPSS Analytical Decision Management. Son necesarios algunos pasos de configuración. Consulte el tema Consideraciones de seguridad en la documentación de instalación.

Retrotracción SQL. Cuando los datos de origen se originan en una base de datos, la ejecución se retrotrae a la base de datos siempre que sea posible (en un proceso llamado "retrotracción SQL"). En general, IBM SPSS Analytical Decision Management soporta el mismo nivel de retrotracción SQL que IBM SPSS Modeler, pero con las limitaciones siguientes:

- Si el proyecto incluye campos adicionales o reglas de selección que tienen referencias a modelos o reglas externos, estas referencias no se retrotraerán a la base de datos, y esto puede inhibir el grado de retrotracción que se puede conseguir para la ruta de proyecto como un todo.
- En aplicaciones distintas a IBM SPSS Modeler Advantage, la retrotracción solo está habilitada para las reglas de selección (selecciones globales y selecciones de miembro en la primera dimensión). El grado de retrotracción se puede aumentar si elige un campo Entero como el ID de entidad en el modelo de datos de proyecto.

Versión 7.0.0.1

IBM SPSS Decision Management 7 Fixpack 1 (7.0.0.1) ha añadido las características siguientes:

Visión general de datos. Está disponible una nueva característica de visión general de datos en la pestaña Datos de todas las aplicaciones, en la pestaña Modelado y la pestaña Puntuación en IBM SPSS Modeler Advantage, y en la pestaña Desplegar de aplicaciones que incluyen la característica *Puntuar ahora*. Consulte el tema "Visión general de datos" en la página 43 si desea más información. Esta característica solo está disponible con IBM SPSS Modeler Server 15 Fixpack 1 (15.0.0.1). No está disponible con la versión 15.

Informe de resumen de puntuación. Tras puntuar un proyecto, una nueva sección *Informe de resumen de puntuación* proporciona detalles como, por ejemplo, dónde se han puntuado los datos, tamaño de archivo (si procede), nombre de base de datos y tabla (si se aplica), nombre de modelo o proyecto, número de registros puntuados, y, en algunos casos, un gráfico y/o una tabla que resume el modelo. La sección Informe de resumen de puntuación está disponible en IBM SPSS Modeler Advantage y en la pestaña Desplegar de aplicaciones que incluyen la característica Puntuar ahora. Tenga en cuenta que el tamaño de archivo solo se visualiza si se ejecuta en IBM SPSS Modeler Server 15 Fixpack 1 (15.0.0.1). No está disponible con la versión 15.

Descargar copia local de archivo puntuado. Si elige enviar el resultado de la puntuación a un archivo sin formato, también puede descargar una copia local del archivo, si lo desea. Los administradores pueden utilizar un IBM SPSS Collaboration and Deployment Services Deployment Manager basado en el navegador para habilitar o inhabilitar la característica, y para aumentar o reducir el límite máximo permitido del tamaño de archivo. Si el botón **Descargar archivo** está inactivo, el tamaño del archivo

excede el límite máximo de tamaño definido por el administrador, o el administrador ha inhabilitado la característica. La característica está inhabilitada de forma predeterminada. Esta característica solo está disponible con IBM SPSS Modeler Server 15 Fixpack 1 (15.0.0.1). No está disponible con la versión 15.

Nuevos valores opcionales para crear modelos. Al crear modelos en IBM SPSS Modeler Advantage, están disponibles nuevos valores opcionales como, por ejemplo, opciones de partición de datos, elegir qué técnicas de modelado (algoritmos) incluir al crear el modelo, especificar el número máximo de modelos para conservar y combinar y definir criterios de beneficios antes de crear el modelo. Las opciones disponibles dependen del tipo de modelo y el objetivo seleccionados. Consulte el tema “Valores de modelo opcionales” en la página 63 si desea más información. Tenga en cuenta que las siguientes características nuevas solo están disponibles con IBM SPSS Modeler Server 15 Fixpack 1 (15.0.0.1). No están disponibles con la versión 15.

- Seleccionar qué técnicas de modelado incluir al crear el modelo
- Definir criterios de beneficios
- Visualizar detalles de nivel superior sobre modelos individuales
- Visualizar detalles adicionales sobre un modelo seleccionado

Visor de árbol. Tras crear un modelo en IBM SPSS Modeler Advantage y visualizar los resultados de modelos individuales, el nombre de cada modelo es un enlace que abre un diálogo nuevo que contiene información sobre el modelo como, por ejemplo, el campo objetivo y los campos de entrada. Si el modelo seleccionado es un *modelo de árbol*, ahora también se visualizará un visor de árbol. Tenga en cuenta que el visor de árbol solo está disponible con IBM SPSS Modeler Server 15 Fixpack 1 (15.0.0.1). No está disponible con la versión 15.

Mejoras de evaluación. Se han realizado varias mejoras en el cuadro de diálogo Evaluar en IBM SPSS Modeler Advantage. Por ejemplo, está disponible un nuevo botón **Modificar selección de modelo**, lo que le permite modificar los modelos para incluir en la evaluación. Tras realizar las selecciones, también puede copiar las selecciones de modelo actuales de nuevo en la tabla de resultados de modelos individuales, si lo desea. Algunas de las opciones nuevas solo están disponibles para determinados tipos de modelo (la mayoría de las opciones nuevas solo son para el tipo de modelo *predictivo* y algunos también son para el tipo de modelo de *regla predictiva*). Tenga en cuenta que las siguientes características nuevas del cuadro de diálogo Evaluar solo están habilitadas con IBM SPSS Modeler Server 15 Fixpack 1 (15.0.0.1). No están disponibles con la versión 15.

- Gráficos de rendimiento para objetivos numéricos
- Simulación de beneficios para un objetivo numérico
- Gráficos de matrices de coincidencias (para todos los objetivos)
- Vista de tabla para todos los gráficos
- El botón **Modificar selección de modelo**
- La opción para **Incluir evaluaciones de modelos individuales para gráficos de rendimiento**

Microsoft Windows 8 e Internet Explorer 10. IBM SPSS Collaboration and Deployment Services 5 Fixpack 1 (5.0.0.1) añade soporte para Windows 8 e Internet Explorer 10. No están soportados con IBM SPSS Collaboration and Deployment Services 4.2.1 o 5.

Mejoras de la interfaz de usuario. Se han realizado varios cambios menores en la interfaz de usuario para mejorar la capacidad de uso.

Versión 7

IBM SPSS Decision Management 7 (7.0.0.0) ha añadido las características siguientes:

Optimización. La finalidad de la optimización es identificar la solución que mejor cumple un objetivo específico, como, por ejemplo, maximizar los ingresos de una campaña de marketing o minimizar el

riesgo de fraude o abandono. El resultado de la optimización es una solución que responde a esta pregunta en todas las combinaciones posibles dentro de la población, de la forma que maximiza mejor el valor que desea optimizar (como el beneficio).

Modelos de agrupación en clúster. Nuevo en IBM SPSS Modeler Advantage, los métodos de modelos de agrupación en clúster ayudan a descubrir grupos de registros de datos con valores o patrones similares. Estas técnicas se utilizan en marketing (segmentación de clientes) y otras aplicaciones empresariales (los registros que están dentro de clústeres de un solo registro pueden contener errores o ser ejemplos de fraude). A menudo, se hace referencia a los modelos de agrupación en clúster como modelos de aprendizaje no supervisado, ya que no utilizan un campo objetivo y no devuelven una predicción específica que se pueda evaluar como verdadera o falsa. IBM SPSS Decision Management le permite crear modelos de clúster automáticamente. También puede combinar, renombrar, modificar o suprimir clústeres, o añadir clústeres personalizados basándose en conocimientos empresariales previos.

Modelos de asociación. Nuevo en IBM SPSS Modeler Advantage, los modelos de asociación le permiten predecir qué elementos tienen más probabilidad de aparecer juntos, y predecir la solidez de la relación entre ellos. Las reglas de asociación asocian una conclusión particular (la compra de un producto concreto, por ejemplo) con un conjunto de condiciones (la compra de otros productos, por ejemplo).

Datos. La pestaña Datos ahora incluye la capacidad de tener varias entradas de datos, definir el modelo de datos de proyecto con una vista de aplicación, hacer referencia a las salidas de modelo como si fueran datos, subir archivos de datos locales y tener datos de una vista previa y de recuento.

Mejoras en la integración de datos de IBM Cognos. Al igual que con IBM SPSS Modeler 15, IBM SPSS Decision Management ahora soporta la importación de informes Cognos construidos previamente, filtros y cálculos Cognos y el control de la disgregación automática.

Varios elementos. Las aplicaciones ahora pueden incluir más de una jerarquía de dimensión. Al lanzar una aplicación, si la aplicación seleccionada se ha configurado para incluir más de un elemento, se le pedirá que seleccione uno o más de los elementos disponibles (jerarquías) para utilizar en la aplicación. El árbol de dimensiones en la pestaña Definir especifica el rango de decisiones o recomendaciones posibles que se pueden devolver. Si una aplicación incluye más de una jerarquía de dimensión, se incluirán en pestañas separadas del área del árbol de dimensiones en la pestaña Definir.

Asignación dinámica. En releases anteriores, para asignar algo, tenía que existir en el árbol de dimensiones. Ahora, para los elemento de dimensión sin hijos, puede leer elementos de una tabla enlazada mediante una clave en el modelo de datos del proyecto. Por ejemplo, podría tener un elemento de dimensión llamado *Coches* sin hijos. Podría configurar asignaciones dinámicas en *Coches* para hacer referencia a campos en una tabla como, por ejemplo, *BMW*, *Audi* y *Ford*.

Varias asignaciones. Una nueva opción **Asignación múltiple** permite la asignación de registros a varios elementos de dimensión (como varias ofertas) para cada segmento. Las asignaciones múltiples se pueden realizar para cada regla, en lugar de una sola asignación por regla.

Mejoras de la interfaz de usuario. Se han realizado muchos cambios e incorporaciones a la interfaz de usuario de IBM SPSS Decision Management para mejorar el uso y la escalabilidad de todas las aplicaciones.

Mejoras de la galería. Se han realizado varias mejoras en la galería como, por ejemplo, la opción **Borrar galería** para eliminar de un elemento de galería a la vez.

Mejoras del editor de reglas. En función de desde dónde se ha lanzado el editor de reglas, ahora los usuarios pueden definir el tipo de regla para las reglas de selección (incluir o excluir), asignar puntos de riesgo para reglas de agregación, seleccionar una dimensión para reglas de asignación y utilice los controles *OR* y *Split OR* para combinar varias expresiones dentro de una regla y crear una expresión *OR*. Al crear reglas, ahora también puede seleccionar una dimensión desde un árbol jerárquico (para

aplicaciones con más de un árbol) o lanzar el editor de expresiones para crear una expresión. Estas funciones no estaban disponibles previamente en el cuadro de diálogo del editor de reglas.

Integración ampliada de IBM Operational Decision Management. En IBM SPSS Collaboration and Deployment Services Deployment Manager, puede utilizar el cuadro de diálogo Añadir nueva regla externa para definir referencias a conjuntos de reglas de IBM Operational Decision Management que se pueden utilizar después en IBM SPSS Decision Management. Anteriormente, solo se podía acceder a estas reglas a través de un servicio web. Con este release, ahora puede hacer referencia a reglas locales en la base de datos de IBM Operational Decision Management. Para la nueva opción local, la regla se ejecuta a través del IBM SPSS Modeler Server (el mismo servidor definido en Deployment Manager para ser utilizado por IBM SPSS Decision Management).

Mejoras de WhatIf?. Se han realizado varias mejoras en las funciones del cuadro de diálogo de WhatIf?. Por ejemplo, para la fecha de simulación ahora puede entrar un rango de fechas en las que puede realizar la evaluación de WhatIf?.

Subir archivos de ruta. Ahora puede abrir cualquier archivo de ruta de IBM SPSS Modeler desde el sistema de archivos local y almacenarlo en el repositorio, si lo desea.

Varias credenciales de IBM SPSS Modeler Server. Si tiene varias credenciales de IBM SPSS Modeler Server definidas en IBM SPSS Collaboration and Deployment Services Deployment Manager, ahora también puede especificar más de una para su uso con IBM SPSS Decision Management. A continuación, cuando el usuario de IBM SPSS Decision Management inicia una sesión, el servidor probará la primera credencial de la lista para ver si el usuario tiene la autoridad, se continuará siguiendo el orden de la lista hasta que se encuentre una credencial con la autoridad adecuada. Si el usuario no tiene autorización para ninguna de las credenciales de la lista ordenada, no podrá utilizar IBM SPSS Decision Management. Esto proporciona un mejor control sobre qué usuarios tienen acceso a qué pares del sistema de archivos de IBM SPSS Modeler Server.

Versión 6.2

IBM SPSS Decision Management 6.2 ha añadido las características siguientes:

Soporte para datos Cognos. Cuando se utiliza con IBM SPSS Modeler Server 14.2 o más reciente, IBM SPSS Decision Management 6.2 ha añadido la capacidad de leer datos de un servidor IBM Cognos Business Intelligence al definir orígenes de datos y de volver a escribir resultados de puntuación en un Cognos BI Server. Tras definir una conexión con un servidor Cognos BI, podrá examinar paquetes, dimensiones, consultas e informes para elegir los campos deseados. Consulte el tema "Origen de IBM Cognos BI" en la página 40 si desea más información.

Integración del sistema de gestión de reglas empresariales ILOG. Se puede hacer referencia a las reglas creadas en un sistema de gestión de reglas empresariales, como ILOG, o se pueden utilizar en aplicaciones IBM SPSS Decision Management, siempre que se hayan desarrollado para soportar el mismo modelo de datos utilizado en el proyecto actual de IBM SPSS Decision Management. Estas reglas externas se pueden seleccionar y utilizar de la misma forma que otros modelos y reglas dentro de la aplicación IBM SPSS Decision Management. Consulte el tema "Utilización de reglas de IBM Operational Decision Management" en la página 55 si desea más información. Tenga en cuenta que ahora se hace referencia a ILOG como IBM Operational Decision Management.

Soporte para Linux en System Z. SUSE Enterprise Server 10 está soportado en el IBM System z10 de 64 bits.

Capítulo 2. Ejemplo de modelado de respuestas de cliente

IBM SPSS Modeler Advantage

IBM SPSS Modeler Advantage es una aplicación fácil de utilizar que pone la potencia del modelado predictivo a disposición del usuario empresarial. Mediante los modelos predictivos, pueden identificarse patrones a partir de lo ocurrido en el pasado y usarlos para predecir lo que es probable que ocurra en el futuro.

Por ejemplo, se puede usar un modelo para predecir los clientes que con más probabilidad se perderán o responderán a una determinada oferta en función de características como, por ejemplo, ingresos, edad y membresías y organizaciones a las que estén suscritos. Las predicciones resultantes pueden utilizarse para generar listas de clientes objetivo o casos de interés que sirvan como datos de entrada de una planificación estratégica, o pueden integrarse con reglas en el contexto de una aplicación predictiva.

Maximización del beneficio en una campaña de correo directo

En este ejemplo de modelo predictivo se comercializan productos y servicios financieros a clientes de un banco ficticio. En una campaña de prueba reciente, se ofreció a una serie de clientes uno de varios servicios nuevos y se registraron sus respuestas. Sin embargo, las tasas de respuesta fueron inferiores a las esperadas y la campaña no dio beneficios.

Con IBM SPSS Modeler Advantage, el objetivo es convertir esa pérdida en una ganancia dirigiéndose a los mejores clientes disponibles teniendo en cuenta los ingresos y costes esperados, y otros factores.

El resultado de la campaña de prueba se resume a continuación.

Tabla 1. Resultado de la campaña de prueba

Resultado de la campaña de prueba	
Recuento de correos	3000
Recuento de respuestas	111
Tasa de respuesta	3,7 %
Coste	6750
Ingresos	6660
Beneficio total (pérdida)	(90)

Este ejemplo usa los archivos de datos *bank_customer_data.txt* y *bank_response_data.txt*, que se distribuyen e instalan con IBM Analytical Decision Management. En el archivo *bank_response_model.str* también se proporciona una versión completa de este ejemplo. Póngase en contacto con el administrador para obtener los detalles de la instalación y los archivos de ejemplo si es necesario. Consulte el tema “Archivos de muestra” para obtener más información.

Archivos de muestra

Los datos y otros archivos de muestra se incluyen con IBM Analytical Decision Management que se puede utilizar para replicar los ejemplos mostrados en las guías de usuario.

Datos

Durante una instalación, se instala un número de archivos de datos de muestra en el directorio `\Demos` de IBM SPSS Modeler Server (por ejemplo, `C:\Program Files\IBM\SPSS\ModelerServer\18\Demos`). Estos datos se pueden utilizar con las aplicaciones de muestra creadas previamente e incluyen:

- *bank_customer_data.txt*
- *bank_response_data.txt*
- *customer_contact_history.txt*
- *customer_summary_data.txt*
- *insurance_claim_data.txt*
- *insurance_fraud_data.txt*
- *inventory_orders.txt*
- *inventory_stock.txt*
- *machine_log_data.txt*
- *machine_summary_data.txt*
- *retail_purchase_data.txt*

Los archivos de aplicación de muestra, modelos y reglas que utilizan estos archivos de datos presuponen la ubicación predeterminada indicada arriba. Si los archivos de datos se instalan en una unidad o ubicación diferente, las vías de acceso de origen de datos se deben actualizar en las aplicaciones de muestra, en consecuencia.

Muestras

Se proporciona un número de archivos de aplicación de muestra, modelos y reglas para su uso con las aplicaciones IBM Analytical Decision Management creadas previamente. Las muestras se proporcionan en un archivo disponible para su descarga electrónica. Complete los pasos siguientes para importar archivos de muestra para una o más aplicaciones construidas previamente en IBM SPSS Collaboration and Deployment Services Repository.

1. Descargue el archivo *ADM_18_Demo.zip* en una ubicación temporal y extraiga el archivo.
2. Abra IBM SPSS Collaboration and Deployment Services Deployment Manager e inicie una sesión en la conexión de Content Server.
3. Pulse con el botón derecho del ratón en la carpeta de Content Repository y seleccione **Importar**.
4. En el cuadro de diálogo Importar, pulse **Examinar** y seleccione el archivo *DecisionManagementExamples.pes* que ha copiado o extraído en una ubicación temporal en el paso 1. Pulse **Aceptar**.
5. En el diálogo Conflictos de importación, pulse **Aceptar**. Cuando se completa la importación, verá las carpetas nuevas siguientes que contienen proyectos de ejemplo que se pueden abrir en IBM Analytical Decision Management.

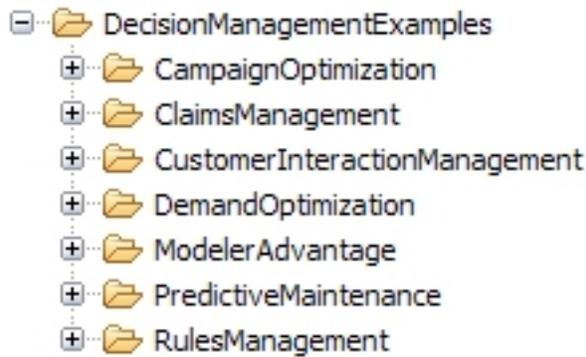


Figura 1. Carpetas y artefactos de muestra

Tenga en cuenta que todos los archivos de aplicación de muestra, modelos y reglas presuponen que los archivos de datos están instalados en el directorio *Demos* de IBM SPSS Modeler Server (por ejemplo, *C:\Program Files\IBM\SPSS\ModelerServer\18\Demos*). Si los archivos de datos están instalados en una ubicación diferente, las vías de acceso de origen de datos se deben actualizar en las aplicaciones de muestra para la configuración local.

Lanzamiento de la aplicación

La página de lanzamiento de IBM Analytical Decision Management incluye enlaces a las aplicaciones disponibles en el sitio. Un administrador puede configurar dichos enlaces; puede seleccionar aplicaciones adicionales pulsando en el enlace **Añadir aplicación** en la esquina superior derecha.

1. Para empezar a trabajar con el ejemplo, abra la página de lanzamiento de IBM Analytical Decision Management. Para facilitar el acceso en el futuro, puede añadir un marcador a esta página en el navegador o crear un atajo en el escritorio.
2. Para crear un modelo, seleccione **Nuevo** en la lista desplegable de la aplicación en el grupo de IBM SPSS Modeler Advantage y pulse **Ir**. Se mostrará la página de inicio de IBM SPSS Modeler Advantage.



Figura 2. IBM SPSS Modeler Advantage Página de inicio

Cada icono de la página de inicio representa un paso distinto en el proceso de modelado, desde la definición de los datos hasta la construcción y evaluación del modelo. Para empezar, pulse en el icono Datos de la izquierda.

Configuración de un origen de datos

La pestaña Datos le permite configurar orígenes de datos de para los archivos de datos de muestra que se utilizan en el ejemplo. El administrador puede configurar conjuntos de datos o puede añadir los suyos propios.

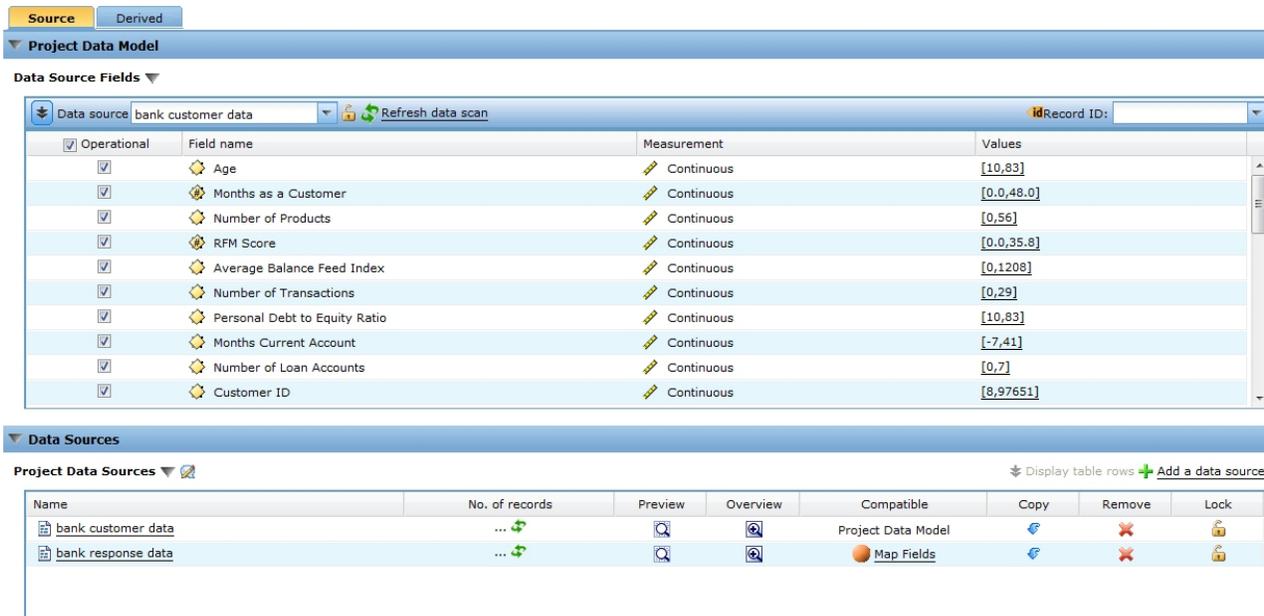


Figura 3. Data tab

El modelo de datos del proyecto define los campos que están disponibles en toda la aplicación, y pueden ser necesarios para las reglas y los modelos que se van a evaluar. Todos los demás orígenes de datos se deben correlacionar con el modelo de datos para ser utilizados con la aplicación.

Este ejemplo utiliza dos archivos de datos, ambos se distribuyen con la aplicación:

- **bank_customer_data.txt** incluye información sobre clientes actuales, incluyendo el ID de cliente, el historial de pagos, el número de productos que poseen, el número de meses que han sido cliente, así como información demográfica como la edad y los ingresos. El modelo de datos del proyecto se basa en este archivo.
- **bank_response_data.txt** contiene información sobre clientes a quienes se les ha ofrecido uno o varios servicios en campañas anteriores y los registros de sus respuestas. Analizando estos datos, la aplicación puede identificar qué clientes tienen más probabilidad de responder a ofertas similares en futuras campañas.

De forma predeterminada, los archivos de datos de ejemplo se instalan en la carpeta *Demos* de IBM SPSS Modeler Server a la que se puede acceder desde la aplicación (por ejemplo, *C:\Program Files\IBM\SPSS\ModelerServer\18\Demos*). Póngase en contacto con el administrador para obtener ayuda para instalar estos archivos, si es necesario.

Para definir el modelo de datos del proyecto :

1. En el panel Orígenes de datos de proyecto, seleccione **Añadir un origen de datos** desde el desplegable **Origen de datos**.
2. Especifique *datos de cliente bancarios* como nombre de origen de datos.
3. Seleccione **Archivo** y seleccione *bank_customer_data.txt*.
4. En **Especificar campos de entrada**, asegúrese de que el tipo de medición para el campo *Tiene un registro de pagos negativo* está establecido en **Marca**. El valor de marca indica que el campo tiene dos valores posibles, que indican *verdadero* y *falso* (normalmente codificados como *T* y *F*, o 0 y 1 en un campo numérico), para indicar la presencia o ausencia de dicho atributo.
5. En **Especificar campos de entrada**, asegúrese también de que el tipo de medición para el campo *Género* está establecido en **Nominal**. (Los valores de este campo simplemente indican el género como hombre o mujer, y no *verdadero/falso*, como sería en el caso de una marca.)

6. Pulse **Enlaces**. Se visualizará el diálogo Enlaces.
7. Pulse **Guardar**.

Puesto que este es el primer origen de datos definido para el proyecto, se selecciona automáticamente para el modelo de datos de proyecto, y se utiliza para definir los campos necesarios para la aplicación. Todos los demás conjuntos de datos se correlacionan en relación con este origen.

Para definir el origen de datos de respuestas:

El conjunto de datos de respuestas contiene muchos de los mismos campos utilizados en el modelo de datos del proyecto, así como dos campos adicionales llamados *ID de campaña* y *Respuesta*. Estos campos indican qué oferta se ha realizado a cada cliente, y si han contestado o no, y se pueden utilizar para generar un modelo que predice qué clientes tienen más probabilidades de responder en el futuro.

1. En el panel Orígenes de datos de proyecto, seleccione **Añadir un origen de datos** desde el desplegable **Origen de datos**.
2. Especifique *datos de respuestas bancarios* como nombre de origen de datos.
3. Seleccione **Archivo** y seleccione *bank_response_data.txt*.
4. En **Especificar campos de entrada**, asegúrese de que el tipo de medición para el campo *Respuesta* está establecido en **Marca**.
5. En **Especificar campos de entrada**, deseleccione la casilla de verificación junto al campo *ID de campaña* para indicar que este campo no se debe utilizar como entrada al generar el modelo. (Este campo se puede utilizar para indicar para qué campañas se ha seleccionado un cliente en el pasado, pero no es útil para predecir respuesta futuras.)
6. Pulse **Guardar**.

Puesto que algunos de los campos del modelo de datos no se han encontrado en el conjunto de datos de respuestas, se visualiza un icono naranja para este origen de datos en la columna **Compatible** en la pestaña Datos, que indica que algunos campos coinciden, pero no todos. El conjunto de datos es adecuado para la generación de modelos, pero no para la simulación, puntuación o despliegue. (Para poder utilizarse con estas finalidades, deben estar presentes todos los campos del modelo de datos.)

Construcción del modelo

La pestaña Modelado permite especificar el origen de datos, el campo objetivo y otras opciones de construcción y evaluación del modelo.

Model type
 Predictive model Change Model...

Data source
 bank response data + +

Records used to build: ... ↻

Target 🎯
 Response

Evaluate
Test

Optional Settings ▶

▶ Build Model

▼ Predictive Model Results

Display: Combined model results Individual model results

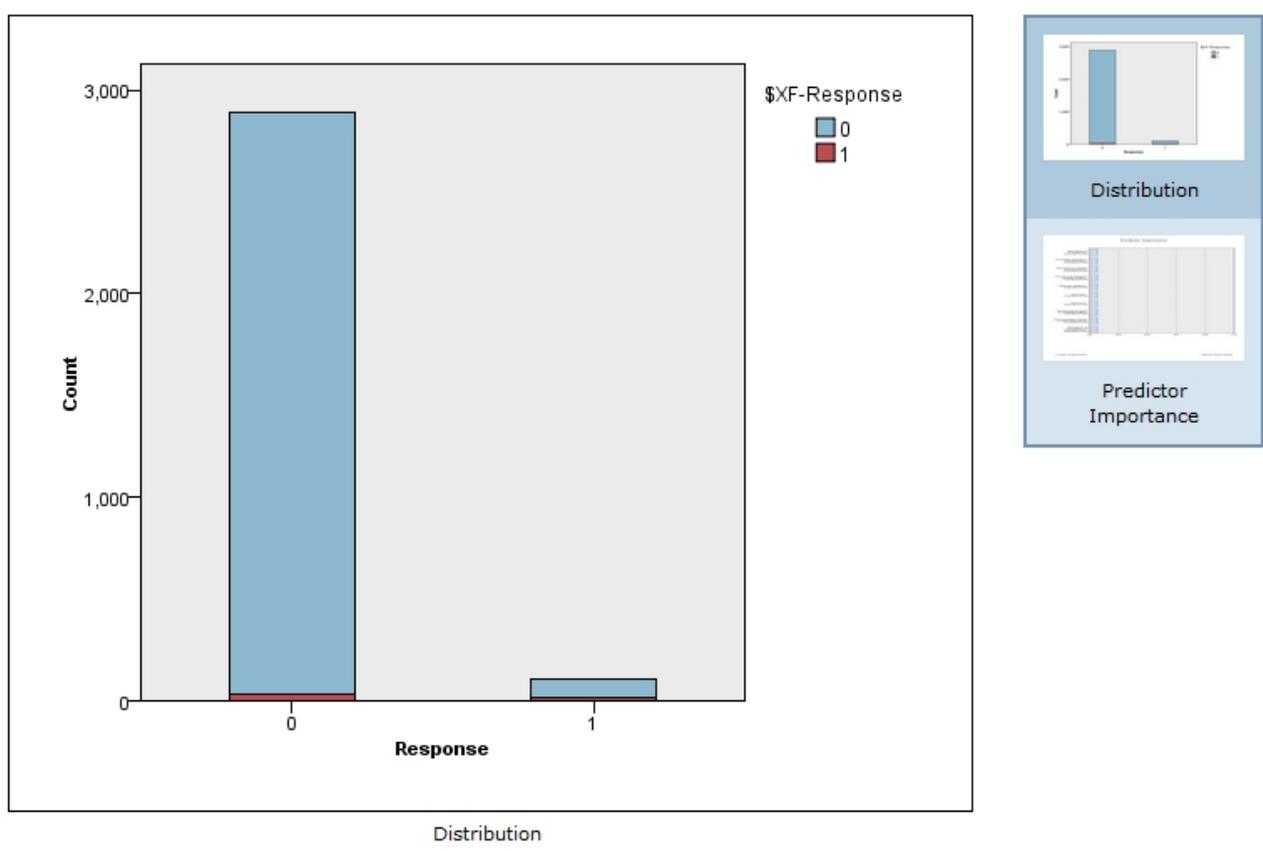


Figura 4. Model tab

1. Seleccione *datos de respuestas del banco* para el origen de datos y seleccione *Respuesta* como el campo objetivo.
2. Si aparece un mensaje **Examen de datos necesario**, pulse el icono para realizar un examen de datos. Después, pulse **Generar modelo**.

Cuando el modelo se completa, se muestra un gráfico de distribución. El gráfico muestra los valores observados en el eje horizontal, con un valor de 1 que indica una respuesta positiva.

- La altura de las barras refleja el índice de respuesta general. De los 3000 clientes del conjunto de datos de respuestas, 111 han contestado, para una tasa de aciertos general de un 3,8 %.
- Las áreas coloreadas dentro de cada barra representan las respuestas predichas por el modelo, con el rojo que indica una predicción positiva, el azul que indica una respuesta negativa, tal como se indica

en la leyenda del extremo superior derecho. Debido al índice de respuesta general bajo, hay relativamente pocas respuestas predichas (áreas rojas). Sin embargo, la barra del valor falso (*Respuesta = 0*) es abrumadoramente azul, lo que indica que la mayoría de estos clientes pronosticados como que no iban a responder, realmente no lo han hecho.

Para hacerse una mejor idea de cómo se va a comportar el modelo, vaya a la sección siguiente (“Evaluación del modelo”).

Evaluación del modelo

En la pestaña Modelado, pulse **Evaluar** para ver cómo se comporta el modelo cuando se aplica a los datos de clientes y para comprender cómo la utilización del modelo puede afectar al balance final.

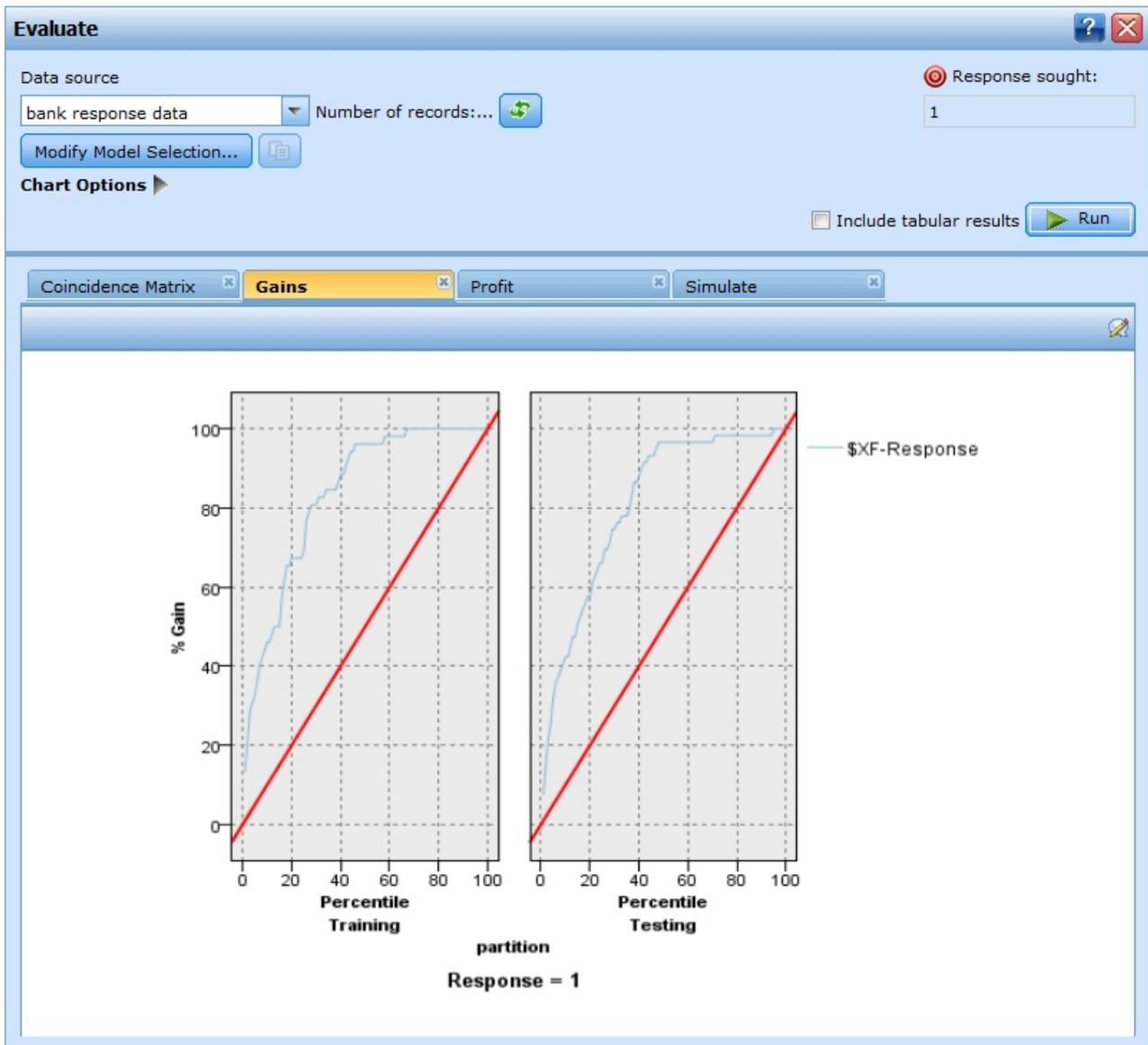


Figura 5. Evaluating the model

1. En el origen de datos, seleccione *datos de respuesta del banco*.
2. Seleccione **Gráfico de ganancias** y **Gráfico de beneficios** en los tipos de gráfico a construir.

3. Seleccione **Simular beneficio máximo** y especifique *3000* para la población. Esto coincide con el número de la campaña de prueba.
4. Especifique *2,25* en el coste y *60* en los ingresos. Estos valores indican el coste que le supone a usted cada oferta realizada (por ejemplo, costes de producción, tramitación y envío de correos) así como los ingresos esperados de cada respuesta favorable.
5. En **Respuesta buscada**, seleccione *1*.
6. Pulse **Ejecutar**.

El gráfico de ganancias muestra la mejora o "elevación" que se puede obtener utilizando el modelo, como se muestra en la figura anterior. Puesto que estaba seleccionada la opción **Particionar los datos automáticamente** (bajo Valores opcionales), al construirse el modelo se visualizan gráficos independientes para las particiones de entrenamiento y prueba. Cada uno muestra la proporción de coincidencias (respuestas positivas) en cada percentil respecto del número total de coincidencias de la muestra. La línea plana diagonal representa la respuesta esperada si no se usara ningún modelo (para recopilar la mitad de coincidencias, habría que contactar con la mitad de personas). La línea curva muestra la mejora obtenida al usar el modelo). Dirigiéndose al 60% de los clientes que tienen mayor probabilidad de responder, puede lograr casi todas las coincidencias con una envío de correos mucho menor.

El gráfico de beneficios muestra las ganancias en términos de beneficio global a partir de los valores de coste y beneficio especificados. En la muestra usada para entrenar el modelo, el beneficio parece alcanzar un máximo justo por debajo del percentil trigésimo.

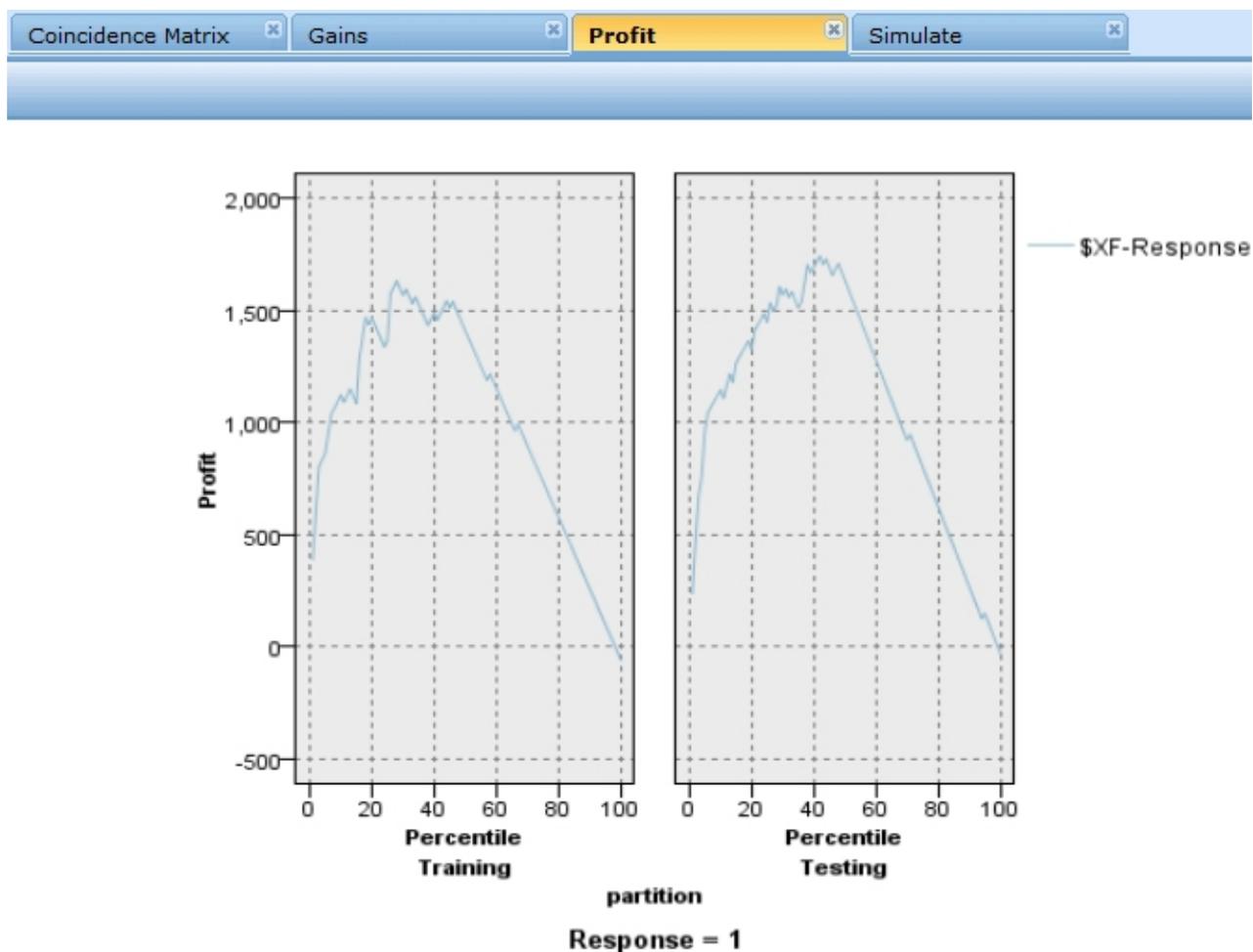


Figura 6. Profit chart

Para comprender cómo funciona esto, vaya a la pestaña Simular para ver el beneficio máximo esperado para el tamaño de población especificado.

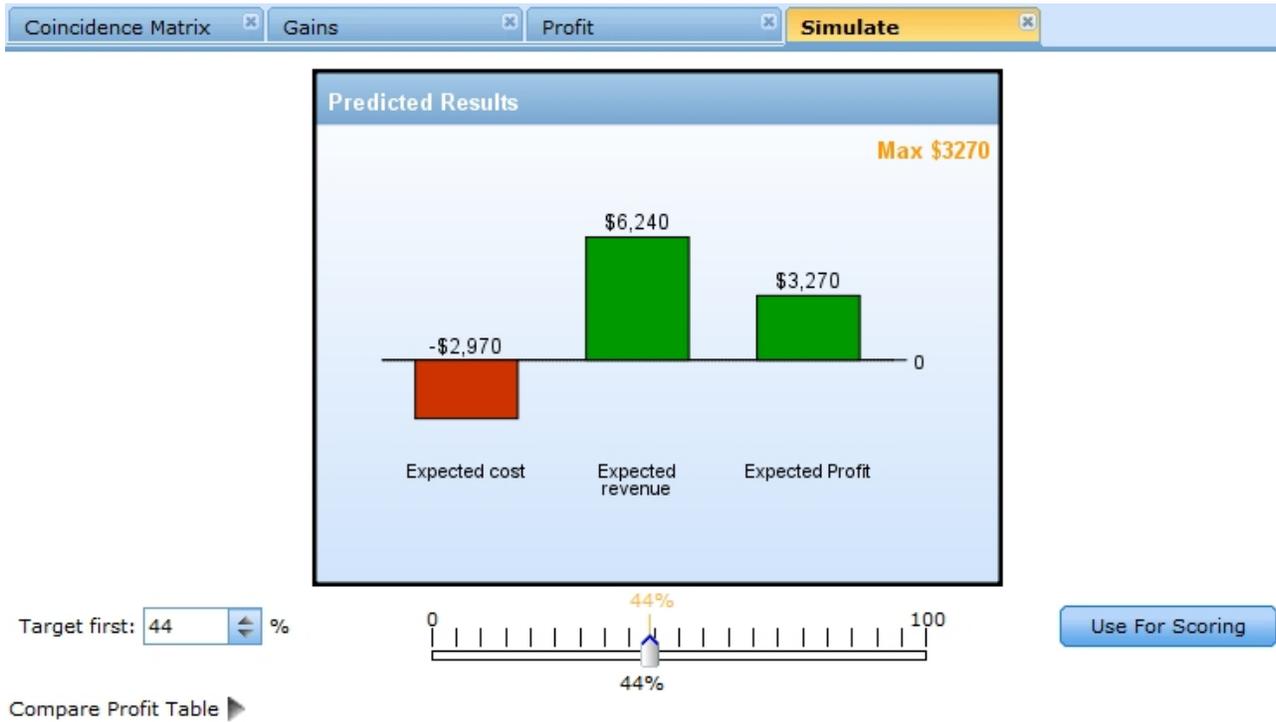


Figura 7. Profit simulation

El beneficio esperado se calcula mediante la fórmula siguiente:

$$\text{beneficio esperado} = (\text{probabilidad de respuesta} * \text{ingresos}) - \text{coste}$$

Donde la probabilidad de respuesta se determina para cada cliente en función del modelo.

Dirigiéndose a aquellos que el modelo identifica como que tienen una mayor probabilidad de responder, puede maximizar el beneficio por cliente. A medida que incluya más clientes, el beneficio total aumentará hasta un punto determinado, incluso si el retorno por cliente disminuye gradualmente a medida que se incluyen clientes con tasas de respuesta inferiores. La simulación ayuda a decidir dónde trazar la línea para obtener el beneficio máximo.

Compare Profit Table ▼

	Targeted 44%	Random 44%	Everyone
Offers made	1320	1320	3000
Expected response	104	48.8	111
Expected response(%)	7.9	3.7	3.7
Expected cost	\$2970	\$2970	\$6750
Expected revenue	\$6240	\$2930.4	\$6660
Expected Profit	\$3270	-\$39.6	-\$90

Figura 8. Profit comparison

La tala de comparación de beneficios resume este resultado con más detalle.

- **44% objetivo.** Esta columna muestra el resultado de dirigirse al 44% de clientes que tienen las tasas más altas según la predicción del modelo. Al incluir de manera selectiva los 1320 clientes que tienen una mayor probabilidad de responder, se maximiza el beneficio, dando lugar a un beneficio esperado de 3270 \$.
- **44% aleatorio.** Muestra el resultado de extraer una muestra aleatoria del mismo tamaño. En este caso se enviaría el mismo número de ofertas (1320), pero la respuesta esperada sería inferior (coincidiendo con la tasa de la campaña de prueba) porque no se está aprovechando la información adicional que proporciona el modelo. Esto se traduce en una pérdida, porque la tasa global de respuestas no es lo suficientemente elevada como para compensar el coste del envío por correo de cada oferta.
- **Todos.** Muestra el resultado de incluir los 3000 clientes del ejemplo, coincidiendo exactamente con el resultado de la campaña de prueba. De nuevo, esto se traduce en pérdidas porque no se aprovecha el modelo.

Si se fija en los resultados, puede ver cómo la utilización del modelo puede mejorar la tasa de respuestas de forma que una campaña con pérdidas se convierta en una con beneficios. La diferencia entre una pérdida de 90,00 \$ y una ganancia de 3277,50 \$ representa el beneficio obtenido a través de modelado predictivo. Este retorno se basa en los 3000 clientes de la campaña de prueba, pero si se aumenta el tamaño de la población objetivo, los beneficios experimentan el correspondiente aumento.

7. Después de analizar los resultados de la simulación, seleccione **Usar en puntuación** bajo el gráfico Resultados pronosticados para actualizar los valores de la pestaña Puntuación en función del porcentaje esperado (44% en este caso). Consulte el tema “Puntuación del modelo” en la página 19 para obtener más información.
8. Si lo desea, guarde el proyecto como *bank_response_model.str*. En la sección siguiente, se cambia el tipo de modelo y se crea un proyecto.

Modelado de reglas predictivo

Aunque los modelos predictivos pueden ofrecer toda la potencia predictiva necesaria, en ocasiones puede ser necesario un control y una comprensión mayores del funcionamiento interno del modelo. En tales casos puede usarse un modelado de reglas predictivo (interactivo). En un modelo de reglas predictivo pueden crearse reglas de forma manual para segmentar los datos o buscar automáticamente segmentos con concentraciones altas o bajas del valor del campo seleccionado. Por ejemplo, podrían buscarse clientes que hayan dado una respuesta positiva a la campaña y luego identificar los segmentos con una probabilidad más alta de responder.

1. Vuelva a la pestaña Modelado y pulse **Cambiar modelo**, y seleccione **Modelo de reglas predictivo**.
2. Seleccione *datos de respuesta del banco* como origen de datos y seleccione *Respuesta* como campo de destino.
3. Especifique 1 como respuesta buscada, lo que indica que se buscan los clientes que han dado una respuesta positiva.
4. Seleccione **Iniciar construcción**.
5. Seleccione **Aumentar modelo** para identificar los segmentos que tengan mayor probabilidad de responder. Se crearán y visualizarán las tres mejores reglas que identifiquen los segmentos que tengan la mayor probabilidad de responder.
6. Pulse en los iconos **Incluir/Excluir** para seleccionar si se excluyen o incluyen los clientes seleccionados por cada regla. Por ejemplo, los segmentos de regla segundo y tercero muestran tasas de respuesta algo superiores al 8%, que siguen siendo más altas que la tasa de campaña de prueba global del 3,7%, pero que ni siquiera se acercan a la tasa del 15.65% del primer segmento. Para ver el impacto de la exclusión o inclusión de estas dos reglas en el beneficio global, puede ejecutarse la simulación de beneficio en ambos sentidos y comparar el resultado de esta forma:
7. Con los tres segmentos de regla establecidos a **Incluir**, pulse **Evaluar**.
8. Seleccione **Gráfico de ganancias** y **Gráfico de beneficios** en los tipos de gráfico a construir.

9. Seleccione **Simular beneficio máximo** y especifique **3000** para la población, **2,25** para el coste y **60** para los ingresos. Son las mismas especificaciones usadas al evaluar el modelo predictivo. La respuesta buscada se completa de forma automática.
10. Pulse **Ejecutar** y, cuando se construyan los gráficos, vaya a la pestaña Simular.
El beneficio máximo simulado de este modelo es \$3178,90, un poco inferior al del modelo automatizado en este caso. De forma opcional, puede cerrar el diálogo Evaluar, cambiar el valor de los segmentos de regla segundo y tercero a **Excluir** y repetir la simulación de beneficios para ver cómo cambia el resultado. El verdadero beneficio del modelado de reglas predictivo (interactivo) son el control y comprensión mayores que ofrece.
11. Si lo desea, guarde el proyecto como *bank_predictive_rule_model.str*.

Prueba del modelo

Cuando haya creado el modelo, puede probar una selección de registros de cliente para ver con qué probabilidad responden a la campaña.

1. Para obtener una vista previa de cómo procesará la puntuación del modelo registros concretos, seleccione **Probar** en la pestaña Modelado. Aparecerá el diálogo Probar.
2. En el área **Registros**, incremente la cifra de Devolución a 100.
3. En el área **Seleccione los campos a visualizar**, seleccione *ID de cliente*. Además se incluirán los campos *\$XF-Response* y *\$XFRP-Response*, porque indican la respuesta pronosticada (0 para no, 1 para sí en este caso) y la puntuación de propensión de cada cliente.

De forma opcional, se pueden incluir campos adicionales como *Edad*, *Semanas desde la última oferta* y *Tiene historial de impagos* que pueden ser útiles para comprender cómo o por qué el modelo puntúa a un cliente determinado de una manera determinada.

Puntuación del modelo

Puntuar un modelo significa aplicarlo al origen de datos en el que se está interesado. Por ejemplo, si ha generado un modelo que predice qué clientes tienen mayor probabilidad de responder, ahora podrá puntuar clientes potenciales para identificar a quién desea incluir en la campaña. En el cuadro de diálogo Evaluar, determinó el porcentaje de registros por puntuar a fin de maximizar los beneficios e indicó el porcentaje de registros que deseaba puntuar en función de este resultado.

1. Vaya a la pestaña Puntuar.
2. En **Origen de datos por puntuar**, asegúrese de que *datos de respuesta del banco* esté seleccionado.
3. Especifique el destino del resultado de la puntuación (si solo desea realizar una vista previa de las puntuaciones, puede saltarse este paso).
4. Bajo **Campos de salida**, asegúrese de que el campo *ID de cliente* esté seleccionado.
5. Los valores de **Puntuación** ya deberían estar configurados para puntuar el mismo porcentaje de registros que en la simulación de beneficios ejecutada al evaluar el modelo. Si no pulsó el botón **Usar para puntuación** al visualizar el resultado de la simulación, puede actualizar los valores aquí.
6. En la parte inferior del diálogo, pulse **Vista previa** para visualizar una muestra del resultado que obtendrá la puntuación.
7. Pulse **Descripción general de los datos** para obtener una visión general de los datos que va a usar la puntuación. Consulte el tema “Visión general de datos” en la página 43 para obtener más información.

Puesto que ha optado por puntuar en función de la propensión, solo se incluirán los clientes que tengan la propensión más alta a responder. Cruzando el ID de cliente con la base de datos de clientes, podrá generar una lista de distribución a partir de este resultado.

8. Pulse **Puntuar**; el modelo se aplicará a los datos de cliente y aparecerá un mensaje de confirmación en la parte inferior del diálogo.

Ejemplo de modelado de clúster

En un modelo de clúster pueden categorizarse los registros en grupos con características similares. Esto puede ayudar a encontrar grupos naturales en los datos. Por ejemplo, los clientes podrían segmentarse en función de sus características demográficas o de su comportamiento de compra.

Este ejemplo utiliza los datos del archivo *bank_customer_data.txt*, que se distribuye e instala con el producto. En el archivo *bank_cluster_model.str* también se proporciona una versión completa de este ejemplo. Póngase en contacto con el administrador para obtener los detalles de la instalación y los archivos de ejemplo si es necesario. Consulte el tema “Archivos de muestra” en la página 9 para obtener más información.

1. Vuelva a la página de lanzamiento de aplicaciones y cree un proyecto de IBM SPSS Modeler Advantage.
2. En la pestaña Datos, en el panel Orígenes de datos del proyecto, añada un nuevo origen de datos denominado *datos de clientes del banco* como se describe en “Configuración de un origen de datos” en la página 11.
3. En la pestaña Modelado, pulse **Cambiar modelo**, seleccione **Modelo de clústeres** y pulse **Guardar**.
4. En el origen de datos, seleccione *datos de clientes del banco*.
5. Expanda la sección **Valores opcionales**. En los modelos de clústeres, existen valores opcionales adicionales de **Opciones de clúster automático** para seleccionar un campo de evaluación o definir un rango deseado de clústeres a buscar.

Cuando se construye un modelo de clústeres, en realidad se crea internamente una serie de modelos de clústeres distintos. Luego dichos modelos se comparan y clasifican conforme a alguna medida de calidad, y se selecciona el mejor modelo para su uso en el modelo de clúster automático. Por tanto, estos dos valores de opciones de clúster automático proporcionan cierto control sobre cuál de los modelos se selecciona. Si se especifica aquí un campo de evaluación, se seleccionará el modelo que mejor diferencie los valores del campo de evaluación. Y si se define el rango deseado, se descartarán los modelos que encuentren una serie de clústeres fuera del rango especificado.

6. En este ejemplo, seleccione la opción **Establecer un rango deseado para el número de clústeres encontrados** y utilice el mínimo de **3** y el máximo de **15** predeterminados. Esto garantizará que el modelo de clústeres no tenga un número excesivamente grande o pequeño de clústeres.
7. El resto de los valores opcionales son los mismos que en otros tipos de modelado. Deseleccione la opción **Limpiar y preparar de forma automática los datos para una construcción de modelos fiable** para que el ejemplo de agrupación en clústeres automática sea más fácil de entender.
8. Contraiga el panel **Valores opcionales** y expanda el panel **Clústeres**.

En el panel Clústeres, hay una sección para clústeres manuales y otra para clústeres automáticos. Los clústeres manuales permiten definir clústeres a partir del conocimiento de los datos. En este ejemplo se define un clúster manual para *clientes de alto valor* y otro clúster manual para *clientes solteros jóvenes* de la manera siguiente.

9. Pulse en el icono **Crear una regla**. Llámela *clientes de alto valor*, añada lo siguiente a las expresiones y pulse **Aceptar**.

```
Months as a Customer > 12  
Income > 65000
```

10. Cree otra regla llamada *clientes solteros jóvenes* que tenga las expresiones siguientes.
Age < 35
Marital Status = U

11. Las dos nuevas reglas se listarán bajo la sección **Clústeres manuales**. Para ver cuántos clientes se capturan con estos clústeres manuales, pulse **Recuento de registros de clúster**. Luego vuelva a expandir los clústeres manuales y fíjese en la columna Recuento para ver cuántos clientes se han encontrado en cada clúster. Verá que se han encontrado 67 clientes de alto valor y 33 clientes solteros jóvenes.

12. Luego pulse **Buscar clústeres automáticos** para encontrar de forma automática otros clústeres que puedan ser de interés. Cuando haya terminado, expanda la sección **Clústeres automáticos**. Verá que se han encontrado tres clústeres automáticos.

13. Pulse **Ver clústeres automáticos** para ver más detalles. Se mostrará el Visor de resultados de clúster automático.

La página *Resumen de modelo* proporciona información básica sobre el algoritmo de modelo de clúster automático usado (*Bietápico* en este ejemplo), el número de clústeres automáticos encontrado y una idea básica de la calidad del clúster.

Existen otros gráficos. Puede pasar el cursor del ratón por las diversas áreas de cada gráfico para ver más detalles e interactuar con algunos de los gráficos (como, por ejemplo, el de *Importancia de predictor*).

Con la vista *Clústeres* (el último gráfico), también se pueden comparar clústeres para identificar patrones de interés. En una implementación de la vida real, se estudiarían los clústeres automáticos y se les daría un nombre más significativo (pulsando en cada etiqueta de clúster en la vista *Clústeres*). Una vez actualizados los nombres de etiqueta de clúster, los nuevos nombres se usarán en todos los demás gráficos y en la pestaña Modelado. Luego se procedería a evaluar y probar el modelo y por último se puntuaría cuando se estuviera satisfecho con él.

14. Si lo desea, guarde el proyecto como *bank_cluster_model.str*.

Para obtener detalles adicionales sobre el modelado de clústeres, consulte “Generación de un modelo de agrupación en clúster” en la página 61.

Ejemplo de modelado de asociaciones

El modelado de asociaciones es una técnica importante de la analítica predictiva. Con un modelo de asociaciones se puede predecir qué elementos tienen más probabilidad de aparecer juntos, así como la solidez de las relaciones entre ellos. Por ejemplo, a partir de los artículos en la cesta de compra en un momento dado, podría identificarse la compra siguiente más probable que un cliente realizaría.

Este ejemplo utiliza los datos del archivo *retail_purchase_data.txt*, que se distribuye e instala con el producto. Los datos constan de información de compra de artículos electrónicos como, por ejemplo, televisores, ordenadores y teléfonos inteligentes.

En el archivo *retail_association_model.str* también se proporciona una versión completa de este ejemplo. Póngase en contacto con el administrador para obtener los detalles de la instalación de los archivos de ejemplo si es necesario. Consulte el tema “Archivos de muestra” en la página 9 para obtener más información.

1. Vuelva a la página de lanzamiento de aplicaciones y cree un proyecto de IBM SPSS Modeler Advantage.
2. En la pestaña Datos, en el panel Orígenes de datos del proyecto, añada un nuevo origen de datos denominado *datos de compra minorista*. Seleccione **Archivo** y seleccione *retail_purchase_data.txt*.
3. En la pestaña Modelado, pulse **Cambiar modelo**, seleccione **Modelo de asociaciones** y pulse **Guardar**.
4. En el origen de datos, seleccione *datos de compra minorista*.
5. En el formato de los datos, seleccione **Tabular**. Un conjunto de datos transaccionales sólo mostraría los elementos que se han incluido y todos los elementos aparecerían en una única columna. Sin embargo, en este conjunto de datos, cada elemento se encuentra en una columna aparte con un distintivo *verdadero* o *falso* que indica si se ha adquirido o no.
6. Pulse **Construir modelo**. Le dará un error en el que se indica que la construcción ha fallado y donde se sugiere que los umbrales podrían ser demasiado altos. Cierre el error.

Cuando se buscan reglas de asociación, suele ser una buena idea buscar reglas que sean de aplicación general, fiables y sencillas.

7. Expanda la sección Valores opcionales. Bajo Opciones de regla de asociación, cambie el primer umbral de 10 a 1 y baje el segundo umbral de 80 a 1. Lo normal es probar distintos umbrales y reconstruir el modelo cada vez para alcanzar un equilibrio entre un número excesivo de reglas y otro insuficiente.

8. Pulse **Construir modelo**.

En el resultado del modelo, cada fila representa una regla de asociación independiente. La primera regla de la lista dice que si un cliente compra un *Teléfono inteligente*, es probable que compre también una *TV de gran pantalla*. La columna **Cobertura (%)** indica que alrededor de un 15% (o 1 de cada 6 clientes) ha comprado el *Teléfono inteligente* y la columna **Confianza (%)** indica que dichos clientes comprarán también una *TV de gran pantalla* con una probabilidad en torno al 30%.

Por tanto, incluso un pequeño conjunto de reglas (239 en este ejemplo) puede usarse en un contexto empresarial para predecir e incluso recomendar lo que los clientes comprarán después.

9. Luego pulse **Probar** y pulse **Ejecutar** para ver cómo se van a realizar las predicciones o recomendaciones.

En este caso, el modelo ha buscado las reglas que arrojan las cifras de confianza más elevadas, ha visto lo que el usuario ya había comprado y, a partir de ambas cosas, ha realizado recomendaciones sobre lo que es más probable que compren los clientes.

10. Pulse en el icono de vista previa junto al primer registro, por ejemplo, para ver que el primer cliente ha comprado una *TV de gran pantalla* y *altavoces*, pero ningún otro producto. Cierre la vista previa.

La recomendación para este usuario consiste en ofrecerle una *TV estándar*, como se aprecia en la segunda columna. El valor de la tercera columna, **0,3**, significa que hay una probabilidad del 30% de que los clientes que ya han adquirido una *TV de gran pantalla* y *altavoces* también adquieran una *TV estándar*.

Sin embargo, ¿cómo puede saberse si se trata de una anomalía o de una tendencia pronosticada que puede ser de aplicación general? Apunte el número de la regla (**115** en este ejemplo) para poder volver y examinarla en más detalle.

11. Cierre el diálogo Prueba.

12. Pase las páginas del modelo hasta llegar al número de regla 115. Observe que, en los modelos con muchas páginas de resultados, también se puede pulsar **Buscar una regla por ID** y escribir el número de regla.

13. Pulse en la flecha que aparece junto a cualquiera de las cabeceras de columna del resultado y seleccione

Columnas > Cobertura (N)

Se mostrará la columna **Coberturas (N)**. Viendo el número de instancias, está claro que no se trata de una anomalía, porque la regla 115 se basa en 58 clientes distintos que han adquirido los tres productos en la misma compra. Esto es suficiente para dar confianza en la utilización de la regla como base para recomendaciones en el futuro.

Nos hemos fijado en una única regla para ver cómo podría usarse en una recomendación. No obstante, IBM Analytical Decision Management también proporciona potentes opciones para personalizar el modelo y hacerlo más relevante a objetivos concretos. Por ejemplo, puede que se desee predecir qué clientes están interesados en comprar una *TV de gran pantalla*:

14. Pulse **Aplicar filtros**. Se mostrará un cuadro de diálogo donde se pueden definir filtros de modo que aparezcan en el modelo las reglas que coincidan con los filtros.

15. Seleccione **Habilitar filtro** y seleccione *TV de gran pantalla* en el artículo.

16. Pulse **Guardar** para volver a los resultados del modelo.

El número de reglas ha bajado de 239 a 34.

Puede que también le interese confirmar que solo quiere usar reglas basadas en un número de clientes razonable. Para ello, podría excluir las reglas que estén basadas en el comportamiento de menos de 10 clientes:

17. Pulse en la flecha situada junto a la cabecera de la columna Cobertura (N) y seleccione **Ordenar ascendentemente**.

18. En la columna ¿Excluir?, seleccione todas las reglas que tengan un valor de Cobertura (N) inferior a 10 instancias.
19. Observe que, aunque ahora muchas de las reglas se han filtrado y excluido, siguen estando en el modelo en este punto. Para eliminarlas permanentemente de modo que no se utilicen cuando se puntúe el modelo, pulse **Suprimir reglas excluidas y filtradas** y pulse **Aceptar**.
El número de reglas devueltas se ha reducido a 20, basadas todas ellas en 10 instancias o más, y prediciendo todas ellas *TV de gran pantalla*.
20. Para ver cuántas de estas reglas se van a aplicar, pulse **Probar** y luego pulse **Ejecutar**. *TV de gran pantalla* se recomienda ahora a algunos clientes más y algunos tienen un nivel de confianza superior al de otros.
21. Cierre el cuadro de diálogo Probar.
22. En la sección Valores opcionales, expanda la sección Opciones de puntuación para visualizar las opciones de puntuación específicas de los modelos de asociaciones.
De forma predeterminada, el número máximo de predicciones es 3. Este se basa en las cifras de confianza más altas, aunque se puede cambiar el criterio de regla para elegir reglas basadas en cobertura, soporte de regla, la mayor cantidad de *elevación* (el incremento de probabilidad de compra de un determinado artículo) o en la capacidad de despliegue. Puede que también le interese asegurarse de no ofrecer una *TV de gran pantalla* a los clientes que ya hayan comprado una, seleccionando para ello **Garantizar que las predicciones no estén presentes**.
Las opciones de puntuación predeterminadas son adecuadas en la mayoría de situaciones. Sin embargo, en algunos casos, la flexibilidad adicional que proporcionan estas opciones puede ser ventajosa.
El siguiente paso del proceso sería usar la pestaña Puntuación para puntuar el modelo. Consulte el tema “Puntuación de modelos en una tabla de base de datos, un archivo, Analytic Server o un servidor IBM Cognos BI” en la página 73 para obtener más información.
23. Si lo desea, guarde el proyecto como *retail_association_model.str*.

Para obtener información adicional sobre el modelado de asociaciones, consulte “Creación de un modelo de asociación” en la página 62.

Capítulo 3. Navegación por el espacio de trabajo

Cambiar el idioma visualizado

Cada usuario puede cambiar fácilmente el idioma de su aplicación, sin tener que reiniciar ningún servidor. Por ejemplo, los usuarios de habla alemana, los de habla inglesa y los de habla japonesa todos pueden estar utilizando la misma aplicación en su propio idioma.

Desde el menú **Herramientas** en el navegador, vaya a **Opciones de Internet** u **Opciones**, cambie al idioma deseado y, después, renueve la pantalla. Si IBM Analytical Decision Management no se proporciona en el idioma que desea, se visualizará en inglés de forma predeterminada, o en el segundo idioma de su elección, si está disponible.

La interfaz de IBM Analytical Decision Management soporta los idiomas siguientes. Tenga en cuenta que esta lista está sujeta a cambios de release a release, y que la documentación no está disponible en todos los mismos idiomas que la interfaz de usuario.

- Inglés
- Alemán
- Español
- Francés
- Italiano
- Japonés
- Coreano
- Portugués (Brasil)
- Ruso
- Chino (China/simplificado)
- Chino (Taiwán/tradicional)

Inicio de aplicaciones

La página de inicio de aplicaciones proporciona enlaces configurables a aplicaciones instaladas. Cada panel representa una aplicación instalada. Pulse el desplegable para elegir entre los proyectos, modelos o reglas guardados recientemente. También puede añadir uno o más atajos personalizados que apuntan a versiones específicas etiquetadas.

- Para acceder a la página de inicio de las aplicaciones, vaya hasta *http://hostname:port/DM*, donde *hostname* es la dirección IP o el nombre de la máquina del repositorio y *port* es el número de puerto del servidor de aplicaciones. Si desconoce esta información, póngase en contacto con el administrador. También puede marcar como favorita esta página en el navegador, o crear un acceso directo en el escritorio.

Si tiene problemas al acceder al URL de IBM Analytical Decision Management o cualquiera de los URL de IBM SPSS Collaboration and Deployment Services, un administrador puede haberlos personalizado. Póngase en contacto con el administrador. Está disponible más información en el capítulo *Raíces de contexto de aplicación* de la *Guía de instalación y configuración del repositorio de IBM SPSS Collaboration and Deployment Services*.

- Para crear un proyecto, modelo o regla nuevo, seleccione **Nuevo** desde la lista desplegable para la aplicación deseada y pulse **Ir**.
- Para abrir un proyecto, modelo o regla guardado, seleccione una versión o atajo que se haya utilizado recientemente en la lista desplegable en el panel de la aplicación, o seleccione **Examinar** en la lista

desplegable para elegir un objeto o versión diferente. Para ver rápidamente la ubicación del repositorio de un elemento de la lista, pase el ratón por encima en el menú desplegable.

- Para añadir un nuevo atajo personalizado que siempre estará disponible en la lista desplegable, pulse el icono de la flecha triangular en el extremo superior derecho del panel de la aplicación. De forma opcional, puede especificar el número de atajos para mostrar.

Otras formas de abrir una aplicación

- Copie o escriba el URL en la barra de direcciones del navegador, o añada el URL a la lista de marcadores o favoritos guardados en el navegador.
- Cree un atajo de escritorio que apunte al URL.
- Haga un enlace a un URL de aplicación desde dentro del sitio o portal de la Intranet de la empresa.

Personalización de la página de inicio

Los controles de la esquina superior derecha de la página de inicio de la aplicación le permite añadir aplicaciones y conmutar la visualización entre la vista del panel y lista. También puede arrastrar paneles para reorganizarlos en la página.

- Para añadir un nuevo panel de aplicación, seleccione **Añadir aplicación** en la esquina superior derecha de la página de inicio de la aplicación. (Si no está disponible ninguna aplicación adicional, este enlace no está disponible.)

Se visualiza una lista de aplicaciones disponibles. De forma opcional, puede crear atajos a una o más versiones etiquetadas.

- Para cambiar la disposición de paneles de aplicación en la página de inicio, pulse en la barra de título de cualquier panel y mantenga pulsado el botón del ratón para arrastrar el panel hasta una nueva ubicación.
- Para eliminar una aplicación, pulse el icono para eliminar (x roja) en la esquina superior derecha del panel de aplicación. Los atajos que apuntan a dicha aplicación también se eliminarán. (El icono para eliminar no estará disponible si el administrador ha bloqueado el panel de la aplicación en la página.)
- Los proyectos, modelos y reglas que no se han guardado pueden aparecer listados en la galería de la página de inicio. Esta galería muestra modelos que se siguen creando, así como aquellos modelos que ya se han completado, pero que todavía no se han guardado. Consulte el tema “Galería” si desea más información.

La página de inicio

La aplicación proporciona un flujo de trabajo paso a paso, tal como se representa mediante los iconos de la página de inicio. Pulse cualquier icono para saltar a dicho paso.

Datos. Defina conjuntos de datos para ser utilizados por la aplicación.

Modelo. Cree modelos utilizando datos analíticos, evalúelos utilizando gráficos de importancia de predictor, distribución, ganancias, etc. para obtener información descubriendo patrones ocultos en los datos y utilice todo esto para refinar más las decisiones.

Puntuación. Puntúe el modelo.

Galería

La galería se visualiza en la página de inicio de aplicaciones. Lista el trabajo no guardado como, por ejemplo, un modelo que se ha dejado para construir por la noche, lo que le permite iniciar el modelo antes de abandonar el trabajo, cerrar el navegador y volver a encontrar los resultados por la mañana. También puede utilizar la galería para recuperar el trabajo después de haber cerrado accidentalmente el navegador. Los elementos de la galería pueden incluir proyectos, reglas, modelos, exploraciones de datos, evaluaciones y otras tareas de larga ejecución, según corresponda a la aplicación.

- Si el nombre del elemento está subrayado, pulse el nombre para abrir el elemento relevante. Por ejemplo, si tiene una exploración de datos sin guardar, pulse el título de la exploración de datos en la galería para abrir la pestaña Datos.
- Si no necesita que aparezca un elemento en la galería, pulse el icono de suprimir para eliminarlo; se visualiza un mensaje de aviso para que confirme que desea eliminar el elemento de forma permanente. O bien, para eliminar de golpe varios elementos de la galería, pulse **Borrar galería**. Aparecerá el cuadro de diálogo Borrar galería que le permite seleccionar en qué pestañas eliminar todos los objetos.
- En la pestaña Proyectos de la galería, el menú desplegable **Creado en** le permite filtrar los elementos visualizados en la pestaña para cada aplicación. Por ejemplo, al trabajar en una aplicación, es posible que no desee ver los elementos de la galería para otras aplicaciones.

Establecimiento de preferencias

Existe una serie de preferencias que puede ajustar para personalizar IBM Analytical Decision Management para sus necesidades. Principalmente, esta personalización consta del establecimiento de sus propias opciones predeterminadas para que no tenga que seleccionar de forma repetitiva las mismas acciones al crear, probar y guardar su trabajo. Las preferencias que defina son exclusivas para el usuario y se aplican a cualquier aplicación que utilice.

Para acceder a los valores de preferencias de usuario, desde dentro de cualquier aplicación, pulse el icono **Definir preferencias de usuario** en la barra de herramientas.



Figura 9. Icono Definir preferencias de usuario

Puede especificar sus propios valores preferidos para cada área, tal como se detalla en las secciones que siguen.

Valores generales

Número predeterminado de objetos abiertos recientemente con atajos de la aplicación. Seleccione el número de elementos que van a estar disponibles en la lista desplegable en la página de inicio de la aplicación.

Formato de fecha. Seleccione el formato para utilizar al visualizar o registrar fechas.

Formato de hora. Seleccione el formato para utilizar al visualizar o registrar la hora.

Moneda. Seleccione la moneda predeterminada para cada nuevo proyecto. Tenga en cuenta que esto no sobrescribirá los datos existentes aportados como una moneda diferente.

Posiciones decimales. Seleccione el número de posiciones decimales para mostrar para los números.

Número de filas que se mostrarán en vista previa. Seleccione el número máximo de registros para mostrar al visualizar vistas previas; por ejemplo, al comprobar un origen de datos al importar.

Número de resultados por página (tablas paginadas). Seleccione el número máximo de resultados para mostrar en cada página al visualizar tablas.

Número máximo de elementos visualizados en un gráfico. A veces, los resultados gráficos no funcionan bien cuando hay muchos elementos para mostrar. Especifique el número máximo de elementos para mostrar en gráficos. Esto forzará a que se visualice una tabla en lugar de un gráfico, si el número de elementos excede este máximo. Este valor solo se aplica a campos categóricos en vistas generales de datos y el gráfico del informe de resumen de puntuación para modelos con un objetivo categórico.

Número de segmentos de datos (mosaicos) que se va a aplicar en gráficos y tablas. Seleccione el número máximo de segmentos de datos para utilizar en toda la aplicación para gráficos y tablas, donde sea aplicable.

Valores de modelo

Tipo de modelo creado de forma predeterminada. Especifique si, de forma predeterminada, desea que la aplicación genere el mejor modelo automáticamente a partir de los datos disponibles (modelo predictivo), si desea utilizar un modelo que identifica segmentos que muestran una probabilidad mayor o menor en relación con un resultado específico (modelo de regla predictiva), si desea utilizar un modelo que categoriza registros en un determinado número de clústeres (modelo de agrupación en clústeres), o si desea utilizar un modelo que predice qué elementos tienen más probabilidad de aparecer juntos y predice la solidez de la relación entre ellos (modelo de asociación). Solo los tipos de modelo configurados para aparecer en la aplicación estarán disponibles en este menú desplegable.

Partir datos automáticamente para habilitar la evaluación del modelo en el origen de datos de generación para su evaluación y prueba. Seleccionar esta opción divide los datos en subconjuntos o muestras separados para la formación y realización de pruebas del modelo. Si crea el modelo en un subconjunto y lo prueba en otro, podrá obtener una idea sobre cómo se generalizará en otros conjuntos de datos. También puede especificar un valor de porcentaje para el *tamaño de partición de entrenamiento* y el *tamaño de partición de prueba* generados aleatoriamente. Estas opciones están disponibles para todos los tipos de modelo, excepto el modelado de asociación.

Limpiar automáticamente y preparar datos para una generación de modelos fiable. Si selecciona esta opción, conseguirá que la aplicación haga muchas de las tareas de preparación de datos. Por ejemplo, reparación de problemas de datos, cribado de datos menos útiles o derivación de nuevos atributos.

Puede elegir si la operación de limpieza deberá dar igual importancia a la velocidad y, también, a la exactitud, dar prioridad a la exactitud sobre la velocidad, o dar prioridad a la velocidad como, por ejemplo, cuando se trabaja con conjuntos de datos muy grandes o se busca un resultado rápido.

Nota: Los valores de preparación de datos automática solo se aplican al crear un modelo nuevo, y el intervalo no se realiza.

Restringir el tiempo de creación de un modelo sencillo. Utilice esta opción para especificar la cantidad predeterminada de tiempo para dedicar a la generación de cada modelo. Esto puede ser especialmente útil al procesar conjuntos de datos grandes, y puede tener un impacto significativo en el tiempo necesario para que se complete la generación del modelo.

Valores de prueba y simulación

Número predeterminado de registros probados. Especifique el número predeterminado de registros para incluir al ejecutar una prueba.

Fecha de prueba/simulación predeterminada. Especifique si la fecha de simulación predeterminada se va a establecer en la fecha actual, o una fecha específica.

Abrir y guardar proyectos, modelos y reglas

Puede abrir en, suprimir de o guardar en IBM SPSS Collaboration and Deployment Services Repository, o descargar en el sistema de archivos local. También puede cargar archivos de ruta de IBM SPSS Modeler desde el sistema de archivos local y guardarlos en IBM SPSS Collaboration and Deployment Services Repository. Guardar en IBM SPSS Collaboration and Deployment Services Repository proporciona importantes ventajas, incluyendo el soporte para varias versiones, etiquetado y seguridad. En función de la configuración local, no todas las opciones pueden estar soportadas.



Crea un modelo nuevo.



Guarda el proyecto actual en su última ubicación conocida en IBM SPSS Collaboration and Deployment Services Repository. De forma opcional, puede aplicar una etiqueta nueva al guardar.



Guarda y etiqueta el proyecto actual, o lo guarda con un nombre nuevo. Donde sea aplicable, puede elegir entre guardar en IBM SPSS Collaboration and Deployment Services Repository y descargar en el sistema de archivos local.



Abre un proyecto desde IBM SPSS Collaboration and Deployment Services Repository o el sistema de archivos local, según corresponda. O bien, donde sea aplicable, también puede cargar un proyecto desde el sistema de archivos local. El acceso al sistema de archivos no está soportado para todas las aplicaciones. Si se carga una ruta IBM SPSS Modeler, puede elegir abrir el archivo o guardarlo en el repositorio..

Abrir y guardar en el repositorio

IBM SPSS Collaboration and Deployment Services Repository proporciona importantes ventajas, incluyendo el soporte para varias versiones, etiquetado y seguridad. Cada vez que se guarda, se crea una nueva versión y podrá volver a una versión anterior cuando lo desee. También puede gestionar versiones utilizando etiquetas y aplicar valores de seguridad para determinar quién puede acceder, ver o eliminar un objeto.

La vista de árbol del panel izquierdo muestra carpetas en el IBM SPSS Collaboration and Deployment Services Repository que está autorizado para ver. Desplácese por el árbol de carpeta o realice una búsqueda para encontrar lo que está buscando. También puede filtrar la lista de archivos mediante etiquetas disponibles en el panel del extremo inferior izquierdo. Por ejemplo, podría pulsar una etiqueta de *producción* para mostrar solo los archivos marcados con dicha etiqueta.

Para ver más detalles sobre un archivo, cuando estén disponibles, pulse el icono de vista previa o el icono de información junto a ella.

Para suprimir un archivo, pulse el icono para suprimir junto a él. Esto es especialmente útil para los usuarios que no tienen acceso al software de cliente IBM SPSS Collaboration and Deployment Services Deployment Manager.

Carpeta. Muestra las carpetas disponibles y le permite navegar por la jerarquía. Elija la carpeta donde guardar. También puede crear una carpeta nueva y cambiar el nombre de una carpeta existente utilizando los iconos encima de la lista de carpetas.

Filtrar por etiqueta. Limita la visualización a los objetos con la etiqueta seleccionada. Al examinar carpetas, solo se muestran los objetos con esta etiqueta.

Al exportar reglas utilizadas en una aplicación, también esta disponible un desplegable **Guardar como o Exportar como**. Seleccione **Conjunto de reglas individuales** o **Reglas individuales** en el desplegable para especificar quién va a guardar la regla. Si solo están disponibles reglas referenciadas, el desplegable estará inhabilitado y la regla se guardará como un conjunto de reglas individuales (las reglas referenciadas no se pueden guardar de forma individual).

Si elige guardar como reglas individuales, se visualizará una tabla para especificar detalles para cada regla individual que desee guardar.

Nombre. Entre un nombre exclusivo para el objeto que está guardando.

Palabras clave. Entre cualquier palabra clave para identificar rápidamente el objeto guardado al buscarlo. Por ejemplo, estas podrían ser *Ventas* y *1-2009*.

Descripción. Entre una descripción más larga del objeto guardado. Por ejemplo, esto podría ser *Análisis de ventas minoristas; primer trimestre del 2009*.

Etiquetas, si es probable que tenga varias versiones del mismo objeto guardado, puede identificarlas con una etiqueta exclusiva. Seleccione una etiqueta existente o cree una nueva.

Trabajar con etiquetas

Las etiquetas se utilizan para identificar una versión específica de un proyecto, modelo o regla. Cuando se comparte una regla entre varios proyectos, por ejemplo, puede especificar que desea utilizar la versión etiquetada como *producción*. Si se crea una versión más nueva de la regla, la utilizada en el proyecto no cambiará, a menos que la etiqueta de *producción* se mueva explícitamente a dicha versión.

Al hacer referencia a objetos compartidos, se recomienda encarecidamente utilizar una versión etiquetada, particularmente en un entorno de producción, ya que los cambios en la última versión pueden tener, de lo contrario, consecuencias no deseadas. En determinados casos como, por ejemplo, al hacer referencia a una regla desde dentro de una regla, se debe utilizar una versión etiquetada. Consulte el tema “Exportación y reutilización de reglas” en la página 54 si desea más información.

Si no se especifica ninguna versión, se utiliza la última versión, lo que significa que se seleccionarán automáticamente los cambios posteriores en la regla. (Piense en *LATEST* como una etiqueta especial que se registra automáticamente en cada versión nueva cuando se crea.)

Cada vez que guarde un nuevo objeto o versión, puede elegir las etiquetas para aplicar a dicha versión. Consulte el tema “Abrir y guardar proyectos, modelos y reglas” en la página 28 si desea más información.

Restricciones en nombres de etiqueta. Al definir etiquetas para su uso con IBM Analytical Decision Management, los caracteres ? # & y : (dos puntos) no se pueden incluir en nombres de etiqueta, ya que pueden impedir que la aplicación interprete correctamente las referencias a objetos que utilizan estas etiquetas. Por ejemplo, los siguientes nombres no son nombres de etiqueta válidos:

Probar y desplegar

DEPLOY:PHASE1

Bloqueo de proyectos

Al abrir un proyecto, modelo o regla, está disponible una casilla de verificación Bloquear en el extremo superior derecho que le permite bloquear el objeto en IBM SPSS Collaboration and Deployment Services Repository de forma que otros usuarios no pueden editarlo.

- Cuando un proyecto está bloqueado, otros usuarios pueden abrirlo para verlo, pero no pueden guardar cambios.
- Cuando se cierra un proyecto bloqueado (por ejemplo, cerrando el navegador), el bloqueo se libera de forma que otros pueden editar el objeto.
- Si elige desbloquear un proyecto sin cerrarlo (deseleccionando la casilla de verificación Bloquear), se le pedirá que guarde los cambios. Los cambios sin guardar se perderán una vez desbloqueado el objeto.
- Tras desbloquear un proyecto, puede seguir realizando cambios, a menos que otro usuario lo bloquee mientras lo había desbloqueado (en cuyo caso, ya no podrá guardar cambios).

Propiedades

El diálogo Propiedades incluye una pestaña **Proyecto** y una pestaña **Repositorio**. Para mostrar el diálogo Propiedades, pulse el icono Propiedades en cualquier aplicación.



Figura 10. Icono Propiedades

En la pestaña **Proyecto**, puede especificar el máximo de miembros permitidos para campos *nominales*. Se utiliza en la pestaña Datos. Seleccione **Máximo de miembros para campos nominales** y especifique un valor. El valor predeterminado es **250**. Esto es útil en situaciones como, por ejemplo, las situaciones:

1. Un usuario crea un nuevo origen de datos y ejecuta una exploración de datos.
2. Un número de campos que contienen información categórica útil aparecen en el modelo de datos de proyecto como *sin tipo*, porque el número de valores distintos de los datos excede el límite de tamaño máximo predeterminado definido en el diálogo Propiedades.
3. El usuario pulsa el icono Propiedades y aumenta el valor **Máximo de miembros para campos nominales** a un tamaño apropiado para sus datos.
4. De nuevo en la pestaña Datos, el usuario pulsa **Renovar análisis de datos**, pero los campos permanecen como *sin tipo*, porque la acción de renovación no altera la medida para conservar cualquier cambio que haya realizado el usuario. En otras palabras, una renovación conserva la medida existente para un campo y solo vuelve a calcular los valores debido a que cualquier cambio que haya realizado el usuario en el modelo de datos se debería conservar.
5. Para que aparezcan los valores en el modelo de datos de proyecto, el usuario debe cambiar manualmente la medida de cada campo de *sin tipo* a *categórico* antes de ejecutar la renovación. Tenga en cuenta que si hay muchos campos afectados, puede ser más rápido volver a crear el origen de datos.

Tenga en cuenta que el valor para el máximo de miembros para campos nominales no se aplica a orígenes de datos del tipo de archivo *.sav*. Muchos archivos *.sav* tienen valores y archivos personalizados de usuario, así que no se verán afectados por este valor. El usuario debe cambiar manualmente los campos.

En la pestaña **Repositorio**, puede añadir palabras claves y una descripción para ayudar a identificar un proyecto, modelo o conjunto de reglas al buscarlo en una fecha posterior:

- **Palabras clave.** Entre cualquier palabra clave para identificar rápidamente el objeto guardado al buscarlo. Por ejemplo, estas podrían ser *Ventas* y *T1-2013*.
- **Descripción.** Entre una descripción más larga del objeto guardado. Por ejemplo, esta podría ser *Análisis de ventas minoristas; primer trimestre de 2013*.

Capítulo 4. Gestión de orígenes de datos

La pestaña Datos se usa para definir orígenes de datos en análisis, simulación y pruebas, puntuación u otros usos. El administrador puede configurar conjuntos de datos o usted puede añadir los suyos propios.

La pestaña incluye una pestaña **Origen** y una pestaña **Derivado**. La pestaña Origen se divide en estas áreas principales:

- **Modelo de datos del proyecto.** Define los campos necesarios para la aplicación. Todos los demás conjuntos de datos se correlacionan respecto de este origen. La sección **Campos de origen de datos** lista los campos de entrada y tipos que utiliza el proyecto. Si lo desea, utilice la pestaña **Derivado** para añadir expresiones o una salida de modelo que proceda de distintos orígenes de datos para ampliar el modelo de datos del proyecto.
- **Orígenes de datos.** La sección **Orígenes de datos de proyecto** lista los orígenes de datos que se han guardado para ser utilizados con el modelo, regla o proyecto actual. La sección **Mis orígenes de datos** lista orígenes de datos que se han definido o guardado. Los conjuntos de datos de esta lista se pueden copiar a o desde distintos proyectos y luego promoverse para su uso como modelo de datos del proyecto, posibilitando la compartición de datos entre aplicaciones.

Modelo de datos del proyecto

El modelo de datos del proyecto actúa como una plantilla que lista los campos de entrada y los tipos que va a utilizar el proyecto.

Origen de datos. Muestra el origen de datos seleccionado para la aplicación.

ID de registro. Si lo desea, puede seleccionar un campo para utilizarlo como ID de *entidad* a fin de habilitar la retrotracción SQL. Esto puede mejorar el rendimiento cuando los datos se originen en una base de datos y el campo seleccionado sea un entero. En IBM SPSS Modeler Advantage, el campo seleccionado aquí se selecciona de forma predeterminada como salida de una puntuación por lotes y se deselecciona de forma predeterminada como entrada en una generación de modelo. Este campo es opcional.

Operativo. Una marca en esta casilla indica que los datos del campo están disponibles para su uso en reglas y expresiones, y en predicciones. Desmarque la casilla de verificación si el campo se va a utilizar como campo objetivo del análisis. Por ejemplo, si está ejecutando una consulta para ver cómo van a responder los clientes a una campaña de envío de correo directo, el campo de los resultados será analítico, porque no se conocen los valores de respuesta de antemano.

Nombre del campo. Muestra el nombre de cada campo tal y como aparece listado en el origen de datos, junto con un icono que identifica su tipo de almacenamiento.

Medición. Muestra el tipo de medición de cada campo en el origen de datos. Si se cambia cualquiera de los tipos de medición, habrá que renovar el examen de datos para garantizar la compatibilidad de datos. Consulte el tema "Niveles de medición" en la página 44 para obtener más información.

Valores. Lista los valores de cada campo en el origen de datos. Por ejemplo, los valores máximo y mínimo de un rango. En los campos de distintivo, mantenga el puntero del ratón sobre el valor para visualizar una ayuda contextual que indica los valores de "verdadero" y "falso" definidos para el campo. También puede pulsar en un valor para editarlo.

Normalmente el modelo de datos del proyecto define las entradas como un conjunto de campos procedentes de una única tabla lógica. Esto siempre es así en el caso de IBM SPSS Modeler Advantage. Sin embargo, en otras aplicaciones, el modelo de datos del proyecto puede consistir en una tabla primaria junto con 0 o más tablas secundarias. Dichas tablas secundarias se definen mediante los orígenes de datos secundarios asociados al origen de datos primario.

Orígenes de datos del proyecto

Los orígenes de datos del proyecto pueden incluir los predefinidos por el administrador o los añadidos por los usuarios. De forma opcional, el administrador puede haber bloqueado uno o más orígenes de datos para impedir a los usuarios modificarlos o eliminarlos, o bloquear todas las opciones de datos de forma que los usuarios no puedan crear orígenes de datos.

Si las entradas de un origen de datos no coinciden directamente con las del modelo de datos del proyecto, puede correlacionar las primeras con las segundas y corregir la discrepancia. Por ejemplo, si el modelo de datos del proyecto requiere un campo llamado *compra* con los valores *Sí* y *No* (distintivo de nivel de medición), cualquier origen de datos utilizado deberá tener un campo equiparable que se pueda correlacionar.

Nombre. Muestra el nombre del origen de datos y un icono que identifica su tipo de archivo.

Número de registros. Pulse en el icono de esta columna para mostrar el recuento de registros de cualquier origen de datos de la tabla. El recuento aparece junto al icono.

Vista previa. Pulse en el icono de esta columna para obtener una vista previa de la muestra de los datos contenidos en el origen. Para obtener más información, consulte “Vista previa de datos” en la página 43.

Descripción general. Pulse el icono de esta columna para obtener una visión general del origen de datos. Para obtener más información, consulte “Visión general de datos” en la página 43.

Compatible. Muestra una nota que indica que el origen de datos se utiliza como modelo de datos del proyecto, o una bola verde, naranja o roja para indicar el grado de compatibilidad del origen de datos con el modelo de datos de proyecto.

- Una bola verde indica que el origen de datos es compatible operativamente con el origen de datos del modelo de datos del proyecto. Un origen de datos operativamente compatible es uno que incluye todos los campos operativos del modelo de datos del proyecto, pero que puede tener campos adicionales. Este origen de datos es apto para operaciones de reglas, puntuación, simulación y pruebas.
- Una bola naranja muestra que el origen de datos tiene al menos un campo que es compatible con el modelo de datos del proyecto, con el mismo nombre y tipo. Este origen de datos también puede tener campos adicionales y es apto para generar y evaluar modelos.
- Una bola roja muestra que el origen de datos no es compatible con el modelo de datos del proyecto y hay que correlacionar los campos antes de que se pueda utilizar en la aplicación. Un origen de datos incompatible es uno que tiene al menos un campo cuyo tipo no es compatible con el tipo equivalente del modelo de datos del proyecto.

Correlacionar campos. Esta opción permite comparar los campos del origen de datos con los del modelo de datos del proyecto y correlacionar o anular la correlación de campos compatibles para que coincidan con los que necesita el modelo de datos del proyecto. Para obtener más información, consulte “Campos de correlaciones” en la página 45.

Copiar. Copia el origen de datos en el área **Mis orígenes de datos**.

Mis orígenes de datos

Los orígenes de datos de esta lista se guardan con la cuenta de usuario, de forma que están disponibles siempre que se inicia sesión y se pueden copiar en cualquier proyecto, modelo o regla que se abra (suponiendo que el administrador haya otorgado la autoridad para hacerlo). Los campos de esta parte de la pestaña funcionan de la misma forma que los del área **Orígenes de datos del proyecto**, si bien no hay ninguna columna **Compatibilidad**.

La columna **Copiar** permite copiar el origen de datos en el área **Orígenes de datos del proyecto**.

Trabajar con orígenes de datos

- Para añadir un nuevo origen de datos a las listas **Orígenes de datos de proyecto** o **Mis orígenes de datos**, seleccione **Añadir un origen de datos**. Para obtener más información, consulte “Creación de un nuevo origen de datos”.
- Para cambiar el nivel de medición, o el tipo, de un campo en el origen de datos del **Modelo de datos de proyecto**, seleccione el nivel relevante. Para obtener más información, consulte “Niveles de medición” en la página 44.
- Para copiar orígenes de datos en o de la lista **Orígenes de datos del proyecto**, pulse en la correspondiente flecha de la columna **Copiar**.
- Para correlacionar nombres de campo para un origen de datos con el modelo de datos de proyecto, pulse el enlace apropiado en la columna **Compatible**. (Una vez que se han correlacionado los nombres de campo, el enlace deja de visualizarse.) Para obtener más información, consulte “Campos de correlaciones” en la página 45.
- Para obtener una vista previa de un origen de datos, pulse en el icono **Vista previa**. Para obtener más información, consulte “Vista previa de datos” en la página 43.
- Para añadir campos adicionales (expresiones, campos que utilizan reglas de segmentación, o una salida de modelo procedente de distintos orígenes de datos) a fin de ampliar el modelo de datos del proyecto, pulse en la pestaña **Derivado**. Para obtener más información, consulte “Derivación de campos” en la página 46.

Creación de un nuevo origen de datos

1. Para acceder al editor de orígenes de datos, seleccione **Añadir un origen de datos** en el área **Orígenes de datos de proyecto** o **Mis orígenes de datos**.
2. Escriba un nombre para el origen de datos.
3. Seleccione el tipo de origen. Tenga en cuenta que la lista de bases de datos disponibles es la lista de las disponibles en la máquina IBM SPSS Modeler.
 - **Archivo**. Están soportados muchos tipos de archivos como, por ejemplo, texto (*.txt, *.csv, etc.), Excel, los formatos *.sav y *.zsav utilizados por IBM SPSS Statistics y archivos SAS. Consulte el tema “Origen de archivo” en la página 36 si desea más información.
 - **Base de datos**. Cualquier base de datos que soporta ODBC, incluyendo Microsoft SQL Server, DB2, Oracle y otros. Consulte el tema “Origen de base de datos” en la página 39 si desea más información.
 - **Vista de datos**. Crea una conexión que se va a utilizar con reglas de IBM Operational Decision Management. Para reglas complejas que utilizan un modelo de datos jerárquico, tenga en cuenta que los administradores pueden restringir el uso de dichas reglas porque pueden degradar el rendimiento. Consulte el tema “Origen de la vista de datos” en la página 39 si desea más información.
 - **Servidor IBM Cognos BI**. Cualquier base de datos que soporta el servidor IBM Cognos BI. Consulte el tema “Origen de IBM Cognos BI” en la página 40 si desea más información.
 - **Analytic Server**. No puede seleccionar este tipo de origen al crear un nuevo origen de datos. En IBM SPSS Modeler 16, se ha introducido IBM SPSS Analytic Server, que permite el uso del sistema de archivos distribuido de Hadoop (HDFS) como origen de datos. Los usuarios de IBM SPSS

Modeler pueden crear rutas utilizando el nodo de origen de Analytic Server del HDFS y generar modelos basándose en el sistema de datos grandes de Hadoop. Los usuarios de IBM Analytical Decision Management pueden abrir estas rutas de modelo en IBM SPSS Modeler Advantage para su visualización (no pueden crear nuevos proyectos utilizando orígenes de datos de Analytic Server).

4. Especifique opciones según convenga para el tipo de origen.

Puede pulsar el icono **Número de registros** para mostrar el número de registros en el origen de datos actual, o pulse **Vista previa** para mostrar los primeros diez registros del origen de datos.

Origen de archivo

La aplicación soporta varios tipos de archivo de datos. Cuando se especifica el nombre de archivo o se busca el archivo, la aplicación detecta automáticamente el tipo y amplía el editor de orígenes de datos para mostrar campos adicionales que se relacionan con dicho tipo. Si el tipo de archivo no es correcto, o requiere una versión específica (por ejemplo, Microsoft Excel 1997-2003), puede cambiar el tipo a una necesaria.

Puede pulsar **Examinar** para seleccionar un archivo en el repositorio, o pulsar **Subir archivo local** para seleccionar un archivo en el sistema de archivos local (si está habilitado).

Si sube un archivo local, se subirá a IBM SPSS Modeler Server. En el diálogo Subir archivo, busque y seleccione el archivo local que desea subir y, después, busque y seleccione el destino de subida de IBM SPSS Modeler Server. Cuando seleccione el destino, podrá cambiar el nombre de archivo, si lo desea.

La prestación **Subir archivo local** está inhabilitada de forma predeterminada. Los administradores pueden utilizar el IBM SPSS Collaboration and Deployment Services Deployment Manager basado en navegador para habilitarlo o inhabilitarlo, y para aumentar o reducir el tamaño máximo de archivo permitido.

Archivos de datos basados en texto

Cuando seleccione un origen de datos basado en texto, se le pedirá que entre más detalles.

La primera fila tiene nombres de columna. Seleccione esto si los nombres de cada columna se incluyen como fila de cabecera en el origen de datos.

Codificación. Especifica el método de codificación de texto utilizado. Puede elegir entre el valor predeterminado del sistema o UTF-8.

Símbolo decimal. Especifica cómo se deben representar los decimales en los datos.

- **Punto (.)**. El carácter de punto se utilizará como separador de decimal.
- **Coma (,)**. El carácter de coma se utilizará como separador de decimal.

Delimitadores. Utilizando las casillas de verificación listadas para este control, puede especificar qué caracteres, como la coma (,), definen límites de campo en el archivo. También puede especificar más de un delimitador como, por ejemplo, " | " para registros que utilizan varios delimitadores. El delimitador predeterminado es la coma.

Nota: Si la coma también está definida como separador de decimal, aquí los valores predeterminados no funcionarán. En los casos donde la coma es a la vez el delimitador de campos y el separador de decimal, seleccione **Otros** en la lista de delimitadores. A continuación, especifique manualmente una coma en el campo de entrada.

Seleccione **Permitir múltiples delimitadores vacíos** para tratar varios caracteres vacíos adyacentes como un único delimitador. Por ejemplo, si un valor de datos viene seguido por cuatro espacios y, después,

otro valor de datos, este grupo se tratará como dos campos, en lugar de cinco.

Opciones avanzadas

Caracteres de comentario de EOL. Especifique caracteres como, por ejemplo, # o !, para indicar anotaciones en los datos. Siempre que aparezca uno de estos caracteres en el archivo de datos, se ignorará todo lo que aparece hasta ese carácter, pero sin incluir el siguiente carácter de nueva línea.

Especificar campos de entrada. Especifique el número de campos de entrada que se van a utilizar en cada registro.

Especifique un formato de datos para los tipos de origen de datos de **archivo** que son del tipo de **campos de longitud variable en texto plano**, puede utilizar esta sección para definir el tipo de almacenamiento de entrada y el formato de cada campo para garantizar que los valores se lean correctamente. Esto es similar a las funciones disponibles en el cliente IBM SPSS Modeler. La opción **Alterar temporalmente** indica si el valor predeterminado se ha alterado temporalmente. Si selecciona **Alterar temporalmente** activa los controles **Almacenamiento** y **Formato de entrada**. Deseleccionar **Alterar temporalmente** cambiará los valores de nuevo a sus valores predeterminados originales. **Formato de entrada** solo se aplica a tipos de almacenamiento reales, de fecha, hora y marca de tiempo.

Omitir caracteres de cabecera. Especifique cuántos caracteres desea ignorar al principio del primer registro.

Líneas que deben escanearse para el tipo. Especifique cuántas líneas escanear para los tipos de datos especificados.

Eliminar espacios anteriores y posteriores. Seleccione descartar espacios anteriores y posteriores en cadenas al importar. Puede elegir eliminar desde la izquierda, la derecha, de ambos lados, o nada.

Caracteres no válidos. Seleccione **Descartar** para eliminar caracteres no válidos del origen de datos. Seleccione **Sustituir con** para sustituir caracteres no válidos con el símbolo especificado (solo un carácter). Los caracteres no válidos son caracteres nulos o cualquier carácter que no existe en el método de codificación especificado.

Comillas. Mediante las listas desplegables se puede especificar cuántas comillas simples y dobles se tratan en la importación. Puede optarse por **Descartar** todas las comillas, **Incluir como texto** incluyéndolas en el valor del campo o **Emparejar y descartar** para emparejar pares de comillas y eliminarlos. Si quedase alguna comilla sin emparejar, daría un mensaje de error. Tanto **Descartar** como **Emparejar y descartar** almacenan el valor del campo (sin comillas) como una cadena.

Nota: Cuando se usa **Emparejar y descartar**, se conservan los espacios. Cuando se usa **Descartar**, los espacios iniciales y finales dentro y fuera de las comillas se eliminan (por ejemplo, ' " ab c" , "d ef " , " gh i " ' quedaría como 'ab c, d ef, gh i'). Cuando se usa **Incluir como texto**, las comillas se tratan como caracteres normales, de modo que los espacios iniciales y finales se eliminan de forma natural.

Especifique campos de entrada. Consulte el tema “Selección de campos de entrada” en la página 43 si desea más información.

Archivos de datos Excel

Al seleccionar un origen de datos Excel, se le pedirá que entre más detalles:

La primera fila tiene nombres de columna. Seleccione esta opción si los nombres de cada columna se incluyen como fila de cabecera en el origen de datos.

Rango con nombre. Le permite seleccionar un rango con nombre de celdas, tal como se ha definido en la hoja cálculo Excel. Si utiliza un rango con nombre, otros valores de hoja de cálculo y datos dejan de aplicarse y, como resultado, se inhabilitan.

Elegir hoja de trabajo. Especifica la hoja de trabajo para importar, ya sea por nombre o índice.

- **Por nombre.** Seleccione el nombre de la hoja de trabajo que desea importar.
- **Por índice.** Especifique el valor de índice para la hoja de trabajo que desea importar, empezando por 0 para la primera hoja de trabajo, 1 para la segunda hoja de trabajo, etc.

Rango en hoja de trabajo. Puede importar datos que empiezan por la primera fila que no está en blanco o con un rango de celdas explícito.

- **El rango comienza en la primera fila que no está en blanco.** Localiza la primera celda que no está en blanco y la utiliza como la esquina superior izquierda del rango de datos.
- **Rango explícito de celdas.** Le permite especificar un rango explícito por fila y columna. Por ejemplo, para especificar el rango de Excel A1:D5, puede entrar A1 en el primer campo y D5 en el segundo (o de forma alternativa, R1C1 y R5C4). Se devuelven todas las filas del rango especificado, incluidas las filas en blanco.

En filas en blanco. Si se encuentra una fila en blanco, puede elegir entre omitir e ignorar la fila o elegir **Devolver filas en blanco** para seguir leyendo todos los datos hasta el final de la hoja de trabajo, incluidas las filas en blanco.

Especifique campos de entrada. Consulte el tema “Selección de campos de entrada” en la página 43 si desea más información.

Archivos de datos IBM SPSS Statistics

Al seleccionar un origen de datos de IBM SPSS Statistics (archivo *.sav* o *.zsav*), se le pedirá que entre más detalles. Si el archivo está protegido con una contraseña, también se le pedirá que entre la contraseña.

Nombres de variable. Seleccione un método para manejar nombres de variable y etiquetas al importar desde un archivo *.sav* o *.zsav* de IBM SPSS Statistics.

- **Leer nombres y etiquetas.** Seleccione leer tanto nombres, como etiquetas de variable; esta es la mejor opción. Las etiquetas se pueden visualizar en gráficos, navegadores de modelo y otros tipos de salida.
- **Leer etiquetas como nombres.** Seleccione leer las etiquetas de variable descriptivas en el archivo *.sav* de IBM SPSS Statistics, en lugar de los nombres de campo abreviados, y utilizar estas etiquetas como nombres de variable.

Valores. Seleccione un método para manejar valores y etiquetas al importar desde un archivo *.sav* o *.zsav* de IBM SPSS Statistics.

- **Leer datos y etiquetas.** Seleccione leer tantos los valores reales, como las etiquetas de valor; esta es la opción predeterminada.
- **Leer etiquetas como datos.** Seleccione si desea utilizar las etiquetas de valor del archivo *.sav* o *.zsav*, en lugar de los códigos numéricos o simbólicos utilizados para representar los valores. Por ejemplo, si selecciona esta opción para datos con un campo de género cuyos valores de 1 y 2 representan realmente *hombre* y *mujer*, respectivamente, se convertirá el campo a una cadena y se importará *hombre* y *mujer* como los valores reales.

Es importante tener en cuenta los valores que faltan en los datos de IBM SPSS Statistics, antes de seleccionar esta opción. Por ejemplo, si un campo numérico utiliza etiquetas solo para valores que faltan (0 = *Sin respuesta*, -99 = *Desconocido*), si selecciona la opción anterior solo se importarán las etiquetas de valor *Sin respuesta* y *Desconocido* y se convertirá el campo a una cadena. En dichos casos, deberá importar los propios valores.

Especifique campos de entrada. Consulte el tema “Selección de campos de entrada” en la página 43 si desea más información.

Origen de base de datos

Cuando seleccione un origen de base de datos, se le pedirá que especifique más detalles:

Base de datos. Seleccione la base de datos que desea utilizar.

Conectar. Seleccione si desea conectarse a la base de datos utilizando credenciales de base de datos o credenciales almacenadas en IBM SPSS Collaboration and Deployment Services Repository. A continuación, especifique el nombre y la contraseña de inicio de sesión de la base de datos, o bien seleccione una credencial en el repositorio y pulse **Conectar** para comprobar que tiene acceso a la base de datos.

Tenga en cuenta que el nombre y la contraseña de la base de datos no se guardan. Así pues, la próxima vez que abra y ejecute el proyecto, se le pedirá que inicie sesión, a menos que el administrador haya configurado el inicio de sesión único (SSO) para las conexiones de IBM SPSS Modeler Server y la base de datos ODBC. Consulte la sección *Consideraciones de seguridad* de la guía de instalación de IBM Analytical Decision Management si desea más información.

Nombre de tabla. Seleccione la tabla de la cual se van a utilizar datos.

Eliminar espacios anteriores y posteriores. Seleccione descartar espacios anteriores y posteriores en cadenas al importar. Puede elegir eliminar desde la izquierda, la derecha, de ambos lados, o nada.

Entrecomillar los nombres de las tablas y las columnas. Especifique si desea que los nombres de tabla y columna aparezcan entrecomillados cuando se envían consultas a la base de datos (si, por ejemplo, contienen espacios o signos de puntuación).

- La opción **Cuando sea necesario** entrecomillará los nombres de tabla y columna *solo* si incluyen caracteres no estándar. Los caracteres no estándar incluyen caracteres no ASCII, caracteres de espacio y cualquier carácter no alfanumérico distinto a un punto final (.).
- Seleccione **Siempre** si desea que *todos* los nombres de tabla y columna aparezcan entre comillas.
- Seleccione **Nunca** si *no* desea que los nombres de tabla y columna aparezcan entre comillas.

Especifique campos de entrada. Consulte el tema “Selección de campos de entrada” en la página 43 si desea más información.

IBM Analytical Decision Management soporta los mismos orígenes de datos de base de datos que IBM SPSS Modeler para leer, escribir y distribuir datos. Si desea más detalles, consulte el informe de compatibilidad del producto para IBM SPSS Modeler 18. Tenga en cuenta que es diferente respecto a las bases de datos soportadas por IBM SPSS Collaboration and Deployment Services 8, que es la base de datos que aloja el repositorio.

Origen de la vista de datos

Un origen de vista de datos crea una conexión que se va a utilizar con reglas de IBM Operational Decision Management. Para reglas complejas que utilizan un modelo de datos jerárquico, tenga en cuenta que los administradores pueden restringir el uso de dichas reglas porque pueden degradar el rendimiento. Si desea más detalles, consulte “Bloqueo de opciones de datos” en la página 83.

Un *nodo de vista de datos* se puede configurar en IBM SPSS Modeler para incluir datos en rutas que se han definido en una vista de datos analíticos de IBM SPSS Collaboration and Deployment Services. Una vista de datos analíticos define una estructura para acceder a datos que describen las entidades utilizadas en modelos predictivos y reglas empresariales. La vista asocia la estructura de datos con orígenes de datos físicos para el análisis.

Si desea más información, consulte la documentación de IBM SPSS Modeler y la sección *Nodos de origen, proceso y salida* de la ayuda de IBM SPSS Modeler.

Cuando seleccione un origen de vista de datos, se le pedirá que entre más detalles:

Origen de datos analíticos. Seleccione el origen de datos analíticos en el repositorio.

Nombre de tabla. Seleccione la tabla en la que se van a utilizar los datos. Esta es una tabla del modelo de datos en la vista de datos analíticos. Cada tabla del modelo de datos representa un concepto o entidad implicada en el proceso de análisis predictivo. Los campos para las tablas corresponden a los atributos de las entidades representadas por las tablas. Por ejemplo, si está analizando pedidos de cliente, su modelo de datos podría incluir una tabla para clientes y una tabla para pedidos. La tabla de clientes podría tener campos para el identificador de cliente, edad, género, estado civil y país de residencia. La tabla de pedidos podría tener campos para el identificador del pedido, número de artículos del pedido, coste total y el identificador para el cliente que ha realizado el pedido. El identificador de cliente se podría utilizar para asociar los clientes de la tabla de clientes con sus pedidos en la tabla de pedidos.

Plan de acceso de datos. Seleccione el plan de acceso de datos a partir del cual se van a utilizar los datos. Un plan de acceso de datos asocia las tablas del modelo de datos en una vista de datos analíticos con orígenes de datos físicos. Una vista de datos analíticos normalmente contiene varios planes de acceso de datos. Cuando cambie el plan de acceso de datos en uso, cambie los datos utilizados por el proyecto. Por ejemplo, si la vista de datos analíticos contiene un plan de acceso de datos para la formación de un modelo y un plan de acceso de datos para probar un modelo, puede cambiar de datos de formación a datos de prueba cambiando el plan de acceso de datos en uso.

Atributos opcionales. Si están disponibles y lo desea, seleccione atributos opcionales para añadir al origen de datos.

Especifique campos de entrada. Consulte el tema “Selección de campos de entrada” en la página 43 si desea más información.

Origen de IBM Cognos BI

Puede importar datos de la base de datos de IBM Cognos BI o informes de una sola lista para su uso en IBM Analytical Decision Management.

Nota: Si IBM SPSS Modeler Server en uso con IBM Analytical Decision Management no soporta el servidor IBM Cognos BI, esta opción no está disponible.

Cuando se selecciona **Servidor IBM Cognos BI** y se pulsa **Conectar** se le pedirá que entre los detalles de conexión de servidor.

1. Escribe el **URL de servidor** del servidor IBM Cognos desde el cual importar o exportar datos. Póngase en contacto con el administrador del sistema Cognos si no sabe qué URL utilizar.
2. Seleccione el **Modo** en el cual se conectará. Para iniciar una sesión como un usuario específico (por ejemplo, un administrador), seleccione **Establecer credenciales** y entre el espacio de nombre, el nombre de usuario y la contraseña de IBM Cognos:
 - Entre el **Espacio de nombres** del proveedor de la autenticación de seguridad de IBM Cognos utilizado para iniciar sesión en el servidor. El proveedor de autenticación se utiliza para definir y mantener usuarios, grupos y roles y para controlar el proceso de autenticación.
 - Entre el **Nombre de usuario** de IBM Cognos con el que iniciará una sesión en el servidor.
 - Entre la **Contraseña** asociada al nombre de usuario especificado.

O

Examine y seleccione una credencial de IBM Cognos almacenada en IBM SPSS Collaboration and Deployment Services Repository.

3. De forma alternativa, para iniciar una sesión sin credenciales de usuario, en cuyo caso no podrá especificar los campos de espacio de nombres, nombre de usuario y contraseña, seleccione **Utilizar conexión anónima**.

Nota: Algunas conexiones de servidor no permiten conexiones anónimas.

Selección de detalles de objeto Cognos

Tras configurar la conexión del servidor IBM Cognos, seleccione el modo. El modo **Datos** le permite examinar y seleccionar un paquete del cual importar datos. El modo **Informe de lista** le permite examinar y seleccionar un informe de lista predefinido del cual importar datos.

Un paquete contiene un modelo de IBM Cognos y todas las carpetas, consultas, informes, vistas, atajos, URL y definiciones de trabajo asociados a dicho modelo. Un modelo de IBM Cognos define reglas empresariales, descripciones de datos, relaciones de datos, dimensiones empresariales y jerarquías y otras tareas administrativas.

1. Para seleccionar un paquete, seleccione el modo **Datos** y pulse **Examinar**. Se muestra un cuadro de diálogo que lista todos los paquetes disponibles de los que podrá importar contenido.
2. O bien seleccione un informe de lista única, seleccione el modo **Informe de lista** y pulse **Examinar**. Se muestra un cuadro de diálogo que lista paquetes y carpetas desde donde podrá seleccionar un informe. *Nota:* Solo se pueden informar informes de lista única; no son compatibles varias listas.
3. Cuando sea necesario, expanda la lista para encontrar el paquete o informe que necesita. Seleccione el elemento y pulse **Aceptar**. Se muestran los detalles del paquete o informe en el editor de orígenes de datos.
4. Una vez seleccionado un paquete, podrá elegir los objetos para importar. Por ejemplo, seleccione los temas de consulta del objeto del paquete (que representan tablas de base de datos) o temas de consulta de campo individual (que representan columnas de tabla) para utilizar desde un paquete. *Nota:* Los datos deben estar en formato UTF-8.
5. Seleccione la opción **Agregar datos antes de importar** si desea importar datos agregados, en lugar de datos en bruto. Por regla general, los datos agregados tiene muchos menos registros que los datos en bruto.

Seleccionar objetos. Muestra los objetos (también conocidos como espacios de nombre) que puede seleccionar en el informe de paquete o lista. Expanda uno o más objetos para mostrar los campos de entrada que puede importar. Si selecciona un objeto, se seleccionan automáticamente todos sus campos de entrada.

Revisar parámetros. Si se han definido parámetros, ya sea para un objeto de datos o un informe, puede especificar valores para estos parámetros antes de importar el objeto o informe. Esta tabla contiene la lista de parámetros necesarios para filtros y cálculos. Estos pueden tener valores predeterminados. Se proporciona una advertencia si uno o varios parámetros no tienen un valor predeterminado. Defina parámetros según sea necesario en la columna **Valor**. Puede utilizar parámetros para realizar ajustes (por ejemplo, para realizar un cálculo parametrizado) antes de importar los datos.

Revisar filtros. Si un paquete tiene filtros definidos, puede importar uno o varios de ellos. Si un filtro que importa está asociado a datos importados, dicho filtro se aplica antes de que se importen los datos. Los filtros pueden estar parametrizados o no.

Revisar campos de entrada. Lista los campos de entrada (también conocidos como objetos de base de datos) que se importarán en IBM Analytical Decision Management para el procesamiento. Si no requiere un campo particular, deselectione la casilla de verificación junto a su nombre de campo.

Iconos de objeto Cognos: Los distintos tipos de objeto que puede importar desde una base de datos Cognos BI se representan mediante distintos iconos, como ilustra la tabla siguiente.

Tabla 2. Iconos de objeto Cognos.

Icono	Objeto
	Paquete
	Espacio de nombres
	Tema de consulta
	Elemento de consulta
	Dimensión de medida
	Medida
	Dimension
	Jerarquía de niveles
	Nivel
	Filtro
	Informe
	Cálculo autónomo

Origen de Analytic Server

Cuando se crea un origen de datos, no se puede seleccionar el tipo del origen de datos de Analytic Server. Solo se usa a efectos de visualización, cuando se usa IBM SPSS Modeler Advantage para abrir un modelo creado en IBM SPSS Modeler contra un origen de datos de Analytic Server.

IBM SPSS Analytic Server se introdujo en IBM SPSS Modeler 16, permitiendo la utilización de un sistema de archivos distribuido Hadoop (HDFS) como origen de datos. Los usuarios de IBM SPSS Modeler pueden crear corrientes usando el nodo de origen HDFS de Analytic Server y construir modelos basados en el sistema de big data Hadoop. Los usuarios de IBM Analytical Decision Management pueden abrir estas corrientes de modelo en in IBM SPSS Modeler Advantage (no se pueden crear modelos usando orígenes de datos de Analytic Server). Tenga en cuenta que no se pueden realizar determinadas acciones como, por ejemplo, cambiar el destino o tipo de modelo.

Al visualizar un origen de datos de Analytic Server en la pestaña Datos, aparece la siguiente información:

Nombre del origen de datos. Nombre dado al origen de datos en IBM SPSS Modeler. Puede cambiar este nombre, si así lo desea. Los campos restantes no son editables.

Origen de Analytic Server. Ruta al origen de datos tal como se ha configurado en IBM SPSS Modeler.

Campos de entrada especificados. Los campos de entrada utilizados.

Selección de campos de entrada

Cuando haya una gran cantidad de datos, es posible que tenga que reducir o ajustar los datos importados del origen (por ejemplo, para utilizar solo los campos que son relevantes para su área de interés).

1. En la parte inferior del diálogo Editor de origen de datos, pulse **Especificar campos de entrada**; el diálogo se amplía para mostrar todos los campos de datos en el origen. Además, todas las expresiones que cree también se visualizan. Si desea detalles sobre cómo crear expresiones, consulte “Creador de expresiones” en la página 47.
2. Seleccione cada uno de los campos y expresiones que desee utilizar.

Además de seleccionar los datos que puede explorar, donde sea necesario, cambie el nivel de medición para uno o varios de los campos de datos. Si desea más información, consulte “Niveles de medición” en la página 44.

Vista previa de datos

Para verificar que ha seleccionado el origen de datos correcto, pulse el icono **Vista previa** para mostrar una muestra de los datos incluidos en el origen.

Visión general de datos

La característica de visión general de datos está disponible en la pestaña Datos de todas las aplicaciones, en la pestaña Modelado y la pestaña Puntuación en IBM SPSS Modeler Advantage y en la pestaña Desplegar de las aplicaciones que incluyen la característica Puntuar ahora.

1. Para ejecutar una visión general de un origen de datos, pulse el icono Visión general de datos disponible en toda la aplicación.



Figura 11. Icono Visión general de datos

2. Aparecerá el diálogo Visión general de datos. Si lo desea, seleccione un campo de superposición en el desplegable de la sección Opciones de visión general.
A continuación, tras ejecutar la visión general, las pestañas estarán disponibles para mostrar resultados solo para el campo seleccionado principal, o para solapar el campo seleccionado principal con el campo especificado en este desplegable. Por ejemplo, en los resultados es posible que desee ver información sobre la **edad** de los clientes y, después, solaparla con otro campo como, por ejemplo, **Género**.
3. Si lo desea, seleccione **Utilizar datos particionados** en la sección Opciones de visión general. Esta opción está disponible si, en Valores opcionales de la pestaña Modelado, ha seleccionado la opción **Partir datos automáticamente para habilitar la evaluación del modelo en el origen de datos de generación para su evaluación y prueba**. Esta opción divide los datos en subconjuntos o muestras individuales para la formación o las pruebas del modelo. Mediante la generación del modelo en un subconjunto y probarlo en otro, podrá tener una idea sobre cómo se generalizará en otros conjuntos de datos.
4. Seleccione los campos de datos para incluir en la visión general y pulse **Ejecutar visión general**. Se listan todos los campos disponibles en el origen de datos. De forma predeterminada, todos los campos de entrada de modelo y el objetivo están seleccionados.
La visión general de datos se ejecutará y aparecerán los resultados. Puede ordenar la información y elegir qué columnas mostrar.
5. Pulse cualquier campo para ver sus detalles. Se abrirá una nueva pestaña de resultados para cada campo que seleccione, lo que le permite ver gráficos y tablas y seleccionar campos de solapamiento, si está disponible.

Niveles de medición

El nivel de medición ayuda a caracterizar la información representada por cada campo de datos, y puede determinar cómo se utiliza un campo determinado en reglas, modelado u otras aplicaciones. El nivel de medición se puede especificar en el modelo de datos de proyecto para un origen de datos; por ejemplo, es posible que desee definir el nivel de medición para un campo de entero con valores de 0 y 1 en *Marca*, para indicar que 0 = *Falso* y 1 = *Verdadero*. De forma alternativa, puede cambiar el nivel en el diálogo **Editor de orígenes de datos** al especificar los campos de entrada que se van a utilizar. Si desea más información, consulte “Selección de campos de entrada” en la página 43.

Los niveles de medición siguientes están disponibles:

- **Continuo**, se utiliza para describir valores numéricos como, por ejemplo, un rango de 0–100 o 0,75–1,25. Un valor continuo puede ser un entero, número real o fecha/hora.
- **Categorico**, se utiliza para valores de cadena cuando se desconoce un número exacto de valores distintos. Esto es un tipo de datos **sin instanciar**, lo que significa que todavía no se conoce toda la información posible sobre el almacenamiento y el uso de los datos. Una vez leídos los datos, el nivel de medición será *Marca*, *Nominal* o *Sin tipo*, en función del número máximo de miembros para campos nominales especificados en el cuadro de diálogo Propiedades de proyecto.
- **Marca**, se utiliza para datos con dos valores distintos que indican la presencia o ausencia de un rasgo, como true y false, Yes y No o 0 y 1. En IBM Analytical Decision Management, el primer valor de la lista de categorías se utilizará como el valor "falso" y el segundo será el valor "verdadero", tal como se indica en una ayuda contextual al pasar el ratón por encima de los valores cuando se visualizan en la pestaña Datos. Los datos se pueden representar como texto, entero, número real, fecha, hora o marca de tiempo.
- **Nominal**, se utiliza para describir datos con varios valores distintos, cada uno se trata como un miembro de un conjunto, como pequeño/medio/grande. Los datos nominales pueden tener cualquier valor numérico de almacenamiento, cadena o fecha/hora. Tenga en cuenta que definir el nivel de medición en *Nominal* no cambia automáticamente los valores a un almacenamiento de cadena. Si desea más información sobre cómo definir el máximo de miembros permitidos para campos nominales, consulte “Propiedades” en la página 31.
- **Ordinal**, se utiliza para describir datos con varios valores distintos que tienen un orden inherente. Por ejemplo, las categorías salariales o clasificaciones de satisfacción se pueden escribir como datos ordinales. El orden se define mediante el orden de clasificación natural de los elementos de datos. Por ejemplo, 1, 3, 5 es el orden de clasificación predeterminado para un conjunto de enteros, mientras que ALTO, BAJO, NORMAL (orden alfabético ascendente) es el orden para un conjunto de cadenas. El nivel de medición ordinal le permite definir un conjunto de datos categoricos como datos ordinales para fines de visualización, generación de modelos y la exportación a otras aplicaciones (como IBM SPSS Statistics) que reconocen datos ordinales como un tipo distinto. Puede utilizar un campo ordinal en cualquier punto donde se pueda utilizar un campo nominal. Además, los campos de cualquier tipo de almacenamiento (real, entero, cadena, fecha, hora, etc.) se pueden definir como ordinal.
- **Sin tipo**, se utiliza para datos que no se ajustan a ninguno de los tipos anteriores, para campos con un solo valor, o para datos nominales donde el conjunto tiene más miembros que el máximo definido. También es útil para casos en los cuales, de lo contrario, el nivel de medición sería un conjunto con muchos miembros (como un número de cuenta). Cuando se selecciona **Sin tipo** para un campo, el rol se establece automáticamente en **Ninguno**, con el **ID de registro** como única alternativa. El tamaño máximo predeterminado para conjuntos es 250 valores exclusivos. Este número se puede ajustar o inhabilitar en el cuadro de diálogo Propiedades de proyecto, al que se puede acceder desde el icono de barra de herramientas.
- **Colección**, se utiliza para identificar datos no geoespaciales que se registran en una lista. Una colección es realmente un campo de lista de profundidad cero, donde los elementos de dicha lista tienen uno de los otros niveles de medición.
- **Geoespacial**, se utiliza con el tipo de almacenamiento de lista para identificar datos geoespaciales. Las listas pueden ser campos Listas de números enteros o Listas de números reales indistintamente con una profundidad de lista que está entre cero y dos, inclusives.

Edición de valores de campo de datos

En tablas de datos, la columna **Valores** lista los valores para cada campo del origen de datos (por ejemplo, los valores máximo y mínimo de un rango). Para campos de marca, mantenga el puntero del ratón encima del valor para mostrar una ayuda contextual que indica los valores "verdadero" y "falso" definidos para el campo. También puede pulsar un valor para editarlo. Se abre un cuadro de diálogo y su contenido varía en función del tipo de campo de datos seleccionado. Por ejemplo:

- Si el tipo de campo es **Continuo**, puede entrar los valores del límite superior e inferior.
- Si el tipo de campo es **Marca**, puede intercambiar los valores verdadero y falso.
- Si el tipo de campo es **Nominal**, puede añadir, eliminar, reordenar y renombrar valores.

Campos de correlaciones

Al añadir un nuevo origen de datos, es posible que tenga que correlacionar todos sus campos para garantizar la compatibilidad con el modelo de datos de proyecto. Por ejemplo, si el modelo de datos del proyecto requiere un campo llamado *compra* con los valores *Sí* y *No* (marca de nivel de medición), cualquier origen de datos utilizado debe tener un campo compatible. Si los nombres del campo no son idénticos, se pueden correlacionar en consecuencia. Tenga en cuenta que el campo de entrada y el campo correlacionado asociado deben tener el mismo tipo de datos.

De forma alternativa, si el modelo de datos del proyecto cambia, puede suceder que muchos orígenes de datos dejen de estar correlacionados correctamente. Si este es el caso, la bola del indicador de compatibilidad del área **Orígenes de datos de proyecto** cambiará al color apropiado para cada uno de estos orígenes de datos; entonces, tendrá que volver a correlacionarlos todos con el nuevo modelo de datos de proyecto.

1. En el área **Orígenes de datos de proyecto**, pulse **Correlacionar campos**. Se abre un diálogo y el sistema realiza la mejor predicción automáticamente en los campos de correlación. En cualquier punto, puede volver a la mejor predicción del sistema pulsando **Restablecer**.
2. Seleccione un campo en la columna **Entradas disponibles** y su campo correspondiente en la columna **Campos de entrada obligatorios** y pulse **Correlacionar** para correlacionar los dos campos. Repita esto hasta que se hayan correlacionado todos los campos en la columna **Entradas disponibles**.
3. Cuando haya acabado, pulse **Aceptar**.

Cuando sea posible, IBM Analytical Decision Management correlacionará automáticamente los tipos de almacenamiento compatibles cuando un usuario añade un origen de datos adicional. Si el tipo no es compatible, el usuario estará autorizado para convertir tipos.

La tabla siguiente lista las correlaciones permitidas.

Tabla 3. Tipos que se pueden convertir

A →						
De ↓	Entero	Real	Serie	Fecha	Hora	Marca de tiempo
Entero	n/d	Auto	Sí	No permitido	No permitido	No permitido
real	Sí (dígitos decimales truncados)	n/d	Sí	No permitido	No permitido	No permitido
Serie	Sí	Sí	n/d	Sí	Sí	Sí
Fecha	No permitido	No permitido	Sí	n/d	No permitido	Sí (hora como 00:00)
Hora	No permitido	No permitido	Sí	No permitido	n/d	Sí (fecha base)

Tabla 3. Tipos que se pueden convertir (continuación)

A →						
De ↓	Entero	Real	Serie	Fecha	Hora	Marca de tiempo
Marca de tiempo	No permitido	No permitido	Sí	Auto (hora truncada)	Auto (fecha truncada)	n/d

Derivación de campos

En la mayoría de las aplicaciones de IBM Analytical Decision Management, puede añadir campos adicionales al modelo de datos de proyecto como expresiones, salida de modelo, campos creados a partir de reglas de segmentación, o campos de tablas secundarias. La pestaña **Derivado** de la pestaña Datos muestra los campos que no se van a incluir como parte de los datos, y los campos se pueden ordenar de forma que se puedan utilizar de forma acumulativa (su orden en la tabla coincidirá con el orden en el cual se calculan; por ejemplo, el segundo campo derivado puede utilizar el primer campo derivado como entrada).

La pestaña **Derivado** no está incluida en la aplicación de IBM SPSS Rules Management.

- En la pestaña Datos, pulse la pestaña **Derivado**.

Añadir una expresión:

- Pulse el icono **Derivar campo utilizando una expresión**. Se visualizará el creador de expresiones. Si desea más información, consulte "Creador de expresiones" en la página 47.

Añadir una salida del modelo:

1. Pulse el icono **Derivar campo utilizando un modelo o regla** y, después, examine y seleccione un modelo. Se visualiza el diálogo **Entradas y salidas** para el modelo seleccionado.
2. Si los campos de entrada de modelo son compatibles, la sección **Asignación de entrada de modelo** del diálogo se contraerá. Si los campos de entrada de modelo no son compatibles, la sección se ampliará y deberá asignar entradas de tabla disponibles a entradas del modelo seleccionado.
3. En la sección **Salidas de modelo**, se listarán todas las salidas de modelo disponibles para el modelo seleccionado. Seleccione los campos de salida del modelo que desea añadir a los datos. La columna **Nombre de salida del modelo** muestra el nombre del campo de salida como lo genera el modelo.

Se puede seleccionar cualquier campo de la salida de modelo. Los campos de salida del modelo son todos los campos que proceden de la ruta que tampoco son entradas para la ruta (todos los campos añadidos por la ruta). Los campos de entrada pueden no incluir todo lo que aparece en el modelo de datos de proyecto, así que es posible que un campo de salida de modelo tenga el mismo nombre que un campo de modelo de datos de proyecto. Así puede, también puede cambiar los nombres de campo de datos aquí, si lo desea.

La columna **Medición** muestra el tipo de medición de la salida del modelo que tiene como prefijo el icono de tipo apropiado, y la columna **Valores** muestra los valores de la salida del modelo. Tenga en cuenta que estas columnas pueden mostrar resultados para algunos de los campos y no otros.

Normalmente, los tipos y valores de medición se conocen para las salidas de modelo reales (valor pronosticado, probabilidad, etc.) sin una exploración de datos, por lo tanto, si se conocen, se mostrarán. Pero los tipos o valores de medición pueden desconocerse para otros campos de salida para una ruta (por ejemplo, un campo derivado) sin una exploración de datos.

4. Se incluirá una sección **Parámetros de modelo** en la parte inferior del diálogo, si la ruta seleccionada incluye parámetros de puntuación. Entre valores para los parámetros del modelo.
5. Pulse **Guardar** cuando haya acabado en el diálogo. Se añadirán las salidas de modelo seleccionadas a la sección **Campos adicionales** de la pestaña Derivado.

Crear un campo utilizando reglas de segmentación:

1. Pulse el icono **Derivar campo utilizando reglas de segmentación**.
2. Cree una nueva regla de segmentación. Los segmentos se definen mediante una o varias expresiones que devuelven un resultado verdadero o falso. Si desea más información, consulte “Definición de reglas de segmentación” en la página 51.
3. Pulse **Aceptar** cuando haya acabado en el diálogo. La nueva regla de segmentación se añadirá a la sección **Campos adicionales** de la pestaña Derivado.

Añadir un campo de una tabla secundaria:

1. Si ha definido una tabla secundaria con una relación unívoca en la pestaña Datos, pulse el icono **Añadir campo de una tabla secundaria**.
2. Seleccione una tabla secundaria en el desplegable y, después, seleccione uno o más campos de tabla.
3. Pulse **Guardar** cuando haya acabado en el diálogo. Los campos se añadirán a la sección **Campos adicionales** de la pestaña Derivada.

Crear de expresiones

El creador de expresiones se puede usar para derivar campos adicionales o atributos que se usen en una aplicación. Por ejemplo, si se están utilizando datos bancarios, puede que desee crear una expresión que muestre la proporción entre los ingresos de un cliente y el número de cuentas de préstamo que tienen.

Las expresiones están disponibles en la aplicación de la misma forma que los datos operativos (por ejemplo, en reglas y en predictores de modelo).

Para mostrar las expresiones existentes o crear una expresión, vaya a la pestaña Datos y pulse en la pestaña **Derivado**.

Creación de una expresión

1. En el diálogo del creador de expresiones, especifique un nombre exclusivo para la expresión.
2. Escriba en el campo **Ecuación de expresión**, utilizando las listas de campos y funciones como referencias
o bien
3. Seleccione en las listas los campos y funciones que desee.
4. Pulse dos veces o pulse en la flecha azul para añadir el campo o la función al campo **Ecuación de expresión**.
5. Utilice los botones de operando en el centro del diálogo para insertar operaciones en la expresión.
6. Especifique una **Descripción** que explique para qué se utiliza la expresión.
7. Pulse **Comprobar** para validar la expresión y luego pulse **Guardar**.

Nota: Se pueden copiar y pegar expresiones desde el cliente de IBM SPSS Modeler, pero tenga en cuenta que los resultados pueden no ser siempre los esperados si la expresión depende de una ruta (por ejemplo, si se pega una función @ soportada o una función de fila, puede que no arroje los resultados deseados, porque puede que el resto de la ruta no esté configurado para soportarla).

Las únicas expresiones que están soportadas oficialmente son las que se proporcionan en la interfaz de usuario de IBM Analytical Decision Management.

Selección de funciones

La lista de funciones muestra todos los operadores y funciones de disponibles. Desplace para seleccionar una función de la lista o, para facilitar la búsqueda, use una lista desplegable para visualizar un subconjunto de funciones u operadores. Las funciones disponibles se agrupan en categorías para facilitar su búsqueda.

A continuación se muestran los detalles de las categorías. Si desea información más detallada, consulte la documentación de IBM SPSS Modeler.

Tabla 4. Categorías de función

Tipo de función	Descripción
Funciones usadas recientemente	Contiene una lista de las funciones utilizadas durante la sesión actual.
Operadores	Lista todos los operadores que se pueden utilizar al generar expresiones. Los operadores también están disponibles en los botones del centro del diálogo
Funciones de información	Se utiliza para comprender mejor los valores de campo. Por ejemplo, la función <code>is_string</code> devuelve un valor verdadero para todos los registros cuyo tipo es una serie.
Funciones de conversión	Se utiliza para construir nuevos campos o convertir el tipo de almacenamiento. Por ejemplo, la función <code>to_timestamp</code> convierte el campo seleccionado a una marca de hora.
Funciones de comparación	Se utiliza para comparar valores de campo entre sí o con una serie especificada. Por ejemplo, <code><=</code> se utiliza para comparar si los valores de dos campos son menores o iguales.
Funciones lógicas	Se utiliza para realizar operaciones lógicas como, por ejemplo, las operaciones <code>if</code> , <code>then</code> , <code>else</code> .
Funciones numéricas	Se utiliza para realizar cálculos numéricos como, por ejemplo, el logaritmo natural de valores de campo.
Funciones aleatorias	Se utiliza para seleccionar elementos o generar números aleatoriamente.
Funciones de cadena	Se utiliza para realizar distintas operaciones en series como, por ejemplo, <code>stripchar</code> , que le permite eliminar un carácter especificado.
Funciones de fecha y hora	Se utiliza para realizar distintas operaciones en los campos de fecha, hora y marca de hora.
Funciones de espacios en blanco y nulos	Se utiliza para acceder, marcar y, con frecuencia, llenar los espacios en blanco especificados por el usuario o los valores que faltan en el sistema. <code>@BLANK</code> no está soportada porque su comportamiento depende de otros nodos como, por ejemplo, el nodo Tipo. <code>@NULL(CAMPO)</code> y <code>undef</code> están soportadas. <code>@NULL(CAMPO)</code> es de tipo booleano y devuelve verdadero si el valor de CAMPO es el ausente en el sistema <code>\$null\$</code> . Devuelve falso para todos los demás valores, incluyendo los vacíos definidos por el usuario. Si desea comprobar ambos, use <code>@BLANK(CAMPO)</code> y <code>@NULL(CAMPO)</code> . <code>undef</code> es de tipo Cualquiera y se suele emplear para introducir un valor <code>\$null\$</code> (por ejemplo, para rellenar valores en blanco con nulos en el nodo Rellenar).
Todas las funciones	Contiene una lista completa de las funciones disponibles.

Selección de campos y valores

La lista de campos muestra todos los campos disponibles. Desplácese para seleccionar un campo en la lista. Pulse dos veces o pulse la flecha azul para añadir un campo a la expresión.

Visualización y selección de valores

Los valores de campo se pueden visualizar desde una serie de lugares de la aplicación. Para ver valores de un campo desde el creador de expresiones, seleccione el campo deseado y pulse el botón de entradas de expresión para abrir un cuadro de diálogo que lista valores para el campo seleccionado. A continuación, puede seleccionar un valor y pulse **Insertar** para pegar el valor en la expresión actual.



Figura 12. Botón de entradas de expresión

Para campos de marca y nominales, se listan todos los valores definidos. Para campos continuos (rango numérico), se visualizan los valores mínimo y máximo.

Comprobación de expresiones

Pulse **Comprobar** en el creador de expresiones (esquina inferior derecha) para validar la expresión. Las expresiones que no se han comprobado se visualizan en rojo. Si se encuentran errores, se muestra un mensaje que indica la causa.

Se comprueban los elementos siguientes:

- Entrecorillado correcto de valores y nombres de campo
- Uso correcto de parámetros y variables globales
- Uso válido de operadores
- Existencia de campos de referencia
- Existencia y definición de globales referenciados

Si encuentra errores en la sintaxis, intente crear la expresión utilizando las listas y los botones de operador, en lugar de escribiendo la expresión manualmente. Este método añade automáticamente las comillas correctas para campos y valores.

Tenga en cuenta las limitaciones siguientes al generar expresiones en IBM Analytical Decision Management. Las expresiones no pueden contener ninguno de los elementos siguientes:

- Una referencia a un parámetro de ruta de IBM SPSS Modeler
- Una referencia a un global de ruta de IBM SPSS Modeler
- Una referencia a una función de base de datos
- Una referencia a unas de las funciones siguientes de campo especial o valor de campo @:
 - @TARGET
 - @PREDICTED
 - @FIELD
 - @PARTITION_FIELD
 - @TRAINING_PARTITION
 - @TESTING_PARTITION
 - @VALIDATION_PARTITION

Uso de expresiones que devuelven un booleano

IBM SPSS Modeler y IBM Analytical Decision Management no tienen un tipo de datos booleano. Una expresión booleana no se puede utilizar para generar directamente un campo nuevo. La utilización de expresiones booleanas se limita a expresiones que esperan una expresión booleana (como, por ejemplo, la sentencia `if`). A continuación se muestran tres ejemplos:

```
if ('RFM Score' > 5) then "T" else "F" endif
```

```
if ('RFM Score' < 5) then "Low" elseif ('RFM Score' >= 5 and 'RFM Score' < 10) then "Medium" else "High" endif  
hassubstring('Phone type','Coolphone')
```

En los dos primeros ejemplos anteriores, las expresiones "RFM Score" (`'RFM Score' > 5`, `'RFM Score' >= 5` y `'RFM Score' < 10`) pueden entenderse como booleanas.

Dichas expresiones no pueden usarse en sí mismas para derivar un campo. Daría un error similar a Tipo no válido para una expresión que no es condición: Booleano. Sin embargo, expresiones como estas se pueden usar en expresiones condicionales, como las de los ejemplos completos, y pueden emplearse expresiones condicionales para derivar campos con medidas categóricas como, por ejemplo, un campo de tipo distintivo. La medida predeterminada de un campo con dos valores es distintivo.

El segundo de los ejemplos anteriores podría generar un campo con tres valores distintos (siempre que se encuentren registros que pertenezcan a cada una de las categorías Low, Medium y High durante la exploración de los datos). Probablemente esto daría lugar a que el nivel de medida del campo tomase un valor predeterminado de nominal. No es probable que el tercero de los ejemplos anteriores devuelva un campo con dos valores distintos.

Consulte “Niveles de medición” en la página 44 para obtener más información sobre los niveles de medición.

Capítulo 5. Trabajar con reglas

Las reglas se utilizan para automatizar decisiones basándose en la lógica empresarial, o en la salida de modelos predictivos. En función de la aplicación, podría utilizar reglas para excluir clientes basándose en el historial de pagos, para consultar reclamaciones de riesgo con un unidad de investigaciones especial, o para dirigirse a aquellos clientes con mayor propensión a abandonar, comprar o cometer fraude, tal como se identifica mediante un modelo predictivo. Mientras que las reglas permiten la automatización, los modelos proporcionan la capacidad para predecir, y la capacidad para combinarlos es una de las características clave que distinguen las aplicaciones predictivas de otras herramientas de análisis.

Las reglas se pueden crear, aplicar y guardar en distintos puntos dentro de cada aplicación. De forma alternativa, puede crear y editar reglas compartidas utilizando IBM SPSS Rules Management.

Creación de reglas

- En el punto de la aplicación donde desea añadir una regla, pulse el icono **Crear una nueva regla**.
- De forma alternativa, pulse el icono **Buscar una regla existente** para examinar las reglas existentes que se pueden añadir o reutilizar en la aplicación actual. Consulte el tema “Exportación y reutilización de reglas” en la página 54 si desea más información.
- Para crear una regla compartida, inicie IBM SPSS Rules Management desde la página de inicio de la aplicación.

Definición de reglas de segmentación

Las reglas de segmentación se utilizan en toda la aplicación para seleccionar y asignar registros en distintas aplicaciones, según corresponda. Los segmentos se definen mediante una o más expresiones que devuelven un resultado de verdadero o falso. Por ejemplo, se podría tener una regla con un único segmento que devuelva verdadero si el valor del campo edad es menor que 18. Los segmentos se pueden definir a partir de campos o expresiones definidos en los datos o en las puntuaciones de propensión o predicciones devueltas por un modelo. En la aplicaciones con una pestaña Definir configuradas para tener varias jerarquías de dimensión (conocidas como *elementos*), también se puede elegir entre otros miembros de dimensión.

Cada segmento se define mediante una o varias expresiones y devuelve un valor de verdadero si se cumplen todas las condiciones. Por ejemplo, un segmento definido por la edad y el género solo incluiría los registros que cumplieran ambas condiciones, y la regla devolvería *falso* para todos los demás.

Cómo crear o editar una regla de segmentación

1. En la aplicación donde desee añadir la regla, pulse en el icono **Crear una regla**.
2. Especifique el nombre de la regla.
3. En la columna de entrada, pulse en el icono de flecha hacia abajo para especificar un campo, una regla o un modelo, según se desee.

Tabla 5. Opciones de la columna de entrada.

Icono	Uso
	Si está habilitada, permite seleccionar un miembro de dimensión de un elemento definido previamente.
	Permite elegir un campo en el conjunto de datos actuales.
	Permite elegir un modelo o una regla existente en el repositorio.

Tabla 5. Opciones de la columna de entrada (continuación).

Icono	Uso
	Permite crear un modelo.
	Permite crear una expresión.

4. Elija la función que desee utilizar. Consulte la tabla siguiente para ver las opciones disponibles.

Tabla 6. Operadores disponibles.

Tipo de valor	Operadores disponibles
Indicación de fecha y hora, Fecha, Fecha y hora	<ul style="list-style-type: none"> • = (igual a) • != (distinto de) • > (mayor que) • < (menor que) • >= (mayor o igual que) • <= (menor o igual que) • BETWEEN (entre) • ISNIL (es nulo) • IS NOT NIL (no es nulo)
Un rango	<ul style="list-style-type: none"> • = (igual a) • != (distinto de) • > (mayor que) • < (menor que) • >= (mayor o igual que) • <= (menor o igual que) • BETWEEN (entre) • ISNIL (es nulo) • IS NOT NIL (no es nulo) • ONE OF (uno de) • NOT ONE OF (no es uno de)
Un conjunto	<ul style="list-style-type: none"> • = (igual a) • != (distinto de) • ISNIL (es nulo) • IS NOT NIL (no es nulo) • ONE OF (uno de) • NOT ONE OF (no es uno de)
Un distintivo	<ul style="list-style-type: none"> • = (igual a) • != (distinto de) • ISNIL (es nulo) • IS NOT NIL (no es nulo)

5. Especifique un valor según convenga para el tipo de entrada. Para los valores no numéricos se proporciona una lista desplegable de opciones.

De forma alternativa, pulse en el icono **Buscar una regla existente** para examinar las reglas ya definidas. Consulte el tema “Exportación y reutilización de reglas” en la página 54 para obtener más información.

Tenga en cuenta que una regla de segmentación como EDAD BETWEEN 50 y 20 no devolverá resultados. Ocurre lo mismo con la expresión $x \leq \text{valor and valor} < y$, que nunca puede cumplirse. Esto es coherente con el comportamiento de SQL Server. Edad BETWEEN 20 y 50 es el formato correcto.

Tenga en cuenta además que una regla de segmentación como, por ejemplo, Género = \$null\$ tampoco devolverá resultado alguno. Si se desea descartar los campos cuyo valor Género esté vacío, puede utilizarse en su lugar una expresión como Género Not One Of (Masculino, Femenino).

Conjuntos de reglas de segmentación

- Las reglas de segmentación también se pueden combinar en conjuntos. Un conjunto de reglas de segmentación devuelve un valor de verdadero si *cualquiera* de los segmentos individuales es verdadero. Tenga en cuenta que esto es distinto de lo que ocurre con las expresiones que hay *dentro* de un segmento, que deben ser todas verdaderas.
- Se pueden exportar conjuntos de reglas de segmentación de cualquier aplicación que utilice reglas o se puede hacer esto en IBM SPSS Rules Management si está disponible. Consulte el tema “Exportación y reutilización de reglas” en la página 54 para obtener más información.

Definición de reglas de selección

Una regla de selección es un conjunto de una o varias reglas de segmentación o conjuntos de reglas con un atributo *Incluir* o *Excluir* asignado a cada una. Las reglas de selección se utilizan en todas las aplicaciones para seleccionar subconjuntos de registros para el modelado, la puntuación u otros procesos.

1. En la aplicación donde desea añadir la regla, pulse el icono **Crear una nueva regla**. Este icono está disponible en puntos donde se pueden aplicar reglas de selección.
2. Especifique un nombre para la regla, seleccione si va a utilizar la regla para la inclusión o la exclusión (o haga esto más tarde), y añada una o varias expresiones, como desee. Consulte el tema “Definición de reglas de segmentación” en la página 51 si desea más información.
3. Pulse **Aceptar** para guardar la regla.
4. Pulse los iconos de la columna **Incluir/Excluir** si desea conmutar este valor para cada regla. Para los conjuntos de reglas de selección referenciados, este valor no se puede cambiar. No puede combinar inclusiones y exclusiones en la misma regla.

De forma alternativa, pulse el icono **Buscar una regla existente** para examinar reglas que ya se han definido. Consulte el tema “Exportación y reutilización de reglas” en la página 54 si desea más información.

Evaluación de reglas de selección

Pulse el icono **Ver lógica de regla de selección** para ver detalles sobre cómo se realizan selecciones (AND lógico versus OR).

- Al evaluar reglas de selección, si se especifican varios segmentos de inclusión, se incluyen los registros que coinciden con *cualquiera* de los segmentos. Por ejemplo, los clientes se podrían incluir basándose en la valoración del crédito o (OR) en la pertenencia.
- Si se especifican varios segmentos de exclusión, se excluyen los registros que coinciden con cualquiera de los segmentos. Por ejemplo, los clientes se podrían excluir basándose en la edad o (OR) la valoración del crédito.
- En el caso de un conflicto, las reglas de exclusión tienen prioridad. Por ejemplo, un cliente que acaba en un segmento excluido se excluirá, independientemente de cualquier segmento para el cual, de lo contrario, estaría cualificado.
- Al crear conjuntos de reglas de selección en IBM SPSS Rules Management, todas las reglas deben ser del mismo tipo, ya sea *Incluir* o *Excluir*. No puede combinar reglas de inclusión y exclusión en el mismo conjunto de reglas.

Exportación y reutilización de reglas

De forma predeterminada, las reglas creadas en una aplicación se guardan con dicha aplicación. De forma alternativa, las reglas se pueden exportar o guardar en IBM SPSS Collaboration and Deployment Services Repository para su uso en varias aplicaciones. En este caso, cada aplicación almacena una referencia a una versión específica etiquetada de la regla compartida; cualquier cambio en la regla se aplica automáticamente en todas las aplicaciones que utilizan dicha versión.

Exportación de reglas

Desde cualquier aplicación que contiene una regla local que desea compartir, pulse el icono **Exportar** para guardar la regla en IBM SPSS Collaboration and Deployment Services Repository donde se puede reutilizar según convenga.

- Al exportar varias reglas, especifique si desea exportarlas todas en un solo conjunto de reglas o en reglas individuales. En el caso de reglas individuales, cada segmento se exporta como una regla separada.
- Al exportar varias reglas de selección, si elige **Almacenar con selecciones**, se exporta un conjunto de reglas de selección. En este caso, puede exportar inclusiones o exclusiones, pero no ambas. Si necesita una combinación de inclusiones y exclusiones, estas se deben exportar en conjuntos de reglas independientes. Si **Almacenar con selecciones** no está seleccionado, se exporta un conjunto de reglas de segmento.
- Al exportar desde un editor de reglas de agregación, si elige **Exportar con puntos de riesgo**, se crea un conjunto de reglas de agregación. Si esta opción no está seleccionada, se exporta un conjunto de reglas de segmento.

Utilización de reglas existentes

- En cualquier aplicación donde desee utilizar una regla compartida, pulse el icono **Buscar una regla existente** para examinar IBM SPSS Collaboration and Deployment Services Repository para buscar reglas existentes que se puede reutilizar en la aplicación actual.
- Para hacer referencia a una versión etiquetada específica, pulse la flecha junto al nombre de la regla y elija la versión deseada. Si no se ha seleccionado ninguna versión, se utiliza la última. (Piense en *LATEST* como una etiqueta especial que se registra automáticamente en cada versión nueva cuando se crea.) Se recomienda utilizar una versión etiquetada específica para evitar resultados inesperados que se pueden producir cuando se crea una nueva versión de una regla.

Las reglas referenciadas se pueden combinar con reglas locales en muchos casos, pero se distinguen visualmente de las reglas locales mediante el texto en negrita, y con un icono que indica el tipo de regla (o conjunto de reglas).

- De forma alternativa, las reglas referenciadas se pueden convertir a reglas locales y guardar en la aplicación actual, si lo desea. Para hacer esto, pulse el nombre de cualquier regla referenciada y, después, seleccione Convertir a reglas individuales en el cuadro de diálogo Vista de previa de regla.

Directrices para compartir reglas

Al examinar y trabajar con reglas compartidas, se aplican las directrices siguientes:

- Las reglas se pueden exportar para reutilizarse desde cualquier aplicación, pero una vez exportadas, se pueden editar en IBM SPSS Rules Management, y no se pueden editar en ninguna otra aplicación. Tenga en cuenta que IBM SPSS Rules Management puede no estar instalado en todos los sitios ni disponible para todos los usuarios.
- El uso de una versión etiquetada específica (distinta a *LATEST*) se recomienda encarecidamente, sobre todo en un entorno de producción, ya que la actual cambia a la última versión, de lo contrario, puede tener consecuencias no deseadas. En determinados casos, como al hacer referencia a una regla, se debe utilizar una versión etiquetada.

- Las reglas individuales se pueden añadir desde IBM SPSS Collaboration and Deployment Services Repository mediante referencia, o convertirse a reglas locales desde el cuadro de diálogo Vista previa de regla.
- Se puede hacer referencia a las reglas individuales y se pueden combinar con reglas locales en cualquier editor de reglas. Si se exportan como un conjunto de reglas, las referencias anidadas en reglas individuales se conservarán.
- Se puede hacer referencia a conjuntos de reglas de segmento y selección y se pueden combinar como reglas individuales en la mayoría de los casos, pero el conjunto de reglas combinadas no se puede exportar sin convertir, en primer lugar, cualquier conjunto de reglas anidadas a reglas individuales. De esta forma, un conjunto de reglas exportado puede contener referencias a reglas anidadas, *pero no a conjuntos de reglas anidadas*.
- También se puede hacer referencia a conjuntos de reglas de agregación, pero no se pueden combinar con otras reglas, y no pueden contener referencias anidadas a otras reglas. Al hacer referencia a un conjunto de reglas de agregación, la regla referenciada sustituye todo el contenido del editor de reglas de agregación. (De forma alternativa, si no se hace referencia a un conjunto de reglas de agregación, el editor de agregación puede contener una combinación de reglas locales, reglas de segmento referenciadas y conjuntos de reglas de segmento.)
- Al utilizar IBM SPSS Rules Management, se aplican las mismas restricciones que para las reglas exportadas. Las reglas anidadas están permitidas, los conjuntos de reglas anidadas no lo están. Esto se debe a que IBM SPSS Rules Management se ha diseñado para crear y editar reglas compartidas (similar a exportar desde cualquiera de las otras aplicaciones).
- Para evitar cualquiera de las restricciones en los conjuntos de reglas referenciadas, ya sea en IBM SPSS Rules Management o en cualquier otro lugar, un conjunto de reglas siempre se puede convertir a reglas individuales locales.
- Las reglas compartidas no pueden contener referencias a reglas externas de IBM Operational Decision Management.

Añadir anotaciones

En una agregación o asignación, pulse el icono **Añadir anotaciones** para añadir una columna de campos de texto libre en la cual puede entrar una nota que se incluirá con la salida de la regla. Por ejemplo, esto puede ser información adicional que indica al operador de un centro de asistencia telefónica con quién contactar cuando se marca una reclamación de seguros para la referencia.

Las anotaciones se pueden utilizar en combinación con el valor **Devolver con** especificado para una dimensión, para proporcionar detalles adicionales. Aunque el valor **Devolver con** está asociado a la dimensión o decisión que se devuelve, la anotación le indica qué reglas se han activado como entrada en dicha decisión.

Utilización de reglas de IBM Operational Decision Management

Se puede hacer referencia reglas creadas en un sistema de gestión de reglas empresariales como, por ejemplo, IBM Operational Decision Management, y se pueden utilizar en aplicaciones IBM Analytical Decision Management. Estas reglas externas se pueden seleccionar y utilizar de la misma forma que otras reglas pulsando el icono **Buscar una regla existente** en el punto deseado de la aplicación. Las reglas externas se distinguen mediante su icono, pero de lo contrario se utilizan en la forma estándar.

Si desea más información sobre cómo crear reglas externas para ser utilizadas en IBM Analytical Decision Management, consulte la *Guía de diseñadores de aplicaciones*.

La regla se inserta en la aplicación de la misma forma que cualquier otra regla. Tenga en cuenta que las salidas de las regla deben coincidir con el contexto en el cual se utiliza la regla; por ejemplo, las asignaciones requieren un valor verdadero/falso o sí/no (booleano) como entrada en la decisión.

Al igual que con otras reglas, pulse el nombre de cualquier regla externa utilizada en la aplicación para ver detalles sobre la regla, incluyendo entrada y campos, junto con la descripción de la regla y otras propiedades.

Si una regla externa tiene varios campos de salida, se muestra un indicador que le permite elegir el campo deseado.

Si se utiliza una regla externa en lugar de un modelo, no se muestra ningún campo objetivo, y se muestran todas las salidas configuradas en la lista desplegable.

Sugerencias para utilizar reglas externas

- Las reglas que devuelven un valor de *verdadero* o *falso* (o *sí/no* o *0/1*) se pueden utilizar en cualquier selección o asignación, como entrada en cualquier regla del editor de reglas, o como un campo **Devolver con** asociado a un elemento de dimensión específico.
- Las reglas que devuelven salidas categóricas (*rojo, blanco, o azul; norte, sur, este, oeste, etc.*) se pueden utilizar en la sección **Utilizar un modelo para decidir** en la pestaña Definir, como una entrada en el editor de reglas, o como un campo **Devolver con** asociado a un elemento de dimensión.
- Las reglas que devuelven salidas numéricas se pueden utilizar en la sección **Utilizar un modelo para decidir** en la pestaña Definir, como entrada en una priorización, como entrada en el editor de reglas, o como un campo **Devolver con** asociado a un elemento de dimensión.
- Los orígenes de datos de la vista de datos analíticos soportan el uso de reglas externas complejas (jerárquicas). Para hacer referencia a una regla externa compleja, el modelo de datos del proyecto debe ser un origen de *vista de datos* y las funciones deben estar habilitadas (en la pestaña Datos, un administrador debe seleccionar la opción **Permitir el uso de reglas ODM basándose en los datos jerárquicos** apropiada).
- No puede crear una regla compartida en IBM Analytical Decision Management que contiene una referencia a una regla externa.

Capítulo 6. Creación de modelos predictivos

Acerca de modelos predictivos

Los modelos se pueden utilizar para predecir lo que probablemente sucederá en el futuro, basándose en patrones de datos del pasado. Por ejemplo, podría utilizar un modelo para predecir qué clientes tienen menos probabilidades de abandonar, o los que tienen más probabilidades de responder a una oferta particular, de acuerdo con características como, por ejemplo, ingresos, edad y las organizaciones y asociaciones a las que están suscritos.

Los modelos se pueden utilizar la misma forma que las reglas, pero mientras que las reglas se pueden basar en políticas corporativas, la lógica empresarial y otros supuestos, los modelos se basan en observaciones reales de resultados del pasado y pueden descubrir patrones que, de lo contrario, pueden no ser aparentes. Mientras que las reglas aportan lógica empresarial común a las aplicaciones, los modelos dan información y capacidad predictiva.

Datos para generar y puntuar modelos

Se utilizan dos tipos diferentes de datos en el proceso de modelado:

- Para generar el modelo, necesitará información sobre lo que desea predecir. Por ejemplo, si desea predecir el abandono, necesitará información sobre los clientes que han abandonado en el pasado. A menudo, a esto se hace referencia como datos históricos o analíticos y debe contener algunos o todos los campos del modelo de datos de proyecto, además de un campo adicional que registra la salida o el resultado que desea predecir. Este campo adicional se utiliza como **objetivo** para la creación de modelos.
- Para utilizar el modelo para predecir futuros resultados, necesitará datos sobre el grupo o población que le interesa como, por ejemplo, posibles clientes o reclamaciones de ingresos, por ejemplo. A menudo, se hace referencia a esto como **datos operativos** o datos de **puntuación**. Normalmente, el modelo de datos del proyecto se basa en estos datos.
- Si el campo objetivo se incluye en el modelo de datos, la columna **Operativo** no debería estar seleccionada para este campo, porque no está disponible al puntuar el modelo. Habiendo construido el modelo mediante datos históricos, el objetivo de la puntuación es aplicar el modelo a nuevos datos donde el resultado todavía no se conoce.

Generación de modelos

La pestaña Modelado le permite especificar un origen de datos, campo objetivo y otras opciones. También puede elegir qué tipo de modelo generar.

Tenga en cuenta que si se han añadido **parámetros** a la aplicación (archivo de ruta) en IBM SPSS Modeler, estará disponible un enlace **Parámetros** en la pestaña Modelado y la pestaña Puntuación. Si pulsa el enlace **Parámetros** se mostrarán los parámetros definidos y sus valores actuales, y le permite cambiar los valores de parámetro para afectar a la generación del modelo o la puntuación.

Generación de un modelo predictivo

Al generar un modelo predictivo, las únicas especificaciones necesarias son un origen de datos y un campo objetivo. También puede elegir opciones para preparar y particionar datos y para seleccionar los campos y registros para utilizar. El modelo se genera aplicando un rango de métodos según sea apropiado para el tipo de objetivo e identificando automáticamente la técnica o combinación específica que funciona mejor en los datos seleccionados.

Cuando se puntúa, el modelo devuelve uno o más campos que contienen predicciones, así como valores de confianza asociados a estas predicciones. Por ejemplo, si el campo objetivo es un campo de marca llamado *Respuesta*, el modelo devuelve un campo llamado *XF-Respuesta* que contiene valores predichos para dicho campo.

1. Especifique un origen de datos. Este puede ser cualquier origen de datos analíticos que registra el resultado que desea predecir.
2. Especifique un campo objetivo que registra el resultado que desea predecir. Por ejemplo, el campo objetivo podría indicar qué clientes han abandonado, contestado a ofertas anteriores, han presentado reclamaciones fraudulentas, etc.

Todos los demás campos se incluirán automáticamente como entradas que pueden ser útiles para predecir este valor.

3. Especifique valores opcionales, como desee. Consulte el tema “Valores de modelo opcionales” en la página 63 si desea más información.
4. Si lo desea, pulse el icono Visión general de datos para ver una descripción general de los datos que se utilizará para crear el modelo actual. Consulte el tema “Visión general de datos” en la página 43 si desea más información.
5. Pulse **Generar modelo**.

Puede cerrar el navegador o trabajar en otras tareas mientras se genera el modelo. Se puede acceder al modelo desde la galería una vez completada la generación. Consulte el tema “Galería” en la página 26 si desea más información. Si el modelo está tardando más tiempo en generarse de lo esperado, pulse el botón **Detener** para revertir a la versión guardada previamente.

Cuando se completa la generación del modelo, se visualizan los resultados. Puede elegir ver los **Resultados de modelos combinados** (gráficos de rendimiento para el modelo combinado que se ha creado actualmente) o los **Resultados de modelos individuales** (detalles de nivel superior de los modelos individuales que conforman el modelo combinado).

Al visualizar *resultados de modelos individuales*, el nombre de cada modelo es un enlace que abre un nuevo diálogo que contiene información sobre el modelo como, por ejemplo, el campo objetivo y campos de entrada. Si el modelo seleccionado es un *modelo de árbol*, también se visualizará un visor de árbol.

6. De forma opcional, utilice Evaluar y probar características para ver cómo funciona el modelo en los datos de muestra. Consulte el tema “Evaluación de modelos” en la página 67 si desea más información.

Resultados de modelado predictivo

Cuando se completa un modelo predictivo, la sección de resultados le permite cambiar entre las vistas de **Resultados de modelos combinados** y **Resultados de modelos individuales**. Los resultados de modelos combinados incluyen gráficos de distribución y de importancia de variable para el modelo combinado que se ha creado, que proporciona una evaluación visual rápida del rendimiento del modelo y de la importancia de distintos predictores para obtener el resultado final. Los resultados del modelo individual listan información sobre todos los modelos individuales que conforman el modelo combinado, lo que le permite pulsar cada modelo individual para ver sus detalles.

Gráficos de distribución

Los gráficos de distribución del modelo muestran valores observados en el eje horizontal, con valores predichos en el eje vertical axis o superpuestos. Esto le permite ver con qué frecuencia se ha predicho cada valor o respuesta, y con qué frecuencia era correcta dicha predicción, lo que proporciona una evaluación visual rápida de la precisión del modelo.

- Para objetivos de marca o conjunto, la distribución se muestra utilizando un gráfico de barras, con una barra para cada categoría del campo objetivo. La altura de cada barra indica cuántas veces ha predicho el modelo dicho resultado. Las regiones coloreadas dentro de cada barra indican cuántas veces se ha observado realmente el resultado coincidente en los datos de prueba. Si el modelo es razonablemente preciso, el área mayor dentro de cada barra debería coincidir con la respuesta correspondiente.

- Para objetivos de rango, la distribución se representa mediante un diagrama de dispersión agrupado. De nuevo, los valores predichos se trazan en el eje horizontal, y los valores observados se tratan en el eje vertical. Idealmente, los puntos deben agruparse a lo largo de la línea diagonal. Cuanto más dispersos están los puntos (resultados) y más lejos están de una distribución lineal, menos preciso es el modelo.

Gráficos de importancia de predictor

Es normal desear centrar los esfuerzos de modelado en los campos predictores que importan más y considerar descartar o ignorar aquellos con menos importancia. El gráfico de importancia de predictor le ayuda a hacer esto indicando la importancia relativa de cada predictor en la estimación del modelo. Puesto que los valores son relativos, la suma de los valores para todos los predictores en la visualización es 1.0. La importancia de predictor no está relacionada con la precisión del modelo. Solo está relacionada con la importancia de cada predictor al realizar una predicción, no con si la predicción es precisa o no.

El cálculo de la importancia de predictor puede tardar mucho más que la generación del modelo, especialmente, si se utilizan conjuntos de datos grandes.

Si la opción de partición automática de datos estaba seleccionada cuando se creó el modelo, la partición de prueba se seleccionará automáticamente en la página Evaluar.

Si el particionamiento no estaba activado, tendrá que seleccionar un origen de datos para evaluar el modelo. Se recomienda seleccionar un origen de datos distinto al utilizado para crear el modelo. Tenga en cuenta que el particionamiento está activado de forma predeterminada, así pues, es posible que esto no sea necesario, a menos que haya cambiado este valor.

Resultados de modelos individuales

Los resultados de modelos individuales muestran información sobre cada modelo individual que conforma el modelo combinado. El nombre de cada modelo es un enlace que abre un nuevo diálogo que contiene detalles completos sobre dicho modelo como, por ejemplo, el campo objetivo y los campos de entrada. Si el modelo seleccionado es un *modelo de árbol*, también se visualizará un visor de árbol especial. El visor de árbol es interactivo y proporciona las funciones siguientes:

- Acercar y alejar el árbol
- Mostrar los elementos de árbol como tablas, gráficos o tablas y gráficos combinados
- Cambiar la orientación del árbol de arriba a abajo, de izquierda a derecha o de derecha a izquierda
- Utilice la *ventana de mapa del árbol* y los iconos de flecha para desplazarse por árboles grandes más fácilmente

Creación de un modelo de regla predictiva

Los modelos de regla predictiva le permite identificar segmentos que muestran una probabilidad mayor o menor con respecto a un resultado específico. Por ejemplo, podría buscar clientes que tienen menos probabilidades de abandono, o más probabilidades de decir sí a una oferta particular. El modelo resultante se parece a una lista de reglas de selección, con un atributo **Incluir** o **Excluir** que especifica cómo se puntúa cada segmento.

Los modelos de regla predictiva se parecen a reglas de selección, pero se diferencian en cómo se puntúan. En el caso de una regla de selección, los registros excluidos se eliminan, lo que genera menos registros en la salida. En el caso de un modelo, no se elimina ningún registro, pero se añaden campos adicionales que indican los valores de predicción y propensión para cada registro. Los segmentos incluidos se puntúan como *Verdadero* (o *Falso*, o cualquier resultado que esté intentando predecir). Los segmentos excluidos se puntúan como *NULL*, al igual que cualquier registro que esté en el resto porque no coinciden con ninguna regla.

A cada registro también se le da una puntuación de propensión que indica la probabilidad de la respuesta objetivo, lo que hace posible identificar los clientes que tienen más probabilidad de responder. Como tal, los modelos de regla predictiva son muy apropiados para generar listas de distribución y se usan con frecuencia en la gestión de relaciones con los clientes, incluyendo las aplicaciones de centro de atención telefónica o marketing.

- La creación de modelos de regla predictiva está disponible solo para objetivos categóricos.
- Puede ver, modificar y reorganizar segmentos y elegir qué incluir o excluir para fines de puntuación. Por ejemplo, puede elegir excluir un grupo de clientes de futuras ofertas e incluir a otros e inmediatamente ver cómo afecta esto en su tasa de aciertos general.
- Al igual que con las reglas de selección, puede añadir o editar manualmente segmentos, según sea necesario. A menudo, es más fácil dejar que el modelo los busque automáticamente, sin embargo, se proporciona una serie de opciones para ayudarle a ajustar de forma precisa el resultado.
- Las reglas se aplican en el orden listado, con la primera regla coincidente que determina el resultado para un registro determinado. Si se toman de forma independiente, las reglas o condiciones se pueden solapar, pero el orden de las reglas resuelve la ambigüedad. Si no coincide ninguna regla, el registro se asigna a la regla restante.

Obtener un modelo de regla predictiva

1. Especifique un origen de datos. Este puede ser cualquier origen analítico que registra el resultado que desea predecir.
2. Especifique un campo objetivo que registra el resultado que desea predecir.
3. Especifique el valor o la respuesta específico que desea encontrar. Por ejemplo, si el campo objetivo se llama *respuesta*, podría elegir *Verdadero* o *Falso* para encontrar los clientes que han contestado y los que no, en consecuencia.
Por ejemplo, el objetivo podría indicar qué clientes: han abandonado, han contestado a ofertas pasadas, han presentado reclamaciones fraudulentas, etc. Todos los demás campos se incluirán automáticamente como entradas que pueden ser útiles para predecir este valor.
4. Especifique valores opcionales, como desee. Consulte el tema “Valores de modelo opcionales” en la página 63 si desea más información.
5. Si lo desea, pulse el icono Visión general de datos para ver una descripción general de los datos que se utilizará para crear el modelo actual. Consulte el tema “Visión general de datos” en la página 43 si desea más información.
6. Pulse **Iniciar generación**.
7. Especifique si desea encontrar segmentos con una probabilidad alta o baja en relación con el objetivo especificado. Por ejemplo, podría buscar grupos o clientes con mayor probabilidad de abandonar, o con menos probabilidad, e incluirlos en consecuencia.
8. Especifique el número máximo de segmentos que se van a encontrar. Normalmente, este número se mantiene pequeño en beneficio de la simplicidad y el rendimiento.
9. Pulse **Ampliar modelo** para obtener una lista de una o más reglas de segmento que identifican grupos con mayor o menor probabilidad.
10. Pulse los iconos **Incluir/Excluir** para especificar cómo se van a utilizar las reglas y cómo se van a puntuar los segmentos relevantes.
Los segmentos incluidos reciben una puntuación que coincide con el valor objetivo (por ejemplo *response=true*). Los segmentos excluidos se puntúan como *NULL*, pero no se eliminan del archivo.
11. De forma opcional, puede añadir o editar manualmente segmentos, según sea necesario. Consulte el tema “Definición de reglas de segmentación” en la página 51 si desea más información.
12. De forma opcional, utilice Evaluar y probar características para ver cómo funciona el modelo en los datos de muestra.

Sugerencias:

- Utilice las flechas arriba/abajo para subir o bajar reglas. Cada registro se puntúa basándose en el primer acierto, de forma que el orden de las reglas cambiará el resultado. Por ejemplo, un registro que coincide con varias reglas se puntuará basándose en la primera que se aplica.
- Cambie una o más opciones y vuelva a pulsar **Ampliar modelo** para encontrar reglas adicionales. Por ejemplo, podría reducir el umbral para un intervalo mínimo de segmento o confianza para poder identificar segmentos adicionales.
- Pulsar **Ampliar modelo** no borra la tabla ni vuelve a crear el modelo de cero, solo busca más reglas de segmento para añadir a la tabla.

Generación de un modelo de agrupación en clúster

Los modelos de agrupación en clúster le permiten categorizar registros en un determinado número de clústeres. Esto puede ayudarle a identificar grupos naturales en los datos.

Los modelos de agrupación en clúster se centran en la identificación de grupos de registros similares y en el etiquetado de los registros de acuerdo con el grupo al que pertenecen. Esto se realiza sin la ventaja de disponer conocimientos previos sobre los grupos y sus características. De hecho, es posible que ni siquiera sepa exactamente cuántos grupos buscar. Esto es lo que distingue los modelos de agrupación en clúster de otras técnicas de aprendizaje de máquina, no hay ningún campo de salida u objetivo predefinido para el modelo que se va a predecir. A menudo, se hace referencia a estos modelos como modelos de **aprendizaje sin supervisor**, puesto que no hay ningún estándar externo a través del cual juzgar el rendimiento de la clasificación del modelo. No hay respuestas *correctas* ni *incorrectas* para estos modelos. Su valor se determina mediante su capacidad de capturar agrupaciones interesantes en los datos y proporcionar descripciones útiles de estas agrupaciones.

Los métodos de agrupación en clúster se basan en la medición de distancias entre registros y clústeres. Los registros se asignan a clústeres de una forma que tiende a minimizar la distancia entre registros que pertenecen al mismo clúster.

A menudo, los modelos de agrupación en clúster se utilizan para crear clústeres o segmentos que, después, se utilizan como entradas en análisis posteriores. Un ejemplo común de esto lo ilustran los segmentos de mercado utilizados por los usuarios de marketing para particionar su mercado global en subgrupos homogéneos. Cada segmento tiene características especiales que afectan al éxito de los esfuerzos de marketing orientados a ello. Si está utilizando la minería de datos para optimizar la estrategia de marketing, normalmente, puede mejorar el modelo de forma significativa identificado los segmentos apropiados utilizando la información de dicho segmento en los modelos predictivos.

Obtener un modelo de agrupación en clúster

1. Especifique un origen de datos. Este puede ser cualquier origen analítico que registra el resultado que desea predecir.
2. Especifique valores opcionales, como desee. Consulte el tema “Valores de modelo opcionales” en la página 63 si desea más información.
3. Si lo desea, pulse el icono Visión general de datos para ver una descripción general de los datos que se utilizará para crear el modelo actual. Consulte el tema “Visión general de datos” en la página 43 si desea más información.
4. Pulse **Buscar clústeres**.
5. De forma opcional, puede añadir clústeres manuales. Consulte el tema “Utilización de clústeres manuales” si desea más información.
6. De forma opcional, utilice Evaluar y probar características para ver cómo funciona el modelo en los datos de muestra.

Utilización de clústeres manuales

Si lo desea, puede definir manualmente clústeres y, después, utilizar la agrupación en clústeres automática de registros no identificados por ninguno de los clústeres definidos manualmente. Cree

clústeres manuales a través del conjunto de reglas de asignación basada en el segmento del "primer acierto". Las reglas que cree definirán los clústeres y el resto se utilizará para el modelado de clúster automático.

Tenga en cuenta que los clústeres manuales no están incluidos en los gráficos de resumen de resultados.

1. En la sección **Clúster manual**, pulse el icono **Crear una nueva regla** y añada una o varias reglas para definir clústeres manuales.
2. Si lo desea, cambie los nombres de etiqueta del clúster.
3. Pulse el icono **Registros en clústeres manuales** para realizar un recuento de registros del número total de registros incluidos en los clústeres manuales. La columna **Recuento** se actualizará para cada clúster manual definido.

Creación de un modelo de asociación

Los modelos de asociación le permiten predecir qué elementos tienen más probabilidad de aparecer juntos y predecir la solidez de la relación entre ellos.

Las **reglas de asociación** asocian una conclusión particular (la compra de un producto particular, por ejemplo) con un conjunto de condiciones (la compra de otros productos, por ejemplo). Por ejemplo, la regla

```
cerveza <= verduras_en_conserva & comida_congelada (173, 17,0%, 0,84)
```

indica que *cerveza* a menudo aparece cuando aparecen juntos *verduras_en_conserva* y *comida_congelada*. El regla es fiable en un 84 % y se aplica al 17 % de los datos, o 173 registros. Los algoritmos de regla de asociación encuentran automáticamente las asociaciones que podría encontrar manualmente utilizando técnicas de visualización, como el nodo web.

La ventaja de los algoritmos de regla de asociación con respecto a los algoritmos de árbol de decisión más estándar (árboles C5.0 y C&R) es que las asociaciones pueden existir entre *cualquiera* de los atributos. Un algoritmo de árbol de decisión creará reglas solo con una única conclusión, mientras que los algoritmos de asociación intentan encontrar muchas reglas, cada una de las cuales puede tener una conclusión diferente.

La desventaja de los algoritmos de asociación es que intentan encontrar patrones en un espacio de búsqueda potencialmente muy grande y, por lo tanto, pueden necesitar mucho más tiempo para ejecutarse que un algoritmo de árbol de decisión. Los algoritmos utilizan un método de **generación y prueba** para encontrar reglas, inicialmente, se generan las reglas sencillas y estas se validan con respecto al conjunto de datos. Las reglas buenas se almacenan y todas las reglas, sujetas a distintas restricciones, se especializan después. **La especialización** es el proceso de añadir condiciones a una regla. Estas reglas nuevas se validan con respecto a los datos y el proceso almacena de forma repetitiva las mejores reglas o las más interesantes que se encuentran. Normalmente, el usuario proporciona algún límite al número posible de antecedentes permitidos en una regla, y se utilizan distintas técnicas basadas en la teoría de la información o los esquemas eficientes de indexación para reducir el espacio de búsqueda posiblemente grande.

Obtener un modelo de asociación

1. Seleccione un origen de datos. Este puede ser cualquier origen analítico que registra el resultado que desea predecir.
2. Especifique un formato de datos. Consulte el tema "Datos tabulares versus transaccionales" en la página 63 si desea más información.
3. Especifique valores opcionales, como desee. Consulte el tema "Valores de modelo opcionales" en la página 63 si desea más información.
4. Si lo desea, pulse el icono Visión general de datos para ver una descripción general de los datos que se utilizará para crear el modelo actual. Consulte el tema "Visión general de datos" en la página 43 si desea más información.

5. Pulse **Generar modelo**.
6. De forma opcional, utilice las características de prueba para ver cómo actúa el modelo en los datos de muestra. Tenga en cuenta que la evaluación no está disponible para el tipo de modelo de asociación.

Con otros tipos de modelo, puede elegir entre distintas técnicas de modelado. Para los modelos de asociación, se utiliza el tipo de modelo *A priori*.

Datos tabulares versus transaccionales

Cuando se crea un modelo de asociación, se dispone de los formatos de datos **tabular** y **transaccional**. Los datos tabulares se seleccionan de forma predeterminada. Si selecciona **Transaccional**, seleccione los campos **ID** y **Contenido** en los menús desplegables que aparecen.

Formato transaccional

Los datos transaccionales tienen un registro diferente para cada transacción o elemento. Si un cliente realiza varias compras, por ejemplo, cada una sería un registro aparte con los elementos asociados vinculados mediante un ID de cliente. Esto se conoce a veces como formato **rollo de caja registradora (till-roll)**.

Cliente	Compra
1	mermelada
2	leche
3	mermelada
3	bread
4	mermelada
4	bread
4	leche

Datos en formato de tabla

Los datos tabulares (también conocidos como datos **de la cesta** o **de la tabla de verdad**) tienen elementos representados por marcadores diferentes, donde cada campo de marcas representa la presencia o ausencia de un elemento específico. Cada registro representa un conjunto completo de elementos asociados. Los campos de marcas pueden ser categóricos o numéricos, aunque ciertos modelos pueden tener requisitos más específicos.

Cliente	Atasco	Pan	Leche
1	J	1	1
2	1	1	J
3	J	J	1
4	J	J	J

Valores de modelo opcionales

Al crear modelos, puede elegir entre los valores opcionales siguientes. Tenga en cuenta que en función del tipo de modelo y el objetivo seleccionados, es posible que no vea todos estos valores. Si cambia alguno de estos valores para un modelo existente, el modelo se debe volver a crear para que se aplique los cambios.

Limpiar automáticamente y preparar datos para una generación de modelos fiable. Identifica y repara problemas de datos para que la creación de modelos sea más rápida, más predecible y más fiable. Los

campos de pantallas que son problemáticos o que probablemente no serán útiles, por ejemplo, manejar valores ausentes y extremos, derivar atributos nuevos cuando proceda y mejorar el rendimiento a través de técnicas inteligentes de muestreo y cribado. La primera vez que se crea un modelo con un origen de datos nuevo, el análisis se realiza para identificar problemas y soluciones, lo cual puede ralentizar el rendimiento solo en el primer pase. En las ejecuciones posteriores, se aplican los arreglos, pero el análisis no se repite, a menos que cambie el origen de datos. Este valor puede estar inhabilitado para algunos modelos, incluidos aquellos con valores de preparación de datos personalizados especificados por un usuario experto.

Nota: Los valores de preparación de datos automática solo se aplican al crear un modelo nuevo, y el intervalo no se realiza.

Partir datos automáticamente para habilitar la evaluación del modelo en el origen de datos de generación para su evaluación y prueba. Seleccionar esta opción divide los datos en subconjuntos o muestras separados para la formación y realización de pruebas del modelo. Si crea el modelo en un subconjunto y lo prueba en otro, podrá obtener una idea sobre cómo se generalizará en otros conjuntos de datos. También puede especificar un valor de porcentaje para el *tamaño de partición de entrenamiento* y el *tamaño de partición de prueba* generados aleatoriamente.

El tamaño mínimo de partición de entrenamiento es 1 y el máximo es 100. El tamaño mínimo de la partición de prueba es 0 y el máximo es 100. El tamaño mínimo de la partición de validación es 0 y el máximo es 100. El porcentaje total debe ser 100 o menos para que el modelo se pueda crear correctamente.

Estas opciones de particionamiento están disponibles para todos los tipos de modelo, excepto los modelos de asociación.

Elija las técnicas del modelo para la generación de modelos. Si lo desea, pulse **Seleccionar** para elegir qué técnicas de generación de modelos incluir al crear el modelo. Basándose en el objetivo que ha seleccionado, se listarán todas las técnicas de generación de modelos disponibles. Las técnicas de generación de modelos disponibles (y aquellas seleccionadas de forma predeterminada) variarán en función del objetivo que ha seleccionado. Por ejemplo, el algoritmo *Lista de decisiones* solo aparecerá para objetivos *de marca*. Esta característica solo está disponible para modelos predictivos.

Si desea detalles completos sobre las técnicas de generación de modelos, consulte la *IBM SPSS Modeler Algorithms Guide* y otra documentación que se suministra con el producto IBM SPSS Modeler y que está disponible en la web.

Número máximo de modelos para combinar. Le permite definir el número máximo de modelos que se van a conservar y combinar. Si define esta opción en 1, se creará y conservará un solo modelo. Si hay menos modelos creados que el valor especificado aquí, todos los modelos creados se conservarán y combinarán. Esta opción solo está disponible para modelos predictivos.

Criterios de beneficios (utilizado para clasificar modelos con destinos binarios). Le permite definir criterios de beneficios antes de crear el modelo. Los valores definidos aquí se definirán como el valor predeterminado para evaluar el modelo. Esta opción solo está disponible para modelos predictivos con un objetivo de marca seleccionado.

Especificar entradas para utilizar. Le permite seleccionar los campos que desea utilizar. Normalmente, estos sería aquellos que tienen alguna relación práctica con lo que está intentando predecir como, por ejemplo, la edad o los ingresos. Si tiene un conjunto de datos grande, limitar el número de campos es una forma de simplificar el modelo. Si los datos contienen campos como, por ejemplo, ID de cliente o número de contacto, normalmente estos no serían útiles para crear modelos y no se deberían seleccionar. Los campos que duplican otros datos también se deben excluir.

Si pulsa un campo de entrada enlazado (una expresión), se abre el visor de expresión para dicha expresión. Para editar una expresión, consulte la pestaña Datos. Consulte el tema “Creador de expresiones” en la página 47 si desea más información.

Especificar selecciones para utilizar. Especifica qué registros incluir o excluir al crear modelos. Puede buscar reglas existentes, o crear reglas nuevas, según proceda. Consulte el tema “Definición de reglas de selección” en la página 53 si desea más información.

Además, si se han definido selecciones globales, se visualizarán aquí, y puede especificar si también se deberán aplicar durante la generación de modelos. Puede elegir utilizar todas o ninguna de las reglas de selección global; no puede elegir un subconjunto.

Opciones de segmentación

Al crear modelos de regla predictiva, están disponibles una serie de opciones de segmentación. Estas se suman a las opciones disponibles con otros tipos de modelos.

Tamaño mínimo de segmento. Estos valores dictan el tamaño mínimo del segmento, y el mayor de los dos valores tiene prioridad. Por ejemplo, si el valor de porcentaje es igual a un número mayor que el valor absoluto, el valor del porcentaje tiene prioridad.

- **Porcentaje de segmento anterior.** Especifica el tamaño mínimo del grupo como un porcentaje de registros. El valor mínimo permitido es 0; el valor máximo permitido es 99,9.
- **Valor absoluto.** Especifica el tamaño mínimo del grupo como un número absoluto de registros. El valor mínimo permitido es 1; no hay ningún valor máximo.

Número máximo de atributos. Especifica el número máximo de condiciones por regla de segmentación. El valor mínimo permitido es 1; no hay ningún valor máximo.

Permitir reutilización de atributos. Cuando está habilitado, cada ciclo puede considerar todos los atributos, incluso aquellos que se han utilizado en ciclos anteriores. Las condiciones para un segmento se crean en ciclos, donde cada ciclo añade una nueva condición. El número de ciclos se define utilizando el valor **Número máximo de atributos**.

Intervalo de confianza para nuevas condiciones. Especifica el nivel de confianza para probar la significación del segmento. Este valor juega un rol importante en el número de segmentos (si los hay) que se devuelven, así como la regla de número-de-condiciones-por-segmento. Cuanto mayor es el valor, menor es el conjunto de resultados devuelto. El valor mínimo permitido es 50; el valor máximo permitido es 99,9.

Opciones de clúster automático

Al generar modelos de agrupación en clúster, están disponibles una serie de opciones de clúster automático. Estas se suman a las opciones disponibles con otros tipos de modelos.

Campo de evaluación. Especifica el campo para utilizar para la evaluación. Si está seleccionado un campo de evaluación, el método de clasificación en el nodo de generación de clúster automático de IBM SPSS Modeler estará establecido en *Importancia de evaluación*. Si no está seleccionado ningún campo de evaluación, el método de clasificación se dejará como valor predeterminado. Tenga en cuenta que aquí se puede seleccionar un campo analítico como el campo de evaluación.

Establecer un rango deseado para el número de clústeres encontrados. Especifica el rango mínimo y máximo deseados para el número de clústeres encontrados por el modelo de clúster automático.

Opciones de regla de asociación y puntuación

Al crear modelos de asociación, están disponibles una serie de opciones de regla de asociación y puntuación. Dichas opciones son adicionales a algunas de las opciones disponibles con otros tipos de modelos.

Opciones de regla de asociación.

- **Cobertura (%)**. Especifica el porcentaje mínimo de cobertura para controlar qué grado de aplicación general tendrán las reglas. Por ejemplo, la definición de un valor de 80 % significa que todas las reglas se aplicarán al menos al 80 % de los casos. Tenga en cuenta que los valores más altos generarán menos reglas o puede que incluso ninguna.
- **Confianza (%)**. Especifica el porcentaje mínimo de confianza.
- **Número máximo de condiciones previas**. Especifica el número máximo de condiciones previas.
- **Los datos ya están ordenados por ID (formato transaccional únicamente)**. Esta opción se selecciona para inhabilitar la ordenación que tiene lugar de forma predeterminada al inicio de la creación del modelo para garantizar que los identificadores sean correlativos. La selección de esta opción reducirá el tiempo de procesamiento. Esta opción no tiene ningún impacto si el formato de datos es tabular.

Opciones de puntuación.

- **Número máximo de predicciones**. Especifica el número máximo de predicciones incluidas. Este valor se utiliza junto con el desplegable del criterio de regla para generar las predicciones "de nivel superior", donde nivel superior indica el nivel más alto de confianza, soporte, elevación, etc.
- **Criterio de regla**. Selecciona la medida utilizada para determinar la fuerza de las reglas. Las reglas se ordenan según la fuerza de los criterios aquí seleccionados a fin de devolver las mejores predicciones de un conjunto de elementos. Los criterios disponibles son:
 - **Confianza** muestra la proporción existente entre el soporte de regla en y el soporte de condición de regla. Esto indica la proporción de registros con la(s) condición(es) de regla especificada(s) para la(s) cual(es) la(s) predicción(es) de la regla también es/son verdadera(s). Por ejemplo, si el 50 % de los datos de entrenamiento contienen la palabra "pan" (indicando soporte de condición de regla), pero solo el 20 % contiene las palabras "pan" y "queso" (indicando soporte de regla), la predicción de la regla pan -> queso sería Soporte de regla / Soporte de condición de regla o, en este caso, el 40 %.
 - **Cobertura (%)**, muestra el soporte de condición de regla, es decir, la proporción de identificadores para los cuales las condiciones de regla son verdaderas, basándose en los datos de entrenamiento. Por ejemplo, si el 50 % de los datos de entrenamiento incluyen la compra de pan, la regla pan ->queso tendrá un soporte de condición de regla del 50 %.
 - **Soporte de regla (%)** muestra la proporción de registros para los cuales toda la regla, la(s) condición(es) de regla y la(s) predicción(es) de regla son verdaderas. Por ejemplo, si el 20 % de los datos de entrenamiento contienen tanto pan como queso, el soporte de regla de la regla pan -> queso será del 20 %.
 - **Elevación** muestra la proporción de confianza de la regla respecto de la probabilidad anterior de tener la predicción de regla. Por ejemplo, si el 10 % de toda la población compra pan, un regla que prediga si las personas van a comprar pan con una confianza del 20 % tendrá una elevación de $20/10 = 2$. Si otra regla indica que las personas comprarán pan con una confianza del 11 %, la regla tendrá una elevación cercana a 1, lo que significa que tener la(s) condición(es) de regla no implica una gran diferencia en la probabilidad de tener la predicción de regla. En general, las reglas con una elevación distinta de 1 serán más interesantes que las reglas con una elevación cercana a 1.
 - **Desplegabilidad (%)** es una medida de qué porcentaje de los datos de entrenamiento cumple la(s) condición(es) de regla pero no cumple la predicción de la regla. En términos de compra de productos, significa básicamente qué porcentaje de la base total de clientes posee (o ha adquirido) la(s) condición(es) pero aún no ha adquirido la(s) predicción(es) de la regla. The deployability statistic is defined as $((\text{Rule Condition Support in \# of Records} - \text{Rule Support in \# of Records}) / \text{Number of Records}) * 100$, where Rule Condition Support means the number of records for which the rule conditions are true and Rule Support means the number of records for which both rule conditions and the prediction are true.
- **Comprobar toda la transacción**. Seleccione entre las opciones siguientes:
 - **Garantizar que las predicciones no estén presentes** garantiza que las predicciones de regla no estén presentes. Por ejemplo, si el propósito de la puntuación es recomendar mobiliario para el hogar, no es probable que una cesta que ya contiene una mesa de comedor recomiende comprar otra. En tal

caso, debería seleccionar esta opción. Por otra parte, si los productos son perecederos o desechables (como el queso, un biberón o un pañuelo), las reglas en las que el consecuente ya se esté presente en la cesta pueden ser de valor. En este último caso, la opción más útil podría ser **No comprobar predicciones**.

- **Garantizar que las predicciones estén presentes** garantiza que las predicciones de regla también estén presentes en la cesta. Este método es útil cuando se intenta comprender mejor los clientes o las transacciones existentes. Por ejemplo, puede que se desee identificar las reglas con la mayor elevación y luego explorar qué clientes se ajustan a dichas reglas.
- **No comprobar predicciones** incluye todas las reglas al puntuar, independientemente de la presencia o ausencia de predicciones de regla en la cesta.
- **Permitir repetición de predicciones.** Seleccione esta opción para incluir varias reglas con la misma predicción de regla al puntuar. Por ejemplo, seleccionar esta opción permitiría que se puntuaran las reglas siguientes:
pan & queso -> vino
queso & fruta -> vino
Desactive esta opción para excluir las predicciones repetidas a la hora de puntuar.
- **Ignorar elementos sin coincidencias.** Seleccione esta opción para ignorar la presencia de elementos adicionales en el conjunto de elementos. Por ejemplo, cuando esta opción está seleccionada para una cesta que contiene [tienda & saco de dormir & tetera], la regla tienda & saco de dormir -> estufa de gas se aplicará a pesar del elemento adicional (tetera) presente en la cesta.

Evaluación de modelos

Tras crear o abrir un modelo completado, pulse **Evaluar** para obtener gráficos como, por ejemplo, de distribución, distribución de errores, matriz de coincidencias, ganancias, elevación, respuestas, beneficios y rentabilidad de la inversión, y para simular el máximo beneficio posible. Los tipos de gráfico y otras opciones disponibles variarán en función del tipo de modelo que se haya seleccionado. El diálogo Evaluar solo está disponible si el modelo seleccionado está completo. Tenga en cuenta que la evaluación no está disponible para el tipo de modelo de asociación.

1. En el panel superior, seleccione un conjunto de datos de evaluación. Al igual que con la creación del modelo, este debe ser un conjunto de datos analíticos para los cuales se conocen respuestas reales. (De lo contrario, no tiene forma de comparar las respuestas reales con las predichas por el modelo.)
2. Si desea modificar los modelos para incluir en la evaluación, pulse **Modificar selección de modelo**. Una vez realizadas las selecciones, también puede pulsar el botón siguiente **Volver a copiar selecciones de modelos actuales en la tabla de resultados del modelo individual**, si lo desea. Estas opciones solo están disponibles para el tipo de modelo predictivo.



Figura 13. Volver a copiar selecciones de modelos actuales en la tabla de resultados del modelo individual

3. Seleccione los gráficos que desea crear.
4. Si está simulando los beneficios, seleccione el cuadro **Simular beneficio máximo** y entre la población en la que se basará el cálculo del beneficio, el presupuesto y el valor mínimo. Consulte el tema "Simulación de beneficio máximo" en la página 69 si desea más información.
5. Si solicita un gráfico de beneficios o de rentabilidad de inversión, o si está realizando una simulación del beneficio, especifique los valores de costes e ingresos. Estos valores indican el coste para cada oferta y la cantidad de ingresos que espera obtener de cada respuesta favorable, respectivamente.
6. Seleccione el valor de respuesta que se considera un "acierto". Este debe ser uno de los valores para el campo objetivo utilizado para crear el modelo, como *response = true* para indicar una respuesta positiva o favorable.

7. Si solicita un gráfico de rendimiento, también puede seleccionar la opción **Incluir evaluaciones de modelos individuales para los gráficos de rendimiento**.
8. Si lo desea, seleccione **Incluir resultados en formato de tabla**. Esto requiere más tiempo para que se devuelvan los resultados de evaluación.
9. Pulse **Ejecutar** para ver los resultados.

Gráficos de distribución (todos los objetivos)

Los gráficos de distribución trazan los valores observados respecto a los previstos y se visualizan de forma predeterminada cuando se completa el modelo. El cuadro de diálogo Evaluar le permite obtener gráficos de distribución para un conjunto de datos distinto al conjunto de datos de prueba. Para unos objetivos de rango, el gráfico de distribución es el único tipo de gráfico disponible.

Puede elegir entre los gráficos adicionales siguientes:

Gráficos de distribución de errores

Los gráficos de distribución de errores muestran un histograma de los errores en los resultados del modelo, donde el error es la diferencia entre el valor pronosticado y el valor real. Un error positivo surge cuando el valor pronosticado es mayor que el valor real y un error negativo surge cuando es menor. Este gráfico solo está disponible para objetivos continuos con el tipo de modelo predictivo.

Gráficos de matriz de coincidencias

Una matriz de coincidencias traza los recuentos de resultados *predichos* en comparación con resultados *reales*. Para un objetivo numérico, el rango se divide entre el número de intervalos que haya especificado.

Gráficos de ganancias

Las ganancias se utilizan para mostrar la ganancia o "elevación" que se puede obtener utilizando el modelo, y se definen como la proporción de coincidencias en cada incremento en relación con el número total de coincidencias disponibles. La línea plana diagonal traza la respuesta prevista para toda la muestra, si no se ha utilizado ningún modelo. En este caso, el índice de respuesta sería constante, puesto que una persona tiene la misma probabilidad de responder que cualquier otra. Para duplicar el resultado, tendría que preguntar dos veces al mismo número de personas. La línea curva indica cuánto puede mejorar este resultado seleccionando a aquellas personas que tienen más probabilidad de responder. Cuanto más pronunciada es la curva, mayor es la ganancia.

Gráficos de elevación

El gráfico de elevación traza el porcentaje de registros en cada incremento que son coincidencias con el porcentaje general de acierto en el conjunto de datos de formación.

Gráficos de respuestas

El gráfico de respuestas traza el porcentaje de registros en el incremento que son coincidencias.

Gráficos de beneficios

El beneficio es igual a los ingresos para cada registro menos el coste del registro. Los beneficios para un cuantil son simplemente la suma de beneficios para todos los registros del cuantil. Se asume que los ingresos se aplican solo a los aciertos, pero los costes se aplican a todos los registros. Consulte el tema "Simulación de beneficio máximo" en la página 69 si desea más información.

gráficos de rentabilidad de la inversión

El ROI (rendimiento de la inversión) es similar al beneficio en el sentido de que implica definir ingresos y costes. El ROI compara los beneficios con los costes para cada intervalo.

Simulación de beneficio máximo

Basándose en los índices de respuesta predichos por el modelo, puede simular el beneficio máximo que se va a producir para una población determinada. Para ello, tendrá que especificar los valores de coste e ingresos, así como el tamaño de la población.

Para predecir el beneficio máximo:

1. Especifique la población, o el número total de registros, en los que se deberán basar la predicción.
Nota: Al puntuar un modelo, esto se define automáticamente en el número total de registros en el origen de datos o subconjunto.
2. Especifique un valor para el presupuesto si desea que se incluya el parámetro del presupuesto en la simulación del beneficio, y un valor mínimo.
3. Especifique el coste asociado a cada registro. Por ejemplo, esto podría ser el coste de impresión y del envío de correos, por persona, del material promocional para una campaña.
4. Especifique los ingresos esperados que generaría una respuesta positiva. Por ejemplo, el costa de venta de un elemento en una campaña.
5. Seleccione la respuesta en la que está interesado.
6. Pulse **Ejecutar**.

El gráfico de resultados predichos muestra el mejor resultado para el beneficio máximo, e identifica el porcentaje de registros que necesarios para conseguir este objetivo, basándose en la propensión a responder identificada por el modelo.

Debajo del gráfico puede mover una barra de control deslizante de cero a 100 % de los registros en el objetivo para ver cómo afectaría esto al beneficio.

Cuando las cifras son aceptables, puede utilizar los valores para puntuar los registros. Para ello, pulse **Utilizar para se puntuación**. El porcentaje de registros del objetivo se visualizan en el campo % **principales** del área Opciones de propensión de la pestaña Puntuación.

Comparación de beneficios

Para verificar que los registros elegidos son una buena elección, puede comparar el beneficio probable que se va obtendrá a partir de estos registros, un grupo aleatorio del mismo porcentaje de registros, y el conjunto completo de registros.

Los detalles mostrados incluyen el número de registros seleccionados, el índice de respuesta probable, costes, ingresos y beneficios en general. Puede exportar esta información en formato HTML para su uso en otras aplicaciones.

Gráfico de beneficios

Puede mostrar el beneficio acumulado y no acumulado predichos a partir de los registros seleccionados en forma de gráfico.

Nota: Este gráfico se genera cuando se ejecuta la evaluación por primera vez y no se actualiza automáticamente con ninguno de los cambios que realice a través del selector en la sección de simulación de beneficio.

Probar un modelo

El cuadro de diálogo Prueba le permite ver cómo se puntuarán uno o varios registros de muestra. Los resultados son específicos para el contexto en el cual se ejecuta la prueba y puede reflejar las puntuaciones devueltas por un modelo o una regla. A continuación se muestran algunos ejemplos de los tipos de resultados que puede ver, en función del tipo de aplicación:

- Podría consultar unos cuantos clientes específicos y ver qué ofertas se les han asignado a cada uno.
 - Podría consultar unos cuantos registros de cliente y ver qué clientes es más probable que respondan a una campaña de envío de correo directa.
 - Podría consultar unas cuantas reclamaciones de seguro específicas y ver qué acciones se les ha asignado a cada una de ellas.
 - Podría consultar unos cuantos casos específicos y ver cómo se evalúa cada regla.
1. En la pestaña Modelador, pulse **Probar**. Se visualiza el diálogo Probar.
 2. Seleccione el origen de datos que desea utilizar, o seleccione **Especificar datos personalizados** en la sección Registros para entrar registros de prueba manualmente.
 3. Especifique selecciones para utilizar, si hay alguna. Por ejemplo, es posible que desee incluir o excluir determinados registros de la prueba. Puede buscar reglas existentes, o crear reglas nuevas, según proceda. Consulte el tema “Definición de reglas de selección” en la página 53 si desea más información.
 4. Si se utiliza un origen de datos, especifique el número de registros que desea que se devuelvan. Puede devolver los primeros N registros, hasta un máximo de 100, leer desde el principio de los datos o en una muestra aleatoria de registros. Si se prueba un modelo de asociación utilizando datos transaccionales, los datos se entrada se muestrean en el nivel de transacciones, no en el nivel de registro.
 5. Si se utilizan datos personalizados, seleccione **Especificar datos personalizados** y pulse **Añadir registro personalizado**. Todos los campos definidos en el modelo de datos de proyecto se listan, lo que le permite especificar un valor para cada uno. Pulse **Guardar** cuando termine.
 6. En **Seleccionar campos que deben mostrarse**, seleccione los campos que desea que aparezcan en la salida de la prueba.
 7. Pulse **Ejecutar**. El resumen de la prueba muestra una lista de asignaciones correctas para cada registro.
 8. Pulse el icono de información al principio de cada fila para obtener detalles para dicha fila.

Utilización de puntuaciones de modelo en aplicaciones

Puntuar un modelo significa que se aplica a los mismos datos o población para poder obtener un resultado o predicción que se puede utilizar como entrada en decisiones. Por ejemplo, podría puntuar a todos los clientes en la base de datos para identificar posibles clientes para una campaña de correo directo, o puntuar a los clientes de uno en uno cuando contactan con el centro de asistencia telefónica para poder decidir qué productos o servicios ofrecerles. En función de la aplicación, los resultados de puntuación se pueden escribir en una tabla de base de datos o un archivo sin formato, o utilizarse como entradas en reglas de segmentación, selección y asignación que dirigen las decisiones en una aplicación.

Las puntuaciones de modelo se pueden utilizar en las formas siguientes:

- Como entrada en reglas. Consulte el tema “Creación de reglas” en la página 51 si desea más información.
- Para aplicaciones que soportan la puntuación por lotes, las puntuaciones de modelo se pueden escribir en una tabla de base de datos o un archivo sin formato. Consulte el tema “Puntuación de modelos en una tabla de base de datos, un archivo, Analytic Server o un servidor IBM Cognos BI” en la página 73 si desea más información.

Campos de salida del modelo

Por convenio, los nombres de los campos de salida de puntuación se basan en el nombre del campo objetivo, pero con un prefijo personalizado añadido. Por ejemplo, si el campo objetivo se llama *Respuesta*, los campos de salida tendrán nombres como *\$XF-Respuesta* y *\$XFRP-Respuesta*. Los campos específicos dependen del tipo de modelo, y el nivel de medición para el campo objetivo seleccionado, como se detalla a continuación. Consulte el tema "Niveles de medición" en la página 44 si desea más información.

Objetivos de marca. Los modelos automatizados con objetivos de marca tendrán como salida los campos siguientes:

- **\$XF-<objetivo>**. El valor predicho para cada registro, notificado en términos de valores "verdaderos" y "falsos" definidos para el campo.
- **\$XFRP-<objetivo>**. La puntuación de propensión; indica la probabilidad del valor "verdadero" para cada registro. Las puntuaciones de propensión difieren de las puntuaciones de confianza, que normalmente se notifican en términos de la predicción para dicho registro. Por ejemplo, una predicción de "falso" con una confianza alta se traducirá en una propensión baja ya que esto indicaría una probabilidad alta de *no* responder.

Objetivos categóricos. Los modelos automatizados con objetivos categóricos (nominal o ordinal) tendrán como salida los campos siguientes:

- **\$XS-<objetivo>**. Valor predicho para cada registro.
- **\$XSC-<objetivo>**. Valor de confianza asociado a la predicción

Objetivos continuos. Los modelos automatizados con objetivos numéricos continuos tendrán como salida los campos siguientes:

- **\$XR-<objetivo>**. Valor predicho para cada registro.
- **\$XRE-<objetivo>**. Error estándar para la predicción.

Modelos de regla predictiva (interactiva). Los modelos de regla predictiva (solo objetivos de marca o categóricos) tendrá como salida los campos siguientes:

- **\$D-<objetivo>**. Resultado predicho para cada registro, con el valor "verdadero" definido para el campo utilizado para indicar la respuesta deseada, y nulo que indica cualquier otro valor.
- **\$DP-<objetivo>**. La probabilidad del resultado deseado para cada registro; coincide con la probabilidad mostrada para el segmento en la tabla de resultados al generar el modelo. Consulte el tema "Creación de un modelo de regla predictiva" en la página 59 si desea más información.
- **\$DI-<objetivo>**. Número de índice utilizado para identificar el segmento en el cual se coloca el registro (se muestra en la columna más a la izquierda de los resultados de modelado).
- **\$DRP-<objetivo>**. La puntuación de propensión que indica la probabilidad de la respuesta deseada. Disponible solo para objetivos de marca.

Capítulo 7. Modelos de puntuación

Puntuación de modelos en una tabla de base de datos, un archivo, Analytic Server o un servidor IBM Cognos BI

Puntuar un modelo significa aplicarlo a la muestra de datos o población que interese. Por ejemplo, se podrían utilizar los clientes existentes para generar un modelo que prediga qué clientes van a responder a una campaña con mayor probabilidad y luego utilizar dicho modelo para puntuar los registros de personas que en ese momento no sean clientes, pero que quizás se desee incluir en un envío de correos promocionales.

Los pasos básicos de la puntuación de un modelo son:

1. Seleccionar el origen de datos y cualquier subconjunto de estos datos para puntuar.
2. Seleccionar el destino de los datos puntuados.
3. Seleccionar la información de salida y cómo puntuarla.
4. Especificar qué registros puntuar. En el desplegable Puntuación, seleccione si se van a puntuar **Todos los registros**, el **N% superior**, los **N registros principales**, la **Propensión mín/máx** (disponible solo en los objetivos de *distintivo*) o la **Puntuación mínima** (disponible solo en objetivos *continuos*).
5. Obtenga una vista previa de los resultados esperados y ejecuta una vista preliminar de los datos si lo desea.
6. Pulse **Puntuar ahora** para ejecutar la puntuación contra el modelo utilizando los valores. Cuando haya finalizado, revise la sección del informe de resumen de puntuación. Consulte el tema “Informe de resumen de puntuación” en la página 80 para obtener más información.

Las puntuaciones se escriben en uno o varios campos que se añaden a los datos. Por convenio, los nombres de estos campos se basan en el campo objetivo con un prefijo. Por ejemplo, si el campo objetivo se llama *Respuesta*, los campos de salida tendrán nombres como *\$XF-Respuesta* y *\$XFRP-Respuesta*. Consulte el tema “Selección de campos de salida” en la página 78 para obtener más información.

Nota: Cuando los datos de origen provienen de una base de datos, la ejecución se retrotrae a la base de datos siempre que sea posible (en un proceso llamado "retrotracción SQL"). En general, IBM Analytical Decision Management soporta el mismo nivel de retrotracción SQL que IBM SPSS Modeler, pero con las limitaciones siguientes:

- Si el proyecto incluye campos adicionales o reglas de selección que tienen referencias a modelos o reglas externos, dichas referencias no se retrotraerán a la base de datos, lo que puede inhibir el grado de retrotracción que se puede conseguir para la ruta de proyecto en su conjunto.
- En aplicaciones distintas a IBM SPSS Modeler Advantage, la retrotracción solo está habilitada hasta las reglas de selección (selecciones globales y selecciones de miembro en la primera dimensión). El grado de retrotracción se puede aumentar si elige un campo Entero como ID de entidad en el modelo de datos de proyecto.

IBM SPSS Modeler Advantage

En el caso de **IBM SPSS Modeler Advantage**, la ruta guardada contendrá dos ramas: la rama de creación y la rama de puntuación.

El nodo de origen de la rama de **creación** siempre se basará en el origen de datos del modelo de datos del proyecto. Incluirá campos marcados como operativos o como no operativos.

El nodo de origen de la rama de **puntuación** se basará siempre también en el origen de datos del modelo de datos del proyecto. No obstante, solo incluirá campos que estén marcados como operativos. Los

campos marcados como no operativos no se incluirán, ya que no son necesarios para la puntuación. Por ejemplo, el campo de destino de un modelo predictivo suele estar marcado como no operativo y no es necesario como entrada de puntuación.

De forma predeterminada, la rama de puntuación de la ruta guardada tendrá un nodo de terminal de tabla. Dicho nodo de tabla se utilizará para indicar la rama de puntuación en los valores de despliegue de la ruta. El servicio de puntuación de IBM SPSS Collaboration and Deployment Services se valdrá de esta información para determinar qué nodo terminal usar al realizar la puntuación en tiempo real.

Si está configurada, la pestaña **Puntuación** tendrá una sección para la **Puntuación por lotes**. En **Puntuación por lotes**, hay una sección para las **Opciones de puntuar ahora**. Estos valores solo se aplican cuando se puntúa en la interfaz de usuario de IBM Analytical Decision Management pulsando el botón **Puntuar ahora**. No se reflejarán en la ruta generada que se guarda. Por tanto, no se aplican cuando se puntúa a través del cliente de IBM SPSS Modeler ni cuando se planifica un trabajo mediante el cliente de IBM SPSS Collaboration and Deployment Services Deployment Manager.

La puntuación por lotes también tiene controles para seleccionar el destino y los campos de salida. Al guardar la ruta, se creará un nodo terminal adicional en la rama de puntuación. Se usará la configuración del destino y del campo de salida para crear un nodo terminal de exportación en la rama de puntuación. Estos valores se aplican cuando se ejecuta la ruta desde el cliente de IBM SPSS Collaboration and Deployment Services Deployment Manager o cuando se planifica un trabajo mediante el cliente IBM SPSS Collaboration and Deployment Services Deployment Manager.

Proyectos que no son de IBM SPSS Modeler Advantage

En los proyectos que **no son de IBM SPSS Modeler Advantage**, la ruta guardada solo tendrá la rama de puntuación. Al igual que la rama de puntuación de IBM SPSS Modeler Advantage, el nodo de origen de la rama de puntuación siempre se basará en el origen de datos del modelo de datos del proyecto. Al igual que antes, solo incluirá los campos que estén marcados como operativos. Si se usan tablas secundarias, se crearán nodos de origen adicionales. Dichos nodos de origen secundarios se definen mediante los orígenes de datos secundarios asociados al origen de datos primario.

La pestaña **Desplegar**, si se ha configurado, puede tener secciones de **Puntuación en tiempo real** y **Puntuación por lotes**. Estos valores son independientes entre sí.

Al guardar la ruta, se usarán los valores de **Puntuación en tiempo real** para crear un nodo terminal de tabla en la rama de puntuación. Este nodo de tabla se usará para indicar la ruta de puntuación en los valores de despliegue de la ruta guardada. El servicio de puntuación de IBM SPSS Collaboration and Deployment Services se valdrá de esta información para determinar qué nodo terminal usar al realizar la puntuación en tiempo real.

En **Puntuación por lotes**, hay una sección para las **Opciones de puntuar ahora**. Estos valores solo se aplican cuando se puntúa en la interfaz de usuario de IBM Analytical Decision Management pulsando el botón **Puntuar ahora**. No se reflejarán en la ruta generada que se guarda. Por tanto, no se aplican cuando se puntúa a través del servicio de puntuación de IBM SPSS Collaboration and Deployment Services ni cuando se planifica un trabajo usando un cliente de IBM SPSS Collaboration and Deployment Services Deployment Manager.

La puntuación por lotes también tiene controles para seleccionar el destino y los campos de salida. Al guardar la ruta, estos valores se usarán para crear un nodo terminal de exportación en la rama de puntuación. Estos valores se aplican al planificar un trabajo utilizando el cliente de IBM SPSS Collaboration and Deployment Services Deployment Manager o al ejecutar la ruta desde el cliente de IBM SPSS Modeler .

Todos los proyectos

Tanto los proyectos de IBM SPSS Modeler Advantage como los proyectos que no son de IBM SPSS Modeler Advantage pueden generar nodos terminales distintos en las puntuaciones en tiempo real y por lotes. Cuando planifique trabajos usando el cliente de IBM SPSS Collaboration and Deployment Services Deployment Manager, asegúrese de configurar el paso de trabajo de forma que solo se ejecute el nodo de exportación designado para la puntuación por lotes.

Selección de datos y subconjuntos

En la lista **Origen de datos para puntuar**, seleccione el origen de datos necesario.

Subconjuntos de datos

Si no desea puntuar todo el origen de datos, puede seleccionar un subconjunto de los datos utilizando una regla existente, o creando una regla nueva que solo se utilizará para este modelo. Consulte el tema "Definición de reglas de selección" en la página 53 si desea más información.

Selección de registros para puntuar

Para los modelos que devuelven puntuaciones de propensión, puede utilizar estos valores para seleccionar los registros que se van a puntuar. Por ejemplo, podría puntuar el 30 % de los clientes con la mayor propensión a responder y utilizar estos para generar una lista de distribución. Las puntuaciones de propensión indican la probabilidad del valor "verdadero" para el objetivo y están disponibles para modelos predictivos con objetivos de marca, así como modelos de regla predictiva (interactivos). (En el caso de modelos de regla predictiva, la respuesta seleccionada es el valor "verdadero" y todos los demás son falsos. Para otros tipos de modelos, estas opciones no están disponibles.)

Existen tres formas para elegir registros basándose en la propensión:

- **El porcentaje superior.** De forma predeterminada, esto selecciona el diez por ciento superior. Esta opción le permite simular los resultados de puntuación de acuerdo con los beneficios. Consulte el tema "Simulación de beneficio máximo" en la página 69 si desea más información.
- **El número superior de registros.** De forma predeterminada, esto selecciona los 1000 registros superiores, donde está disponible.
- **Mediante la especificación de valores de propensión mínimo y máximo.** Por ejemplo, todos los registros que tienen una propensión a responder entre el 75 % y el 100 %.

Selección del destino de puntuación

Puede elegir colocar la salida de los datos de puntuación en una base de datos, archivo, Analytic Server o servidor Cognos BI. Los tipos de archivo soportados incluyen texto (*.txt, *.csv), Excel, el formato *.sav utilizado por IBM SPSS Statistics, SAS y los soportados por Data Collection. Consulte el tema "Destino de archivo" en la página 76 si desea más información.

Destino de base de datos

Cuando elija enviar la salida de la puntuación a una base de datos, se le pedirá que entre más detalles.

1. Entre el **Nombre de base de datos** o busque la que necesita.
2. Especifique cómo se van a almacenar las puntuaciones. Puede optar por crear una tabla nueva, en cuyo caso se le pedirá que entre un nombre para la tabla. De forma alternativa, puede añadirlos a una tabla existente o sobrescribir un origen de datos existente, en cuyo caso se le pedirá que seleccione la tabla o el origen de datos relevante.
3. Cambie el tamaño de la serie predeterminado, si lo desea.

Al añadir a, o sobrescribir, una base de datos existente, podrá correlacionar los campos de puntuación con campos del destino existente. Consulte el tema "Correlacionar campos de puntuación" en la página 78 si desea más información.

Destino de archivo

Las puntuaciones pueden ser salidas de distintos tipos de archivo. Para cada tipo, busque un archivo existente o cree uno nuevo en el que se añadirán los datos de puntuación.

Al añadir a, o sobrescribir, un archivo existente puede correlacionar los campos de puntuación con los campos del destino existente. Consulte el tema “Correlacionar campos de puntuación” en la página 78 si desea más información.

Archivo Excel

Tipo. Seleccione el tipo de archivo Excel que está exportando.

Incluir nombres de campo. Especifica si los nombres de campo se deben incluir en la primera fila de la hoja de trabajo.

Archivo sin formato

Cuando elija enviar la salida de puntuación a un archivo sin formato, puede especificar más detalles para controlar el contenido del archivo.

Modo de escritura. Si está seleccionado **Sobrescribir**, cualquier dato existente en el archivo especificado se sobrescribirá. Si está seleccionado **Añadir**, la salida se añadirá al final del archivo existente, conservando los datos que contiene.

- **Incluir nombres de campo.** Si esta opción está seleccionada, los nombres de campo se escribirán en la primera línea del archivo de salida. Esta opción está disponible solo para el modo de escritura **Sobrescribir**.

Nueva línea después de cada registro. Si esta opción está seleccionada, cada registro se escribirá en una línea nueva en el archivo de salida.

Separador de campo. Especifica el carácter para insertar entre valores de campo en el archivo de texto generado. Las opciones son **Coma**, **Tabulador**, **Espacio** y **Otros**. Si selecciona **Otros**, especifique los caracteres delimitadores deseados en el cuadro de texto.

Comillas para símbolos. Especifica el tipo de entrecomillado para utilizar para valores de campos simbólicos. Las opciones son **Ninguna** (los valores no están entrecomillados), **Simple (')**, **Doble (")**, y **Otras**. Si selecciona **Otras**, especifique los caracteres de entrecomillado deseado en el cuadro de texto.

Codificación. Especifica el método de codificación de texto utilizado. Puede elegir entre el valor predeterminado del sistema o UTF-8.

Símbolo decimal. Especifica cómo se deben representar los decimales en los datos.

- **Punto (.)**. El carácter de punto se utilizará como separador de decimal.
- **Coma (,)**. El carácter de coma se utilizará como separador de decimal.

Archivo de datos de IBM SPSS Statistics

Exportar nombres de campo. Especifica un método de manejo de nombres de variable y etiquetas al exportar a un archivo *.sav* de IBM SPSS Statistics. Tenga en cuenta que los nombres no válidos se corrigen automáticamente sustituyendo caracteres no válidos.

- **Nombres y etiquetas de variable.** Seleccione exportar tanto los nombres de campo, como las etiquetas de campo. Los nombres se exportan como nombres de variable de IBM SPSS Statistics, mientras que las etiquetas se exportan como etiquetas de variable de IBM SPSS Statistics.

- **Nombres como etiquetas de variable.** Seleccione utilizar los nombres de campo como etiquetas de variable en IBM SPSS Statistics. Para evitar la posibilidad de crear nombres no válidos de IBM SPSS Statistics, seleccione en su lugar **Nombres y etiquetas de variable**.

Archivo de datos SAS

Tipo. Seleccione el formato de archivo SAS que se va a crear. Puede elegir entre tres formatos de archivo SAS: SAS para Windows/OS2 (*.sd2), SAS para UNIX (*.ssd) o SAS Versión 7/8/9 (*sas7bdat).

Exportar nombres de campo. Seleccione opciones para exportar nombres de campo y etiquetas para ser utilizados con SAS.

- **Nombres y etiquetas de variable.** Seleccione exportar tanto los nombres de campo, como las etiquetas de campo. Los nombres se exportan como nombres de variable SAS, mientras que las etiquetas se exportan como etiquetas de variable SAS.
- **Nombres como etiquetas de variable.** Seleccione utilizar los nombres de campo como etiquetas de variable en SAS. Para evitar la posibilidad de crear nombres SAS no válidos, seleccione en su lugar **Nombres y etiquetas de variable**.

Destino de Analytic Server

Cuando elija enviar una salida de puntuación a un Analytic Server se le pedirá que entre más detalles.

1. Especifique el **Origen de base de datos**.
2. Especifique cómo se van a almacenar las puntuaciones. Puede añadirlas a una tabla existente o sobrescribir un origen de datos existente, en cuyo caso, se le pedirá que seleccione la tabla o el origen de datos relevante.
3. Seleccione campos de salida para incluir con el resultado de puntuación.
4. Especifique qué registros puntuar. Desde el desplegable Puntuación, seleccione si se van a puntuar **Todos los registros**, los **N% primeros**, los **N registros principales**, **La propensión mín/máx** (solo para los objetivos de *marca*), o la **Puntuación mínima** (solo disponible para objetivos *continuos*).

Destino del servidor Cognos BI

Cuando elija enviar la salida de puntuación a un servidor Cognos BI, se le pedirá que entre más detalles.

Nota: Puede exportar solo datos relacionales, no datos OLAP.

Para exportar datos a Cognos BI, debe especificar lo siguiente:

- **Conexión Cognos.** La conexión con el servidor Cognos BI.
- **Conexión ODBC.** La conexión con el servidor de datos que utiliza el servidor Cognos BI.

Las conexiones deben apuntar a la misma base de datos y el nombre de usuario y la contraseña para que se conecte el servidor Cognos a ODBC deben ser iguales que los detalles de ODBC.

Puede exportar los datos reales al servidor de datos y los metadatos de paquete al servidor Cognos BI.

1. Seleccione el **Destino** del Cognos BI Server y pulse **Conectar**; se le pedirá que entre los detalles de conexión del servidor.
 - a. Escriba el **URL de servidor** del servidor IBM Cognos desde el cual importar o exportar datos. Póngase en contacto con el administrador del sistema Cognos si no sabe qué URL utilizar.
 - b. Seleccione el **Modo** en el cual se conectará. Para iniciar una sesión como un usuario específico (por ejemplo, un administrador), seleccione **Establecer credenciales** y entre el espacio de nombre, el nombre de usuario y la contraseña de IBM Cognos:
 - Entre el **Espacio de nombres** del proveedor de la autenticación de seguridad de IBM Cognos utilizado para iniciar sesión en el servidor. El proveedor de autenticación se utiliza para definir y mantener usuarios, grupos y roles y para controlar el proceso de autenticación.
 - Entre el **Nombre de usuario** de IBM Cognos con el que iniciará una sesión en el servidor.

- Entre la **Contraseña** asociada al nombre de usuario especificado.

O

Examine y seleccione una credencial de IBM Cognos almacenada en IBM SPSS Collaboration and Deployment Services Repository.

- c. De forma alternativa, para iniciar una sesión sin credenciales de usuario, en cuyo caso no podrá especificar los campos de espacio de nombres, nombre de usuario y contraseña, seleccione **Utilizar conexión anónima**.

Nota: Algunas conexiones de servidor no permiten conexiones anónimas.

2. Seleccione el **Origen de datos**.
3. Escriba la vía de acceso y el nombre de la **carpeta** en el servidor Cognos BI donde se va a crear el paquete de exportación. De forma alternativa, puede buscar el que necesite.
4. Escriba el **nombre de paquete** del paquete que va a contener los metadatos exportados. Este debe ser un paquete nuevo; no puede exportar a un paquete existente. Consulte el tema “Selección de detalles de objeto Cognos” en la página 41 si desea más información.
5. Para seleccionar el destino ODBC, especifique el **nombre de base de datos** o busque el que necesita.
6. Especifique cómo se van a almacenar las puntuaciones. Puede optar por crear una tabla nueva, en cuyo caso se le pedirá que entre un nombre para la tabla. De forma alternativa, puede añadirlos a una tabla existente o sobrescribir un origen de datos existente, en cuyo caso se le pedirá que seleccione la tabla o el origen de datos relevante.

Al añadir a, o sobrescribir, una base de datos existente, podrá correlacionar los campos de puntuación con campos del destino existente. Consulte el tema “Correlacionar campos de puntuación” si desea más información.

Correlacionar campos de puntuación

Al añadir o sobrescribir puntuaciones en una tabla o archivo de base de datos existente, es posible que tenga que correlacionar los campos de puntuación con campos de la tabla o el archivo existente. Tenga en cuenta que la salida y el campo correlacionado asociado deben tener el mismo tipo de datos.

- Al escribir puntuaciones en una tabla de base de datos existente, todos los campos se deben correlacionar con campos existentes para que pueda continuar la puntuación, ya sea sobrescribiendo o añadiendo las puntuaciones nuevas. Si es necesario, pulse **Correlacionar campos** para especificar salidas para cualquier campo que quede.
- Al escribir en un archivo existente, la correlación solo es necesaria si está seleccionado **Añadir** en las opciones del archivo de salida. Si se sobrescribe un archivo existente, la correlación no es necesaria ya que el archivo nuevo simplemente sustituirá el antiguo.

1. Pulse **Correlacionar campos**. Se abre un diálogo y el sistema realiza la mejor predicción automáticamente en los campos de correlación. En cualquier punto, puede volver a la mejor predicción del sistema pulsando **Restablecer**.
2. Seleccione entre los destinos de salida disponibles y pulse **Correlacionar** para correlacionarlos con los campos de puntuación necesarios.
3. Cuando haya acabado, pulse **Guardar**.

Selección de campos de salida

En la pestaña Puntuar, seleccione **Especificar campos de salida** para elegir los campos para incluir con la salida de puntuación. Además de las puntuaciones reales, normalmente querrá al menos un campo que le permita identificar cada registro como, por ejemplo, un campo de id de cliente. De forma opcional, algunos o todos los campos de entrada utilizados en la creación de modelos también se pueden incluir.

Campos de salida del modelo

Por convenio, los nombres de los campos de salida de puntuación se basan en el nombre del campo objetivo, pero con un prefijo personalizado añadido. Por ejemplo, si el campo objetivo se llama *Respuesta*,

los campos de salida tendrán nombres como *\$XF-Respuesta* y *\$XFRP-Respuesta*. Los campos específicos dependen del tipo de modelo, y el nivel de medición para el campo objetivo seleccionado, como se detalla a continuación. Consulte el tema "Niveles de medición" en la página 44 si desea más información.

Objetivos de marca. Los modelos automatizados con objetivos de marca tendrán como salida los campos siguientes:

- **\$XF-<objetivo>**. El valor predicho para cada registro, notificado en términos de valores "verdaderos" y "falsos" definidos para el campo.
- **\$XFRP-<objetivo>**. La puntuación de propensión; indica la probabilidad del valor "verdadero" para cada registro. Las puntuaciones de propensión difieren de las puntuaciones de confianza, que normalmente se notifican en términos de la predicción para dicho registro. Por ejemplo, una predicción de "falso" con una confianza alta se traducirá en una propensión baja ya que esto indicaría una probabilidad alta de *no* responder.

Objetivos categóricos. Los modelos automatizados con objetivos categóricos (nominal o ordinal) tendrán como salida los campos siguientes:

- **\$XS-<objetivo>**. Valor predicho para cada registro.
- **\$XSC-<objetivo>**. Valor de confianza asociado a la predicción

Objetivos continuos. Los modelos automatizados con objetivos numéricos continuos tendrán como salida los campos siguientes:

- **\$XR-<objetivo>**. Valor predicho para cada registro.
- **\$XRE-<objetivo>**. Error estándar para la predicción.

Modelos de regla predictiva (interactiva). Los modelos de regla predictiva (solo objetivos de marca o categóricos) tendrá como salida los campos siguientes:

- **\$D-<objetivo>**. Resultado predicho para cada registro, con el valor "verdadero" definido para el campo utilizado para indicar la respuesta deseada, y nulo que indica cualquier otro valor.
- **\$DP-<objetivo>**. La probabilidad del resultado deseado para cada registro; coincide con la probabilidad mostrada para el segmento en la tabla de resultados al generar el modelo. Consulte el tema "Creación de un modelo de regla predictiva" en la página 59 si desea más información.
- **\$DI-<objetivo>**. Número de índice utilizado para identificar el segmento en el cual se coloca el registro (se muestra en la columna más a la izquierda de los resultados de modelado).
- **\$DRP-<objetivo>**. La puntuación de propensión que indica la probabilidad de la respuesta deseada. Disponible solo para objetivos de marca.

Modelos de puntuación

Una vez que haya especificado los detalles de puntuación, podrá tener una vista previa de una muestra de las puntuaciones de modelo. Esto es especialmente útil cuando tiene un conjunto de datos grande, porque le permite comprobar rápidamente que los resultados de la puntuación aparecen como espera, sin tener que esperar a que se puntúe todo el conjunto de datos.

Nota: Si se ha encontrado que los datos no son válidos o no están completos, se visualiza una advertencia cuando intente tener una vista previa.

La vista previa muestra los datos como una tabla que muestra los registros por fila y las variables por columnas.

Puntuación

Para establecer cuántos registros se puntuarán, pulse el icono para renovar. Esto proporciona un recuento del número de registros que se utilizarán para generar el modelo, teniendo en cuenta el número total de registros en el origen de datos y cualquier valor opcional que especifique.

Cuando se pulsa **Puntuar ahora**, se muestra una barra de progreso en la parte inferior de la pestaña. Puede detener la puntuación en cualquier momento mientras se visualiza.

Notas:

Si se ha encontrado que los datos no son válidos o están incompletos, se visualiza una advertencia cuando se intentan puntuar.

Cuando se puntuó en una base de datos, la ejecución se retrotrae a la base de datos, siempre que sea posible (a menudo, se hace referencia como "retrotracción SQL"). Tenga en cuenta que si el proyecto incluye una regla que contiene una referencia a un modelo, la ejecución no se retrotraerá a la base de datos.

Una vez que los datos se han puntuado correctamente, un mensaje le informa dónde se han almacenado los resultados, ya sea en la base de datos o el archivo que haya especificado, y la sección Informe de resumen de puntuación proporciona un resumen. Consulte el tema "Informe de resumen de puntuación" si desea más información. Si elige enviar la salida de puntuación a un archivo sin formato, y si un administrador ha habilitado las funciones, puede descargar una copia local del archivo, si lo desea.

Informe de resumen de puntuación

Tras puntuar un proyecto, la sección del informe de resumen de puntuación proporciona detalles como, por ejemplo, dónde se han puntuado los datos, tamaño de archivo (si es aplicable), base de datos y nombre de table (si es aplicable), nombre de modelo o nombre de proyecto, número de registros puntuados y, en algunos casos, un gráfico y/o tabla que resume el modelo. La sección Informe de resumen de puntuación está disponible en IBM SPSS Modeler Advantage y en la pestaña Desplegar de aplicaciones que incluyen la característica Puntuar ahora.

Si opta por enviar la salida de puntuación a un archivo sin formato, también puede descargar una copia local del archivo, si lo desea. Los administradores pueden utilizar un IBM SPSS Collaboration and Deployment Services Deployment Manager basado en el navegador para habilitar o inhabilitar la característica, y para aumentar o reducir el límite máximo permitido del tamaño de archivo. Si el botón **Descargar archivo** está inactivo, el tamaño del archivo excede el límite máximo de tamaño definido por el administrador, o el administrador ha inhabilitado la característica. La característica está inhabilitada de forma predeterminada.

Capítulo 8. Administración de aplicaciones

Los administradores pueden controlar la forma de interactuar los usuarios empresariales con aplicaciones IBM Analytical Decision Management. Este nivel de control hace que los trabajos de los usuarios empresariales sean más fáciles, ya que pueden utilizar de forma más rápida y eficaz la aplicación para resolver problemas empresariales. También permite a los administradores restringir el acceso a determinadas funciones.

Utilice IBM SPSS Collaboration and Deployment Services Deployment Manager para otorgar autoridad administrativa a usuarios o grupos específicos, como sea necesario. Solo los usuarios a los que se ha otorgado la acción *Administración de Decision Management* tienen acceso a las características descritas en esta sección.

1. Abra IBM SPSS Collaboration and Deployment Services Deployment Manager.
2. Desde el menú Herramientas, seleccione **Administración de servidores**.
3. En la pestaña Administración de servidores, conéctese al servidor.
4. En **Usuarios y grupos**, seleccione **Repositorio de usuario local**.
5. Pulse **Nuevo grupo** para crear un grupo nuevo para administradores y un grupo nuevo para usuarios empresariales. Por ejemplo:
 - Administradores de Decision Management
 - Usuarios de Decision Management

También puede crear usuarios nuevos en este momento, si lo desea. Al crear los grupos, añada los usuarios apropiados a cada grupo. Por ejemplo, a miembros del grupo de administradores se les otorgará la autorización para administrar aplicaciones IBM Analytical Decision Management en la organización.

6. Vaya a Roles y pulse **Nuevo rol** para crear un nuevo rol administrativo y un nuevo rol de usuario empresarial. Por ejemplo:
 - Administrador de IBM Analytical Decision Management
 - Usuario de IBM Analytical Decision Management

Al crear los nuevos roles, asígneles las acciones siguientes, como mínimo.

Tabla 7. Acciones de rol.

Rol	Acciones necesarias
Administrador de Decision Management	Acceso a contenido y carpetas
	Administración de Decision Management
	Edición de trabajos
	Ejecutar informe dinámicamente
	Modelo de puntuación
	Enviar trabajo
	Mostrar todas las versiones
Usuario de Decision Management	Acceso a contenido y carpetas
	Edición de trabajos
	Ejecutar informe dinámicamente
	Modelo de puntuación
	Enviar trabajo
	Mostrar últimas O Mostrar todas las versiones*

**Nota:* Es necesaria la acción *Mostrar últimas* o *Mostrar todas las versiones* si los usuarios necesitan acceder a versiones de proyecto que no han creado. De lo contrario, los proyectos se abrirán en el modo de solo lectura si el usuario actual no tiene acceso a la última versión.

7. Seleccione el rol de administrador de Decision Management y pulse **Editar usuarios y grupos**. Añada el grupo de administradores de Decision Management que acaba de crear.
8. Seleccione el rol de usuario de Decision Management y pulse **Editar usuarios y grupos**. Añada el grupo de usuarios de Decision Management que acaba de crear.

Opciones generales

Están disponibles determinadas opciones generales para los administradores en todas las pestañas de la interfaz de usuario de IBM Analytical Decision Management. Estas opciones permiten a los administradores ocultar o bloquear pestañas de los usuarios empresariales.

- La parte superior de cada pestaña tiene la opción **Ocultar pestaña <nombre pestaña> a no administradores**, donde **<nombre pestaña>** es el nombre de una pestaña. Así se oculta toda la pestaña a los usuarios empresariales. Por ejemplo, si no desea que los usuarios vean la pestaña Datos para una aplicación, inicie una sesión en la aplicación, vaya a la pestaña Datos y seleccione **Ocultar pestaña Datos a no administradores**.
- La parte superior de cada pestaña tiene la opción **Bloquear todas las opciones de <nombre pestaña>**, donde **<nombre pestaña>** es el nombre de la pestaña. Así se bloquean todos los controles de la pestaña, de forma que son de solo lectura para el usuario empresarial. Por ejemplo, si desea que los usuarios empresariales puedan ver los valores de datos actuales, pero no cambiarlos, podrá definir esta opción en la pestaña Datos de una aplicación. Este valor también es muy útil para situaciones en las que desee definir determinadas cosas de antemano para simplificar el trabajo del usuario empresarial.
- Los administradores también pueden bloquear y desbloquear valores individuales en cada pestaña pulsando el icono de candado, cuando está disponible.



Por ejemplo, en aplicaciones con una pestaña Definir, los administradores pueden bloquear el árbol de jerarquía de dimensión, de forma que los usuarios empresariales no pueden cambiarlo. Si pulsa el icono de candado, se bloquea o desbloquea su control asociado. Los usuarios empresariales verán el icono de candado cerrado para los controles que bloquea el administrador, pero nunca verán un icono de candado abierto.

Los elementos individuales que están bloqueados también se pueden visualizar con un fondo en gris o colores apagados para indicar que no están disponibles para la selección; se muestra un mensaje de error si el usuario intenta seleccionarlos. Cuando está bloqueada toda una pestaña, se muestra un mensaje a tal efecto en la parte superior de la pestaña.

Control de atajos de aplicación

La página de inicio de aplicaciones proporciona una forma fácil para que los usuarios empresariales abran aplicaciones y añadan atajos personalizados a versiones etiquetadas específicas de aplicaciones. Los administradores pueden especificar qué atajos de aplicación se mostrarán en la página de inicio para todos los usuarios. Tras iniciar la sesión con la autorización de administrador:

- Si lo desea, pulse el icono del candado en un atajo de aplicación, de forma que los usuarios empresariales no pueden eliminarlo de su página de inicio. Tenga en cuenta que los usuarios empresariales pueden seguir añadiendo y eliminando otros atajos de aplicaciones. Los usuarios verán el icono de candado cerrado, en lugar del icono para cerrar para los atajos que están bloqueados. Consulte el tema “Inicio de aplicaciones” en la página 25 si desea más información.

Ocultar la página de inicio

Cada aplicación proporciona un flujo de trabajo paso a paso, tal como se representa mediante los gráficos de la página de inicio. Si pulsa un gráfico saltará a dicha sección de la interfaz de usuario. Sin embargo, es posible que los usuarios empresariales no siempre tengan que acceder a cada paso del flujo de trabajo de una aplicación.

Los administradores pueden ocultar la página de inicio a los usuarios empresariales. Por ejemplo, en una aplicación determinada, los usuarios empresariales podrían solo utilizar la pestaña Informes. En este caso, un administrador podría ocultar la página de inicio y todas las demás pestañas a los usuarios empresariales. Así pues, cuando el usuario empresarial inicia una sesión o va directamente a la pestaña Informes, nunca ve ninguna de las otras pestañas.

1. En la página de inicio, seleccione **Ocultar página de inicio de aplicación**.
2. En el desplegable, especifique qué pestaña mostrar cuando los usuarios abren por primera vez la aplicación.

Tenga en cuenta que si, posteriormente, oculta esta pestaña seleccionada, se muestra la siguiente pestaña no oculta de la aplicación.

Bloqueo de opciones de datos

Datos

La pestaña Datos define los conjuntos de datos que utiliza la aplicación para el análisis, la simulación y realización de pruebas, la puntuación y otros usos operativos. Estos orígenes de datos se definen en relación con la instalación del servidor; por lo tanto, todas las vías de acceso de los archivos de datos, orígenes ODBC y etc. están todos relacionados con el servidor.

Como administrador, es posible que desee configurar algunos o todos los datos antes de que los usuarios empresariales empiecen a utilizar sus aplicaciones. Los administradores pueden:

- Seleccionar **Bloquear origen de datos** al crear un nuevo origen de datos para bloquearlo una vez creado.
- Bloquear y desbloquear orígenes de datos individuales pulsando el icono de candado junto a un origen de datos de la tabla, o en el diálogo del editor de origen de datos.
- Definir y bloquear el modelo de datos utilizado en la aplicación seleccionando un modelo de datos en el desplegable **Modelo de datos de proyecto** y pulsando el icono del candado.

Tenga en cuenta que los usuarios empresariales pueden copiar un origen de datos bloqueado de los conjuntos de datos de proyecto en su sección Mis conjuntos de datos, pero permanece bloqueado en ambos conjuntos. Consulte el tema Capítulo 4, “Gestión de orígenes de datos”, en la página 33 si desea más información.

- Si el modelo de datos de proyecto es una *Vista de datos*, utilice el desplegable **Permitir el uso de reglas ODM basándose en los datos jerárquicos** para controlar el uso de las reglas IBM Operational Decision Management (ODM) que tienen un modelo de datos jerárquico. El uso de dichas reglas puede provocar un bajo rendimiento. Las opciones siguientes están disponibles en el desplegable:
 - Se recomienda **No permitir** para obtener el mejor rendimiento. Impide que se utilicen reglas ODM con una estructura de datos compleja en cualquier lugar del proyecto. Se visualiza un error si un usuario selecciona una regla compleja en cualquier punto de la aplicación.
 - **Solo en la pestaña Modelado** permite que se haga referencia a las reglas complejas al crear modelos en la pestaña Modelado. Esto permite una flexibilidad añadida, con un acierto de rendimiento. Esta opción solo se aplica a IBM SPSS Modeler Advantage.

Bloqueo de opciones de destino de puntuación

Para aplicaciones configuradas para utilizar la sección Puntuar ahora (pestaña Desplegar o Puntuar), hay una sección que permite a los usuarios empresariales especificar el destino para almacenar datos de puntuación (en una base de datos o archivo). Por ejemplo, si se almacenan las puntuaciones en una base de datos, especifique el nombre de la base de datos y el método para almacenar las puntuaciones (como, por ejemplo, la creación de una tabla nueva en la base de datos).

Como administrador, puede bloquear las opciones de destino para la puntuación. Por ejemplo, es posible que desee que las puntuaciones solo se añadan a una tabla existente de una determinada base de datos, en lugar de dar a los usuarios empresariales la opción de sobrescribir datos existentes, añadir datos nuevos a una tabla existente, crear una tabla nueva o almacenar datos en un archivo.

- Para bloquear todas las opciones de destino de puntuación, pulse el icono del candado junto al desplegable Destino en la pestaña Desplegar. Los usuarios empresariales podrán ver las opciones de destino, pero no editarlas.

Descarga de metadatos de proyecto para su uso en la creación de reglas externas

Se puede hacer referencia a las reglas creadas en sistema de gestión de reglas empresariales como, por ejemplo, IBM Operational Decision Management, y se pueden utilizar en aplicaciones IBM Analytical Decision Management, siempre que se hayan desarrollado para dar soporte al mismo modelo de datos utilizado en el proyecto IBM Analytical Decision Management actual, y se pueden desplegar como un servicio web para ser utilizado por la aplicación. Para ello, un administrador de IBM Analytical Decision Management debe descargar los metadatos relevantes. Existen dos formas para descargar los metadatos, y la complejidad del modelo de datos y la complejidad de la posible regla resultante difieren en consecuencia:

Descarga de metadatos de proyecto desde IBM Analytical Decision Management

Un administrador puede pulsar el icono **Descargar metadatos** en IBM Analytical Decision Management para descargar un archivo *.ZIP* que contiene metadatos para el proyecto actual. El archivo descargado se guarda localmente en el sistema utilizado por el administrador para desencadenar la descarga.



Figura 14. Icono Descargar metadatos

Los metadatos descargados de esta forma soportan la creación de reglas basándose en el modelo de datos del proyecto. El modelo de datos del proyecto será plano (no jerárquico), así que las reglas creadas según estos metadatos solo pueden utilizar un modelo de datos sin formato. Se puede hacer referencia a una regla creada de acuerdo con este modelo de datos sin formato en un proyecto de IBM Analytical Decision Management, independientemente del tipo de origen de datos del modelo de datos de proyecto.

Los archivos siguientes se incluyen en el archivo *.ZIP*:

- **Definición de esquema XML (*.XSD)**. Contiene definiciones de los tipos de campo en el modelo de datos para el proyecto actual de IBM Analytical Decision Management. Este archivo se puede importar en una herramienta de desarrollo externa como, por ejemplo, IBM Operational Decision Management Rule Designer, para poder desarrollar reglas para ser utilizadas con el proyecto actual.
- **Plantilla del lenguaje de descripción de servicio web (*.WSDL)**. Se incluye para ayudar en el desarrollo de servicios web que utilizan el modelo de datos. El archivo *.WSDL* se proporciona como una plantilla que se puede utilizar para crear servicios web responsables para mediar mensajes entre la

aplicación IBM Analytical Decision Management y el servicio de regla externa. El integrador puede proporcionar manualmente detalles sobre la mediación del servicio web.

Descarga de los metadatos de modelo de datos para una vista de datos analíticos de IBM SPSS Collaboration and Deployment Services

En IBM SPSS Collaboration and Deployment Services Deployment Manager, puede exportar metadatos para una vista de datos analíticos particular.

Los metadatos descargados de esta forma soportan la creación de reglas basándose en el modelo de datos lógico de una vista de datos analíticos. El modelo de datos lógico de una vista de datos analíticos puede ser jerárquico. Un origen de vista de datos de IBM Analytical Decision Management soporta la generación de una representación XML de los datos jerárquicos para cada registro, que se pueden pasar después a reglas externas durante la ejecución. Las reglas creadas basándose en estos metadatos pueden ser complejas y utilizar el modelo de datos jerárquico.

En IBM Analytical Decision Management, solo se puede hacer referencia a una regla externa basada en un modelo de datos jerárquico en un proyecto donde el modelo de datos de proyecto utiliza un origen de vista de datos. Para garantizar la compatibilidad, el origen de la vista de datos y la regla externa se deben basar en el mismo modelo de datos lógico de la vista de datos analíticos. Solo el origen de la vista de datos puede generar los datos jerárquicos basándose en la instancia de la vista de datos analíticos a los que hace referencia. En otras palabras, para utilizar reglas externas, el modelo de datos del proyecto del proyecto de IBM Analytical Decision Management debe coincidir con el que se ha utilizado para crear la regla, y la vista de datos analíticos hace que esto sea posible.

Para descargar un archivo .ZIP con metadatos para el modelo de datos lógico de la vista de datos analíticos:

1. En IBM SPSS Collaboration and Deployment Services Deployment Manager, abra una vista de datos analíticos desde el repositorio y vaya a la pestaña **Lógica**.
2. Pulse con el botón derecho del ratón una tabla y exporte los metadatos XOM.
3. Seleccione una ubicación para guardar el archivo y pulse **Guardar**.

Los archivos siguiente se incluyen en el archivo ZIP:

Definiciones de esquema XML (*.XSD). Se incluyen dos o más definiciones de esquema XML. Uno contiene una definición que puede ser útil para un tipo de respuesta de regla. Los otros esquemas contienen las definiciones de las tablas, incluyendo detalles de la relación jerárquica y tipos de campo en el modelo de datos lógico de la vista de datos analíticos.

Puede importar estos archivos de esquema en una herramienta de desarrollo externa como, por ejemplo, IBM Operational Decision Management Rule Designer, para poder desarrollar reglas para ser utilizadas con un proyecto cuyo modelo de datos utiliza un origen de vista de datos basada en la misma instancia de vista de datos analíticos.

Si desea más información sobre cómo crear reglas externas para ser utilizadas en IBM Analytical Decision Management, consulte la *Guía de diseñadores de aplicaciones*.

Tenga en cuenta que los administradores pueden restringir el uso de estas reglas externas complejas porque pueden degradar el rendimiento. Si desea más detalles, consulte “Bloqueo de opciones de datos” en la página 83.

Edición de la definición de esquema XML

El archivo .XSD descargado desde IBM Analytical Decision Management contiene dos tipos de datos complejo: uno que representa el modelo de datos de proyecto de IBM Analytical Decision Management, que se debe utilizar para el parámetro de entrada de proyectos de reglas de IBM Operational Decision

Management; y uno que representa un tipo de devolución de reglas, que se debe utilizar para el parámetro de salida para proyectos de reglas de IBM Operational Decision Management.

Los metadatos de la vista de datos analíticos de IBM SPSS Collaboration and Deployment Services Deployment Manager contiene dos o más archivos *.XSD*: uno que está relacionado con el modelo de datos lógicos de una vista de datos analíticos, que se debe utilizar para el parámetro de entrada de proyectos de reglas de IBM Operational Decision Management; y uno o varios que contienen un tipo completo que representa un tipo de retorno de reglas, que se debe utilizar para el parámetro de salida de proyectos de reglas de IBM Operational Decision Management.

Los usuarios expertos pueden optar por editar la definición de esquema *.XSD* generada para adaptarla para la regla o reglas específicas que desean crear. Existen circunstancias cuando el usuario desea editar la entrada y/o los tipos relacionados con la respuesta definidos en la definición de esquema *.XSD* relacionada antes de crear sus reglas. Por ejemplo, el usuario experto puede tener que editar el esquema relacionado con el parámetro de entrada para permitir que sus reglas creadas manejen los posibles campos de entrada con valores nulos. El usuario experto puede hacer esto indicando a través de la definición del esquema *.XSD* qué elementos/campos de parámetro de entrada son necesarios y cuáles son opcionales.

Edición del tipo de parámetro de entrada

El número de posibles apariciones de cada elemento relacionado con el campo no está establecido explícitamente en el esquema *.XSD* (es decir, los atributos *minOccurs* y *maxOccurs* no se establecen y el valor predeterminado para ambos es 1). Como tal, si el esquema no se ha editado antes de utilizarlo para generar reglas de IBM Operational Decision Management, todos los campos del tipo de parámetro de entrada serán necesarios para las reglas generadas.

Edición del tipo de parámetro de entrada relacionado con modelo de datos de proyecto

Si el esquema que define el tipo de parámetro de entrada se ha generado descargando metadatos de un proyecto de IBM Analytical Decision Management, el modelo de datos del parámetro de entrada no tendrá formato. En este caso, IBM Analytical Decision Management creará directamente la carga útil necesaria para el servicio de reglas en ejecución. Esto proporciona al usuario experto más libertad al editar el tipo de parámetro de entrada y seguir manteniendo la compatibilidad (es decir, el usuario experto puede eliminar campos/elementos que no necesitará para la creación de reglas, o indicar campos/elementos que son opcionales).

Edición del tipo de parámetro de entrada relacionado con el modelo de datos lógico de la vista de datos analíticos

Si el esquema que define el tipo de parámetro de entrada se ha generado exportando un modelo de datos lógicos de la vista de datos analíticos desde el cliente IBM SPSS Collaboration and Deployment Services Deployment Manager, el modelo de datos del parámetro de entrada puede ser jerárquico. Si es jerárquico, IBM Analytical Decision Management *no* creará directamente la carga útil que necesita el servicio de reglas en la ejecución. El origen de la vista de datos analíticos generará la parte principal de la carga útil. Esto limita la forma en la que el usuario experto puede editar el tipo de parámetro de entrada y mantener la compatibilidad (por ejemplo, el usuario experto solo puede indicar campos/elementos que son opcionales).

Información general

Antes de invocar a una regla creada en un sistema de gestión de reglas empresariales como, por ejemplo, IBM Operational Decision Management, IBM Analytical Decision Management pregunta al servicio de la regla si su modelo de datos de parámetro es jerárquico o sin formato. Si el parámetro de entrada no tiene formato, también comprueba qué entradas necesita el servicio de la regla.

Si el modelo de datos del parámetro de entrada no tiene formato y el registro para puntuar contiene un valor nulo para cualquiera de los campos necesarios del servicio de la regla, IBM Analytical Decision Management no invocará la regla, y los campos de salida/respuesta de la regla se llenan con valores predeterminados. El valor depende del tipo del campo de salida/respuesta, como se indica a continuación:

Tabla 8. Valores para tipos de campo de respuesta.

Tipo de campo de respuesta	Valor
xsd:string	""
xsd:boolean	""
xsd:int	0
xsd:long	0
xsd:byte	0
xsd:double	0.0
xsd:float	0.0
xsd:date	0.0
xsd:time	0.0
xsd:datetime	0.0

Esto puede afectar a los resultados de los registros de puntuación a través de un servicio de regla, y se debe estar muy atento al editar el parámetro de entrada en el esquema *.XSD*.

Si el modelo de datos del parámetro de entrada es jerárquico, y el registro que se va a puntuar contiene un valor nulo para cualquiera de los campos necesarios del servicio de regla, IBM Analytical Decision Management invocará la regla, pero el campo o campos con un valor nulo se omitirán de la carga útil creada por el nodo de origen de la vista de datos. Si el campo omitido era un campo necesario, esto provocará que falle la invocación de la regla, así que se debe volver a extremar las precauciones al editar el parámetro de entrada en el esquema *.XSD*.

La edición del esquema requiere conocimientos de los servicios de regla que se crearán. Si la regla no necesita un campo de entrada particular y el modelo de datos del parámetro de entrada de la regla **no tiene formato**, el elemento relacionado con el campo se puede **eliminar** del parámetro de entrada en el esquema o **marcarse como opcional**. Si un campo particular no es necesario para la regla, y el parámetro de entrada de la regla es **jerárquico**, el elemento relacionado con el campo **no se puede eliminar**; pero se puede **marcar como opcional**. Esto se debe a que IBM Analytical Decision Management no genera directamente la carga útil para reglas basadas en un modelo de datos jerárquico.

Cuando se elimina un campo de parámetro de entrada, se ignora el valor del campo cuando se llama a la regla y la regla no puede hacer referencia al campo. Al convertir un campo de parámetro de entrada en opcional, está permitido un valor nulo y la regla se debe escribir a la defensiva para comprobarlo. Por ejemplo, si el modelo de datos del parámetro de entrada de la regla no tiene formato y sabe que la regla no utilizará el campo Edad, puede eliminar el elemento relacionado para indicar que no es un campo necesario. En el siguiente ejemplo de esquema *.XSD*, si el campo Edad no es necesario, se puede eliminar toda la sección `<xsd:element name="Age">` del esquema:

```
<xsd:complexType name="drugs">
  <xsd:annotation>
    <xsd:appinfo>
      <dmname>drugs</dmname>
    </xsd:appinfo>
  </xsd:annotation>
  <xsd:sequence>
    <xsd:element name="Age">
```

```

<xsd:annotation>
  <xsd:appinfo>
    <dmname>Age</dmname>
  </xsd:appinfo>
</xsd:annotation>
<xsd:simpleType>
  <xsd:restriction base="xsd:long">
    <xsd:minInclusive value="15"/>
    <xsd:maxInclusive value="74"/>
  </xsd:restriction>
</xsd:simpleType>
</xsd:element>
<xsd:element name="Sex">
  <xsd:annotation>
    <xsd:appinfo>
      <dmname>Sex</dmname>
    </xsd:appinfo>
  </xsd:annotation>
  <xsd:simpleType>
    <xsd:restriction base="xsd:string">
      <xsd:enumeration value="F"/>
      <xsd:enumeration value="M"/>
    </xsd:restriction>
  </xsd:simpleType>
</xsd:element>
...
</xsd:sequence>
</xsd:complexType>

```

O bien, si el modelo de datos del parámetro de entrada de la regla no tiene formato o es jerárquico, podría desea convertir un elemento de campo de entrada en opcional si las reglas no lo van a necesitar, definiendo el valor del atributo `minOccurs` del elemento en 0:

```

<xsd:complexType name="drugs">
  <xsd:annotation>
    <xsd:appinfo>
      <dmname>drugs</dmname>
    </xsd:appinfo>
  </xsd:annotation>
  <xsd:sequence>
    <xsd:element name="Age" minOccurs="0">
      <xsd:annotation>
        <xsd:appinfo>
          <dmname>Age</dmname>
        </xsd:appinfo>
      </xsd:annotation>
      <xsd:simpleType>
        <xsd:restriction base="xsd:long">
          <xsd:minInclusive value="15"/>
          <xsd:maxInclusive value="74"/>
        </xsd:restriction>
      </xsd:simpleType>
    </xsd:element>
    <xsd:element name="Sex">
      <xsd:annotation>
        <xsd:appinfo>
          <dmname>Sex</dmname>
        </xsd:appinfo>
      </xsd:annotation>
      <xsd:simpleType>
        <xsd:restriction base="xsd:string">
          <xsd:enumeration value="F"/>
          <xsd:enumeration value="M"/>
        </xsd:restriction>
      </xsd:simpleType>
    </xsd:element>
    ...
  </xsd:sequence>
</xsd:complexType>

```

Ejemplo de modelo de datos jerárquico:

```

<xsd:complexType name="Customer">
  <xsd:sequence>
    <xsd:element name="FirstName" type="xsd:string" minOccurs="0"/>
    <xsd:element name="Surname" type="xsd:string" minOccurs="0"/>
    <xsd:element name="Age" type="xsd:long"/>
    <xsd:element name="Sex" type="xsd:string"/>
    <xsd:element name="CustID" type="xsd:string"/>
    <xsd:element maxOccurs="unbounded" minOccurs="0" name="Orders" type="xom:Order"/>
  </xsd:sequence>
</xsd:complexType>
<xsd:complexType name="Order">
  <xsd:sequence>
    <xsd:element name="CustID" type="xsd:string"/>
    <xsd:element name="NoOfItems" type="xsd:long"/>
    <xsd:element name="TotalCost" type="xsd:long" minOccurs="0"/>
    <xsd:element name="Date" type="xsd:date"/>
  </xsd:sequence>
</xsd:complexType>

```

Las reglas también se pueden escribir a la defensiva para poder manejar valores nulos, en cuyo caso el campo se puede marcar como opcional en el esquema. A continuación aparece un ejemplo de una regla de IBM Operational Decision Management llamada *currentOfferIsReplacePhone* que se ha escrito para tratar con valores nulos:

```

if
  the CURRENT OFFER of 'the customer' is not null
  and the CURRENT OFFER of 'the customer' is "Replace Phone"
then
  make it true that 'the response' is value ;
else
  make it false that 'the response' is value ;

```

Edición del tipo de parámetro de respuesta

Tenga en cuenta que posible crear y utilizar reglas de IBM Operational Decision Management que devuelven una variedad de tipos a IBM Analytical Decision Management (*booleano* para su uso en selecciones, *numérico* para su uso en reglas de agregación, por ejemplo, y *serie*). Sin embargo, para crear cualquier cosa distinta a una regla de respuesta de datos de *serie*, la definición de esquema XML (.XSD) que contiene el tipo de respuesta se debe editar manualmente.

Además, un proyecto de reglas puede contener varias reglas, cada una de las cuales devuelve tipos diferentes. Es necesario mantener varias versiones del .XSD para dar soporte a esto (por ejemplo, una versión donde la variable de respuesta de datos se ha definido en numérica, otra para serie, otra para booleano, etc.).

El archivo .XSD descargado desde IBM Analytical Decision Management contiene dos tipos de datos complejos: uno que representa el modelo de datos de proyecto de IBM Analytical Decision Management que se debe utilizar para el parámetro de entrada de proyectos de reglas de IBM Operational Decision Management, y uno que representa un tipo de retorno de reglas que se debe utilizar para el parámetro de salida de proyectos de regla de IBM Operational Decision Management.

El archivo .XSD relacionado con la respuesta que se ha descargado desde IBM SPSS Collaboration and Deployment Services Deployment Manager contiene la definición del tipo de respuesta y solo contendrá un tipo de datos complejo.

En ambos casos, el tipo de retorno (cuyo nombre siempre tendrá un sufijo "respuesta") define un elemento de valor que toma de forma predeterminada una *serie*. Si desea crear un proyecto de reglas con un tipo distinto de valor de retorno, deberá editar manualmente el .XSD descargado antes de importar el proyecto de reglas de IBM Operational Decision Management, de forma que el elemento valor del tipo de retorno se define como se desea. A continuación, aparecen algunos ejemplos:

```

<xsd:complexType name="claim_dataresponse">
  <xsd:annotation>
    <xsd:appinfo>
      <dmname>claim_dataresponse</dmname>
    </xsd:appinfo>
  </xsd:annotation>
</xsd:complexType>

```

```
        </xsd:appinfo>
    </xsd:annotation>
    <xsd:sequence>
        <xsd:element name="value" type="xsd:string"/>
    </xsd:sequence>
</xsd:complexType>
```

```
<xsd:complexType name="claim_dataresponse">
    <xsd:annotation>
        <xsd:appinfo>
            <dmname>claim_dataresponse</dmname>
        </xsd:appinfo>
    </xsd:annotation>
    <xsd:sequence>
        <xsd:element name="value" type="xsd:boolean"/>
    </xsd:sequence>
</xsd:complexType>
```

```
<xsd:complexType name="claim_dataresponse">
    <xsd:annotation>
        <xsd:appinfo>
            <dmname>claim_dataresponse</dmname>
        </xsd:appinfo>
    </xsd:annotation>
    <xsd:sequence>
        <xsd:element name="value" type="xsd:long"/>
    </xsd:sequence>
</xsd:complexType>
```

Capítulo 9. Obtener ayuda

Se proporcionan tres niveles de asistencia al usuario:

- La ayuda personalizada se proporciona para cada aplicación preconstruida empaquetada con IBM Analytical Decision Management. Para lanzar la ayuda, pulse el icono de ayuda en cualquier pantalla o diálogo:



- Texto de ayuda, que se muestra en una ventana emergente pequeña, que proporciona detalles sobre la aplicación actual o un problema empresarial. El texto de ayuda proporciona una capa adicional de ayuda según contexto personalizada que se puede configurar fácilmente para un sitio o aplicación específico. El texto de ayuda aparece en toda la interfaz de usuario cuando se pulsa el icono pequeño de la texto de ayuda:



- Ayudas contextuales estándar, que se proporciona para controles individuales dentro de una página o diálogo. Para mostrar una ayuda contextual, pase el ratón por encima del botón, enlace u otro control.

Documentación disponible

La documentación de IBM Analytical Decision Management está disponible en dos formatos. Cuando se accede a la ayuda de la aplicación, se inicia el IBM Knowledge Center en línea con el correspondiente tema de ayuda visualizado. Si tuviera algún problema al iniciar la ayuda, póngase en contacto con el administrador si carece de la autoridad necesaria para acceder al URL externo de Knowledge Center.

El Knowledge Center es una recopilación en línea centralizada de toda la documentación. Siempre está disponible, incluso cuando el producto no ejecuta. Le recomendamos que lo marque como favorito en el navegador. También contiene documentación de otros productos de IBM como, por ejemplo, IBM SPSS Modeler, IBM SPSS Collaboration and Deployment Services e Statistics.

Cuando no tenga conexión a Internet, o cuando el Knowledge Center no esté disponible de forma temporal por mantenimiento, también podrá disponer de toda la documentación en formato *.pdf (Portable Document Format). Los archivos PDF están disponibles en los archivos de instalación descargados y también se instalan en el directorio de instalación de IBM SPSS Collaboration and Deployment Services (por ejemplo, *C:\Program Files\IBM\SPSS\Deployment\8.0\Server\documentation\DecisionManagement*).

La documentación en PDF también está disponible en la Web, en <http://www-01.ibm.com/support/docview.wss?uid=swg27047202>, y las notas del release están disponibles en <http://www-01.ibm.com/support/docview.wss?uid=swg27047203>.

Apéndice A. Compartir rutas

Los modelos y proyectos creados en IBM Analytical Decision Management se guardan como archivos de ruta, y se pueden modificar y utilizar en IBM SPSS Modeler si están disponibles. Por ejemplo, puede que desee crear una ruta de modelo o proyecto utilizando IBM SPSS Modeler Advantage u otra aplicación IBM Analytical Decision Management y, después, personalizar la ruta en IBM SPSS Modeler antes de su despliegue en IBM Analytical Decision Management. De forma alternativa, las rutas creadas en IBM SPSS Modeler se pueden abrir en IBM Analytical Decision Management, siempre que los valores de despliegue se haya especificado tal como se describe en este apéndice.

Los modelos que siguen el formato descrito en este apéndice se pueden ver, editar, puntuar y renovar utilizando IBM SPSS Modeler Advantage. Otros modelos también se pueden puntuar en IBM SPSS Modeler Advantage, pero no se visualizan correctamente y no se pueden renovar.

Abrir rutas en IBM Analytical Decision Management

Al abrir rutas en IBM Analytical Decision Management o IBM SPSS Modeler Advantage:

- Si el tipo de despliegue no se especifica (ninguno), la ruta se puede abrir, pero solo está disponible la pestaña Datos.
- Si el tipo de despliegue es **Solo puntuación**, la pestaña Puntuación también puede estar disponible, siempre que la rama de puntuación empiece con un nodo de origen de datos único de un tipo soportado por IBM SPSS Modeler Advantage o IBM Analytical Decision Management. En otras palabras, todos los tipos de origen soportados por IBM SPSS Modeler también están soportados por IBM SPSS Modeler Advantage y IBM Analytical Decision Management, siempre que solo haya un nodo de origen en la ruta (no conjuntos de datos fusionados).
- Si el tipo de despliegue es **Actualización de modelo**, la pestaña Modelado también puede estar disponible, siempre que la rama de modelado se inicie con un nodo de origen de datos único de un tipo soportado por IBM SPSS Modeler Advantage o IBM Analytical Decision Management, y el nodo de modelado seleccionado sea un nodo de modelado automático (soportado para el modelado predictivo), un nodo de lista de decisiones (soportado para el modelado de reglas predictivas), un nodo de agrupación en clúster (soportado para el modelo de cluster), o un nodo de regla de asociación (soportado para el modelado de asociación). Además, debe estar presente un nodo tipo que define el campo objetivo, así como otra información utilizada en el modelado.

Especificación de valores de despliegue en IBM SPSS Modeler

En IBM SPSS Modeler, especifique valores de despliegue en el cuadro de diálogo Propiedades de ruta, del modo siguiente:

1. Desde los menús de IBM SPSS Modeler, elija:
Herramientas > Propiedades de ruta > Despliegue
2. Desde la lista **Tipo de despliegue**, elija **Solo puntuación** o **Actualización de modelos**, según proceda.
3. Seleccione un **Nodo de puntuación**, **Nodo de modelado** y un **Nodo de modelo**, según proceda.
4. Seleccione **Comprobar** para confirmar que todas las opciones necesarias se han especificado, o seleccione **Almacenar** para almacenar la ruta en IBM SPSS Collaboration and Deployment Services Repository.

Para las rutas creadas en IBM SPSS Modeler Advantage o IBM Analytical Decision Management, la información de despliegue se define automáticamente.

Consejos para utilizar rutas de IBM SPSS Modeler en IBM Analytical Decision Management

- Puede que sea más fácil crear rutas utilizando IBM SPSS Modeler Advantage o IBM Analytical Decision Management y, después, modificar en IBM SPSS Modeler, en lugar de crearlas de cero utilizando IBM SPSS Modeler. Las rutas creadas de esta forma se estructurarán de una forma que les permite ser totalmente funcionales en IBM Analytical Decision Management.
- Para todas las rutas creadas en IBM SPSS Modeler Advantage, los supernodos se incluyen en determinadas ubicaciones dentro de la rama que puede personalizar un usuario experto. Se recomienda encarecidamente que cualquier personalización o incorporación en la ruta se añada dentro de estos supernodos para asegurarse de que se conservan en la ruta y que los reconoce IBM SPSS Modeler Advantage.
- Cualquier nodo de regla local que contiene una regla de expresión inmediatamente en sentido descendente de cada nodo de origen tendrá sus campos incluidos en el diálogo Añadir/Editar campos adicionales. Tenga en cuenta que las reglas locales solo se pueden crear dentro de IBM Analytical Decision Management o IBM SPSS Modeler Advantage y no mediante IBM SPSS Modeler.
- Si hay un nodo de partición en sentido ascendente del nodo de modelado (y en sentido descendente del nodo tipo), la opción **Partir datos automáticamente para habilitar la evaluación del modelo en el origen de datos de generación para su evaluación y prueba** está habilitada en IBM SPSS Modeler Advantage, y el estado de la casilla de verificación se define en función de si el nodo de partición está habilitado o no.
- Si hay un nodo de preparación automática de datos (ADP) en sentido descendente del nodo tipo, la opción **Limpiar automáticamente y preparar datos para una generación de modelos fiable** está habilitada y el estado de la casilla de verificación se define en función de si el nodo ADP está habilitado o no.
- Si los nodos de regla o selección están presentes en sentido ascendente del nodo de modelado, estos se utilizan para llenar la sección **Especificar selecciones que deben usarse**.
- Si los parámetros de ruta se han definido (pestaña Parámetros en el cuadro de diálogo Propiedades de ruta), se visualizará un enlace de parámetros en las pestañas Modelado y Puntuar en IBM SPSS Modeler Advantage. El usuario puede pulsar este enlace para especificar estos valores.

Lanzamiento de IBM SPSS Modeler Advantage desde IBM SPSS Modeler

Desde la pantalla inicial de IBM SPSS Modeler, puede elegir crear un modelo con IBM SPSS Modeler Advantage.

Apéndice B. Accesibilidad

Las características de accesibilidad ayudan a los usuarios que tienen alguna discapacidad como, por ejemplo, movilidad restringida o visión limitada, para utilizar correctamente productos de tecnología de información. IBM se esfuerza por proporcionar productos con un acceso utilizable para todo el mundo, independientemente de la edad o aptitudes. Este producto utiliza las teclas de navegación web estándar.

Puesto que IBM Analytical Decision Management es una aplicación basada en navegador, las opciones de accesibilidad para usuarios discapacitados como, por ejemplo, usuarios con problemas de visión, se controlan desde la configuración del navegador web. Si desea detalles sobre acciones de interfaz de usuario, incluyendo atajos de teclado, lectores de pantalla, etc., consulte la documentación del navegador web.

Software de proveedor

IBM Analytical Decision Management puede requerir el uso del algún software de proveedor que no está cubierto bajo el acuerdo de licencia de IBM. IBM no es responsable de las características de accesibilidad de estos productos. Póngase en contacto con el proveedor para la información de accesibilidad sobre sus productos.

IBM y la accesibilidad

Visite [IBM Human Ability and Accessibility Center](#) si desea más información sobre el compromiso de IBM con la accesibilidad.

Avisos

Esta información se ha desarrollado para los productos y servicios ofrecidos en Estados Unidos. Este material puede estar disponible en IBM en otros idiomas. Sin embargo, puede que necesite estar en posesión de una copia del producto o de una versión del producto para poder acceder a él.

Es posible que IBM no ofrezca en otros países los productos, servicios o características de los que se informa en este documento. Consulte a su representante local de IBM si desea más información sobre los productos y servicios disponibles actualmente en su zona. Cualquier referencia hecha a un producto, programa o servicio de IBM no implica ni pretende indicar que sólo pueda utilizarse dicho producto, programa o servicio de IBM. En su lugar, se puede utilizar cualquier producto, programa o servicio funcionalmente equivalente que no infrinja ninguno de los derechos de propiedad intelectual de IBM. No obstante, será responsabilidad del usuario evaluar y verificar el funcionamiento de cualquier programa, producto o servicio que no sea de IBM.

Puede que IBM tenga patentes o solicitudes de patente pendientes que cubran la materia descrita en este documento. La entrega de este documento no le otorga ninguna licencia sobre dichas patentes. Puede enviar sus consultas sobre licencias por escrito a:

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

Para consultas sobre licencias relativas a 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.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japón*

INTERNATIONAL BUSINESS MACHINES CORPORATION PROPORCIONA ESTA PUBLICACIÓN "TAL CUAL" SIN GARANTÍA DE NINGUNA CLASE, SEA EXPRESA O IMPLÍCITA, INCLUIDAS, PERO SIN LIMITARSE A ELLAS, LAS GARANTÍAS IMPLÍCITAS DE NO CONTRAVENCIÓN, COMERCIALIZACIÓN O ADECUACIÓN A UN PROPÓSITO DETERMINADO. Algunas jurisdicciones no permiten la renuncia a garantías explícitas o implícitas en determinadas transacciones y, por tanto, puede que esta declaración no se aplique a su caso.

Esta información puede contener imprecisiones técnicas o errores tipográficos. Periódicamente se realizan cambios en la información contenida en la presente. Dichos cambios se incorporarán en las nuevas ediciones de la publicación. IBM puede realizar mejoras y/o cambios en los productos y/o programas descritos en esta publicación en cualquier momento y sin previo aviso.

Las referencias en esta información a sitios web que no son de IBM se proporcionan solo por comodidad y de ningún modo suponen un aval de dichos sitios web. Los materiales de dichos sitios web no forman parte de los materiales de este producto de IBM, por lo que la utilización de dichos sitios web corre por cuenta y riesgo del usuario.

IBM puede usar o distribuir en modo que le parezca conveniente la información que usted le proporcione sin que por ello se derive obligación alguna hacia usted.

Los titulares de licencias de este programa que deseen obtener información sobre el mismo con el fin de permitir: (i) el intercambio de información entre programas creados independientemente y otros programas (incluido este) y (ii) el uso mutuo de información que se haya intercambiado, deben ponerse en contacto con:

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

Dicha información puede estar disponible conforme a los correspondientes términos y condiciones, incluyendo en algunos casos el pago de una tarifa.

IBM proporciona el programa sujeto a licencia descrito en este documento y todo el material sujeto a licencia disponible al mismo conforme a los términos del Acuerdo de Cliente de IBM, el Acuerdo de Licencia de Programa Internacional de IBM o cualquier acuerdo equivalente entre las partes.

Los datos de rendimiento y los ejemplos de clientes citados se presentan solamente a efectos ilustrativos. Los resultados reales de rendimiento pueden variar dependiendo de las configuraciones concretas y de las condiciones de operación.

La información relativa a productos que no sean de IBM se ha obtenido de los proveedores de dichos productos, de sus anuncios publicados o bien de otras fuentes públicas. IBM no ha sometido a pruebas dichos productos y no puede confirmar la precisión del rendimiento, la compatibilidad ni ninguna otra afirmación relacionada con productos que no sean de IBM. Las preguntas sobre las prestaciones de productos que no sean de IBM deben dirigirse a los proveedores de dichos productos.

Las declaraciones relativas a la dirección o la intención futuras 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 que se utilizan en el funcionamiento diario de la empresa. Para ilustrarlos de la mejor manera posible, los ejemplos incluyen nombres de personas, empresas, marcas y productos. Todos estos nombres son ficticios y cualquier parecido con personas o empresas comerciales reales es pura coincidencia.

Marcas registradas

IBM, el logotipo de IBM e ibm.com son marcas registradas o marcas comerciales registradas de International Business Machines Corp., registradas en muchas jurisdicciones en todo el mundo. Puede que otros nombres de producto y servicio sean marcas registradas de IBM u otras empresas. En Internet hay disponible una lista actualizada con las marcas registradas de IBM, en "Copyright and trademark information", en la dirección www.ibm.com/legal/copytrade.shtml.

Adobe, el logotipo de Adobe, PostScript y el logotipo de PostScript son marcas registradas de Adobe Systems Incorporated en Estados Unidos o en 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 registradas de Intel Corporation o de sus filiales en Estados Unidos y otros países.

Linux es una marca registrada de Linus Torvalds en Estados Unidos y/o en otros países.

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

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

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

Términos y condiciones de la documentación del producto

Los permisos para utilizar estas publicaciones se conceden en virtud de los siguientes términos y condiciones.

Aplicabilidad

Estos términos y condiciones son adicionales a los términos de uso del sitio web de IBM.

Uso personal

Estas publicaciones se pueden reproducir para uso personal no comercial siempre que se conserven todos los avisos de propiedad. No se pueden distribuir, visualizar o realizar trabajos derivados de estas publicaciones, o de partes de las mismas, sin el consentimiento expreso de IBM.

Uso comercial

Puede reproducir, distribuir y mostrar estas publicaciones exclusivamente en su empresa siempre que se conserven todos los avisos de propiedad. No puede realizar trabajos derivados de estas publicaciones, ni reproducir, distribuir o visualizar estas publicaciones, ni de partes de las mismas, fuera de la empresa sin el consentimiento expreso de IBM.

Derechos

Salvo cuando se conceda explícitamente en este permiso, no se concede ningún otro permiso, licencia ni derecho, implícita o explícitamente, a las publicaciones ni a ninguna información, datos, software ni otros elementos sujetos a los derechos de propiedad intelectual contenidos en la presente.

IBM se reserva el derecho de retirar los permisos aquí concedidos cuando, a su criterio, el uso de las publicaciones vaya en detrimento de sus intereses o cuando, según determine IBM, las instrucciones anteriores no se sigan correctamente.

No puede descargar, exportar ni volver a exportar esta información excepto en el caso de cumplimiento total con todas las leyes y regulaciones vigentes, incluyendo todas las leyes y regulaciones sobre exportación de Estados Unidos.

IBM NO GARANTIZA EL CONTENIDO DE ESTAS PUBLICACIONES. LAS PUBLICACIONES SE PROPORCIONAN "TAL CUAL" Y SIN GARANTÍA DE NINGUNA CLASE, NI EXPLÍCITA NI IMPLÍCITA, INCLUYENDO, PERO SIN LIMITARSE A, LAS GARANTÍAS IMPLÍCITAS DE COMERCIALIZACIÓN, NO VULNERACIÓN Y ADECUACIÓN A UN FIN DETERMINADO.

Índice

A

- abrir archivos 28
- accesibilidad 95
- administración
 - atajos de aplicación 82
 - descripción general 81
 - opciones de datos 83
 - opciones de destino de puntuación 84
 - opciones generales 82
 - página de inicio 83
- anotaciones
 - añadir a reglas 55
 - añadir campos 46
 - añadir campos adicionales 46
 - añadir campos utilizando reglas de segmentación 46
 - añadir/editar campos adicionales 46
 - añadir/editar campos derivados 46
 - añadir expresiones 46
 - añadir salidas de modelo 46
- aplicaciones
 - inicio 11
 - pruebas 70
- aplicaciones de muestra 10
- aplicaciones preconstruidas 10
- archivos de datos 10
- archivos de datos basados en texto 36
- archivos de datos de IBM SPSS Statistics 36
- archivos de datos Excel 36
- archivos de muestra 9
- archivos para aplicaciones de muestra 9, 10
- artefacto de demo 9
- artefactos para aplicaciones de muestra 9
- atajos
 - aplicación 25
- atajos de aplicación 25, 82
- ayuda 91
- ayuda en línea 91
- ayudas contextuales 91

B

- bloquear opciones de datos 83
- bloquear opciones de destino de puntuación 84
- bloquear proyectos 30
- booleano 49

C

- cambiar idioma 25
- cambiar idioma del navegador 25
- campo objetivo
 - generar modelos 57
- campo target
 - creación de modelos 13

- campos
 - correlacionar 45
 - derivación 47
 - en expresiones 48
- campos de correlación 78
- campos de correlaciones 45
- campos de entrada
 - correlacionar 45
 - definir 43
- campos de salida en modelos 78
- cargar archivos de ruta 28
- cargar rutas 28
- clúster manual 61
- compartir rutas entre IBM Analytical Decision Management e IBM SPSS Modeler 93
- comprobar expresiones 49
- controlar
 - opciones generales 82
- controlar atajos de aplicación 82
- correlacionar campos de puntuación 78
- creación de modelos
 - campo target 13
- creador de expresiones 46, 47, 49
- crear campos utilizando reglas de segmentación 46

D

- datos de Analytic Server 42
- datos de Hadoop 42
- datos de muestra 10
- datos HDFS 42
- datos para aplicaciones de muestra 10
- derivar campos 47
- descripción 31
- descripción general
 - administración 81
- descripción general de administración 81
- destino de Analytic Server en modelos 77
- destino de archivo en modelos 76
- destino de base de datos en modelos 75
- destino de puntuación en modelos 75
- destino del servidor Cognos BI en modelos 77
- documentación disponible 91

E

- editar campos adicionales 46
- editar campos derivados 46
- editar valores de campo 45
- editor de expresiones 47
- ejemplo de correo directo 9
- ejemplo de modelado de asociaciones 21
- ejemplo de modelado de clúster 20
- elementos no finalizados
 - galería 26

- etiquetar versiones 30
- etiquetas de versión 30
- evaluación de modelos 15, 67
 - modelos de clúster 20, 21
 - modelos de lista interactivos 18
 - modelos de reglas predictivos 18
- exportar reglas 54
- expresiones que devuelven booleano 49

F

- fecha de simulación 27
- formatos de fecha 27
- formatos de hora 27
- formatos de moneda 27
- funciones
 - en expresiones 47

G

- galería
 - trabajo no finalizado 26
 - trabajos de larga ejecución 26
- gráfico de distribución 58
- gráfico de importancia de predictor 58
- gráfico de importancia de variable 58
- gráficos de beneficios 15, 67
- gráficos de coincidencias 67
- gráficos de distribución 67
- gráficos de distribución de errores 67
- gráficos de elevación 67
- gráficos de ganancias 15, 67
- gráficos de rentabilidad de la inversión 67
- gráficos de respuestas 67
- gráficos en modelos 58
- guardar archivos 28, 29, 30

I

- IBM SPSS Collaboration and Deployment Services Repository
 - guardar en 28, 29, 30
- IBM SPSS Modeler Advantage
 - Ejemplo 9
- iconos, IBM Cognos BI 41
- idioma 25
- informe de resumen de puntuación 80

M

- máximo de miembros para campos nominales 31
- mis orígenes de datos 33
- modelo de datos del proyecto 33
- modelos
 - agrupación en clúster 61
 - clúster manual 61
 - asociación 62

modelos (*continuación*)

- campo objetivo 57
- campo target 13
- compilación 13
- correlacionar campos de puntuación 78
- creación de tiempo de restricción 27
- en prueba 19
- establecer preferencias 27
- evaluar 15, 67
- generar 57
- gráficos 58
- informe de resumen de puntuación 80
- lista interactiva 59
- predictivo 57
- pruebas 70
- puntuación 73, 79, 80
- puntuar 19
- regla predictiva 59
- resultados 58
- selección de campo de salida 78
- selección de datos 75
- selección de destino de Analytic Server 77
- selección de destino de archivo 76
- selección de destino de base de datos 75
- selección de destino de puntuación 75
- selección de destino del servidor Cognos BI 77
- selección de registro 75
- simular beneficio 69
- utilizar en reglas 51
- valores opcionales 63, 65

modelos de agrupación en clúster 61, 65

modelos de asociación 62, 65

- datos de transacciones 63
- datos tabulares 63
- datos tabulares frente a datos transaccionales 63

modelos de lista interactiva 59, 65

modelos de puntuación 19, 73, 79

- campos de correlación 78
- destino de Analytic Server 77
- destino de archivo 76
- destino de base de datos 75
- destino del servidor Cognos BI 77
- selección de destino de puntuación 75
- seleccionar campos de salida 78
- seleccionar datos 75

modelos de regla predictiva 59, 65

modelos predictivos 57

N

- niveles de medición 44
- nodo de origen IBM Cognos BI
 - iconos 41
- novedades 1, 2, 3
- nuevas características 3
- nuevas funciones 1, 2

O

- obtener ayuda 91
 - documentación disponible 91
- ocultar la página de inicio 83
- opciones de datos 83
- opciones de destino de puntuación 84
- opciones generales 82
- operadores
 - en expresiones 47
- Operational Decision Management
 - reglas 55
- orígenes de base de datos 39
- orígenes de Cognos BI 40
- orígenes de datos
 - Analytic Server 42
 - añadir campos 46
 - añadir expresiones 46
 - añadir salidas de modelo 46
 - basado en texto 36
 - base de datos 39
 - campos de correlaciones 45
 - campos de entrada 43
 - creador de expresiones 47, 49
 - crear campos utilizando reglas de segmentación 46
 - definición 33
 - definir 35
 - derivar campos 46
 - editar valores de campo 45
 - Excel 36
 - Hadoop 42
 - HDFS 42
 - IBM Cognos BI 40
 - IBM SPSS Statistics 36
 - niveles de medición 44
 - orígenes de archivo 36
 - pestaña Derivado 46
 - visión general de datos 43
 - vista de datos 39
 - vista previa 43
- orígenes de datos de proyecto 33
- orígenes de IBM Cognos BI 40

P

- página de inicio 25, 26, 82, 83
 - añadir aplicaciones 25
 - personalizar 25
- página de inicio de aplicación 82
 - añadir aplicaciones 25
 - personalizar 25
- página de inicio de aplicaciones 25
- página de inicio de la aplicación 26
- página de lanzamiento 11
- palabras clave 28, 31
- parámetros
 - en expresiones 48
- pestaña Derivado 46
- preferencias 27
- preferencias de usuario 27
- probar un modelo 70
- probar una aplicación 70
- profit
 - maximizar 9
 - simular 15

propiedades

- descripción 31
- máximo de miembros para campos nominales 31
- palabras clave 31
- proyecto 31
- repositorio 31

 prueba de los modelos 19

R

- reglas 51
 - anotaciones 55
 - compartido 54, 55
 - crear 51
 - exportar 54
 - externas 55
 - IBM Operational Decision Management 55
 - lógica de selección 53
 - reutilización 54, 55
 - segmento 51
 - selección 53
 - utilizar modelos con 51
- reglas compartidas 54
- reglas de IBM Operational Decision Management
 - utilizar en aplicaciones 55
- reglas de segmentación
 - utilizar modelos con 51
- reglas de selección 53
- reglas empresariales 51
- reglas Excluir 53
- reglas externas
 - utilizar en aplicaciones 55
- reglas Incluir 53
- rutas
 - cargar 28

S

- selección de datos en modelos 75
- selección de registro en modelos 75
- simulación
 - fecha predeterminada 27
- simulación de beneficio en modelos 69

T

- texto de ayuda 91
- tipos de datos 44

U

- URL
 - aplicación 25
- URL de aplicación 25

V

- valores globales
 - en expresiones 48
- visión general de datos 43
- vista de datos 39
- vista previa de datos 43

