

IBM® Analytical Decision Management
Versão 18 Release 0

IBM® SPSS® Modeler Advantage

IBM

Nota

Antes de usar estas informações e o produto suportado por elas, leia as informações em “Avisos” na página 97.

Informações sobre o produto

Esta edição se aplica a versão 18, liberação 0, modificação 0 de IBM Analytical Decision Management e a todas as liberações e modificações subsequentes até que seja indicado em novas edições.

© Copyright IBM Corporation 2010, 2015.

Índice

Prefácio	v	Verificando expressões	48
		Usando expressões que retornam um booleano	49
Capítulo 1. Sobre IBM Analytical Decision Management	1	Capítulo 5. Trabalhando com regras	51
Visão geral do produto	1	Criando regras	51
O que há de novo na versão 18	1	Definindo regras de segmento	51
Versões anteriores	2	Definindo regras de seleção	53
Capítulo 2. Exemplo de Modelagem de Resposta do Cliente	9	Exportando e reutilizando regras	54
IBM SPSS Modeler Advantage	9	Incluindo anotações	55
Maximizando Lucro para Campanhas de Mala Direta	9	Usando regras do IBM Operational Decision Management	55
Arquivos de amostra	9	Capítulo 6. Construindo modelos preditivos	57
Ativando o Aplicativo	11	Sobre modelos preditivos	57
Configurando origens de dados	11	Construindo modelos	57
Construindo o Modelo	13	Construindo um modelo preditivo	57
Avaliando o Modelo	15	Resultados da modelagem preditiva	58
Modelagem de Regra de Previsão	18	Construindo um modelo de regra preditivo	59
Testando o Modelo	19	Construindo um modelo de armazenamento em cluster	61
Pontuando o Modelo	19	Construindo um modelo de associação	62
Exemplo de Modelagem do Cluster	19	Configurações de modelo opcionais	63
Exemplo de Modelagem de Associação	21	Avaliando modelos	67
		Simulando lucro máximo	69
Capítulo 3. Navegando pela área de trabalho	25	Testando o modelo	70
Alterando o idioma exibido	25	Usando escores de modelo nos aplicativos	70
Ativando aplicativos	25	Capítulo 7. Modelos de escoragem	73
A página inicial	26	Escoragem de modelos em uma tabela de base de dados, arquivo, Analytic Server ou servidor IBM Cognos BI	73
Galeria	26	Selecionando dados e subconjuntos	75
Configurando preferências	27	Selecionando registros para serem escorados	75
Abrindo e salvando projetos, modelos e regras	28	Selecionando o destino de escoragem	75
Abrindo e salvando no repositório	29	Selecionando campos de saída	78
Trabalhando com rótulos	30	Modelos de escoragem	79
Bloqueando projetos	30	Escoragem do relatório sumarização	80
Propriedades	31	Capítulo 8. Administrando aplicativos	81
Capítulo 4. Gerenciando origens de dados	33	Opções gerais	82
Criando uma nova origem de dados	35	Controlando atalhos do aplicativo	82
Origem do arquivo	36	Ocultando a página inicial	82
Origem do banco de dados	38	Opções de bloqueio de dados	83
Origem da Visualização de Dados	39	Bloqueando opções de destino de escoragem	83
Origem do IBM Cognos BI	40	Fazendo download dos metadados do projeto para uso na criação de regras externas	84
Origem do Analytic Server	42	Editando a definição de esquema XML	85
Selecionando campos de entrada	42	Capítulo 9. Obtendo ajuda	91
Visualizando dados	43	Documentação disponível	91
Visão geral de dados	43		
Níveis de medição	43		
Editando valores de campo de dados	44		
Mapeando campos	45		
Derivando campos	45		
Construtor de Expressões	47		
Selecionando funções	47		
Selecionando campos e valores	48		

Apêndice A. Compartilhando fluxos . . .	93	Termos e condições da documentação do produto	99
Apêndice B. Acessibilidade	95	Índice Remissivo	101
Avisos	97		
Marcas Comerciais	98		

Prefácio

Sobre o IBM Business Analytics

O software IBM Business Analytics entrega informações completas, consistentes e precisas nas quais os tomadores de decisão confiam para melhorar o desempenho dos negócios. Um portfólio abrangente de business intelligence, análise preditiva, gerenciamento de estratégia e desempenho financeiro e aplicativos analíticos fornece insights claros, imediatos e acionáveis para o atual desempenho e capacidade de prever futuros resultados. Combinadas com ricas soluções para segmentos do mercado, práticas comprovadas e serviços profissionais, as organizações de cada porte podem conduzir a mais alta produtividade, automatizar decisões com segurança e entregar os melhores resultados.

Como parte deste portfólio, o software IBM SPSS Predictive Analytics ajuda as organizações a prever eventos futuros e agir proativamente nesse insight para conduzir os melhores resultados de negócios. Clientes comerciais, do governo e acadêmicos no mundo todo contam com a tecnologia IBM® SPSS como uma vantagem competitiva para atrair, reter e aumentar o número de clientes, enquanto reduz fraude e minimiza riscos. Ao incorporar o software IBM SPSS em suas operações diárias, as organizações se tornam empresas preditivas - aptas para dirigir e automatizar decisões para atender a objetivos de negócios e alcançar vantagem competitiva mensurável. Para obter informações adicionais ou para entrar em contato com um representante, visite <http://www.ibm.com/spss>.

Suporte técnico

O suporte técnico está disponível para clientes de manutenção. Os clientes podem entrar em contato com o Suporte Técnico para obterem assistência com o uso de produtos IBM Corp. ou para obterem ajuda com a instalação de um dos ambientes de hardware suportados. Para entrar em contato com o Suporte Técnico, consulte o website do IBM Corp. em <http://www.ibm.com/support>. Você deve estar preparado para se identificar, identificar sua organização e sua concordância de suporte ao solicitar assistência.

Capítulo 1. Sobre IBM Analytical Decision Management

Visão geral do produto

IBM Analytical Decision Management traz o benefício da análise preditiva para problemas de negócios reais, permitindo a construção de aplicativos customizados padronizados para seus clientes ou mercado. Enquanto normalmente os aplicativos são configurados para resolver problemas bem específicos, todos eles são baseados em um conjunto comum de capacidades:

- Automatizar decisões usando regras de negócios.
- Incluir insight usando modelos preditivos.
- Usar priorização, otimização ou simulação para atingir a melhor decisão com base nos itens acima.

Inúmeros aplicativos do pacote estão disponíveis, customizados para resolverem problemas de negócios específicos. Entre em contato com o representante de vendas para obter mais informações.

O que há de novo na versão 18

O IBM Analytical Decision Management 18 (18.0.0.0) inclui os novos recursos a seguir:

A seguir está uma descrição das principais novas variáveis:

Tempo de execução do cenário. Há um ícone de novas propriedades no diálogo Gerenciar Cenários, que permite visualizar os detalhes sobre um cenário, como o horário de início, o horário de encerramento e o tempo de execução.

Log CPLEX. Se desejar salvar o log CPLEX em um arquivo no sistema de arquivos IBM SPSS Modeler Server para propósitos de resolução de problemas, uma nova opção **Redirecionar log do mecanismo CPLEX para o arquivo** estará disponível no diálogo Opções Avançadas para aplicativos que têm uma guia Otimizar. Observe que em ocasião rara, um aplicativo CPLEX é implementado no Serviço de Escoragem e você deve desativar a criação de log para melhorar o desempenho.

Carregamento em massa e tamanho de confirmação em lote. Na guia Implementar, ao executar a saída para um banco de dados, agora é possível configurar o tamanho do lote e ativar o carregamento em massa. Isso melhora o desempenho ao escorar uma grande tabela e gravar os resultados em um banco de dados.

Visão geral de dados no contexto A Visão Geral de Dados é exibida agora no contexto do local em que o campo está sendo usado no projeto. Por exemplo, qualquer regra criada anteriormente no projeto é levada em consideração.

Nova restrição de otimização. Uma nova restrição foi incluída no aplicativo IBM Analytical Decision Management for Campaign Optimization para o *Número mínimo de ofertas que devem ser aceitas para esse tipo de oferta*.

Melhorias de memória. Foram feitas melhorias na memória para o aplicativo IBM Analytical Decision Management for Campaign Optimization. Para obter mais informações, consulte **Melhores práticas de desempenho > Otimização de CPLEX > Requisitos mínimos de memória** no *Guia de Desempenho e Resolução de Problemas*.

Versões anteriores

Versão 17

IBM Analytical Decision Management 17 (17.0.0.0) inclui as seguintes novas variáveis:

A seguir há uma descrição dos principais recursos novos que foram incluídos na versão 17:

Construtor de Expressão. O Construtor de Expressão foi aprimorado para permitir expressões mais complexas e poderosas. O Construtor de Expressão no IBM Analytical Decision Management agora corresponde melhor ao Construtor de Expressão no cliente IBM SPSS Modeler. Consulte http://www-01.ibm.com/support/knowledgecenter/SS6A3P_17.0.0/adm_user_guide_ddita/configurableapps/dm_expressions.dita.

Guia derivada. Na maioria dos aplicativos IBM Analytical Decision Management, é possível incluir campos adicionais em seu modelo de dados do projeto, como expressões, saída do modelo, campos criados de regras de segmento ou campos de tabelas secundárias. A nova subguia **Derivada** da guia Dados exibe quaisquer campos que devem ser incluídos como parte dos dados. Consulte http://www-01.ibm.com/support/knowledgecenter/SS6A3P_17.0.0/adm_user_guide_ddita/configurableapps/dm_app_derived.dita.

Seção Tabelas secundárias. Uma nova seção Tabelas secundárias foi incluída na maioria dos aplicativos IBM Analytical Decision Management. Isso substitui os vários links **Incluir/editar tabelas adicionais** em liberações anteriores. Consulte http://www-01.ibm.com/support/knowledgecenter/SS6A3P_17.0.0/adm_user_guide_ddita/configurableapps/dm_app_additional_fields.dita.

Autenticação de multifatores. A funcionalidade que suporta a autenticação de multifatores (MFA) foi incluída em IBM SPSS Collaboration and Deployment Services 7, IBM SPSS Modeler 17 e IBM Analytical Decision Management 17. Esses aprimoramentos podem melhorar a segurança para ambientes corporativos de grande escala. Por exemplo, produtos SPSS podem ser integrados à infraestrutura de Kerberos Single Sign On existente de uma empresa e infraestruturas SmartCard (quando os funcionários de uma empresa usam badges de segurança e leitores de cartão para autenticação de máquina e aplicativo, por exemplo). Consulte http://www-01.ibm.com/support/knowledgecenter/SS6A3P_17.0.0/adm_install_guide_ddita/configurableapps/dms_install_security.dita.

Mapeamento de origem de dados. Quando possível, o IBM Analytical Decision Management automaticamente mapeará tipos de armazenamento compatíveis quando um usuário incluir uma origem de dados adicional. Se o tipo não for compatível, o usuário terá permissão de converter os tipos.

Tipos de armazenamento de arquivo. Os usuários agora podem configurar o tipo de armazenamento e o formato ao ler em arquivos de dados de formato fixo e variável, para assegurar que os valores sejam lidos corretamente. Por exemplo, se o tipo de origem de dados for **Arquivo** e o formato de arquivo for **Campos de duração variável em texto simples** ao incluir uma nova origem de dados no IBM Analytical Decision Management, o diálogo Origem de dados terá uma nova seção **Especificar formato de dados** que permite que o usuário especifique o tipo de entrada e o formato de cada campo. Isso é semelhante à funcionalidade disponível no cliente IBM SPSS Modeler.

Resumo de projeto. Para aplicativos com uma guia Implementar, uma nova seção **Resumo de projeto** foi incluída. Use essa seção para validar o projeto e exibir informações importantes de resumo do projeto, incluindo quaisquer erros localizados que possam impactar a implementação. O resumo do projeto exibe informações, como todas as referências de atributo e todas as referências do repositório de atributo e podem ser exportadas como um arquivo *.csv. Também é possível ir para outras guias do aplicativo para corrigir erros e retornar ao resumo do projeto na guia Implementar.

Referência de modelo. Ao referenciar modelos em um aplicativo IBM Analytical Decision Management, o comportamento do link de modelo sofreu mudanças. Ao clicar em um link de referência de modelo, o diálogo de visualizador de modelo é aberto, permitindo que visualize detalhes sobre o modelo. Ou é

possível agora clicar com o botão direito na referência de modelo para selecionar a opção **Referência** ou **Compilação**. Ao clicar em uma referência de regra, o visualizador de regra será aberto (como nas liberações anteriores). Ou é possível clicar com o botão direito em uma referência de regra e selecionar **Referência**.

Sufixo de credencial do servidor analítico. Os administradores podem usar o IBM SPSS Collaboration and Deployment Services Deployment Manager baseado em navegador para configurar um sufixo para anexar a credenciais do Analytic Server. Isso derivará o nome da credencial que será usada para conexões do Servidor analítico no IBM Analytical Decision Management.

Credenciais do IBM Cognos. Ao importar dados de banco de dados IBM Cognos BI ou relatórios de lista única para uso no IBM Analytical Decision Management, é possível selecionar uma credencial IBM Cognos armazenada no Repositório do IBM SPSS Collaboration and Deployment Services em vez de inserir as informações de nome de usuário e senha ou usar uma conexão anônima.

Versão 8

O IBM SPSS Analytical Decision Management 8 estendeu a capacidade do usuário de implementar análise nos processos de negócios diários, suportando uma variedade maior de decisões com mais insight e controle do que nunca.

- Ajuste e visualize o impacto combinado de regras, modelos e otimização em resultados de negócios com capacidade de simulação aprimorada.
- Resolva uma grande variedade de problemas com novos aplicativos criados previamente e novas técnicas de otimização, como otimização de preço e otimização de sortimento.
- Melhora na integração com o IBM Operational Decision Management e suporte a Big Data via integração do IBM SPSS Analytic Server.

A seguir está uma descrição das principais novas variáveis que foram incluídas na versão 8:

Nova tecnologia de instalação. A instalação do produto foi simplificada e melhorada. Agora ela usa tecnologia Gerenciador de Instalação. Para obter mais informações, consulte o *IBM SPSS Analytical Decision Management Installation Guide*.

Aprimoramentos da interface com o usuário. Muitas mudanças foram feitas na interface com o usuário para melhorar a usabilidade.

IBM Analytical Decision Management for Predictive Maintenance. IBM SPSS Analytical Decision Management for Predictive Maintenance é um novo aplicativo criado previamente que usa otimização para prever e evitar falhas de equipamentos antes de elas ocorrerem. É possível minimizar um tempo de inatividade caro e fazer melhor uso de recursos e infraestrutura escassa prevendo falhas e planejando ações corretivas no momento certo. Você pode prever quais partes têm maior probabilidade de falhar primeiro e substituí-las no momento apropriado. Prever quando as máquinas vão falhar ajuda a maximizar os lucros e evitar perdas decorrentes de máquinas com falha e falta de ação.

IBM Analytical Decision Management for Demand Optimization. IBM SPSS Analytical Decision Management for Demand Optimization é um novo aplicativo criado previamente que usa otimização para evitar excesso de estoque e deficiências alocando a quantidade certa de cada produto para cada loja. É possível maximizar lucros e evitar perdas devido a deficiências de inventário e taxas de reabastecimento alocando a quantidade certa de inventário para cada outlet.

Aprimoramentos de simulação. Novas capacidades de simulação combinam a funcionalidade do botão **Simular** anterior (antigamente na guia Definir) com a funcionalidade **E se?** anterior (antigamente na guia Otimizar, Priorizar ou Combinar) em uma única variável. A funcionalidade de simulação nova e melhorada tem sua própria barra de ferramentas no canto superior direito da interface com o usuário e permite que cenários sejam salvos com o projeto e persistam nas sessões.

Implementação e escoragem. Vários aprimoramentos foram feitos na guia Implementar/Escorar.

URL do aplicativo. IBM SPSS Collaboration and Deployment Services incluiu a capacidade de customizar a URL para aplicativos baseados na web, incluindo a URL para ativar o IBM SPSS Analytical Decision Management. Entre em contato com o administrador para obter detalhes. Há mais informações disponíveis no capítulo *Raízes de contexto do aplicativo* do *IBM SPSS Collaboration and Deployment Services Repository Installation and Configuration Guide*.

Aprimoramentos no arquivo de estatísticas. Agora arquivos de dados protegidos por senha e compactados do IBM SPSS Statistics (.sav e .zsav) são suportados no IBM SPSS Analytical Decision Management.

Aprimoramentos do visualizador de árvore Durante a visualização de resultados de modelos individuais no IBM SPSS Modeler Advantage, o nome de cada modelo é um link que é aberto em um novo diálogo contendo informações sobre o modelo, como o campo de resposta e os campos de entrada. Se o modelo selecionado for um *modelo de árvore*, um visualizador de árvore também será exibido. O visualizador de árvore foi aprimorado. Agora é possível executar ações como aumentar e diminuir o zoom, navegar em árvores grandes e expandir ou reduzir nós individuais dentro da árvore.

Opções avançadas na guia Otimizar. Para aplicativos que usam complex mathematical optimization (CPLEX), como 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, agora há opções avançadas disponíveis relacionadas ao algoritmo de otimização CPLEX. Agora é possível exportar um arquivo de modelo Optimization Programming Language (OPL) e incluir configurações customizadas.

Origem da Visualização de Dados. Um novo nó *Visualização de Dados* foi incluído no IBM SPSS Modeler 16 para incluir dados em fluxos que foram definidos em uma **visualização de dados analítica** do IBM SPSS Collaboration and Deployment Services. A nova origem de dados *Visualização de Dados* também pode ser usada na guia *Dados* no IBM SPSS Analytical Decision Management. É possível criar uma conexão para usar com regras do IBM Operational Decision Management. Para regras complexas que usam um modelo de dados hierárquicos, observe que os administradores podem restringir o uso de tais regras porque elas podem comprometer o desempenho.

Antes da versão 8, regras externas requeriam parâmetros de entrada simples que correspondessem ao modelo de dados do projeto do IBM SPSS Analytical Decision Management. Porém, com este novo suporte de visualização de dados analítica, agora as regras externas podem ter parâmetros de entrada hierárquicos. Consulte os tópicos “Origem da Visualização de Dados” na página 39 e “Opções de bloqueio de dados” na página 83 para obter informações e o *Capítulo 5* do *Application Designer’s Guide*.

Integração do IBM SPSS Analytic Server. Em IBM SPSS Modeler 16, IBM SPSS Analytic Server foi introduzido, permitindo o uso de Hadoop Distributed File System (HDFS) como uma origem de dados. Os usuários do IBM SPSS Modeler podem criar fluxos usando o nó de origem HDFS Analytic Server e os modelos de construção baseados no sistema Big Data Hadoop. Os usuários do IBM SPSS Analytical Decision Management podem abrir esses fluxos de modelo no IBM SPSS Modeler Advantage para visualização (eles não podem criar novos projetos usando origens de dados do Analytic Server).

Novo capítulo no Application Designer’s Guide. Um novo capítulo foi incluído no *IBM SPSS Analytical Decision Management Application Designer’s Guide*. Ele fornece um exemplo de configuração de um aplicativo de otimização customizado. O aplicativo IBM SPSS Analytical Decision Management for Campaign Optimization padrão usa um arquivo de modelo Optimization Programming Language (OPL) pré-configurado chamado *CampaignOptimization.mod*. Em alguns casos, o aplicativo Campaign Optimization padrão e seu modelo OPL podem não ser adequados para os propósitos do problema de negócios atual. O novo capítulo percorre um exemplo de um modelo OPL customizado construído usando IBM ILOG CPLEX Optimization Studio e um aplicativo de otimização IBM SPSS Analytical Decision Management customizado para uma empresa fictícia.

Resultados da avaliação tabular. Uma nova opção **Incluir resultados tabulares** está disponível durante a execução de avaliações. O cancelamento da seleção da opção melhora o desempenho porque isso leva mais tempo para executar avaliações do que a inclusão de resultados tabulares.

Nova configuração para número máximo de campos nominais. Uma nova configuração foi incluída no diálogo Propriedades para a configuração do número máximo de membros permitido para campo de dados *Nominal*. Ela é usada na guia Dados. Para obter mais informações, consulte “Propriedades” na página 31.

Conexão única. Conexões ODBC agora podem usar conexão única (SSO) para conexões com banco de dados IBM SPSS Modeler Server e ODBC com o IBM SPSS Analytical Decision Management. Alguns passos de configuração são obrigatórios. Consulte o tópico Considerações de segurança na documentação de instalação.

Retrocesso de SQL. Quando os dados de origem são originados em um banco de dados, a execução retrocede para o banco de dados sempre que possível (em um processo conhecido como "retrocesso de SQL"). Em geral, o IBM SPSS Analytical Decision Management suporta o mesmo nível de retrocesso SQL que o IBM SPSS Modeler, mas com as seguintes limitações:

- Se o projeto incluir regras de seleção ou campos adicionais que fazem referência a modelos ou regras externas, essas referências não retrocederão para o banco de dados e isso pode inibir o grau de retrocesso executável para o fluxo de projeto como um todo.
- Em aplicativos diferentes do IBM SPSS Modeler Advantage, o retrocesso é ativado somente até as regras de seleção (seleções globais e seleções de membros na primeira dimensão). O grau de retrocesso pode ser aumentado se você escolher um campo Número Inteiro como ID de entidade no modelo de dados do projeto.

Versão 7.0.0.1

IBM SPSS Decision Management 7 Fix Pack 1 (7.0.0.1) inclui as seguintes variáveis:

Visão Geral de Dados. Uma nova variável de visão geral está disponível na guia Dados de todos os aplicativos, na guia Modelagem e na guia Score no IBM SPSS Modeler Advantage e na guia Implementar dos aplicativos que incluem a variável *Escorar Agora*. Consulte o tópico “Visão geral de dados” na página 43 para obter mais informações. Essa variável está disponível somente com IBM SPSS Modeler Server 15 Fix Pack 1 (15.0.0.1). Ela não está disponível com a versão 15.

Relatório Sumarização de Escoragem. Após a escoragem de um projeto, uma nova seção *Relatório Sumarização de Escoragem* fornece detalhes, como onde os dados foram escorados, tamanho do arquivo (se aplicável), nome da tabela e do banco de dados (se aplicável), nome do modelo ou nome do projeto, número de registros escorados; em alguns casos, um gráfico e/ou tabela resumindo o modelo. A seção Relatório Sumarização de Escoragem está disponível em IBM SPSS Modeler Advantage e na guia Implementar dos aplicativos que inclui a variável *Escorar Agora*. Observe que o tamanho dos arquivos só será exibido na execução no IBM SPSS Modeler Server 15 Fix Pack 1 (15.0.0.1). Ele não estará disponível com a versão 15.

Download de cópia local do arquivo escorado. Se você escolher enviar a saída da escoragem para um arquivo simples, também é possível fazer download de uma cópia local do arquivo, se desejado. Os administradores podem usar IBM SPSS Collaboration and Deployment Services Deployment Manager baseada em navegador para ativar ou desativar a variável e para aumentar ou diminuir o limite de tamanho de arquivo máximo permitido. Se o botão **Download de Arquivo** estiver inativo, o tamanho do arquivo excede o limite de tamanho máximo configurado pelo administrador ou o administrador desativou a variável. A variável fica desativada por padrão. Essa variável está disponível somente com IBM SPSS Modeler Server 15 Fix Pack 1 (15.0.0.1). Ela não está disponível com a versão 15.

Novas configurações opcionais para construção de modelos. Durante a construção de modelos no IBM SPSS Modeler Advantage, novas configurações opcionais estão disponíveis, como opções de particionamento, escolha de quais técnicas de modelagem (algoritmos) incluir durante a construção do

modelo, especificação do número máximo de modelos para reter e combinar, além da configuração de critérios de lucro antes da construção do modelo. As opções disponíveis dependem do tipo de modelo e da resposta selecionada. Consulte o tópico “Configurações de modelo opcionais” na página 63 para obter mais informações. Observe que as novas variáveis a seguir só estão disponíveis com IBM SPSS Modeler Server 15 Fix Pack 1 (15.0.0.1). Elas não estão disponíveis com a versão 15.

- Selecionando quais técnicas de modelagem incluir durante a construção do modelo
- Configurando critérios de lucro
- Visualizando detalhes de alto nível sobre modelos individuais
- Visualizando detalhes adicionais sobre um modelo selecionado

Visualizador de árvore. Após a construção de um modelo no IBM SPSS Modeler Advantage e a visualização de resultados do modelo individual, o nome de cada modelo é um link que é aberto em um novo diálogo contendo informações sobre o modelo, como o campo de resposta e os campos de entrada. Se o modelo selecionado for um *modelo de árvore*, um visualizador de árvore também será exibido. Observe que o novo visualizador de árvore só está disponível com IBM SPSS Modeler Server 15 Fix Pack 1 (15.0.0.1). Ele não está disponível com a versão 15.

Aprimoramentos de avaliação. Vários aprimoramentos foram feitos na caixa de diálogo Avaliar no IBM SPSS Modeler Advantage. Por exemplo, um novo botão **Modificar Seleção de Modelo** está disponível, permitindo que você modifique os modelos para incluir na avaliação. Após fazer as seleções, também é possível copiar as atuais seleções de modelo de volta na tabela de resultados do modelo individual, se desejar. Algumas das novas opções só estão disponíveis para certos tipos de modelo (a maioria das novas opções é somente para o tipo de modelo *Preditivo* e algumas também são para o tipo de modelo *Regra Preditiva*). Observe que as novas variáveis a seguir na caixa de diálogo Avaliar só são ativadas com IBM SPSS Modeler Server 15 Fix Pack 1 (15.0.0.1). Elas não estão disponíveis com a versão 15.

- Gráficos de desempenho para resposta numérica
- Simulação de lucro para resposta numérica
- Gráficos de matriz de coincidência (para todas as respostas)
- Visualização de tabela de todos os gráficos
- O botão **Modificar Seleção de Modelo**
- A opção para **Incluir avaliações de modelo individual para gráficos de desempenho**

Microsoft Windows 8 e Internet Explorer 10. IBM SPSS Collaboration and Deployment Services 5 Fix Pack 1 (5.0.0.1) inclui suporte para Windows 8 e Internet Explorer 10. Eles não são suportados com IBM SPSS Collaboration and Deployment Services 4.2.1 ou 5.

Aprimoramentos da interface com o usuário. Muitas pequenas mudanças foram feitas na interface com o usuário para melhorar a usabilidade.

Versão 7

IBM SPSS Decision Management 7 (7.0.0.0) inclui as seguintes variáveis:

Otimização. O objetivo da otimização é identificar a solução que melhor atenda a um objetivo específico, como maximizar a renda de uma campanha de marketing ou minimizar o risco de fraude ou perda de clientes. O resultado da otimização é uma solução que responde a essa pergunta em todas as combinações possíveis dentro de sua população da maneira que melhor maximize o valor que você deseja otimizar (como lucro).

Modelos de armazenamento em cluster. Novos no IBM SPSS Modeler Advantage, os métodos de modelagem de cluster ajudam a descobrir grupos de registros de dados com valores ou padrões semelhantes. Essas técnicas também são usadas em marketing (segmentação de cliente) e outros aplicativos de negócios (registros que se encaixam em clusters de registro único podem conter erros ou ser instâncias de fraude). Os modelos de armazenamento em cluster geralmente são chamados de modelos de aprendizado sem supervisão, já que não usam um campo de destino e não retornam uma

predição específica que possa ser avaliada como verdadeira ou falsa. IBM SPSS Decision Management permite a construção de modelos de cluster automaticamente. Também é possível combinar, renomear, modificar ou excluir clusters ou incluir clusters customizados com base no conhecimento de negócios de informações a priori.

Modelos de associação. Nova no IBM SPSS Modeler Advantage, a modelagem associação permite prever quais itens têm maior probabilidade de aparecerem juntos e prever a intensidade do relacionamento entre eles. As regras de associação associam uma determinada conclusão (a compra de um certo produto, por exemplo) a um conjunto de condições (a compra de vários outros produtos, por exemplo).

Dados. A guia Dados agora inclui a capacidade de ter várias entradas de dados, definir o modelo de dados do projeto com uma Visualização do Aplicativo, referir-se a saídas de modelo como se elas fossem dados, fazer upload de arquivos de dados locais e visualizar e contar dados.

Melhorias na integração de dados do IBM Cognos. Assim como acontece com o IBM SPSS Modeler 15, o IBM SPSS Decision Management agora suporta a importação de relatórios Cognos pré-construídos, filtros e cálculos Cognos e o controle de desagregação automática.

Elementos múltiplos. Agora os aplicativos podem incluir mais de uma hierarquia da dimensão. Durante a ativação de um aplicativo, se o aplicativo selecionado estiver configurado para incluir mais de um elemento, será solicitado que você selecione um ou mais dos elementos disponíveis (hierarquias) para usar no aplicativo. A árvore de dimensão na guia Definir especifica o intervalo de recomendações ou decisões possíveis que podem ser retornadas. Se um aplicativo incluir mais de uma hierarquia da dimensão, elas serão incluídas em guias separadas da área da árvore de dimensão na guia Definir.

Alocação dinâmica. Em liberações anteriores, para a alocação de algo, ela tinha que existir na árvore de dimensão. Agora, para itens de dimensão sem filhos, é possível ler itens de uma tabela vinculada por uma chave no modelo de dados do projeto. Por exemplo, você pode ter um item de dimensão chamado *Carros* sem filhos. Você poderia configurar alocações dinâmicas em *Carros* para fazer referência a campos em uma tabela como *BMW*, *Audi* e *Ford*.

Alocações múltiplas. Uma nova opção **Alocações múltiplas** permite a alocação de registros para vários itens de dimensão (como várias ofertas) para cada segmento. Várias alocações podem ser feitas para cada regra em vez de uma única alocação por regra.

Melhorias na interface com o usuário. Foram feitas muitas mudanças e adições na interface com o usuário do IBM SPSS Decision Management para melhorar a usabilidade e a escalabilidade de todos os aplicativos.

Aprimoramentos de galeria. Foram feitos vários aprimoramentos na galeria, como a opção **Limpar Galeria**, para remover mais de um item da galeria por vez.

Aprimoramentos do editor de regras. Dependendo do local a partir do qual o editor de regras é ativado, agora os usuários podem configurar o tipo de regra para regras de seleção (inclusão ou exclusão), designar pontos de risco para regras de seleção, selecionar uma dimensão para regras de alocação e usar os controles *OR* e *Split OR* para combinar diversas expressões dentro de uma regra e criar uma expressão *OR*. Durante a criação de regras, agora também é possível selecionar uma dimensão em uma árvore de hierarquia (para aplicativos com mais de uma árvore) ou ativar o editor de expressão para criar uma expressão. Essa funcionalidade não estava disponível na caixa de diálogo do editor de regras anteriormente.

Integração do IBM Operational Decision Management estendida. No IBM SPSS Collaboration and Deployment Services Deployment Manager, é possível usar a caixa de diálogo Incluir Nova Regra Externa para definir referências aos conjuntos de regras do IBM Operational Decision Management que depois possam ser usadas no IBM SPSS Decision Management. Antes, essas regras só podiam ser acessadas por meio de um serviço da web. Com esta liberação, agora é possível fazer referência a regras

locais no banco de dados do IBM Operational Decision Management. Para a nova opção local, a regra é executada via IBM SPSS Modeler Server (o mesmo servidor definido no Deployment Manager para ser usado pelo IBM SPSS Decision Management).

Aprimoramentos no WhatIf?. Vários aprimoramentos foram feitos na funcionalidade da caixa de diálogo WhatIf?. Por exemplo, para a data de simulação, agora é possível inserir um intervalo de datas no qual executar a avaliação WhatIf?.

Upload de arquivos de fluxo. Agora é possível abrir qualquer arquivo de fluxo do IBM SPSS Modeler a partir do sistema de arquivos local e armazená-lo no repositório se desejar.

Credenciais múltiplas do IBM SPSS Modeler Server. Se você tiver várias credenciais do IBM SPSS Modeler Server definidas no IBM SPSS Collaboration and Deployment Services Deployment Manager, agora é possível especificar mais de uma para usar com o IBM SPSS Decision Management. Depois, quando o usuário do IBM SPSS Decision Management efetuar logon, o servidor tentará a primeira credencial na lista para ver se o usuário tem autoridade, prosseguindo pela lista em ordem até que seja localizada uma credencial com a autoridade adequada. Se o usuário não tiver autoridade a nenhuma das credenciais na lista ordenada, ele não poderá usar o IBM SPSS Decision Management. Isso oferece maior controle sobre quais usuários têm acesso a quais partes do sistema de arquivos IBM SPSS Modeler Server.

Versão 6.2

IBM SPSS Decision Management 6.2 inclui as seguintes variáveis:

Suporte para dados do Cognos. Quando usado com IBM SPSS Modeler Server 14.2 ou mais novo, o IBM SPSS Decision Management 6.2 inclui a capacidade de leitura de dados a partir do servidor IBM Cognos Business Intelligence durante a definição de origens de dados e de gravação dos resultados em um Cognos BI Server. Após definir uma conexão com um servidor Cognos BI, você pode navegar por pacotes, dimensões, consultas e relatórios para escolher os campos desejados. Consulte o tópico “Origem do IBM Cognos BI” na página 40 para obter mais informações.

Integração do ILOG Business Rule Management System. Regras criadas em um Business Rules Management System, como ILOG, podem ser referenciadas e usadas em aplicativos IBM SPSS Decision Management, contanto que tenham sido desenvolvidas para suportar o mesmo modelo de dados usado no atual projeto IBM SPSS Decision Management. Essas regras externas podem ser selecionadas e usadas da mesma maneira que outros modelos e regras dentro do aplicativo IBM SPSS Decision Management. Consulte o tópico “Usando regras do IBM Operational Decision Management” na página 55 para obter mais informações. Observe que agora o ILOG é referido como IBM Operational Decision Management.

Suporte para Linux on System Z. SUSE Enterprise Server 10 é suportado no sistema IBM System z10 de 64 bits.

Capítulo 2. Exemplo de Modelagem de Resposta do Cliente

IBM SPSS Modeler Advantage

IBM SPSS Modeler Advantage é um aplicativo fácil de usar que coloca o poder da modelagem preditiva nas mãos dos usuários corporativos. Usando modelos preditivos, é possível identificar padrões baseados naquilo que aconteceu no passado e usá-los para prever o que provavelmente acontecerá no futuro.

Por exemplo, você pode usar um modelo para prever quais clientes têm a menor probabilidade de migrarem para o concorrente, ou a maior probabilidade de responderem a uma determinada oferta, com base em características como renda, idade e organizações e associações em que estão inscritos. As previsões resultantes podem ser usadas para gerar listas de clientes-alvo ou casos de seu interesse, conforme entrada para planejamento estratégico, ou podem ser integradas a regras no contexto de um aplicativo preditivo.

Maximizando Lucro para Campanhas de Mala Direta

Para este exemplo de um modelo de previsão, você está fazendo marketing de produtos e serviços financeiros para clientes em um banco fictício. Em uma campanha de teste recente, a um número de clientes foi oferecido um dos diversos novos serviços e suas respostas, registradas. No entanto, as taxas de respostas foram menores que o esperado e a campanha não conseguiu transformar em lucro.

Ao utilizar um IBM SPSS Modeler Advantage, seu objetivo é transformar essa perda em lucro, visando os melhores clientes disponíveis, fornecendo seus custos esperados, receita e outros fatores.

Resultados da campanha de teste são resumidos abaixo.

Tabela 1. Resultados da Campanha de Teste

Resultados da Campanha de Teste	
Contagem de correio	3000
Contagem de resposta	111
Taxa de resposta	3.7%
Custo	6750
Receita	6660
Lucro total (perda)	(90)

Este exemplo usa os arquivos de dados *bank_customer_data.txt* e *bank_response_data.txt*, que são distribuídos e instalados com IBM Analytical Decision Management. Uma versão completa deste exemplo também é fornecida no arquivo *bank_response_model.str*. Entre em contato com seu administrador para obter detalhes sobre instalação e arquivos de amostra se necessário. Consulte o tópico “Arquivos de amostra” para obter mais informações.

Arquivos de amostra

Dados e outros arquivos de amostra são incluídos com IBM Analytical Decision Management que podem ser usados para replicar os exemplos mostrados nos guias de usuário.

Dados

Durante a instalação, vários arquivos de dados de amostra são instalados no diretório IBM SPSS Modeler Server \Demos (por exemplo, C:\Program Files\IBM\SPSS\ModelerServer\18\Demos). Estes dados podem ser usados com aplicativos de amostra pré-construídos e incluir:

- *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*

Arquivos de aplicativo de amostra, modelos e regras que usam esses arquivos de dados assumem a localização padrão indicada acima. Se os arquivos de dados forem instalados em uma unidade ou localização diferente, os caminhos de origem de dados deverão ser atualizados nos aplicativos de amostra de acordo.

Amostras

Inúmeros arquivos de aplicativo de amostra, modelos e regras são fornecidos para serem usados com aplicativos IBM Analytical Decision Management criados previamente. As amostras são fornecidas em um arquivo disponível para download eletrônico. Conclua os passos a seguir para importar arquivos de amostra para um ou mais aplicativos criados previamente no Repositório do IBM SPSS Collaboration and Deployment Services.

1. Faça o download do arquivo *ADM_18_Demo.zip* em um local temporário e extraia-o.
2. Abra IBM SPSS Collaboration and Deployment Services Deployment Manager e efetue logon na conexão do Content Server.
3. Clique com o botão direito na pasta Repositório de Conteúdo e selecione **Importar**.
4. Na caixa de diálogo Importar, clique em **Navegar** e selecione o arquivo *DecisionManagementExamples.pes* que copiou ou extraiu em um local provisório na etapa 1. Clique em **OK**.
5. No diálogo Conflitos de Importação, clique em **OK**. Quando a importação for concluída, você verá as seguintes novas pastas contendo projetos de exemplo que podem ser abertos no IBM Analytical Decision Management.

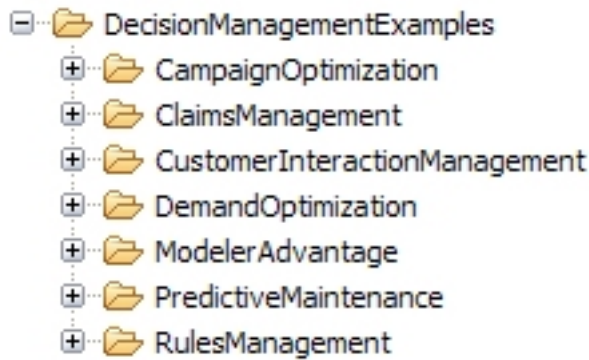


Figura 1. Pastas para artefatos de amostra

Observe que todos os arquivos de aplicativo de amostra, modelos e regras assumem que os arquivos de dados estão instalados no diretório IBM SPSS Modeler Server's Demos (por exemplo, *C:\Program Files\IBM\SPSS\ModelerServer\18\Demos*). Se os arquivos de dados forem instalados em um local diferente, os caminhos de origem de dados deverão ser atualizados nos aplicativos de amostra para a configuração local.

Ativando o Aplicativo

A página de ativação do IBM Analytical Decision Management inclui links para aplicativos disponíveis em seu site. Estes links podem estar configurados por seu administrador, você pode ser capaz de escolher a partir de aplicativos adicionais clicando no link **Incluir aplicativo** no canto superior direito.

1. Para começar a trabalhar por meio do exemplo, abra a página de ativação do IBM Analytical Decision Management. Para facilitar o acesso no futuro, é possível incluir um marcador nesta página em seu navegador ou criar um atalho em seu desktop.
2. Para criar um novo modelo, selecione **Novo** a partir da lista suspensa de aplicativos no grupo do IBM SPSS Modeler Advantage e clique em **Acessar**. A página Inicial do IBM SPSS Modeler Advantage será exibida.



Figura 2. IBM SPSS Modeler Advantage - Página Inicial

Cada ícone na página Inicial representa uma etapa diferente no processo de modelagem, de definição de dados a construção e pontuação do modelo. Para começar, clique no ícone Dados à esquerda.

Configurando origens de dados

A guia Dados permite que você configure origens de dados do para o arquivo de dados de amostra, uma origem de dados do usado no exemplo. Os conjuntos de dados podem ser configurados pelo administrador, ou você pode incluir suas próprias configurações.

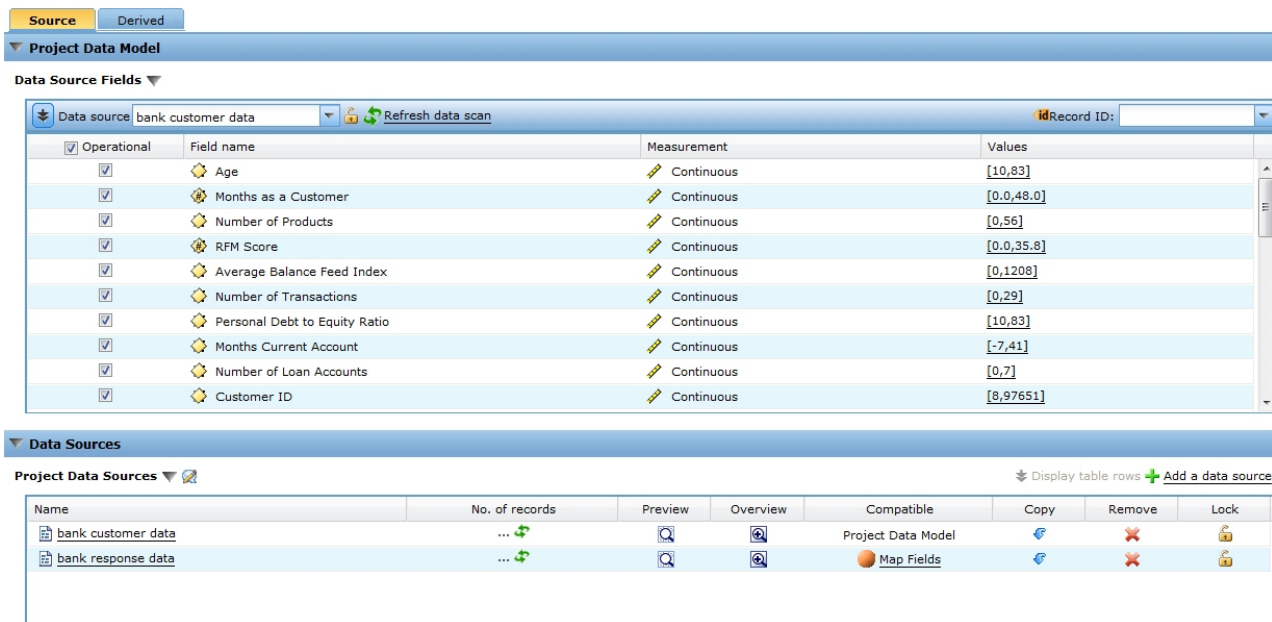


Figura 3. Data tab

O modelo de dados do projeto define os campos que estão disponíveis em todo o aplicativo e pode ser necessário para regras e modelos a serem avaliados. Todas as outras origens de dados devem ser mapeadas para o modelo de dados para serem usadas com o aplicativo.

Este exemplo utiliza dois arquivos de dados, ambos os quais são distribuídos com o aplicativo:

- O **bank_customer_data.txt** inclui informações sobre clientes atuais, incluindo o ID do cliente, histórico de pagamento, o número de produtos que o cliente possui, o tempo (número de meses) que são clientes, bem como informações demográficas, como idade e de renda. O modelo de dados do projeto é baseado nesse arquivo.
- O **bank_response_data.txt** contém informações sobre os clientes aos quais foram oferecidos um ou mais serviços em campanhas anteriores e registros de suas respostas. Ao analisar esses dados, o aplicativo pode identificar quais clientes têm maior probabilidade de responder a ofertas semelhantes em futuras campanhas.

Por padrão, os mesmo arquivos de dados de amostra são instalados na pasta IBM SPSS Modeler Server Demos que é acessível do seu aplicativo (por exemplo, *C:\Program Files\IBM\SPSS\ModelerServer\18\Demos*). Entre em contato com o administrador para obter assistência ao instalar esses arquivos, se necessário.

Para definir o projeto do modelo de dados:

1. No painel Origens de dados do projeto, selecione **Incluir uma origem de dados** no menu suspenso **Origem de dados**.
2. Especifique *dados do cliente do banco* como o nome de origem de dados.
3. Selecione **Arquivo** e selecione *bank_customer_data.txt*.
4. Em **Especificar campos de entrada**, assegure-se de que o tipo de medição para o campo *Tem registro de pagamento inválido* está configurado como **Flag**. A configuração de flag indica que o campo pode ter dois valores possíveis, indicando *true* e *false* (tipicamente codificados como *T* e *F*, ou 0 e 1 em um campo numérico), para indicar a presença ou ausência desse atributo.
5. Em **Especificar campos de entrada**, também assegure-se de que o tipo de medição para o campo *Sexo* está configurado como **Nominal**. (Os valores deste campo simplesmente indicam sexo masculino ou feminino e não *true/false* como seria o caso com um flag.)

6. Clique em **Links**. O diálogo Links será exibido.
7. Clique em **Salvar**.

Como esta é a primeira origem de dados definida para o projeto, ela é automaticamente selecionada para o modelo de dados do projeto, e usada para definir os campos necessários pelo aplicativo. Todos os outros conjuntos de dados são mapeados para esta origem.

Para definir a origem de dados da resposta:

O conjunto de dados de resposta contém muitos dos mesmos campos utilizados no modelo de dados do projeto, bem como dois campos adicionais denominados *ID de campanha* e *Resposta*. Esses campos indicam qual oferta foi feita para cada cliente e se eles responderam e podem ser usados para construir um modelo que prevê quais clientes têm maior probabilidade de responder no futuro ou não.

1. No painel Origens de dados do projeto, selecione **Incluir uma origem de dados** no menu suspenso **Origem de dados**.
2. Especifique *dados de resposta do banco* como o nome de origem de dados.
3. Selecione **Arquive** e selecione *bank_response_data.txt*.
4. Em **Especificar campos de entrada**, assegure-se de que o tipo de medição para o campo *Resposta* está configurado como **Flag**.
5. Em **Especificar campos de entrada**, desmarque a caixa de seleção ao lado do campo *ID de campanha* para indicar que este campo não deve ser utilizado como uma entrada quando a modelagem. (Este campo pode ser utilizado para indicar para quais campanhas um cliente foi selecionado no passado, mas não é útil na previsão de resposta futura.)
6. Clique em **Salvar**.

Como alguns dos campos no modelo de dados não são encontrados no conjunto de dados de resposta, um ícone alaranjado é exibido para esta origem de dados na coluna **Compatíveis** na guia Dados, indicando que alguns campos correspondem, mas não todos. O conjunto de dados é apropriado para modelagem, mas não para simulação, escoragem ou a implementação. (Para ser usado para esses fins, todos os campos do modelo de dados precisam estar presentes.)

Construindo o Modelo

A guia Modelagem permite que você especifique a origem de dados, campo de destino e outras opções para construir e avaliar o 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

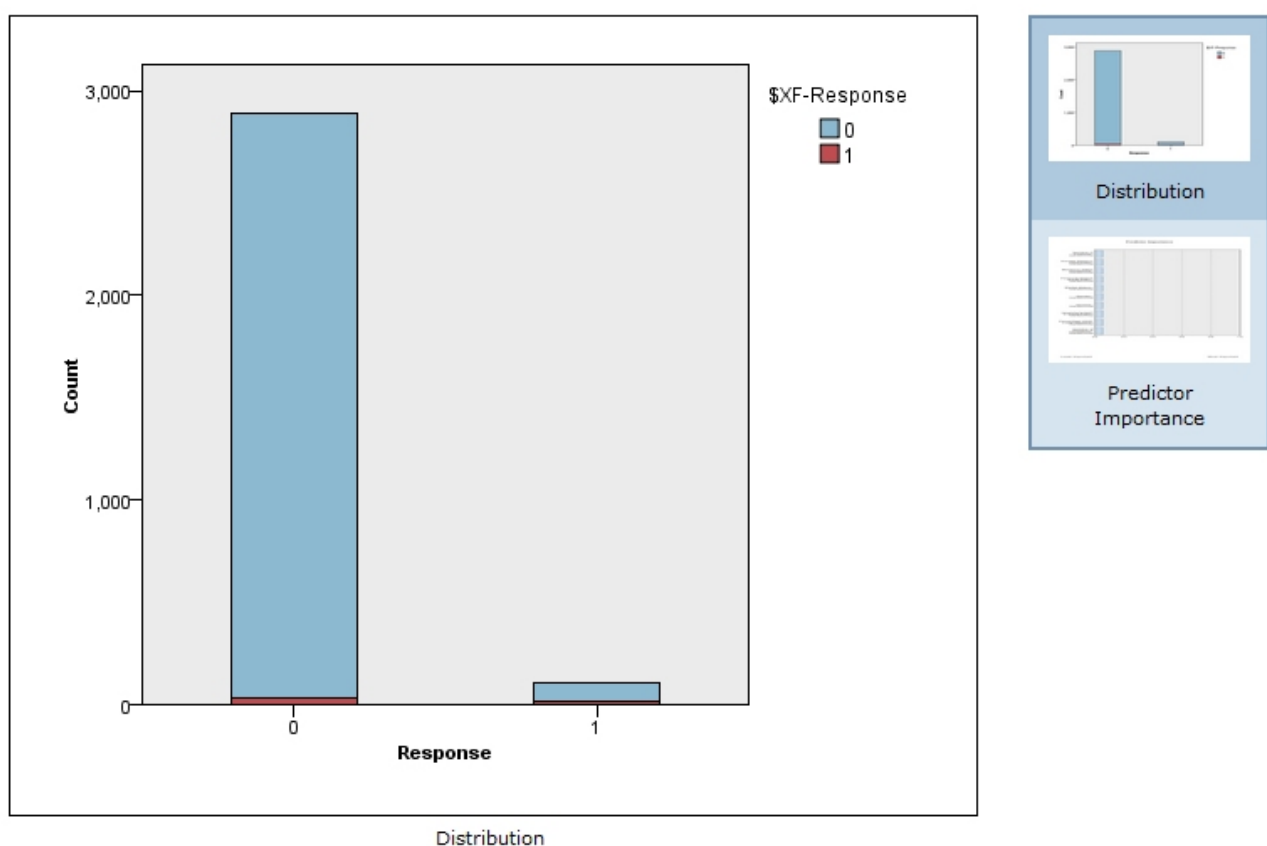


Figura 4. Model tab

1. Selecione *banco de dados de resposta* para a origem de dados e selecione *Resposta* como o campo de destino.
2. Se uma mensagem de **Varredura de dados necessária** aparecer, clique no ícone para executar uma varredura de dados. Então, clique em **Construir modelo**.

Quando o modelo for concluído, um gráfico de distribuição é exibido. O gráfico mostra os valores observados no eixo horizontal, com um valor de 1 indicando uma resposta positiva.

- A altura das barras reflete a taxa de resposta geral. Dos 3000 clientes no conjunto de dados de resposta, 111 responderam, para uma taxa de ocorrência total de cerca de 3,8%.
- As áreas coloridas em cada barra representam a resposta previstas pelo modelo, vermelho indica uma previsão positiva, e azul indica uma resposta negativa, conforme mostrado na legenda no canto

superior direito. Devido à taxa de resposta geral baixa, há relativamente poucas respostas previstas (áreas em vermelho). No entanto, a barra de falso (*Resposta = 0*) é esmagadoramente azul, indicando que a esmagadora maioria dos preditos responderam que não.

Para obter uma ideia melhor sobre como o modelo será executado, continue com a próxima seção (“Avaliando o Modelo”).

Avaliando o Modelo

Na guia Modelagem, clique em **Avaliar** para ver como o modelo é executado quando aplicado aos dados do seu cliente e para entender como o uso do modelo pode impactar seus resultados.

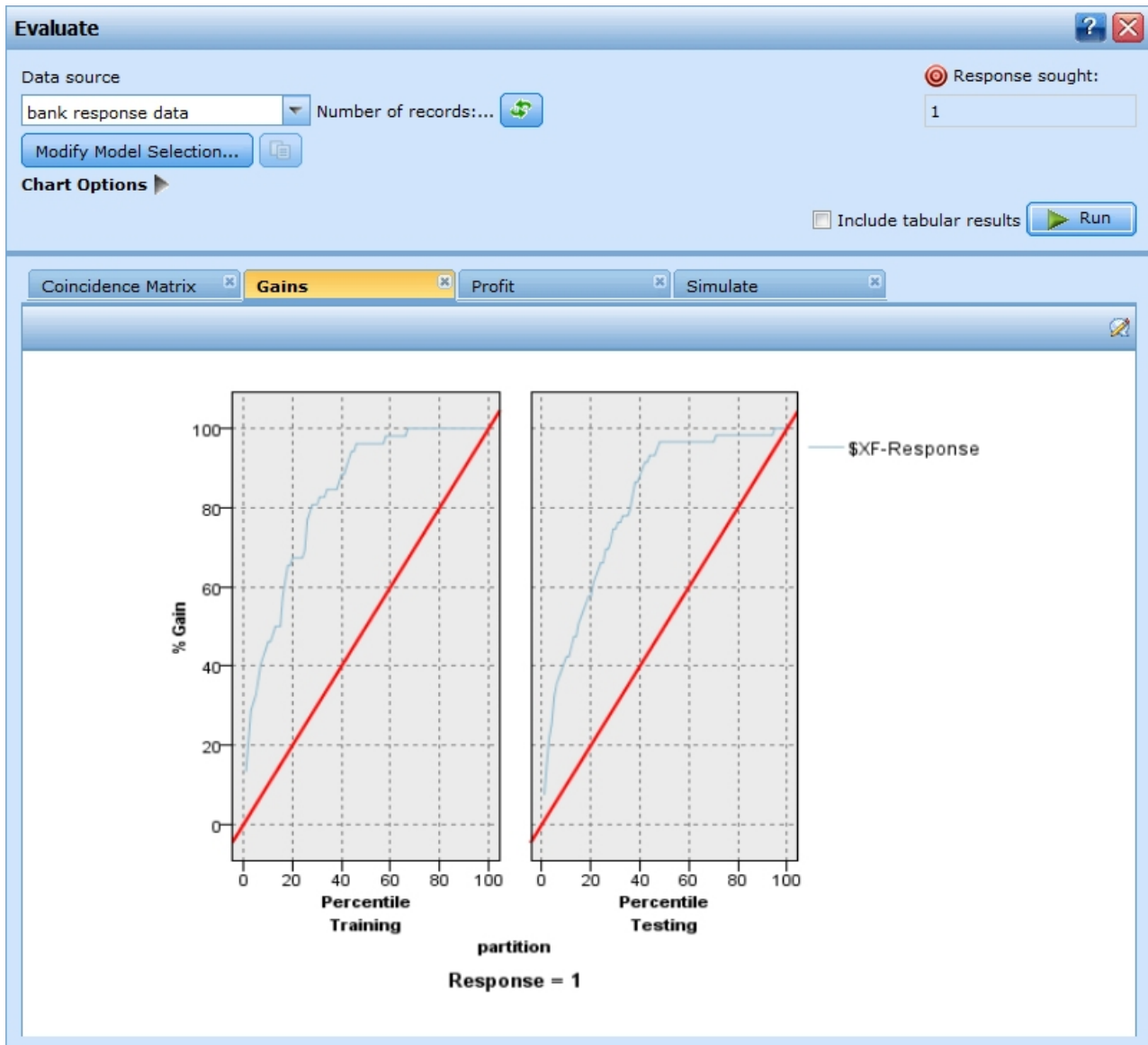


Figura 5. Evaluating the model

1. Para a origem de dados, selecione *dados de resposta do banco*.
2. Selecione **Gráfico de ganhos** e **Gráfico de Lucros** para tipos de gráficos a serem construídos.
3. Selecione **Simular lucro máximo** e especifique *3000* para o preenchimento. Isto corresponde ao número de sua campanha de teste.

4. Especifique 2,25 para custo e 60 para receita. Estes valores indicam o custo para você de cada oferta apresentada (como custos de produção, tratamento e correspondência), bem como a receita esperada de cada resposta favorável.
5. Para **Resposta buscada**, selecione 1.
6. Clique em **Executar**.

O gráfico de ganhos mostra a melhoria ou a "elevação" que pode ser obtida usando o modelo, conforme visto na figura anterior. Como a opção **dados de partição automaticamente** foi selecionada (em configurações Opcionais) ao construir o modelo, gráficos separados são exibidos para as partições de treinamento e teste. Cada um mostra a proporção de ocorrências (respostas positivas) em cada percentil relativo ao número total de ocorrências na amostra. A linha simples diagonal representa a resposta esperada se nenhum modelo foi usado--para reunir metade das ocorrências, seria necessário entrar em contato com metade das pessoas. A linha curva mostra a melhoria utilizando o modelo--direcionando 60% de clientes com maior probabilidade de resposta, é possível ganhar quase todas as ocorrências com uma correspondência muito menor.

O gráfico de lucro mostra os ganhos nos termos de lucro global, com base nos valores de custo e receita especificados. Para a amostra utilizada para treinar o modelo, o lucro parece estar acima, no percentil 30.

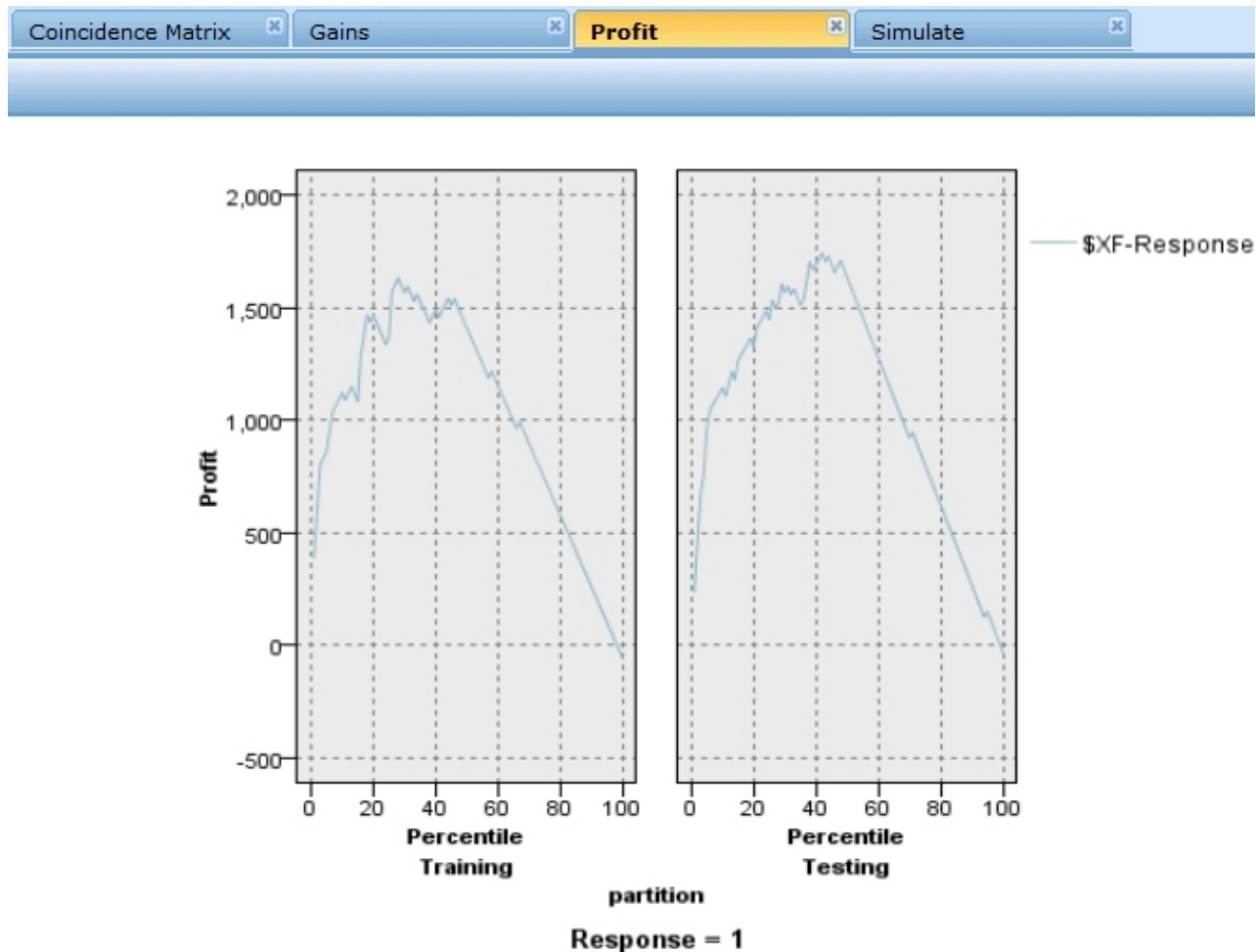


Figura 6. Profit chart

Para entender como isso funciona, acesse a guia Simular para ver o lucro máximo esperado para o tamanho da população especificado.

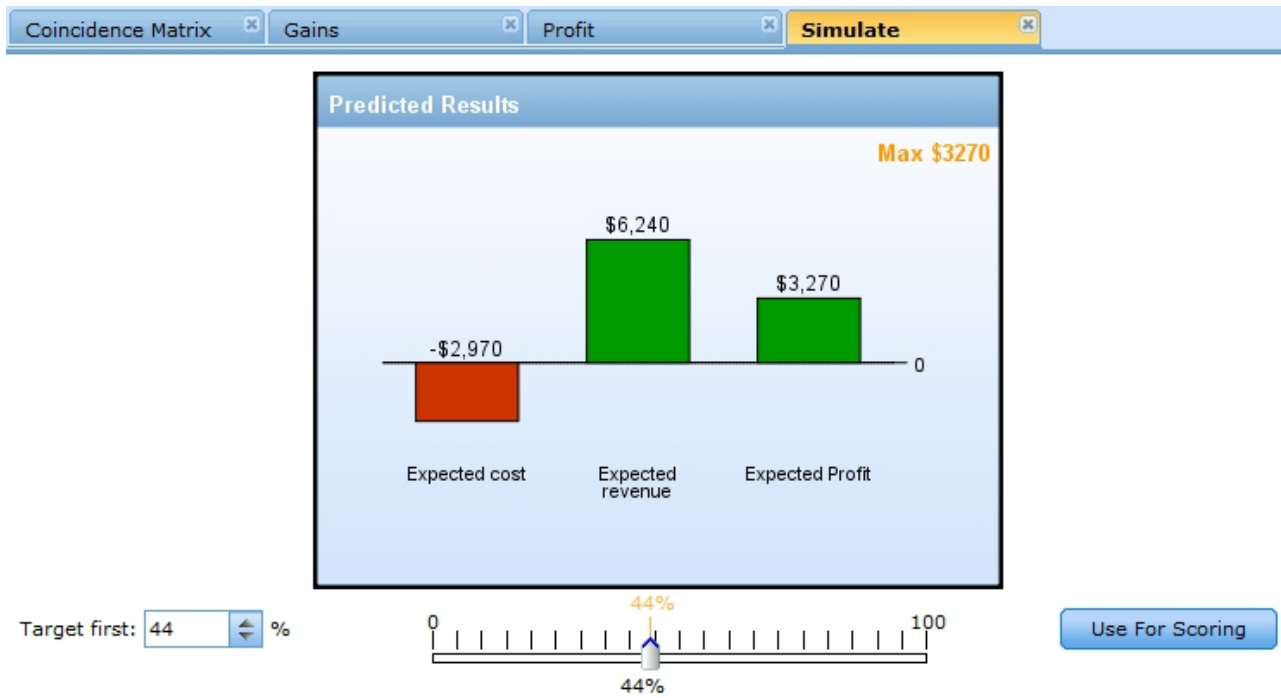


Figura 7. Profit simulation

O lucro esperado é calculado usando a seguinte fórmula:

$$\text{expected profit} = (\text{probability to respond} * \text{revenue}) - \text{cost}$$

Onde a probabilidade de resposta é determinada para cada cliente com base no modelo.

Ao responder a esses identificados pelo modelo como maior probabilidade de resposta, é possível maximizar seu lucro por cliente. À medida que mais clientes forem incluídos, o lucro total aumentará até um determinado ponto, mesmo que o retorno por cliente diminua gradualmente, conforme os clientes com taxas menores de resposta são incluídos. A simulação ajuda você a decidir onde desenhar a linha para obter o lucro máximo.

Compare Profit Table ▼

Export

	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

A Tabela Comparar Lucro resume este resultado com mais detalhes.

- **Destinado 44%.** Essa coluna mostra o resultado ao destinar os 44% dos clientes com as taxas mais altas, conforme previsto pelo modelo. Ao incluir seletivamente os 1320 clientes que provavelmente mais respondem, é possível maximizar seu retorno, resultando em um lucro esperado de \$3270.

- **Aleatório 44%**. Mostra o resultado se uma amostra aleatória foi desenhada no mesmo tamanho. Nesse caso, você enviaria o mesmo número de ofertas (1320), mas a resposta esperada seria inferior (correspondendo à taxa da campanha de teste), porque você não está sendo beneficiado a partir das informações extras fornecidas pelo modelo. Isto resulta em uma perda, porque a taxa de resposta geral não é alta o suficiente para retornar o custo de correspondência de cada oferta.
- **Todos**. Mostra o resultado se você incluiu todos os 3000 clientes na amostra, exatamente o que corresponde ao resultado de sua campanha de teste. Novamente, uma perda é retornada, porque você não está sendo beneficiado a partir do modelo.

Verificando os resultados, é possível observar como a utilização do modelo pode melhorar a taxa de resposta para transformar uma campanha perdida em uma rentável. A diferença entre uma perda de \$90,00 e um lucro de \$3277,50 representa o benefício ganho por meio da modelagem de previsão. Este retorno é baseado nos 3000 clientes de sua campanha de teste, mas ao aumentar o tamanho de sua população alvo, os benefícios aumentam em conformidade.

7. Após revisar os resultados da simulação, selecione **Usar para Escoragem** sob o gráfico Resultados Preditos para atualizar as configurações da guia Escorar com base na porcentagem especificada (44% nesse caso). Consulte o tópico “Pontuando o Modelo” na página 19 para obter mais informações.
8. Se desejar, salve o projeto como *bank_response_model.str*. Na próxima seção, você alterará o tipo de modelo e crie um novo projeto.

Modelagem de Regra de Previsão

Embora os modelos de previsão possam oferecer todo o poder preditivo necessário, há momentos em que você deseja ganhar um pouco mais de controle e insight sobre o funcionamento interno do modelo. Neste caso, a modelagem da regra de previsão (interativa) pode ser usada. Com modelos de regra de previsão, é possível criar manualmente regras para segmentar seus dados ou localizar automaticamente segmentos com concentrações altas ou baixas do valor do campo escolhido. Por exemplo, talvez você procure clientes que deram uma resposta positiva à sua campanha e então identifique segmentos com uma alta probabilidade de resposta.

1. Retornando à guia Modelagem, clique em **Alterar Modelo** e selecione **Modelo de regra de previsão**.
2. Selecione *banco de dados de resposta* para a origem de dados e selecione *Resposta* como o campo de destino.
3. Especifique **1** como a resposta buscada, indicando que deseja procurar por clientes que deram uma resposta positiva.
4. Selecione **Iniciar Construção**.
5. Selecione **Aumentar Modelo** para identificar segmentos com uma alta probabilidade de resposta. As três melhores regras que identificam segmentos com a maior probabilidade de resposta são criadas e exibidas.
6. Clique nos ícones **Incluir/Excluir** para selecionar entre incluir ou excluir clientes selecionados por cada regra. Por exemplo, o segundo e terceiro segmentos de regra mostram taxas de resposta apenas acima de 8%, o que é ainda maior do que a taxa de campanha de teste geral de 3,7%, mas não tão alta como a taxa de 15.65% para o primeiro segmento. Para verificar como a inclusão ou exclusão destas duas regras podem impactar seu lucro geral, é possível executar a simulação do lucro das duas maneiras e comparar os resultados, da seguinte forma:
7. Com todos os três segmentos do conjunto de regras para **Incluir**, clique em **Avaliar**.
8. Selecione **Gráfico de ganhos** e **Gráfico de Lucros** para tipos de gráficos a serem construídos.
9. Selecione **Simular o lucro máximo** e especifique **3000** para o preenchimento, **2,25** para custo e **60** para receita. Estas são as mesmas especificações usadas ao avaliar o modelo preditivo. A resposta buscada é automaticamente concluída.
10. Clique em **Executar** e, quando os gráficos são construídos, acesse a guia Simular.

O lucro máximo simulado para este modelo é \$3.178,90, um pouco menos do que o modelo automatizado, neste caso. Como opção, é possível fechar o diálogo Avaliar, alterar a configuração do

segundo e terceiro segmentos da regra para **Excluir**, e repita a simulação do lucro para verificar como isto altera o resultado. O real benefício da modelagem da regra de previsão (interativa) é o maior controle e insight que ele oferece.

11. Se desejar, salve o projeto como *bank_predictive_rule_model.str*.

Testando o Modelo

Quando tiver construído seu modelo, será possível testar uma seleção de registros do cliente para ver como eles são suscetíveis de responder à sua campanha.

1. Para visualizar como os registros específicos serão processados pela pontuação do modelo, selecione **Testar** a partir da guia Modelagem. O diálogo Teste é exibido.
2. Na área **Registros**, aumente o valor Retorno para 100.
3. Na área **Selecionar Campos para Exibição**, selecione *ID do Cliente*. Além disso, os campos *\$XF-Response* e *\$XFRP-Response* serão incluídos, já que eles indicam a resposta prevista (0 para não, 1 para sim neste caso) e a pontuação de propensão para cada cliente.

Como opção, é possível incluir campos adicionais como *Idade*, *Semana Desde a Última Oferta* e *Possui Registro de Pagamento Inválido* que pode ser útil para entender como ou por que um determinado cliente será pontuado de uma determinada forma pelo modelo.

Pontuando o Modelo

Pontuar um modelo significa aplicá-lo na origem de dados no qual você está interessado. Por exemplo, tendo gerado um modelo que prevê quais clientes terão maior probabilidade de responder, será possível agora pontuar clientes em potencial para identificar quem você deseja incluir em sua campanha. Na caixa de diálogo Avaliar, você determinou a porcentagem de registros a serem pontuados a fim de maximizar o lucro e indicou a porcentagem de registros que desejou pontuar com base neste resultado.

1. Acesse a guia Pontuar.
2. Para **Origem de dados para pontuação**, certifique-se de que *dados de resposta do banco* esteja selecionado.
3. Especifique um destino para os resultados de pontuação. (Se você deseja apenas visualizar as pontuações, é possível ignorar esta etapa.)
4. Em **Campos de Saída**, certifique-se de que o campo *ID do Cliente* esteja selecionado.
5. As configurações de **Pontuação** já devem estar definidas para pontuar a mesma % de registros como a simulação de lucro executado durante a avaliação do modelo. (Se você não selecionou o botão **Utilizar para Pontuação** ao visualizar os resultados de simulação, será possível atualizar as configurações aqui.)
6. Na parte inferior do diálogo, clique em **Visualizar** para exibir uma amostra dos resultados que sua pontuação obterá.
7. Clique em **Visão Geral de Dados** para ter uma visão geral dos dados que sua pontuação irá utilizar. Consulte o tópico “Visão geral de dados” na página 43 para obter mais informações.
Como você escolheu pontuar baseado na propensão, somente clientes com maior propensão responderão serão incluídos. Ao vincular o ID do cliente de volta para o banco de dados do cliente, é possível gerar uma lista de distribuição com base neste resultado.
8. Clique em **Pontuar**; o modelo é aplicado nos dados do cliente e uma mensagem é exibida na parte inferior do diálogo para confirmar isso.

Exemplo de Modelagem do Cluster

Com modelos de armazenamento em cluster, é possível categorizar registros em grupos com características semelhantes. Isso pode ajudar a localizar grupos naturais em seus dados. Por exemplo, é possível segmentar clientes com base em características demográficas ou comportamento de compra.

Este exemplo usa o arquivo de dados *bank_customer_data.txt*, que é distribuído e instalado com o produto. Uma versão completa deste exemplo também é fornecida no arquivo *bank_cluster_model.str*. Entre em contato com seu administrador para obter detalhes sobre instalação e arquivos de amostra se necessário. Consulte o tópico “Arquivos de amostra” na página 9 para obter mais informações.

1. Retorne à página Ativação do aplicativo e crie um novo projeto do IBM SPSS Modeler Advantage.
2. Na guia Dados, no painel Origens de Dados do Projeto, inclua uma nova origem de dados chamada *dados do cliente do banco*, conforme descrito no “Configurando origens de dados” na página 11.
3. Na guia Modelagem, clique em **Alterar Modelo**, selecione **Modelo de armazenamento em cluster** e clique em **Salvar**.
4. Para a origem de dados, selecione *dados do cliente do banco*.
5. Expanda a seção **Configurações Opcionais**. Para modelos de armazenamento de cluster, as configurações opcionais **Opções de Cluster Automático** adicionais estão disponíveis para selecionar um campo de avaliação ou configurar um intervalo desejado de clusters a serem localizados.
Quando um modelo de armazenamento em cluster é criado, nos bastidores, vários modelos diferentes de armazenamento em cluster são realmente criados. Estes modelos então são comparados e classificados por alguma medida de qualidade e o melhor modelo é selecionado para uso pelo modelo de cluster automático. Assim, estas duas configurações de Opções do Cluster Automático fornecem algum controle sobre quais modelos são realmente selecionados. Se um campo de avaliação estiver especificado aqui, o modelo que melhor diferenciar valores do campo de avaliação será selecionado. E se o intervalo desejado for configurado, todos os modelos que localizarem um número de clusters fora do intervalo especificado serão descartados.
6. Para este exemplo, selecione a opção **Configurar um intervalo desejado para o número de clusters localizados** e use o mínimo de **3** e o máximo de **15**. Isso irá garantir que o modelo de armazenamento em cluster não tenha poucos ou muitos clusters.
7. O restante das configurações opcionais é o mesmo para outros tipos de modelagem. Desmarque a opção **Limpar automaticamente e preparar dados para construir um modelo confiável** apenas para facilitar a compreensão do exemplo de armazenamento em cluster automático.
8. Reduza o painel **Configurações Opcionais** e expanda o painel **Clusters**.
No painel Clusters, existe uma seção para clusters manuais e uma seção para clusters automáticos. Clusters manuais permitem que você defina clusters com base em seu conhecimento dos dados. Para este exemplo, vamos definir um cluster manual para *clientes de valor elevado* e um cluster manual para *clientes únicos jovens*, da forma abaixo.
9. Clique no ícone **Criar uma nova regra**. Nomeie a regra *clientes de valor elevado*, inclua em expressões e clique em **OK**.
Meses como um Cliente > 12
Receita > 65000
10. Crie uma nova regra chamada *clientes jovens únicos* com as expressões a seguir.
Idade < 35
Estado Civil = U
11. As duas novas regras serão listadas na seção **Clusters Manuais**. Para verificar quantos clientes estão sendo capturados por estes clusters manuais, clique em **Contagem do registro de clusters**. Em seguida, expanda os clusters manuais novamente e examine a coluna Contagem para verificar quantos clientes foram localizados para cada cluster. Você verá que 67 clientes de valor elevado foram localizados e 33 clientes jovens únicos foram localizados.
12. Em seguida, clique em **Localizar Clusters Automáticos** para localizar automaticamente qualquer outro cluster que seja de seu interesse. Quando terminar, expanda a seção **Clusters Automáticos**. É possível verificar que três clusters automáticos foram localizados.
13. Clique em **Visualizar Clusters Automáticos** para ver mais detalhes. O Visualizador de Resultados de Cluster Automático será exibido.

A página *Resumo do Modelo* fornece informações básicas sobre o algoritmo do modelo de cluster automático que foi utilizado (*TwoStep*, neste exemplo), o número de clusters automáticos localizado e uma ideia básica de qualidade do cluster.

Outros gráficos também estão disponíveis. É possível passar o mouse sobre diversas áreas em cada gráfico para ver mais detalhes e interagir com alguns gráficos (como o gráfico *Importância do Preditor*).

Usando a visualização *Clusters* (o último gráfico), é possível também comparar clusters para procurar padrões interessantes. Em uma implementação real, você pode estudar os clusters automáticos e dar a eles nomes mais significativos (clcando em cada rótulo de cluster na visualização *Clusters*). Após atualizar os nomes dos rótulos dos clusters, os novos nomes serão usados em todos os outros gráficos e voltarão à guia Modelagem. Em seguida, poderá continuar a avaliar e testar o modelo, e finalmente pontuá-lo, uma vez que estiver satisfeito com ele.

14. Se desejado, salve o projeto como *bank_cluster_model.str*.

Para obter mais detalhes sobre a modelagem do cluster, consulte “Construindo um modelo de armazenamento em cluster” na página 61.

Exemplo de Modelagem de Associação

Modelagem de associação é uma técnica importante em análise preditiva. Com modelos de associação, é possível prever quais itens têm maior probabilidade de aparecerem juntos, bem como o fortalecimento dos relacionamentos entre eles. Por exemplo, talvez você identifique a próxima compra provável de um cliente com base em itens de sua cesta atual.

Este exemplo usa o arquivo de dados *retail_purchase_data.txt*, que é distribuído e instalado com o produto. Os dados consistem em informações sobre compra de itens eletrônicos como televisores, computadores e smartphones.

Uma versão completa deste exemplo também é fornecida no arquivo *retail_association_model.str*. Entre em contato com seu administrador para obter detalhes sobre a instalação de arquivos de amostra, se necessário. Consulte o tópico “Arquivos de amostra” na página 9 para obter mais informações.

1. Retorne à página Ativação do aplicativo e crie um novo projeto do IBM SPSS Modeler Advantage.
2. Na guia Dados, no painel Origens de Dados do Projeto, inclua uma nova origem de dados chamada *dados de compra no varejo*. Selecione **Arquivo** e selecione *retail_purchase_data.txt*.
3. Na guia Modelagem, clique em **Alterar Modelo**, selecione **Modelo de associação**, e clique em **Salvar**.
4. Para a origem de dados, selecione *dados de compra no varejo*.
5. Para o formato de dados, selecione **Tabular**. Um conjunto de dados transacionais mostrariam itens que foram incluídos e todos os itens apareceriam em uma única coluna. Mas neste conjunto de dados, cada item está em uma coluna separada com um sinalizador *true* ou *false* para indicar se ele foi adquirido ou não.
6. Clique em **Construir modelo**. Você receberá um erro informando que a construção falhou e sugere que os limites possam estar muito altos. Feche o erro.
Ao procurar regras de associação, geralmente é uma boa ideia procurar regras que normalmente sejam aplicáveis, confiáveis e simples.
7. Expanda a seção Configurações Opcionais. Em Opções de Regra de Associação, altere o primeiro limite de 10 para 1 e reduza o segundo limite de 80 para 1. Geralmente, você pode testar com diversos limites, reconstruindo o modelo a cada vez, para atingir um equilíbrio entre não obter muitas regras e não obter o suficiente.
8. Clique em **Construir modelo**.

Nos resultados do modelo, cada linha representa uma regra de associação separada. A primeira regra na lista diz que se um cliente comprar um *Smartphone*, provavelmente eles também comprem uma

TV Grande. A coluna **Cobertura (%)** mostra que cerca de 15% (ou cerca de 1 em 6 clientes) comprou o *Smartphone*, e a coluna **Confiança (%)** mostra que desses clientes, há cerca de 30% de probabilidade que eles também comprem uma *TV Grande*.

Assim, mesmo que um pequeno conjunto de regras (239 neste exemplo), talvez você consiga usá-las em um contexto de negócios para prever ou mesmo recomendar o que clientes comprarão a seguir.

9. Em seguida, clique em **Testar** e clique em **Executar** para ver como as previsões ou recomendações serão feitas.

Neste caso, o modelo observou as regras para os valores de confiança mais elevados, verificou o que o usuário já havia comprado, e com base nesses dois juntos, fez algumas recomendações sobre o que os clientes provavelmente iriam comprar.

10. Clique no ícone de visualização ao lado do primeiro registro, por exemplo, para verificar se o primeiro cliente comprou uma *TV Grande* e *Alto-falantes*, mas não comprou mais nenhum produto. Feche a visualização.

A recomendação para este cliente é oferecer uma *TV Padrão*, conforme verificado na segunda coluna. O valor da terceira coluna, **0,3**, significa que existe 30% de chance que os clientes que já compraram uma *TV Grande* e *Auto-falantes* comprarão uma *TV Padrão*.

No entanto, como podemos saber se esta é uma anomalia ou uma tendência prevista que pode ser geralmente aplicável? Anote o número da regra (**115** neste exemplo), assim poderemos voltar e observar a regra com mais detalhes.

11. Feche o diálogo Testar.
12. Percorra os resultados do modelo até você ver o número da regra *115*. Observe que para os grandes modelos com muitas páginas de resultados, é possível também clicar em **Localizar uma regra por ID** e digite o número da regra.
13. Clique na seta ao lado para qualquer um dos cabeçalhos da coluna de resultados e selecione **Colunas > Cobertura (N)**

A coluna **Coberturas (N)** será exibida. Com base no número de instâncias, é claro que esta não é uma anomalia, porque a regra 115 é baseada em 58 clientes diferentes que compraram os três produtos em conjunto. Isto é suficiente para nos dar confiança sobre o uso da regra como base para futuras recomendações.

Observamos em uma única regra para verificar como ela pode ser usada para uma recomendação. Mas o IBM Analytical Decision Management também fornece opções sólidas para customizar o modelo, a fim de torná-lo mais relevante para nossos propósitos específicos. Por exemplo, podemos desejar prever quais clientes estão interessados em comprar uma *TV Grande*:

14. Clique em **Aplicar Filtros**. Isso exibe uma caixa de diálogo onde é possível definir filtros, de forma que somente regras que correspondam aos filtros sejam mostrados no modelo.
15. Selecione **Ativar filtro** e selecione *TV Grande* para o item.
16. Clique em **Salvar** para retornar aos resultados do modelo.

O número de regras caiu de 239 para 34.

Você talvez deseje também confirmar que está usando somente regras com base em um número razoável de clientes. Para conseguir isso, talvez você exclua regras que se baseiem no comportamento de menos de 10 clientes:

17. Clique na seta ao lado do cabeçalho da coluna Cobertura (N) e selecione **Classificação Ascendente**.
18. Na coluna Excluir?, selecione todas as regras com um valor de Cobertura (N) de menos de 10 instâncias.
19. Observe que mesmo que muitas regras tenham agora sido filtradas e excluídas, elas ainda estarão no modelo neste ponto. Para removê-las permanentemente, para que não sejam usadas quando o modelo for pontuado, clique em **Excluir regras filtradas e excluídas** e clique em **OK**.

O número de regras retornadas é reduzido para 20, com todas elas baseadas em 10 ou mais instâncias e todas prevendo *TV Grande*.

20. Para ver quantas regras serão aplicadas, clique em **Testar** e em seguida, clique em **Executar**. A *TV Grande* está agora sendo recomendada para diversos clientes e alguns possuem um nível mais alto de confiança do que outros.
21. Feche a caixa de diálogo Testar.
22. Na seção Configurações Opcionais, expanda a seção Opções de Pontuação para exibir opções de pontuação específicas para modelos de associação.
Por padrão, o número máximo de previsões é 3. Isto será baseado nos números mais elevados de confiança, ou é possível alterar o critério de regra para escolher regras com base na cobertura, no suporte à regra, na maior quantidade de *elevação* (o aumento da probabilidade de um determinado item sendo adquirido) ou implementabilidade. Você também talvez queira garantir que não ofereça uma *TV Grande* a clientes que já adquiriram uma selecionando **Garantir previsões não presentes**.
As opções de pontuação padrão são adequadas para a maioria das situações. Mas em alguns casos, a flexibilidade incluída que estas opções oferecem pode ser vantajosa.
A etapa a seguir no processo seria usar a guia Pontuação para pontuar o modelo. Consulte o tópico “Escoragem de modelos em uma tabela de base de dados, arquivo, Analytic Server ou servidor IBM Cognos BI” na página 73 para obter mais informações.
23. Se desejado, salve o projeto como *retail_association_model.str*.

Para obter mais informações sobre a modelagem de associação, consulte “Construindo um modelo de associação” na página 62.

Capítulo 3. Navegando pela área de trabalho

Alterando o idioma exibido

Cada usuário pode alterar facilmente o idioma de seu aplicativo sem precisar reiniciar nenhum servidor. Por exemplo, usuários que falam alemão, usuários que falam inglês e usuários que falam japonês podem usar o mesmo aplicativo em seu próprio idioma.

No menu **Ferramentas** em seu navegador, acesse **Opções da Internet** ou **Opções**, alterne para o idioma desejado e atualize a tela. Se o IBM Analytical Decision Management não for fornecido no idioma desejado, ele será exibido em inglês por padrão ou no segundo idioma de sua escolha, caso ele esteja disponível.

A interface do IBM Analytical Decision Management suporta os seguintes idiomas. Observe que esta lista está sujeita a mudanças de uma liberação para outra, e a documentação não está disponível em todos os mesmos idiomas da interface com o usuário.

- Inglês
- Alemão
- Espanhol
- Francês
- Italiano
- Japonês
- Coreano
- Português (Brasil)
- Russo
- Chinês (China/simplificado)
- Chinês (Taiwan/tradicional)

Ativando aplicativos

A página Ativação de Aplicativos fornece links configuráveis para seus aplicativos instalados. Cada painel representa um aplicativo instalado. Clique no menu suspenso para escolher entre projetos, modelos ou regras salvos recentemente. Também é possível incluir um ou mais atalhos customizados que apontem para versões rotuladas específicas.

- Para acessar a página Ativação de Aplicativos, navegue para *http://hostname:port/DM*, em que *hostname* é o endereço IP ou nome da máquina do repositório e *port* é o número da porta do servidor de aplicativos. Se você não souber essas informações, entre em contato com o administrador. Também é possível incluir um favorito nesta página em seu navegador ou criar um atalho na área de trabalho. Se tiver problemas para acessar a URL do IBM Analytical Decision Management ou qualquer URL do IBM SPSS Collaboration and Deployment Services, é possível que um administrador as tenha customizado. Entre em contato com seu administrador. Mais informações estão disponíveis no capítulo *Raízes de contexto do aplicativo* do *IBM SPSS Collaboration and Deployment Services Guia de instalação e configuração do repositório*.
- Para criar um novo projeto, modelo ou regra, selecione **Novo** na lista suspensa para o aplicativo desejado e clique em **Ir**.
- Para abrir um projeto, modelo ou regra salvos, selecione um atalho ou uma versão usados recentemente na lista suspensa no painel de aplicativo ou selecione **Navegar** na lista suspensa para escolher uma versão ou um objeto diferente. Para ver rapidamente a localização do repositório de um item na lista, passe o mouse sobre o menu suspenso.

- Para incluir um novo atalho customizado que ficará sempre disponível na lista suspensa, clique no ícone de seta triangular no canto superior direito do painel do aplicativo. Opcionalmente, é possível especificar o número de atalhos para exibir.

Outras maneiras de se abrir um aplicativo

- Copie ou digite a URL na barra de endereço em seu navegador ou inclua a URL na lista de favoritos salvos ou favoritos em seu navegador.
- Crie um atalho da área de trabalho apontando para a URL.
- Estabeleça um link com uma URL de aplicativo de dentro do portal ou site intranet da empresa.

Customizando a página de ativação

Os controles no canto superior direito da página Ativação de Aplicativo permite incluir aplicativos e alternar a exibição entre o painel e a visualização de lista. Também é possível arrastar painéis para reorganizá-los na página.

- Para incluir um novo painel de aplicativo, selecione **Incluir Aplicativo** no canto superior direito da página Ativar Aplicativo. (Se nenhum aplicativo adicional estiver disponível, esse link não está disponível.)
Uma lista de aplicativos disponíveis é exibida. Opcionalmente, é possível criar atalhos para uma ou mais versões rotuladas.
- Para alterar a organização dos painéis de aplicativo na página de ativação, clique na barra de título de qualquer painel e mantenha o botão do mouse pressionado para arrastar o painel para uma nova localização.
- Para remover um aplicativo, clique no ícone remover (X vermelho) no canto superior direito do painel de aplicativo. Todos os atalhos que apontarem para esse aplicativo também serão removidos. (O ícone remover não estará disponível se o painel de aplicativo tiver sido bloqueado na página pelo administrador.)
- Projetos, modelos e regras não salvos podem ser listados na galeria na página de ativação. Essa galeria exibe modelos que ainda estão em construção, bem como aqueles cuja construção foi concluída, mas que ainda não foram salvos. Consulte o tópico “Galeria” para obter mais informações.

A página inicial

O aplicativo fornece um fluxo de trabalho passo a passo, conforme representado pelos ícones na página inicial. Clique em qualquer ícone para ir para esse passo.

Dados. Defina os conjuntos de dados a serem usados pelo aplicativo.

Modelo. Construa modelos usando dados analíticos, avalie-os usando gráficos de importância do preditor, distribuição, ganhos, entre outros, para obter insight descobrindo padrões ocultos em dados e use isso para definir melhor suas decisões.

Escorar. Escore o modelo.

Galeria

A galeria é exibida na página Ativação do Aplicativo. Ela lista o trabalho não salvo, como um modelo que foi deixado para ser construído durante a noite, permitindo que você inicie o modelo antes de deixar o trabalho, feche seu navegador e retorne para localizar os resultados na manhã. Também é possível usar a galeria para recuperar seu trabalho após fechar o navegador acidentalmente. Itens na galeria podem incluir projetos, regras, modelos, varreduras de dados, avaliações e outras tarefas de longa execução, conforme apropriado para o aplicativo.

- Se o nome do item estiver sublinhado, clique nele para abrir o item relevante. Por exemplo, se você tiver uma varredura de dados não salva, clique no título da varredura de dados na galeria para abrir a guia Dados.
- Se você não quiser que nenhum item seja mostrado na galeria, clique no ícone de exclusão para removê-lo; uma mensagem de aviso será exibida para confirmar se você deseja excluir o item permanentemente. Ou, para remover vários itens da galeria de uma vez, clique em **Limpar Galeria**. A caixa de diálogo Limpar Galeria será exibida, permitindo que você selecione de quais guias remover todos os objetos.
- Na guia Projetos da galeria, o menu suspenso **Criado em** permite filtrar os itens exibidos na guia para cada aplicativo. Por exemplo, durante o trabalho em um aplicativo, você pode não ver os itens da galeria para outros aplicativos.

Configurando preferências

Há inúmeras preferências que podem ser ajustadas para customizar o IBM Analytical Decision Management para suas necessidades. Primeiramente, essa customização consiste em configurar suas próprias opções padrão para evitar a necessidade de selecionar repetidamente as mesmas ações enquanto você cria, testa e salva seu trabalho. As preferências configuradas são exclusivas para você e se aplicam a qualquer aplicativo que você usar.

Para acessar as configurações de preferência do usuário, de dentro de qualquer aplicativo, clique no ícone **Configurar preferências do usuário** na barra de ferramentas.



Figura 9. Ícone configurar preferências do usuário

É possível especificar suas próprias configurações preferenciais para cada área, conforme detalhado nas seções a seguir.

Configurações gerais

Número padrão de objetos abertos recentemente em atalhos do aplicativo. Selecione o número de itens a serem disponibilizados na lista suspensa na página Ativação de Aplicativo.

Formato de data. Selecione o formato a ser usado durante a exibição ou gravação de datas.

Formato de horário. Selecione o formato a ser usado durante a exibição ou gravação de horários.

Moeda. Selecione a moeda padrão para cada novo projeto. Observe que isso não sobrescreverá dados existentes trazidos como uma moeda diferente.

Casas decimais. Selecione o número de casas decimais para exibir para números.

Número de linhas para mostrar na visualização. Selecione o número máximo de registros a ser mostrado durante a exibição das visualizações; por exemplo, durante a verificação de uma origem de dados na importação.

Número de resultados por página (tabelas paginadas). Selecione o número máximo de resultados para mostrar em cada página durante a exibição de tabelas.

Número máximo de itens exibidos em um gráfico. Às vezes os resultados do gráfico não funcionam bem quando há muitos itens para exibir. Especifique o número máximo de itens para exibir nos gráficos. Isso forçará uma tabela a ser exibida em vez de um gráfico, caso o número de itens exceda esse máximo.

Essa configuração aplica-se apenas aos campos categóricos nas visões gerais de dados e no gráfico de relatório sumarização de escoragem para modelos com uma variável resposta categórica.

Número de segmentos de dados (ladrilhos) a serem aplicados a gráficos e tabelas. Selecione o número máximo de segmentos de dados a ser usado em todo o aplicativo para gráficos e tabelas, onde aplicável.

Configurações de modelo

Tipo de modelo construído por padrão. Especifique se, por padrão, você deseja que o aplicativo construa o melhor modelo automaticamente a partir dos dados disponíveis (modelo preditivo), se você deseja usar um modelo que identifique segmentos que mostrem uma maior ou menor probabilidade relativa a um resultado específico (modelo de regra preditivo), se você deseja usar um modelo que categorize registros em um certo número de clusters (modelo de armazenamento em cluster) ou se você deseja usar um modelo que preveja quais itens têm maior probabilidade de aparecerem juntos e preveja a força do relacionamento entre eles (modelo de associação). Somente tipos de modelo configurados para aparecer no aplicativo estarão disponíveis neste menu suspenso.

Particionar dados automaticamente para ativar a avaliação de modelo na origem de dados de construção para avaliação e teste. A seleção dessa opção divide os dados em subconjuntos separados ou amostras para treinamento e teste de modelo. Ao construir o modelo em um subconjunto e testá-lo em outro, é possível se ter uma ideia de como ele será generalizado para outros conjuntos de dados. Também é possível especificar um valor de porcentagem para o *tamanho da partição de treinamento* e o *tamanho da partição de teste* gerados aleatoriamente. Essas opções estão disponíveis para todos os tipos de modelo, exceto modelagem de associação.

Limpar e preparar automaticamente os dados para a construção de modelo confiável. A seleção dessa opção lhe permite deixar o aplicativo fazer uma grande parte da preparação dos dados. Por exemplo, reparar problemas de dados, exibir dados menos úteis ou derivar novos atributos.

É possível escolher se a operação de limpeza deverá dar um balanceamento igual à velocidade e à precisão, dar prioridade à precisão e não à velocidade ou dar prioridade à velocidade, como quando você está trabalhando com conjuntos de dados extremamente grandes ou procurando um resultado rápido.

Nota: As configurações de preparação de dados automáticos só se aplicam ao criar um novo modelo e a categorização não é executada.

Restringir o tempo gasto com a criação de um único modelo. Use essa opção para especificar a quantidade padrão de tempo a ser gasto construindo cada modelo. Isso pode ser especialmente útil ao processar conjuntos grandes de dados e pode ter um impacto significativo no tempo necessário para a construção de modelo a ser concluída.

Configurações de simulação e teste

Número padrão de registros testados. Especifique o número padrão de registros a serem incluídos durante a execução de um teste.

Data de teste/simulação padrão Especifique se a data de simulação padrão é configurada como a data atual ou uma data específica.

Abrindo e salvando projetos, modelos e regras

É possível abrir, excluir ou salvar no Repositório do IBM SPSS Collaboration and Deployment Services ou fazer download no sistema de arquivos local. Também é possível fazer upload de arquivos de fluxo do IBM SPSS Modeler a partir do sistema de arquivos local e salvar no Repositório do IBM SPSS Collaboration and Deployment Services. O salvamento no Repositório do IBM SPSS Collaboration and

Deployment Services fornece benefícios importantes, incluindo suporte para várias versões, rotulagem e segurança. Dependendo da configuração local, nem todas as opções podem ser suportadas.



Cria um novo modelo.



Salva o projeto atual em sua última localização conhecida no Repositório do IBM SPSS Collaboration and Deployment Services. Opcionalmente, é possível aplicar um novo rótulo ao salvar.



Salva e rotula o projeto atual ou salva com um novo nome. Onde for aplicável, é possível escolher entre salvar no Repositório do IBM SPSS Collaboration and Deployment Services e fazer download sistema de arquivos local.



Abre um projeto a partir do Repositório do IBM SPSS Collaboration and Deployment Services ou sistema de arquivos local, conforme aplicável. Ou, onde for aplicável, também é possível fazer upload de um projeto a partir do sistema de arquivos local. O acesso de sistema de arquivos não é suportado para todos os aplicativos. Se você estiver fazendo upload de um fluxo do IBM SPSS Modeler, é possível escolher abrir o arquivo ou salvá-lo no repositório..

Abrindo e salvando no repositório

O Repositório do IBM SPSS Collaboration and Deployment Services fornece importantes benefícios, incluindo suporte para diversas versões, rótulo e segurança. Cada vez que salva, você cria uma nova versão, e é possível reverter para uma versão anterior a qualquer momento. Também é possível gerenciar versões usando rótulos e aplicar configurações de segurança para determinar quem pode acessar, visualizar ou remover um objeto.

A visualização em árvore na área de janela esquerda exibe pastas no Repositório do IBM SPSS Collaboration and Deployment Services que você tem autoridade para ver. Navegue pela árvore de pasta ou execute uma procura para localizar o que você está procurando. Também é possível filtrar a lista de arquivos por rótulos disponíveis na área de janela inferior esquerda. Por exemplo, você pode clicar em um rótulo *produção* para exibir somente os arquivos marcados com tal rótulo.

Para ver mais detalhes sobre um arquivo, quando disponível, clique no ícone de visualização ou no ícone de informações ao lado dele.

Para excluir um arquivo, clique no ícone de exclusão ao lado dele. Isso é útil principalmente para usuários que não têm acesso ao software cliente IBM SPSS Collaboration and Deployment Services Deployment Manager.

Pasta. Exibe as pastas disponíveis e permite navegar na hierarquia. Escolha a pasta na qual salvar. Também é possível criar uma nova pasta ou renomear uma pasta existente usando os ícones acima da lista de pastas.

Filtrar por Rótulo. Limita a exibição a objetos com o rótulo selecionado. Durante a navegação nas pastas, somente os objetos com esse rótulo são mostrados.

Quando você exporta regras usadas em um aplicativo, um menu suspenso **Salvar como** ou **Exportar como** também fica disponível. Selecione **Conjunto de Regras Único** ou **Regras Individuais** do menu suspenso para especificar como salvar a regra. Se somente regras de referência estiverem disponíveis, o menu suspenso será desativado e a regra será salva como um único conjunto de regras (regras de referência não podem ser salvas individualmente).

Se você escolher salvar como regras individuais, uma tabela será exibida para inserir detalhes para cada regra individual que você deseja salvar.

Nome. Insira um nome exclusivo para o objeto sendo salvo.

Palavras-chave. Insira quaisquer palavras-chave para identificar rapidamente o objeto salvo ao procurá-lo. Por exemplo, pode ser *Vendas e 1-2009*.

Descrição. Insira uma descrição mais longa do objeto salvo. Por exemplo, pode ser *Análise de vendas a varejo; primeiro trimestre de 2009*.

Rótulo(s) Se provavelmente você tiver várias versões do mesmo objeto salvo, é possível identificar cada um com um rótulo exclusivo. Selecione um rótulo existente ou crie um novo.

Trabalhando com rótulos

Rótulos são usados para identificar uma versão específica de um produto, modelo ou regra. Quando você compartilha uma regra entre diversos projetos, por exemplo, é possível especificar que você deseja usar a versão rotulada *produção*. Se uma versão mais nova da regra for criada, a usada no projeto não mudará, a menos que o rótulo *produção* seja explicitamente movido para tal versão.

Durante a referência de objetos compartilhados, é altamente recomendado usar uma versão rotulada, principalmente em um ambiente de produção, já que quaisquer mudanças na versão mais recente podem ter consequências indesejadas. Em certos casos, como quando se faz referência a uma regra de dentro de uma regra, uma versão rotulada deve ser usada. Consulte o tópico “Exportando e reutilizando regras” na página 54 para obter mais informações.

Se nenhuma versão for especificada, a versão mais recente será usada, o que significa que mudanças subsequentes na regra serão selecionadas automaticamente. (Pense em *LATEST* como um rótulo especial que é passado automaticamente para cada nova versão criada.)

Cada vez que você salva um novo objeto ou versão, é possível escolher os rótulos para aplicar em tal versão. Consulte o tópico “Abrindo e salvando projetos, modelos e regras” na página 28 para obter mais informações.

Restrições em nomes de rótulos. Quando se define rótulos para usar com IBM Analytical Decision Management, os caracteres ? # & e : (dois pontos) não podem ser incluídos em nomes de rótulo, já que podem impedir o aplicativo de interpretar corretamente as referências aos objetos que usam esses rótulos. Por exemplo, a seguir estão nomes de rótulos não válidos:

Test & Deploy

DEPLOY:PHASE1

Bloqueando projetos

Quando você abre um projeto, modelo ou regra, uma caixa de seleção Bloquear fica disponível no canto superior direito, permitindo bloquear o objeto no Repositório do IBM SPSS Collaboration and Deployment Services que outros usuários não podem editar.

- Quando um projeto está bloqueado, outros usuários podem abri-lo para visualização, mas não podem salvar mudanças.
- Quando você fecha um projeto bloqueado (por exemplo, fechando o navegador), o bloqueio é liberado para que outros possam editar o objeto.

- Se você escolher desbloquear um projeto sem fechá-lo (cancelando a seleção da caixa Bloquear), será solicitado que você salve quaisquer mudanças. Mudanças não salvas serão perdidas após o objeto ser desbloqueado.
- Após o desbloqueio de um projeto, é possível continuar fazendo mudanças nele, a menos que outro usuário o bloqueie (em cujo caso você não poderá mais salvar mudanças).

Propriedades

O diálogo Propriedades inclui uma guia **Projeto** e uma guia **Repositório**. Para exibir o diálogo Propriedades, clique no ícone Propriedades em qualquer aplicativo.



Figura 10. Ícone Propriedades

Na guia **Projeto**, é possível especificar o máximo de membros permitido para campos *Nominais*. Ela é usada na guia Dados. Selecione **Máximo de membros para campos nominais** e especifique um valor. O padrão é **250**. Isso é útil em situações como as seguintes:

1. Um usuário cria uma nova origem de dados e executa uma varredura de dados.
2. Inúmeros campos contendo informações categóricas úteis aparecem no modelo de dados do projeto como *Sem Tipo*, pois o número de valores distintos nos dados excede o limite de tamanho máximo padrão configurado no diálogo Propriedades.
3. O usuário clica no ícone Propriedades e aumenta o valor **Máximo de membros para campos nominais** para um tamanho apropriado para seus dados.
4. De volta à guia Dados, o usuário clica em **Atualizar varredura de dados**, mas os campos permanecem como *Sem Tipo*, pois a ação de atualização não altera a medida para preservar as mudanças feitas pelo usuário. Em outras palavras, uma atualização retém a medida existente para um campo e recalcula somente os valores, pois quaisquer mudanças feitas pelo usuário no modelo de dados devem ser preservadas.
5. Para obter os valores no modelo de dados do projeto, o usuário deve alterar manualmente a medida de cada campo de *Sem Tipo* para *Categórico* antes de executar a atualização. Observe que se houver muitos campos afetados, talvez seja mais rápido recriar a origem de dados.

Observe que a configuração Máximo de membros para campos nominais não se aplica às origens de dados do tipo de arquivo *sav*. Muitos arquivos *.sav* têm valores e arquivos customizados pelo usuário, portanto, eles não serão afetados por essa configuração. O usuário deve alterar manualmente os campos.

Na guia **Repositório**, é possível incluir palavras-chave e uma descrição para ajudar a identificar um projeto, modelo ou conjunto de regras durante sua procura em uma data futura:

- **Palavras-chave.** Insira quaisquer palavras-chave para identificar rapidamente o objeto salvo ao procurá-lo. Por exemplo, pode ser *Vendas* e *Q1-2013*.
- **Descrição.** Insira uma descrição mais longa do objeto salvo. Por exemplo, pode ser *Análise de vendas a varejo; primeiro trimestre de 2013*.

Capítulo 4. Gerenciando origens de dados

Use a guia Dados para definir origens de dados para análise, simulação e teste, escoragem ou outro uso. Conjuntos de dados podem ser configurados por seu administrador ou você pode incluir o seu próprio.

A guia inclui uma guia **Origem** e uma guia **Derivado**. A guia Origem é dividida nestas principais áreas:

- **Modelo de Dados do Projeto.** Define os campos requeridos pelo aplicativo. Todos os outros conjuntos de dados são mapeados com relação a esta origem. A seção **Campos de Origem de Dados** lista os tipos e os campos de entrada usados por seu projeto. Se desejar, use a guia **Derivado** para incluir saída do modelo ou expressões de diferentes origens de dados para estender seu modelo de dados do projeto.
- **Origens de Dados.** A seção **Origens de Dados do Projeto** lista as origens de dados que foram salvas para uso com o atual projeto, modelo ou regra. A seção **Minhas Origens de Dados** lista origens de dados que você definiu ou salvou. Conjuntos de dados desta lista podem ser copiados para e de projetos diferentes e depois solicitados para uso como um modelo de dados do projeto, possibilitando o compartilhamento de dados entre aplicativos.

Modelo de dados do projeto

O modelo de dados do projeto age como um modelo listando os tipos e os campos de entrada que serão usados por seu projeto.

Origens de dados. Exibe a origem de dados selecionada para o aplicativo.

ID do registro. Se você desejar, é possível selecionar um campo para usar como o ID da *entidade* para ativar o retrocesso de SQL. Isso pode melhorar o desempenho quando os dados são originados de um banco de dados e se o campo selecionado for um número inteiro. No IBM SPSS Modeler Advantage, o campo selecionado aqui será selecionado por padrão como uma saída de um score de lote e ter sua seleção cancelada por padrão como uma entrada para a construção de um modelo. Esse campo é opcional.

Operacional. Uma marca nesta caixa mostra que os dados no campo estão disponíveis para uso em regras e expressões e para previsões. Remova a marca se o campo for usado como campo de destino para análise. Por exemplo, se você estiver executando uma consulta para saber se os clientes responderão bem a uma campanha de correio direto, o campo de resultados será analítico, já que não é preciso saber valores de resposta antecipadamente.

Nome do campo. Exibe o nome de cada campo, conforme listado na origem de dados, junto com um ícone que identifica seu tipo de armazenamento.

Medição. Exibe o tipo de medição de cada campo na origem de dados. Se você alterar qualquer um dos tipos de medição, será preciso atualizar a varredura de dados para assegurar a compatibilidade dos dados. Consulte o tópico “Níveis de medição” na página 43 para obter mais informações.

Valores. Lista os valores para cada campo na origem de dados. Por exemplo, os valores máximo e mínimo em um intervalo. Para campos de flag, mantenha o ponteiro do mouse pressionado sobre o valor para exibir uma dica de ferramenta indicando os valores "true" e "false" definidos para o campo. Também é possível clicar em um valor para editá-lo.

Normalmente, o modelo de dados do projeto define as entradas como um conjunto de campos provenientes de uma única tabela lógica. Esse é sempre o caso para IBM SPSS Modeler Advantage. No entanto, para outros aplicativos, o modelo de dados do projeto pode consistir em uma tabela primária juntamente com 0 e mais tabelas secundárias. Essas tabelas secundárias são definidas pelas origens de

dados secundários que são associadas à origem de dados primários.

Origens de dados do projeto

Origens de dados do projeto podem incluir aquelas predefinidas pelo administrador ou incluídas pelo usuário. Opcionalmente, o administrador pode ter bloqueado uma ou mais origens de dados para evitar que os usuários as modifiquem ou removam ou bloqueado todas as opções de dados para que os usuários não possam criar novas origens de dados.

Se as entradas da origem de dados não corresponderem diretamente às do modelo de dados do projeto, é possível mapear a mais antiga para a mais recente e corrigir a discrepância. Por exemplo, se o modelo de dados do projeto requerer um campo denominado *compra* com valores *Sim* e *Não* (flag de nível de medição), qualquer origem de dados usada poderá ter um campo comparável que possa ser mapeado de acordo.

Nome. Exibe o nome da origem de dados e mostra um ícone que identifica seu tipo de arquivo.

Número de registros. Clique no ícone nesta coluna para mostrar a contagem de registros para qualquer origem de dados na tabela. A contagem é mostrada ao lado do ícone.

Visualização. Clique no ícone nesta coluna para visualizar uma amostra dos dados contidos na origem. Para obter mais informações, consulte “Visualizando dados” na página 43.

Visão geral. Clique no ícone nesta coluna para obter uma visão geral da origem de dados. Para obter mais informações, consulte “Visão geral de dados” na página 43.

Compatível. Exibe uma nota dizendo que a origem de dados é usada como modelo de dados do projeto ou mostra uma bola verde, laranja ou vermelha para indicar o nível de compatibilidade da origem de dados com o modelo de dados do projeto.

- Uma bola verde mostra que a origem de dados é operacionalmente compatível com a origem de dados do modelo de dados do projeto. Uma origem de dados operacionalmente compatível é aquela que inclui todos os campos operacionais do modelo de dados do projeto, mas pode ter campos adicionais. Essa origem de dados é adequada para regras, escoragem, simulação e operações de teste.
- Uma bola laranja mostra que a origem de dados tem pelo menos um campo compatível com o modelo de dados do projeto, com mesmo nome e tipo. Essa origem de dados também pode ter campos adicionais, além de ser adequada para construção e avaliação de modelos.
- Uma bola vermelha mostra que a origem de dados é incompatível com o modelo de dados do projeto e os campos devem ser mapeados antes de ele ser usado no aplicativo. Uma origem de dados incompatível é aquela que tem pelo menos um campo cujo tipo é incompatível com o tipo de modelo de dados do projeto equivalente.

Mapear campos. Esta opção permite comparar os campos da origem de dados com aqueles no modelo de dados do projeto e mapear ou remover o mapeamento de quaisquer campos compatíveis para corresponderem aos requeridos pelo modelo de dados do projeto. Para obter mais informações, consulte “Mapeando campos” na página 45.

Copiar. Copia a origem de dados na área **Minhas Origens de Dados**.

Minhas origens de dados

As origens de dados nesta lista são salvas com sua conta do usuário, portanto, elas ficam disponíveis sempre que você efetua login e podem ser copiadas em qualquer projeto, modelo ou regra que você abrir (supondo que seu administrador lhe dê autoridade para isso). Os campos nesta parte da guia trabalham da mesma maneira que aqueles na área **Origens de Dados do Projeto**; embora não haja nenhuma coluna **Compatibilidade**.

A coluna **Copiar** permite copiar a origem de dados na área **Origens de Dados do Projeto**.

Trabalhando com origens de dados

- Para incluir uma nova origem de dados nas listas **Origens de Dados do Projeto** ou **Minhas Origens de Dados**, selecione **Incluir uma origem de dados**. Para obter mais informações, consulte “Criando uma nova origem de dados”.
- Para alterar o nível de medição, ou tipo, de um campo na origem de dados no **Modelo de Dados do Projeto**, selecione o nível relevante. Para obter mais informações, consulte “Níveis de medição” na página 43.
- Para copiar origens de dados na e da lista **Origens de Dados do Projeto**, clique na seta apropriada na coluna **Copiar**.
- Para mapear nomes de campo para uma origem de dados para o modelo de dados do projeto, clique no link apropriado sob a coluna **Compatível**. (Após os nomes de campo serem mapeados, o link não será mais exibido.) Para obter mais informações, consulte “Mapeando campos” na página 45.
- Para visualizar uma origem de dados, clique no ícone **Visualizar**. Para obter mais informações, consulte “Visualizando dados” na página 43.
- Para incluir campos adicionais (expressões, campos usando regras de segmento ou saída de modelo de diferentes origens de dados) para estender seu modelo de projeto, clique na guia **Derivado**. Para obter mais informações, consulte “Derivando campos” na página 45.

Criando uma nova origem de dados

1. Para acessar o Editor de Origem de Dados, selecione **Incluir uma origem de dados** sob a área **Origens de Dados do Projeto** ou **Minhas Origens de Dados**.
2. Digite um nome para a origem de dados.
3. Selecione o tipo de origem. Observe que a lista de bancos de dados disponíveis é daqueles disponíveis na máquina IBM SPSS Modeler.
 - **Arquivo**. Muitos tipos de arquivo são suportados, como texto (*.txt, *.csv, etc.), Excel, os formatos *.sav e *.zsav usados pelo IBM SPSS Statistics, além dos arquivos SAS. Consulte o tópico “Origem do arquivo” na página 36 para obter mais informações.
 - **Banco de Dados**. Qualquer banco de dados que suporte ODBC, incluindo Microsoft SQL Server, DB2, Oracle e outros. Consulte o tópico “Origem do banco de dados” na página 38 para obter mais informações.
 - **Visualização de Dados**. Cria uma conexão a ser usada com regras do IBM Operational Decision Management. Para regras complexas que usam um modelo de dados hierárquicos, observe que os administradores podem restringir o uso de tais regras porque elas podem comprometer o desempenho. Consulte o tópico “Origem da Visualização de Dados” na página 39 para obter mais informações.
 - **Servidor IBM Cognos BI**. Qualquer banco de dados que suporte o servidor IBM Cognos BI. Consulte o tópico “Origem do IBM Cognos BI” na página 40 para obter mais informações.
 - **Analytic Server**. Não é possível selecionar este tipo de origem durante a criação de uma nova origem de dados. Em IBM SPSS Modeler 16, IBM SPSS Analytic Server foi introduzido, permitindo o uso de Hadoop Distributed File System (HDFS) como uma origem de dados. Os usuários do IBM SPSS Modeler podem criar fluxos usando o nó de origem HDFS Analytic Server e os modelos de construção baseados no sistema Big Data Hadoop. Os usuários do IBM Analytical Decision Management podem abrir estes fluxos de modelo no IBM SPSS Modeler Advantage para visualização (eles não podem criar novos projetos usando origens de dados do Analytic Server).
4. Especifique as opções conforme apropriado para o tipo de origem.

É possível clicar no ícone **Número de registros** para exibir o número de registros na atual origem de dados ou clicar em **Visualizar** para exibir os dez primeiros registros da origem de dados.

Origem do arquivo

O aplicativo suporta vários tipos de arquivo de dados. Quando você insere o nome do arquivo ou navega pelo arquivo, o aplicativo detecta o tipo automaticamente e expande o Editor de Origem de Dados para exibir campos adicionais relacionados a esse tipo. Se o tipo de arquivo estiver incorreto, ou requerer uma versão específica (por exemplo, Microsoft Excel 1997-2003), é possível alterar o tipo para o requerido.

É possível clicar em **Navegar** para selecionar um arquivo do repositório ou clicar **Fazer upload do arquivo local** para selecionar um arquivo do sistema de arquivos local (se ativado).

Se você fizer upload de um arquivo local, ele será transferido por upload para IBM SPSS Modeler Server. No diálogo Upload de Arquivo, navegue para e selecione o arquivo local que deseja transferir por upload e navegue para e selecione o destino de upload do IBM SPSS Modeler Server. Quando você seleciona um destino, é possível alterar o nome do arquivo se desejado.

A capacidade **Fazer upload de arquivo local** fica desativada por padrão. Os administradores podem usar IBM SPSS Collaboration and Deployment Services Deployment Manager baseado em navegador para ativá-lo ou desativá-lo e para aumentar ou diminuir o limite de tamanho de arquivo máximo permitido.

Arquivos de dados baseados em texto

Quando seleciona uma origem de dados baseada em texto, é solicitado que você insira detalhes adicionais.

A primeira linha contém nomes de coluna. Selecione esta opção se os nomes de cada coluna estiverem incluídos como uma linha de título na origem de dados.

Codificação. Especifica o método de codificação de texto usado. É possível escolher entre o padrão do sistema ou UTF-8.

Símbolo decimal. Especifica como decimais devem ser representados nos dados.

- **Ponto (.)**. O caractere de ponto será usado como separador decimal.
- **Vírgula (,)**. O caractere de vírgula será usado como separador decimal.

Delimitadores. Usando as caixas de seleção listadas para este controle, é possível especificar quais caracteres, como a vírgula(,), definem limites de campo no arquivo. Também é possível especificar mais de um delimitador, como ", |" para registros que usam vários delimitadores. O delimitador padrão é a vírgula.

Nota: Se a vírgula também estiver definida como separador decimal, as configurações padrão aqui não funcionarão. Nos casos em que a vírgula for o delimitador de campo e o separador decimal, selecione **Outros** na lista Delimitadores. Depois, especifique uma vírgula manualmente no campo de entrada.

Selecione **Permitir vários delimitadores em branco** para tratar vários caracteres delimitadores em branco adjacentes como um único delimitador. Por exemplo, se um valor de dados fosse seguido por quatro espaços e outro valor de dados, esse grupo seria tratado como dois campos em vez de cinco.

Opções avançadas

Caracteres de comentário EOL. Especifique caracteres, como # ou !, para indicar anotações nos dados. Independentemente dos caracteres que aparecerem no arquivo de dados, tudo, exceto o próximo caractere de nova linha, será ignorado.

Especifique campos de entrada. Especifique o número de campos de entrada a serem usados em cada registro.

Especifique formato de dados para tipos de origem de dados **Arquivo** que são do tipo **Campos de comprimento de variável em texto simples**, é possível usar esta seção para configurar o tipo de armazenamento de entrada e o formato de cada campo para assegurar que os valores sejam lidos corretamente. Isso é semelhante à funcionalidade disponível no cliente IBM SPSS Modeler. A opção **Substituir** indica se o padrão é substituído. Selecionar **Substituir** permite os controles **Armazenamento e Formato de entrada**. Cancelar a seleção de **Substituir** alterará os valores de volta para seus padrões originais. **Formato de entrada** só se aplica a tipos de armazenamento real, data, tempo e registro de data e hora.

Ignorar caracteres de cabeçalho. Especifique quantos caracteres deseja ignorar no início do primeiro registro.

Linhas para varredura de tipo. Especifique quantas linhas varrer para os tipos de dados especificados.

Remover espaços no início e à direita. Selecione para descartar espaços à esquerda e à direita em sequências de caracteres na importação. É possível escolher remover da esquerda, da direita, de ambos ou lados ou de nenhum dos lados.

Caracteres inválidos. Selecione **Descartar** para remover caracteres inválidos da origem de dados. Selecione **Substituir por** para substituir caracteres inválidos pelo símbolo especificado (somente um caractere). Caracteres inválidos são caracteres nulos ou aqueles que não existem no método de codificação especificado.

Aspas. Usando a lista suspensa, é possível especificar como aspas simples e duplas são tratadas na importação. É possível escolher **Descartar** todas as aspas, **Incluir como Texto** incluindo-as no valor do campo ou **Emparelhar e Descartar** para fazer pares de aspas e removê-los. Se uma aspa ficar sem correspondência, você receberá uma mensagem de erro. Tanto **Descartar** quanto **Emparelhar e Descartar** armazenam o valor do campo (sem aspas) como uma sequência de caracteres.

Nota: Ao utilizar **Emparelhar e descartar**, os espaços são mantidos. Ao utilizar **Descartar**, espaços à direita dentro e fora das aspas são removidos (por exemplo, ' " ab c" , "d ef " , " gh i " ' resultará em 'ab c, d ef, gh i'). Ao utilizar **Incluir como texto**, as aspas são tratadas como caracteres normais, de modo que os espaços à esquerda e à direita sejam removidos naturalmente.

Especificar os campos de entrada. Consulte o tópico “Selecionando campos de entrada” na página 42 para obter mais informações.

Arquivo de dados do Excel

Quando você selecionar uma origem de dados do Excel, será solicitada inserção de detalhes adicionais:

A primeira linha contém nomes de coluna. Selecione esta opção se os nomes de cada coluna estiverem incluídos como uma linha de título na origem de dados.

Intervalos nomeados. Permite que você selecione um intervalo nomeado de células conforme definido na planilha do Excel. Se você usar um intervalo nomeado, outras configurações de intervalo de dados e planilha não são mais aplicáveis e são desativadas como resultado.

Escolher planilha. Especifica a planilha a ser importada pelo nome ou pelo índice.

- **Por nome.** Selecione o nome da planilha que você deseja importar.
- **Por índice.** Especifique o valor de índice para a planilha que você deseja importar, começando com 0 para a primeira planilha, 1 para a segunda planilha e assim por diante.

Intervalo na planilha. É possível importar dados que começam com a primeira linha preenchida ou com um intervalo de células explícito.

- **O intervalo começa na primeira linha preenchida.** Localiza a primeira célula preenchida e usa como o canto superior esquerdo do intervalo de dados.
- **Intervalo explícito de células.** Permite que você especifique um intervalo explícito por linha e coluna. Por exemplo, para especificar o intervalo do Excel A1:D5, você pode digitar A1 no primeiro campo e D5 no segundo (ou, como alternativa, R1C1 e R5C4). Todas as linhas no intervalo especificado são retornadas, incluindo as linhas em branco.

Nas linhas em branco. Se uma linha em branco for encontrada, é possível optar por pular e ignorar a linha ou escolher **Retornar linhas em branco** para continuar lendo todos os dados até o final da planilha, incluindo linhas em branco.

Especificar os campos de entrada. Consulte o tópico “Selecionando campos de entrada” na página 42 para obter mais informações.

Arquivos de Dados IBM SPSS Statistics

Quando você selecionar uma origem de dados (arquivo IBM SPSS Statistics (.sav ou .zsav), será solicitada uma inserção de detalhes adicionais. Se o arquivo for protegido por senha, a senha também será solicitada.

nomes de variáveis. Selecione um método de tratamento de nomes de variáveis e rótulos na importação de um arquivo IBM SPSS Statistics .sav ou .zsav.

- **Ler nomes e rótulos.** Selecione para ler os nomes de variáveis e rótulos. Esta é a opção padrão. Rótulos podem ser exibido em gráficos, navegadores de modelo e outros tipos de saída.
- **Ler rótulos como nomes.** Selecione para os rótulos descritivos das variáveis do IBM SPSS Statistics em seu arquivo .sav em vez dos nomes de campo curto e usar esses rótulos como nomes de variáveis.

Valores. Selecione um método de tratamento valores e rótulos na importação de um arquivo IBM SPSS Statistics .sav ou .zsav.

- **Ler dados e rótulos.** Selecione para ler os valores reais e o valor do rótulo. Esta é a opção padrão.
- **Ler rótulos como dados.** Selecione se deseja usar rótulos de valor do arquivo .sav ou .zsav em vez dos códigos numérico ou simbólico usados para representar os valores. Por exemplo, selecionar essa opção para dados com um campo de gênero cujos valores de 1 e 2 representam de fato *masculino* e *feminino*, respectivamente, converterá o campo em uma sequência de caracteres e importará *masculino* e *feminino* como valores reais.

É importante considerar valores omissos em seus dados do IBM SPSS Statistics antes de selecionar essa opção. Por exemplo, se um campo numérico usar rótulos apenas para valores omissos (0 = *Sem resposta*, -99 = *Desconhecido*), então, selecionar a opção acima importará apenas os rótulos de valores *Sem resposta* e *Desconhecido* e converterá o campo em uma sequência de caracteres. Em tais casos, você deve importar os próprios valores.

Especificar os campos de entrada. Consulte o tópico “Selecionando campos de entrada” na página 42 para obter mais informações.

Origem do banco de dados

Quando seleciona uma origem de banco de dados, é solicitado que você insira detalhes adicionais:

Banco de Dados. Selecione o banco de dados que deseja usar.

Conectar. Selecione se deve conectar ao banco de dados usando credenciais do banco de dados ou credenciais armazenadas no Repositório do IBM SPSS Collaboration and Deployment Services. Então, insira seu nome e senha de login do banco de dados ou selecione uma credencial a partir do repositório e clique em **Conectar** para verificar se você tem acesso ao banco de dados.

Observe que o nome e senha de login do banco de dados não são salvos. Portanto, na próxima vez que abrir e executar o projeto, você será solicitado a efetuar login, a menos que seu administrador tenha configurado o Single Sign On (SSO) para conexões de banco de dados do IBM SPSS Modeler Server e ODBC. Consulte a seção *Considerações de segurança* do guia de instalação do IBM Analytical Decision Management para obter mais informações.

Nome da tabela. Selecione a tabela da qual os dados devem ser usados.

Remover espaços no início e à direita. Selecione para descartar espaços à esquerda e à direita em sequências de caracteres na importação. É possível escolher remover da esquerda, da direita, de ambos ou lados ou de nenhum dos lados.

Tabela de cotação e nomes de coluna. Especifique se deseja que os nomes de tabela e coluna sejam colocados entre aspas quando as consultas forem enviadas para o banco de dados (se, por exemplo, eles contiverem espaços ou pontuação).

- A opção **Conforme necessário** colocará os nomes de tabela e campo entre aspas *somente* se eles incluírem caracteres não padrão. Caracteres não padrão incluem caracteres não ASCII, caracteres de espaço e quaisquer caracteres não alfanuméricos além de ponto final (.).
- Selecione **Sempre** se quiser que *todos* os nomes de tabelas e campos sejam colocados entre aspas.
- Selecione **Nunca** se você não deseja *nunca* que os nomes de tabelas e colunas sejam postos entre aspas.

Especificar os campos de entrada. Consulte o tópico “Selecionando campos de entrada” na página 42 para obter mais informações.

IBM Analytical Decision Management suporta as mesmas origens de dados do banco de dados que IBM SPSS Modeler para leitura, gravação e envio de dados por push. Para obter detalhes, consulte relatório de compatibilidade do produto para IBM SPSS Modeler 18. Observe que isso é diferente dos bancos de dados suportados pelo IBM SPSS Collaboration and Deployment Services 8, que é o banco de dados que hospeda o repositório.

Origem da Visualização de Dados

Uma Origem da Visualização de Dados cria uma conexão para ser usada com regras do IBM Operational Decision Management. Para regras complexas que usam um modelo de dados hierárquicos, observe que os administradores podem restringir o uso de tais regras porque elas podem comprometer o desempenho. Para obter detalhes, consulte “Opções de bloqueio de dados” na página 83.

Uma *Origem de Visualização de Dados* pode ser configurada no IBM SPSS Modeler para incluir dados em fluxos que foram definidos em uma visualização de dados analítica do IBM SPSS Collaboration and Deployment Services. Uma visualização de dados analítica define uma estrutura para acessar dados que descrevem as entidades usadas em modelos preditivos e regras de negócios. A visualização associa a estrutura de dados com origens de dados físicas para a análise.

Para obter mais informações, consulte a documentação do IBM SPSS Modeler e a seção *Nós de Origem, Processo e Saída* da ajuda do IBM SPSS Modeler.

Quando seleciona uma Origem de Visualização de Dados, é solicitado que você insira detalhes adicionais:

Origem de dados analítica. Selecione a origem de dados analítica do repositório.

Nome da tabela. Selecione a tabela da qual os dados serão usados. Essa é uma tabela do modelo de dados na visualização de dados analítica. Cada tabela no modelo de dados representa um conceito ou uma entidade envolvidos no processo de analítica preditiva. Campos para as tabelas correspondem aos atributos das entidades representadas pelas tabelas. Por exemplo, se você estivesse analisando pedidos de cliente, seu modelo de dados poderia incluir uma tabela para clientes e uma tabela para pedidos. A tabela de clientes pode ter campos para identificação de cliente, idade, sexo, estado civil e país de residência. A

tabela de pedidos pode ter campos para a identificação de pedido, número de itens no pedido, custo total e identificação para o cliente que fez o pedido. O campo de identificação do cliente poderia ser usado para associar os clientes na tabela de clientes com seus pedidos na tabela de pedidos.

Plano de Acesso a Dados. Selecione o plano de acesso a dados a partir do qual os dados devem ser usados. Um plano de acesso a dados associa as tabelas de modelo de dados em uma visualização de dados analítica às origens de dados físicas. Uma visualização de dados analítica normalmente contém diversos planos de acesso a dados. Ao alterar o plano de acesso a dados em uso, você altera os dados usados por seu projeto. Por exemplo, se a visualização de dados analítica contiver um plano de acesso a dados para treinar um modelo e um plano de acesso a dados para testar um modelo, é possível alternar do treinamento de dados para o teste de dados alterando o plano de acesso a dados em uso.

Atributos Opcionais. Se estiverem disponíveis, e se você desejar, selecione atributos opcionais para incluir na origem de dados.

Especificar os campos de entrada. Consulte o tópico “Selecionando campos de entrada” na página 42 para obter mais informações.

Origem do IBM Cognos BI

É possível importar dados do banco de dados IBM Cognos BI ou relatórios de lista única para usar no IBM Analytical Decision Management.

Nota: Se o IBM SPSS Modeler Server em uso com o IBM Analytical Decision Management não suportar o servidor IBM Cognos BI, esta opção não estará disponível.

Quando você seleciona **Servidor IBM Cognos BI** e clica **Conectar**, é solicitado que você insira os detalhes da conexão do servidor.

1. Digite a **URL do Servidor** do servidor IBM Cognos a partir do qual importar ou exportar dados. Entre em contato com o administrador do sistema Cognos se você não estiver certo quanto a qual URL usar.
2. Selecione o **Modo** no qual irá se conectar. Para efetuar login como um usuário específico (por exemplo, como um administrador), selecione **Configurar Credenciais** e insira o namespace, nome de usuário e senha do IBM Cognos:
 - Insira o **Namespace** do provedor de autenticação de segurança do IBM Cognos usado para efetuar logon no servidor. O provedor de autenticação é usado para definir e manter usuários, grupos e papéis e para controlar o processo de autenticação.
 - Insira o **Nome de Usuário** do IBM Cognos com o qual efetuar logon no servidor.
 - Insira a **Senha** associada ao nome de usuário especificado.

OU

Navegue e selecione uma credencial IBM Cognos armazenada no Repositório do IBM SPSS Collaboration and Deployment Services.

3. Como alternativa, para efetuar login sem credenciais do usuário, caso você não possa especificar os campos de namespace, nome de usuário e senha, selecione **Usar Conexão Anônima**.

Nota: Algumas conexões do servidor não permitem a conexão anônima.

Selecionando detalhes de objetos Cognos

Após configurar a conexão do servidor IBM Cognos, selecione o modo. O modo **Dados** permite navegar e selecionar um pacote do qual importar dados. O modo **Relatório de Lista** permite navegar e selecionar um relatório de lista predefinido do qual importar dados.

Um pacote contém um modelo IBM Cognos e todas as pastas, consultas, relatórios, visualizações, atalhos, URLs e definições de tarefa associados a esse modelo. Um modelo IBM Cognos define regras de negócios, descrições de dados, relacionamentos de dados, dimensões e hierarquias de negócios e outras tarefas administrativas.

1. Para selecionar um pacote, selecione o modo **Dados** e clique em **Navegar**. Uma caixa de diálogo é exibida listando todos os pacotes disponíveis dos quais importar conteúdo.
2. Ou, para selecionar um único relatório de lista, selecione o modo **Relatório de Lista** e clique em **Navegar**. Uma caixa de diálogo é exibida listando pacotes e pastas dos quais selecionar um relatório.
Nota: Somente relatórios de lista únicos podem ser importados; listas múltiplas não são suportadas.
3. Onde for necessário, expanda a lista para localizar o pacote ou relatório desejado. Selecione o item e clique em **OK**. Os detalhes do pacote ou relatório são mostrados no Editor de Origem de Dados.
4. Após selecionar um pacote, você escolhe os objetos para importar. Por exemplo, você seleciona os assuntos de consulta de objeto de pacote (que representam tabelas de banco de dados) ou itens de consulta de campo individual (que representam colunas da tabela) para usar a partir de um pacote.
Nota: Os dados devem estar no formato UTF-8.
5. Selecione a opção **Agregar dados antes de importar** se desejar importar dados agregados em vez de dados brutos. Geralmente os dados agregados têm muito menos registros do que os brutos.

Selecionar Objetos. Exibe os objetos (também conhecidos como namespaces) que podem ser selecionados do relatório de lista ou pacote. Expandir um ou mais objetos para exibir nos campos de entrada que podem ser importados. A seleção de um objeto seleciona automaticamente todos os seus campos de entrada.

Revisar Parâmetros. Se parâmetros tiverem sido definidos, ou para um objeto de dados ou relatório, é possível especificar valores para esses parâmetros antes de importar o objeto ou relatório. Esta tabela contém a lista de parâmetros necessários para filtros cálculos. Eles podem ter valores padrão. Um aviso será fornecido se um ou mais parâmetros não tiverem um padrão. Configure parâmetros conforme apropriado na coluna **Valor**. É possível usar parâmetros para fazer ajustes (por exemplo, para executar um cálculo parametrizado) antes da importação dos dados.

Revisar Filtros. Se um pacote tiver filtros definidos, é possível importar um ou mais. Se um filtro que você importar estiver associado aos dados importados, o filtro será aplicado antes de os dados serem importados. Os filtros podem ou não ser parametrizados.

Revisar Campos de Entrada. Lista os campos de entrada (também conhecidos como objetos de banco de dados) que serão importados no IBM Analytical Decision Management para processamento. Se você não requerer um campo específico, cancele a seleção da caixa de seleção próxima do nome do campo.

Ícones de objeto do Cognos: Os vários tipos de objetos que podem ser importados de um banco de dados Cognos BI são representados por diferentes ícones, conforme ilustrado na tabela a seguir.

Tabela 2. Ícones de objeto do Cognos.






Ícone	Objeto
	Pacote
	Namespace
	Assunto de consulta
	Item de consulta
	Dimensão de medida

Tabela 2. Ícones de objeto do Cognos (continuação).

Ícone	Objeto
	Medida
	Dimensão
	Nível de hierarquia
	Nível
	Filtro
	Relatório
	Cálculo independente

Origem do Analytic Server

Não é possível selecionar o tipo de origem de dados Analytic Server durante a criação de uma nova origem de dados. Ele tem propósitos somente de exibição quando você usa IBM SPSS Modeler Advantage para abrir um modelo que foi criado no IBM SPSS Modeler com relação a uma origem de dados do Analytic Server.

IBM SPSS Analytic Server foi introduzido no IBM SPSS Modeler 16, permitindo o uso do Hadoop Distributed File System (HDFS) como uma origem de dados. Os usuários do IBM SPSS Modeler podem criar fluxos usando o nó de origem HDFS Analytic Server e os modelos de construção baseados no sistema Big Data Hadoop. Os usuários do IBM Analytical Decision Management podem abrir esses fluxos de modelo no IBM SPSS Modeler Advantage (não é possível criar novos projetos usando origens de dados do Analytic Server). Observe que não é possível executar certas ações, como alterar o tipo de modelo ou resposta.

Durante a visualização de uma origem de dados do Analytic Server na guia Dados, as informações a seguir ficam disponíveis:

Nome da origem de dados. O nome dado à origem de dados no IBM SPSS Modeler. É possível alterar esse nome, se você desejar. O restante dos campos a seguir não é editável.

Origem do Analytic Server. Caminho para a origem de dados conforme configurado em IBM SPSS Modeler.

Especifique campos de entrada. Os campos de entrada usados.

Selecionando campos de entrada

Quando você tem uma grande quantidade de dados, você pode querer reduzir ou ajustar os dados importados de sua origem (por exemplo, para usar somente os campos que são relevantes para sua área de interesse).

1. Na parte inferior do diálogo Editor de Origem de Dados, clique em **Especificar Campos de Entrada**; o diálogo é expandido para exibir todos os campos de dados em sua origem. Além disso, todas as expressões criadas também são exibidas. Para obter detalhes sobre como criar expressões, consulte “Construtor de Expressões” na página 47.
2. Selecione cada um dos campos e expressões que deseja usar.

Além de selecionar dados, você pode, quando necessário, alterar o nível de medição para um ou mais campos de dados. Para obter mais informações, consulte “Níveis de medição”.

Visualizando dados

Para verificar se você selecionou a origem de dados correta, clique no ícone **Visualizar** para exibir uma amostra dos dados contidos na origem.

Visão geral de dados

A variável de visão geral está disponível na guia Dados de todos os aplicativos, na guia Modelagem e na guia Escore no IBM SPSS Modeler Advantage e na guia Implementar dos aplicativos que incluem a variável Escorar Agora.

1. Para executar uma visão geral de uma origem de dados, clique no ícone Visão geral de dados disponível no aplicativo.



Figura 11. Ícone Visão geral de dados

2. O diálogo Visão Geral de Dados aparecerá. Se desejar, selecione um campo de sobreposição no menu suspenso na seção Opções de Visão Geral.

Após a execução da visão geral, as guias ficarão disponíveis para exibir resultados somente para o principal campo selecionado ou para sobrepor o principal campo selecionado com o campo especificado neste menu suspenso. Por exemplo, nos resultados, talvez você queira visualizar informações sobre a **Idade** dos clientes e depois sobrepô-las com outro campo, como **Sexo**.

3. Se desejar, selecione **Usar Dados Particionados** na seção Opções de Visão Geral. Essa opção estará disponível se, em Configurações Opcionais na guia Modelagem, você tiver selecionado a opção **Particionar dados automaticamente para ativar a avaliação de modelo na construção da origem de dados para avaliação e teste**. Essa opção divide os dados em subconjuntos ou amostras separados para treinamento e teste de modelo. Ao construir o modelo em um subconjunto e testá-lo em outro, é possível se ter uma ideia de como ele será generalizado para outros conjuntos de dados.

4. Selecione os campos de dados para incluir na visão geral e clique em **Executar Visão Geral**. Todos os campos disponíveis na origem de dados serão listados. Por padrão, todos os campos de entrada de modelo e respostas são selecionados.

A visão geral de dados será executada e os resultados aparecerão. É possível ordenar as informações e escolher quais colunas exibir.

5. Clique em qualquer campo para ver seus detalhes. Uma nova guia de resultados será aberta para cada campo selecionado, permitindo q visualização de gráficos e tabelas e a seleção de campos de sobreposição, se disponíveis.

Níveis de medição

O nível de medição ajuda a caracterizar as informações representadas por cada campo de dados e determinar como um determinado campo é usado em regras, modelagem e outros aplicativos. O nível de medição pode ser especificado no Modelo de Dados do Projeto para uma origem de dados; por exemplo, você pode querer configurar o nível de medição para um campo de número inteiro com os valores de 0 e 1 para *Flag* para indicar que 0 = *False* e 1 = *True*. Alternativamente, é possível alterar o nível no diálogo **Editor de Origem de Dados** durante a especificação dos campos de entrada a serem usados. Para obter mais informações, consulte “Selecionando campos de entrada” na página 42.

Os níveis de medição a seguir estão disponíveis:

- **Contínuo** Usado para descrever valores numéricos, como um intervalo de 0–100 ou 0,75–1,25. Um valor contínuo pode ser um número inteiro, um número real ou data/hora.

- **Catagórico** Usado para valores de sequência de caracteres quando um número exato de valores distintos é desconhecido. Esse é um tipo de dados **não instanciado**, o que significa que todas as informações possíveis sobre armazenamento e uso dos dados ainda não são conhecidas. Após os dados serem lidos, o nível de medição será *Flag*, *Nominal* ou *Sem Tipo*, dependendo do número máximo de membros para campos nominais especificado na caixa de diálogo Propriedades do Projeto.
- **Flag** Usado para dados com dois valores distintos que indicam a presença ou ausência de um traço, como true e false, Yes e No ou 0 e 1. No IBM Analytical Decision Management, o primeiro valor da lista de categorias será usado como "false" e o segundo como "true", conforme indicado pela dica de ferramenta quando você passar o mouse sobre os valores quando eles forem exibidos na guia Dados. Os dados podem ser representados como texto, número inteiro, número real, data, hora ou registro de data e hora.
- **Nominal** Usado para descrever dados com diversos valores distintos, cada um tratado como um membro de um conjunto, como pequeno/médio/grande. Dados nominais podem ter qualquer armazenamento - numérico, sequência de caracteres ou data/hora. Observe que a configuração do nível de medição para *Nominal* não altera automaticamente os valores para o armazenamento de sequência de caracteres. Para obter informações sobre a configuração do máximo de membros permitido para campos nominais, consulte "Propriedades" na página 31.
- **Ordinal** Usado para descrever dados com diversos valores distintos que têm uma ordem inerente. Por exemplo, categorias de salário ou ranqueamentos de satisfação podem ser tipificados como dados ordinais. A ordem é definida pela ordenação natural dos elementos de dados. Por exemplo, 1, 3, 5 é o padrão de ordenação para um conjunto de números inteiros, enquanto HIGH, LOW, NORMAL (ordem alfabética ascendente) é a ordem para um conjunto de sequências de caracteres. O nível de medição ordinal lhe permite definir um conjunto de dados categóricos como dados ordinais para os propósitos de visualização, construção de modelo e exportação para outros aplicativos (como IBM SPSS Statistics) que reconhecem dados ordinais como um tipo distinto. É possível usar um campo ordinal em qualquer lugar em que um campo nominal possa ser usado. Além disso, campos de qualquer tipo de armazenamento (real, número inteiro, sequência de caracteres, data, hora, entre outros) podem ser definidos como ordinais.
- **Sem Tipo** Usado para dados que não estão em conformidade com nenhum dos tipos acima, para campos com um valor único ou para dados nominais nos quais o conjunto tem mais membros do que o máximo definido. Ele também é útil para casos em que o nível de medição seria algum conjunto como muitos membros (como um número de conta). Quando você seleciona **Sem Tipo** para um campo, o papel é automaticamente configurado para **Nenhum** com **ID de Registro** como a única alternativa. O tamanho máximo padrão para conjuntos é 250 valores exclusivos. Esse número pode ser ajustado ou desativado na caixa de diálogo Propriedades do Projeto, que pode ser acessada no ícone da barra de ferramentas.
- **Coleção** Usado para identificar dados geo-espaciais que estão registrados em uma lista. Uma coleção é efetivamente um campo de lista de espessura zero, em que os elementos nessa lista têm um dos outros níveis de medição.
- **Geo-espacial** Usado com o tipo de armazenamento Lista para identificar dados geo-espaciais. As listas podem ser uma Lista de número inteiro ou uma Lista de campos reais com uma espessura de lista que esteja entre zero e dois, inclusive.

Editando valores de campo de dados

Nas tabelas de dados, a coluna **Valores** lista os valores para cada campo na origem de dados (por exemplo, valores máximo e mínimo em um intervalo). Para campos de flag, mantenha o ponteiro do mouse pressionado sobre o valor para exibir uma dica de ferramenta indicando os valores "true" e "false" definidos para o campo. Também é possível clicar em um valor para editá-lo. Uma caixa de diálogo é aberta e seu conteúdo varia, dependendo do tipo de campo de dados selecionado. Por exemplo:

- Se o tipo de campo for **Contínuo**, é possível inserir valores limites máximo e mínimo.
- Se o tipo de campo for **Flag**, é possível trocar os valores true e false.
- Se o tipo de campo for **Nominal**, é possível incluir, remover, reordenar e renomear valores.

Mapeando campos

Durante a inclusão de uma nova origem de dados, talvez seja necessário mapear todos os seus campos para assegurar a compatibilidade com o modelo de dados do projeto. Por exemplo, se o modelo de dados do projeto requerer um campo denominado *compra* com valores *Sim* e *Não* (flag de nível de medição), qualquer origem de dados usada poderá ter um campo compatível. Se os nomes de campo não forem idênticos, eles poderão ser mapeados de acordo. Observe que a entrada e o campo mapeado associado devem ter o mesmo tipo de dados.

Alternativamente, se o modelo de dados do projeto mudar, talvez muitas origens de dados não sejam mais mapeadas corretamente para ele. Se for esse o caso, a bola do indicador de compatibilidade na área **Origens de Dados do Projeto** mudará para a cor apropriada para cada uma dessas origens de dados; depois você terá que remapear todas elas para o novo modelo de dados do projeto.

1. Na área **Origens de Dados do Projeto**, clique em **Mapear Campos**. Um diálogo é aberto e o sistema faz a melhor suposição de mapeamento de campos para você. Em qualquer ponto, é possível retornar para a melhor suposição do sistema clicando em **Reconfigurar**.
2. Selecione um campo na coluna **Entradas Disponíveis**, e seu campo correspondente na coluna **Campos de Entrada Obrigatórios**, e clique em **Mapear** para mapear os dois campos. Repita isso até que todos os campos na coluna **Entradas Disponíveis** tenham sido mapeados.
3. Ao concluir, clique em **OK**.

Quando possível, o IBM Analytical Decision Management automaticamente mapeará tipos de armazenamento compatíveis quando um usuário incluir uma origem de dados adicional. Se o tipo não for compatível, o usuário terá permissão de converter os tipos.

A tabela a seguir lista os mapeamentos permitidos.

Tabela 3. Tipos que podem ser convertidos

To →						
From ↓	Int	Real	String	Date	Time	timestamp
Int	n/d	Auto	Sim	Não permitido	Não permitido	Não permitido
real	Sim (dígitos decimais truncados)	n/d	Sim	Não permitido	Não permitido	Não permitido
Sequência de caracteres	Sim	Sim	n/d	Sim	Sim	Sim
Data	Não permitido	Não permitido	Sim	n/d	Não permitido	Sim (horário como 00h00)
Horário	Não permitido	Não permitido	Sim	Não permitido	n/d	Sim (data base)
Registro de data e hora	Não permitido	Não permitido	Sim	Auto (tempo truncado)	Auto (data truncada)	n/d

Derivando campos

Na maioria dos aplicativos IBM Analytical Decision Management, é possível incluir campos adicionais em seu modelo de dados do projeto, como expressões, saída do modelo, campos criados a partir de regras de segmento ou campos de tabelas secundárias. A guia **Derivada** da guia Dados exibe quaisquer campos que devam ser incluídos como parte dos dados e os campos podem ser ordenados para que possam ser usados cumulativamente (sua ordem na tabela corresponderá à ordem na qual são calculados; por exemplo, o segundo campo derivado pode usar o primeiro campo derivado como uma entrada).

A guia **Derivada** não está incluída no aplicativo IBM SPSS Rules Management.

- Na guia Dados, clique na guia **Derivada**.

Para incluir uma expressão:

- Clique no ícone **Derivar campo usando uma expressão**. O Construtor de Expressão será exibido. Para obter informações adicionais, consulte “Construtor de Expressões” na página 47.

Para incluir uma saída do modelo:

1. Clique no ícone **Derivar campo usando um modelo ou regra**, pesquise e selecione um modelo. O diálogo **Entradas e saídas** para que o modelo selecionado seja exibido.
2. Se os campos de entrada do modelo são compatíveis, a seção **Mapeamento de entrada do modelo** do diálogo será reduzida. Se os campos de entrada do modelo não forem compatíveis, a seção será expandida e você deve mapear entradas de tabela disponível para entradas do modelo selecionado.
3. Na seção **Saídas do modelo**, todas as saídas de modelo disponíveis para o modelo selecionado são listadas. Selecione os campos de saída do modelo que deseja incluir em seus dados. A coluna **Nome da saída do modelo** exibe o nome do campo de saída conforme gerado pelo modelo.

Todos os campos que o modelo usa como saída podem ser selecionados. Os campos de saída do modelo são todos os campos saindo do fluxo que também não são entradas para o fluxo (todos os campos incluídos pelo fluxo). Os campos de entrada não podem incluir tudo no modelo de dados do projeto, assim é possível, para um campo de saída do modelo, ter o mesmo nome que um campo de modelo de dados do projeto. Portanto, você também pode alterar os nomes dos campos de dados aqui, se desejado.

A coluna **Medição** exibe o tipo de medição de saída do modelo prefixada com o ícone de tipo apropriado, e a coluna **Valores** exibe os valores da saída do modelo. Observe que essas colunas podem exibir os resultados para alguns dos campos e para outros, não. Normalmente, os tipos de medição e os valores são conhecidos pelas saídas do modelo real (valor previsto, probabilidade, etc.) sem uma varredura de dados, portanto, se conhecido, serão exibidos. Mas os tipos de medição ou valores podem não ser conhecidos por outros campos de saída a partir de um fluxo (por exemplo, um campo derivado) sem uma varredura de dados.

4. Uma seção **Parâmetros do modelo** será incluída na parte inferior do diálogo se o fluxo selecionado incluir parâmetros de escoragem. Insira valores para os parâmetros do modelo.
5. Clique em **Salvar** quando tiver concluído o diálogo. Quaisquer saídas de modelo selecionadas serão incluídas na seção **Campos adicionais** da guia Derivada.

Para criar um campo usando regras de segmento:

1. Clique no ícone **Derivar campo usando regras de segmento**.
2. Crie uma nova regra de segmento. Os segmentos são definidos por uma ou mais expressões que retornam um resultado true ou false. Para obter informações adicionais, consulte “Definindo regras de segmento” na página 51.
3. Clique em **OK** quando tiver concluído o diálogo. A nova regra de segmento será incluída na seção **Campos adicionais** da guia Derivada.

Para incluir um campo a partir de uma tabela secundária:

1. Se você definiu uma tabela secundária com um relacionamento um-a-um na guia Dados, clique no ícone **Incluir campo a partir de uma tabela secundária**.
2. Selecione uma tabela secundária no menu suspenso e selecione um ou mais campos de tabela.
3. Clique em **Salvar** quando tiver concluído o diálogo. O(s) campo(s) será(ão) incluído(s) na seção **Campos adicionais** da guia Derivado.

Construtor de Expressões

É possível usar o construtor de expressões para derivar campos adicionais ou atributos para serem usados pelo aplicativo. Por exemplo, se estiver usando dados financeiros, talvez você queira criar uma expressão que mostre a razão entre a renda de um cliente e o número de contas para empréstimo que ele tem.

Expressões estão disponíveis em todo o aplicativo da mesma forma que os dados operacionais (por exemplo, em regras e em preditores de modelo).

Para exibir expressões existentes, ou criar uma nova expressão, acesse a guia Dados e clique na guia **Derivado**.

Criando uma expressão

1. No diálogo Construtor de Expressões, insira um nome exclusivo para a expressão.
2. Digite no campo **Equação de Expressão** usando as listas de campo e função como referências *ou*
3. Selecione as funções e os campos desejados das listas.
4. Dê um clique duplo ou clique na seta azul para incluir o campo ou função no campo **Equação de Expressão**.
5. Use os botões de operando no centro do diálogo para inserir operações na expressão.
6. Insira uma **Descrição** que explique para que a expressão é usada.
7. Clique em **Verificar** para validar a expressão e clique em **Salvar**.

Nota: É possível copiar e colar expressões do cliente IBM SPSS Modeler, mas observe que os resultados podem nem sempre ser como esperado se a expressão for dependente do fluxo (por exemplo, se você colar uma função @ suportada ou função de linha, isso pode não produzir os resultados desejados, porque o resto do fluxo pode não ser configurado para suportá-lo).

As únicas expressões que são suportadas oficialmente são aquelas fornecidas na interface com o usuário do IBM Analytical Decision Management.

Selecionando funções

A lista de funções exibe todas as funções e operadores disponíveis do . Role para selecionar uma função na lista ou, para facilitar a procura, use a lista suspensa para exibir um subconjunto de funções ou operadores. As funções disponíveis são agrupadas em categorias para facilitar a procura.

A seguir estão os detalhes sobre as categorias. Para obter informações mais detalhadas, consulte a documentação do IBM SPSS Modeler.

Tabela 4. Categorias de função

Tipo de Função	Descrição
Funções Usadas Recentemente	Contém uma lista de funções usadas durante a sessão atual.
Operadores	Lista todos os operadores que podem ser usados durante a construção de expressões. Os operadores também estão disponíveis a partir dos botões no centro do diálogo
Funções de Informações	Usado para obter insight dos valores de campo. Por exemplo, a função <code>is_string</code> retorna true para todos os registros cujo tipo é uma sequência de caracteres.
Funções de Conversão	Usado para a construção de novos campos ou para a conversão de tipo de armazenamento. Por exemplo, a função <code>to_timestamp</code> converte o campo selecionado em um registro de data e hora.

Tabela 4. Categorias de função (continuação)

Tipo de Função	Descrição
Funções de Comparação	Usado para comparar valores de campo uns com os outros ou com uma sequência de caracteres especificada. Por exemplo, <= é usado para comparar se os valores de dois campos são menores ou iguais.
Funções Lógicas	Usado para executar operações lógicas, como if, then, else.
Funções Numéricas	Usado para executar cálculos numéricos, como o logaritmo natural dos valores de campo.
Funções Aleatórias	Usado para selecionar itens ou gerar números aleatoriamente.
Funções de Sequência de Caracteres	Usado para executar várias operações em sequências de caracteres, como stripchar, que permite remover um caractere especificado.
Funções Data/Hora	Usado para executar várias operações em campos de data, hora e registro de data e hora.
Funções em Branco e Nulas	Usado para acessar, sinalizar e preencher frequentemente valores omissos do sistema ou em branco especificados pelo usuário. @BLANK não é suportado (porque seu comportamento é dependente de outros nós, como o nó Tipo). @NULL(FIELD) e undef são suportados. @NULL(FIELD) é tipo Booleano e retornará verdadeiro, se o valor de FIELD for \$null\$ ausente do sistema. Retorna falso para todos os outros valores, incluindo os em branco definidos pelo usuário. Se você deseja verificar ambos, use @BLANK(FIELD) e @NULL(FIELD). undef é tipo Any e é usado geralmente para inserir um valor \$null\$, por exemplo, para preencher valores em branco com nulos no nó de Preenchimento.
Todas as Funções	Contém uma lista completa de funções disponíveis.

Selecionando campos e valores

A lista de campos exibe todos os campos disponíveis. Role para selecionar um campo da lista. Dê um clique duplo ou clique na seta azul para incluir um campo na expressão.

Visualizando e selecionando valores

Valores de campo podem ser visualizados a partir de inúmeros locais no aplicativo. Para visualizar valores para um campo no Construtor de Expressões, selecione o campo desejado e clique no botão de entradas de expressão para abrir uma caixa de diálogo listando valores para o campo selecionado. Depois é possível selecionar um valor e clicar em **Inserir** para colar o valor na expressão atual.



Figura 12. Botão de entradas de expressão

Para campos de flag e nominal, todos os valores definidos são listados. Para campos contínuos (intervalo numérico), os valores mínimo e máximo são exibidos.

Verificando expressões

Clique em **Verificar** no Construtor de Expressão (canto inferior direito) para validar a expressão. As expressões que não foram verificadas são exibidas em vermelho. Se algum erro for localizado, uma mensagem indicando a causa será exibida.

Os itens a seguir são verificados:

- Aspas corretas nos valores e nomes de campos
- Uso correto de parâmetros e variáveis globais

- Uso válido de operadores
- Existência de campos referenciados
- Existência e a definição de globais referenciados

Se você encontrar erros na sintaxe, tente criar a expressão usando as listas e os botões de operador em vez de digitar a expressão manualmente. Esse método inclui automaticamente as cotações apropriadas para campos e valores.

Observe as limitações a seguir ao construir expressões no IBM Analytical Decision Management. As expressões não podem conter nenhum dos itens a seguir:

- Uma referência a um parâmetro de fluxo do IBM SPSS Modeler
- Uma referência a um fluxo global do IBM SPSS Modeler
- Uma referência a uma função de banco de dados
- Uma referência a uma das funções de campo especial ou valor de campo @ a seguir:
 - @TARGET
 - @PREDICTED
 - @FIELD
 - @PARTITION_FIELD
 - @TRAINING_PARTITION
 - @TESTING_PARTITION
 - @VALIDATION_PARTITION

Usando expressões que retornam um booleano

O IBM SPSS Modeler e o IBM Analytical Decision Management não possuem um tipo de dados booleano. As expressões booleanas não podem ser usadas para produzir um novo campo diretamente. O uso de expressões booleanas é limitado às construções de expressão que esperam uma expressão booleana (como a instrução `if`). Seguem três exemplos:

```
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')
```

Nos dois primeiros exemplos anteriores, as expressões "Escore de RFM" (`'RFM Score' > 5`, `'RFM Score' >= 5` e `'RFM Score' < 10`) podem ser entendidas como booleanas.

Essas expressões não podem ser usadas sozinhas para derivar um campo. Um erro, como `Invalid type for an expressions which is not a condition: Boolean` seria exibido. Mas expressões como essas podem ser usadas em expressões condicionais, como nos exemplos completos, e as expressões condicionais podem ser usadas para derivar campos com medições categóricas, como um campo do tipo `Flag`. `Flag` é a medição padrão para um campo com dois valores.

O segundo exemplo anterior poderia gerar um campo com três valores distintos (desde que os registros incluídos em cada uma das categorias Baixa, Média e Alta sejam localizados durante a varredura de dados). Provavelmente, isso resultaria no nível de medição do campo derivado padronizado como `Nominal`. É pouco provável que o terceiro exemplo acima retorne um campo com dois valores distintos.

Consulte “Níveis de medição” na página 43 para obter mais informações sobre os níveis de medição.

Capítulo 5. Trabalhando com regras

Regras são usadas para automatizar decisões baseadas em lógica de negócios ou na saída dos modelos preditivos. Dependendo do aplicativo, você pode usar regras para excluir clientes com base no histórico de pagamento, para encaminhar reclamações de risco para uma unidade de investigações especial ou para direcionar os clientes com maior propensão de migrar para o concorrente, comprar ou cometer fraude, conforme identificado por um modelo preditivo. Enquanto regras trazem automação, modelos fornecem poder preditivo, e a capacidade de poder combiná-los é uma das principais variáveis que distinguem aplicativos preditivos de outras ferramentas analíticas.

Regras podem ser criadas, aplicadas e salvas em diferentes pontos dentro de cada aplicativo. Alternativamente, é possível criar e editar regras compartilhadas usando IBM SPSS Rules Management.

Criando regras

- No ponto no aplicativo em que deseja incluir a regra, clique no ícone **Criar uma nova regra**.
- Alternativamente, clique no ícone **Localizar uma regra existente** para navegar pelas regras existentes, que podem ser incluídas ou reutilizadas no aplicativo atual. Consulte o tópico “Exportando e reutilizando regras” na página 54 para obter mais informações.
- Para criar uma regra compartilhada, ative IBM SPSS Rules Management na página Ativação de Aplicativo.

Definindo regras de segmento

Regras de segmento são usadas em todo o aplicativo para selecionar e alocar registros em diferentes aplicativos conforme apropriado. Segmentos são definidos por uma ou mais expressões que retornam um resultado verdadeiro ou falso. Por exemplo, você pode ter uma regra com um único segmento que retorna true se o valor do campo de idade for menor que 18. Segmentos podem ser definidos com base em campos ou expressões definidos nos dados ou nos escores de propensão ou predições retornados por um modelo. Em aplicativos com uma guia Definir configurados para ter várias hierarquias de dimensão (chamadas *elementos*), também é possível escolher entre outros membros de dimensão.

Cada segmento é definido por uma ou mais expressões e retorna um valor de true se todas as condições forem atendidas. Por exemplo, um segmento definido por idade e sexo incluiria somente registros que atendessem a ambas as condições e a regra retornaria *false* para todos os outros.

Para criar ou editar uma regra de segmento

1. No aplicativo no qual você deseja incluir a regra, clique no ícone **Criar uma nova regra**.
2. Especifique um nome para a regra.
3. Para a coluna de entrada, clique no ícone seta para baixo para especificar um campo, uma regra ou um modelo, conforme desejado.

Tabela 5. Opções de coluna de entrada.






Ícone	Uso
	Se ativado, permite selecionar um membro de dimensão de um elemento definido anteriormente.
	Permite escolher um campo do atual conjunto de dados.
	Permite escolher uma regra ou modelo existente do repositório.

Tabela 5. Opções de coluna de entrada (continuação).

Ícone	Uso
	Permite construir um novo modelo.
	Permite construir uma expressão.

4. Escolha a função que deseja usar. Consulte a tabela a seguir para conhecer as opções disponíveis.

Tabela 6. Operadores disponíveis.

Tipo de valor	Operadores disponíveis
Registro de data e hora, Data, Data e hora	<ul style="list-style-type: none"> • = (igual) • != (não igual a) • > (maior que) • < (menor que) • >= (maior que ou igual a) • <= (menor que ou igual a) • BETWEEN • ISNIL • IS NOT NIL
Um intervalo	<ul style="list-style-type: none"> • = (igual) • != (não igual a) • > (maior que) • < (menor que) • >= (maior que ou igual a) • <= (menor que ou igual a) • BETWEEN • ISNIL • IS NOT NIL • ONE OF • NOT ONE OF
Um conjunto	<ul style="list-style-type: none"> • = (igual) • != (não igual a) • ISNIL • IS NOT NIL • ONE OF • NOT ONE OF
Uma sinalização	<ul style="list-style-type: none"> • = (igual) • != (não igual a) • ISNIL • IS NOT NIL

5. Especifique um valor conforme apropriado para o tipo de entrada. Para valores não numéricos, uma lista suspensa de opções é fornecida.

Alternativamente, clique no ícone **Localizar uma regra existente** para navegar pelas regras que já foram definidas. Consulte o tópico “Exportando e reutilizando regras” na página 54 para obter mais informações.

Observe que uma regra de segmento como Age BETWEEN 50 and 20 não retornará resultados. Isso é o mesmo que a expressão $x \leq \text{valor e } \text{valor} < y$, que nunca pode ser true. Isso é consistente com o comportamento do SQL Server. Age BETWEEN 20 and 50 é o formato correto.

Observe também que uma regra de segmento, como Gender = \$null\$ não retornará resultados. Se você desejar filtrar os campos para os quais o valor Gender é vazio, poderá usar de outra forma uma expressão como Gender Not One Of (Male, Female).

Conjuntos de regra de segmento

- Regras de segmento também podem ser combinadas em conjuntos. Um conjunto de regras de segmento retorna um valor de true se *qualquer* um dos segmentos individuais for true. Observe que isso é diferente das expressões *dentro* de um segmento, que devem ser todas true.
- É possível exportar conjuntos de regras de segmento a partir de qualquer aplicativo que use regras ou em IBM SPSS Rules Management, se disponível. Consulte o tópico “Exportando e reutilizando regras” na página 54 para obter mais informações.

Definindo regras de seleção

Uma regra de seleção é um conjunto de uma ou mais regras de segmento ou conjuntos de regras com um atributo *Incluir* ou *Excluir* designado a cada um. Regras de seleção são usadas em todos os aplicativos para selecionar subconjuntos de registros para modelagem, escoragem ou outro processamento.

1. No aplicativo no qual você deseja incluir a regra, clique no ícone **Criar uma nova regra**. Este ícone está disponível em pontos em que as regras de seleção podem ser aplicadas.
2. Especifique um nome para a regra, selecione se deseja usar a regra para inclusão ou exclusão (ou faça isso depois) e inclua uma ou mais expressões conforme desejado. Consulte o tópico “Definindo regras de segmento” na página 51 para obter mais informações.
3. Clique em **OK** para salvar a regra.
4. Clique nos ícones na coluna **Incluir/Excluir** se desejar alternar esta configuração para cada regra. Para conjuntos de regras de seleção referenciada, esta configuração não pode ser alterada. Não é possível combinar inclusões e exclusões na mesma regra.

Alternativamente, clique no ícone **Localizar uma regra existente** para navegar pelas regras que já foram definidas. Consulte o tópico “Exportando e reutilizando regras” na página 54 para obter mais informações.

Avaliando regras de seleção

Clique no ícone **Visualizar lógica de regra de seleção** para ver detalhes sobre como as seleções são feitas (AND lógico versus OR).

- Durante a avaliação de regras de seleção, se vários segmentos de inclusão forem especificados, os registros correspondentes a *qualquer* um dos segmentos serão incluídos. Por exemplo, os clientes podem ser incluídos com base na classificação de crédito OR aforamento.
- Se vários segmentos de exclusão forem especificados, os registros correspondentes a *qualquer* um dos segmentos serão excluídos. Por exemplo, os clientes podem ser excluídos com base na idade OR classificação de crédito.
- Em caso de um conflito, as regras de exclusão têm precedência. Por exemplo, um cliente que está dentro de qualquer segmento excluído será excluído, independentemente de quaisquer segmentos para os quais eles se qualificaria.
- Durante a criação de conjuntos de regras de seleção no IBM SPSS Rules Management, todas as regras devem ser do mesmo tipo - *Inclusão* ou *Exclusão*. Não é possível combinar regras de inclusão e exclusão no mesmo conjunto de regras.

Exportando e reutilizando regras

Por padrão, regras criadas em um aplicativo são salvas com esse aplicativo. Alternativamente, regras podem ser exportadas ou salvas no Repositório do IBM SPSS Collaboration and Deployment Services para serem usadas em vários aplicativos. Nesse caso, cada aplicativo armazena uma referência a uma versão rotulada específica da regra compartilhada; todas as mudanças na regra são automaticamente aplicadas a todos os aplicativos que usam essa versão.

Exportando regras

A partir de qualquer aplicativo contendo uma regra local que você deseja compartilhar, clique no ícone **Exportar** para salvar a regra no Repositório do IBM SPSS Collaboration and Deployment Services em que ele possa ser reutilizado conforme a necessidade.

- Durante a exportação de diversas regras, especifique se deseja exportar todas elas em um único conjunto de regras ou em regras individuais. No caso de regras individuais, cada segmento será exportado como uma regra separada.
- Durante a exportação de diversas regras de seleção, se você escolher **Armazenar com Seleções**, um conjunto de regras de seleção será exportado. Nesse caso, é possível exportar inclusões ou exclusões, mas não ambas. Se você precisar de uma combinação de inclusões e exclusões, elas precisarão ser exportadas em conjuntos de regras separados. Se **Armazenar com Seleções** não estiver selecionada, um conjunto de regras de segmento será exportado.
- Durante a exportação de um editor de regras de agregação, se você escolher **Exportar com pontos de risco**, um conjunto de regras de agregação será criado. Se essa opção não for selecionada, um conjunto de regras de segmentos será exportado.

Usando regras existentes

- Em qualquer aplicativo em que deseja fazer uso de uma regra compartilhada, clique no ícone **Localizar uma regra existente** para navegar no Repositório do IBM SPSS Collaboration and Deployment Services pelas regras existentes que podem ser reutilizadas no aplicativo atual.
- Para fazer referência a uma versão rotulada específica, clique na seta próxima do nome da regra e escolha a versão desejada. Se nenhuma versão for selecionada, a mais recente será usada. (Pense em *LATEST* como um rótulo especial que é passado automaticamente para cada nova versão criada.) O uso de uma versão rotulada específica é recomendado para evitar resultados inesperados que podem ocorrer quando uma nova versão de uma regra é criada.

Regras de referência podem ser combinadas com regras locais em muitos casos, mas são visualmente distinguidas de regras locais por texto em negrito e por um ícone que indica o tipo de regra (ou conjunto de regras).

- Alternativamente, regras de referência podem ser convertidas em regras locais e salvas no aplicativo atual, se desejado. Para isso, clique no nome de qualquer regra de referência e selecione Converter em Regras Únicas na caixa de diálogo Visualização de Regra.

Diretrizes para compartilhamento de regras

Durante a navegação e o trabalho com regras compartilhadas, as diretrizes a seguir se aplicam:

- Regras podem ser exportadas para serem reutilizadas em qualquer aplicativo, mas após a exportação, elas podem ser editadas no IBM SPSS Rules Management e não podem ser editadas em nenhum outro aplicativo. Observe que o IBM SPSS Rules Management pode não estar instalado em todos os sites ou disponível para todos os usuários.
- O uso de uma versão rotulada específica (que não seja *LATEST*) é altamente recomendado, principalmente em um ambiente de produção, já que mudanças contínuas na versão mais recente podem ter consequências indesejadas. Em certos casos, como quando se faz referência a uma regra de dentro de uma regra, uma versão rotulada deve ser usada.

- Regras únicas podem ser incluídas a partir do Repositório do IBM SPSS Collaboration and Deployment Services por referência ou convertidas em regras locais na caixa de diálogo Visualização de Regra.
- Regras únicas podem ser referenciadas e combinadas com regras locais em qualquer editor de regras. Se forem exportadas como um conjunto de regras, as referências aninhadas para regras únicas serão preservadas.
- Conjuntos de regras de seleção e segmento podem ser referenciados e combinados como regras únicas na maioria dos casos, mas o conjunto de regras combinadas não pode ser exportado sem antes converter todos os conjuntos de regras aninhadas em regras únicas. Assim, um conjunto de regras exportado pode conter referências a regras aninhadas, *mas não a conjuntos de regras aninhados*.
- Conjuntos de regras de agregação também podem ser referenciados, mas não combinados com outras regras; eles também não podem conter referências aninhadas a outras regras. Durante uma referência a um conjunto de regras de agregação, a regra de referência substitui o conteúdo inteiro do editor de regra de agregação. (Alternativamente, se não estiver fazendo referência a um conjunto de regras de agregação, o editor de agregação poderá conter uma combinação de regras locais, regras de segmento de referência e conjuntos de regras de segmento.)
- Durante o uso do IBM SPSS Rules Management, as mesmas restrições se aplicam que para regras exportadas. Regras aninhadas são permitidas, conjuntos de regras aninhadas não. Isso porque o IBM SPSS Rules Management é projetado para criar e editar regras compartilhadas (semelhantes à exportação de quaisquer outros aplicativos).
- Para evitar quaisquer restrições em conjuntos de regras de referência, tanto no IBM SPSS Rules Management quanto em outro lugar, um conjunto de regras sempre pode ser convertido em regras locais separadas.
- Regras compartilhadas não podem conter referências a regras externas do IBM Operational Decision Management.

Incluindo anotações

Em uma agregação ou alocação, clique no ícone **Incluir Anotações** para adicionar uma coluna de campos de texto livre em que seja possível inserir uma nota que será incluída com a saída da regra. Por exemplo, pode ser uma informação extra dizendo para um operador da central de atendimento quem contatar quando uma reclamação do seguro for sinalizada para indicação.

As anotações podem ser usadas em combinação com o valor **Retornar com** especificado para uma dimensão para fornecer detalhes adicionais. Enquanto o valor **Retornar com** é associado à dimensão ou decisão que é retornada, a anotação diz quais regras foram disparadas como entrada para tal decisão.

Usando regras do IBM Operational Decision Management

Regras criadas em um Business Rule Management System, como IBM Operational Decision Management, são referenciadas e usadas em aplicativos IBM Analytical Decision Management. Essas regras externas podem ser selecionadas e usadas da mesma maneira que outras regras clicando no ícone **Localizar uma Regra Existente** no ponto desejado no aplicativo. Regras externas são distinguidas por ícone, mas são utilizáveis da maneira padrão.

Para obter informações sobre como criar regras externas para usar com IBM Analytical Decision Management, consulte *Application Designers Guide*.

A regra é inserida no aplicativo da mesma maneira que qualquer outra regra. Observe que as saídas de regra devem corresponder ao contexto em que a regra é usada; por exemplo, alocações requerem um valor true/false ou yes/no (booleano) como entrada na decisão.

Assim como com outras regras, clique no nome de qualquer regra externa usada no aplicativo para visualizar detalhes da regra, incluindo entrada e campos, junto com a descrição da regra e outras propriedades.

Se uma regra externa tiver vários campos de saída, será exibido um prompt permitindo escolher o campo desejado.

Se uma regra externa for usada no lugar de um modelo, nenhum campo de destino será mostrado e todas as saídas configuradas serão mostradas na lista suspensa.

Dicas para usar regras externas

- Regras que retornam um valor de *true* ou *false* (ou *yes/no* ou *0/1*) podem ser usadas em qualquer seleção ou alocação, como entrada para qualquer regra no editor de regras ou como um campo **Retornar com** associado a um item de dimensão específico.
- Regras que retornam saídas categóricas (*vermelho, branco* ou *azul; norte, sul, leste, oeste*, etc.) podem ser usadas na seção **Usar um Modelo para Decidir** na guia Definir, como entrada no editor de regras ou como um campo **Retornar com** associado a um item de dimensão.
- Regras que retornam saídas numéricas podem ser usadas na seção **Usar um Modelo para Decidir** na guia Definir, como entrada no editor de regras ou como um campo **Retornar com** associado a um item de dimensão.
- Origens de dados da visualização de dados analíticos suportam o uso de regras externas complexas (hierárquicas). Para referenciar uma regra externa complexa, o modelo de dados do projeto deve ser uma origem *Visualização de Dados* e a funcionalidade deve ser ativada (na guia Dados, um administrador deve selecionar a opção apropriada **Permitir o uso de regras ODM com base em dados hierárquicos**).
- Não é possível criar uma regra compartilhada no IBM Analytical Decision Management que contenha uma referência a uma regra externa.

Capítulo 6. Construindo modelos preditivos

Sobre modelos preditivos

Modelos podem ser usados para prever qual a probabilidade de algo acontecer no futuro com base em padrões de dados passados. Por exemplo, você pode usar um modelo para prever quais clientes têm a menor probabilidade de migrarem para o concorrente, ou a maior probabilidade de responderem a uma determinada oferta, com base em características como renda, idade e organizações e associações em que estão inscritos.

Modelos podem ser usados da mesma maneira que regras, mas enquanto as regras podem ser baseadas em políticas corporativas, lógica de negócios ou outras suposições, os modelos são baseados em observações reais de resultados passados e podem descobrir padrões que poderiam não estar aparentes de alguma outra forma. Enquanto as regras trazem lógica de negócios comum para os aplicativos, os modelos emprestam insight e poder preditivo.

Dados para a construção e a escoragem de modelos

Dois tipos diferentes de dados são usados no processo de modelagem:

- Para construir um modelo, você precisa de informações sobre aquilo que deseja prever. Por exemplo, se deseja prever uma migração para o concorrente, você precisará de informações sobre os clientes que migraram para o concorrente no passado. Geralmente isso é chamado de dado analítico ou histórico e deve conter alguns ou todos os campos no modelo de dados do projeto, mais um campo adicional registrando o resultado que você deseja prever. Esse campo extra é usado como a **resposta** para a modelagem.
- Para usar o modelo na predição de resultados futuros, você precisa de dados sobre o grupo ou a população de seu interesse, como possíveis clientes ou futuras reclamações, por exemplo. Geralmente isso é chamado de dado **operacional** ou dado de **escoragem**. O modelo de dados do projeto normalmente é baseado nesses dados.
- Se o campo de destino estiver incluído no modelo de dados, a coluna **Operacional** não deverá ser selecionada para esse campo, pois ela não estará disponível na escoragem do modelo. Tendo construído o modelo usando dados históricos, o objetivo na escoragem será aplicar o modelo aos novos dados nos quais o resultado ainda não é conhecido.

Construindo modelos

A guia Modelagem permite especificar uma origem de dados, um campo de destino e outras opções. Também é possível escolher qual tipo de modelo construir.

Observe que se os **parâmetros** foram incluídos no aplicativo (arquivo de fluxo) no IBM SPSS Modeler, um link **Parâmetros** estará disponível na guia Modelagem e guia Escore. Um clique no link **Parâmetros** mostra os parâmetros definidos e seus valores atuais e permite alterar os valores de parâmetro para afetar o escore ou a construção do modelo.

Construindo um modelo preditivo

Durante a construção de um modelo preditivo, as únicas especificações obrigatórias são um campo de origem e destino. Também é possível escolher opções para preparar e particionar dados e para selecionar os campos e registros para usar. O modelo é gerado aplicando uma variedade de métodos, conforme apropriado, no tipo de destino e identificando automaticamente a técnica específica ou uma combinação com o melhor desempenho nos dados selecionados.

Quando escorado, o modelo retorna um ou mais campos contendo predições, bem como valores de confiança associados a essas predições. Por exemplo, se o campo de destino for um campo de flag denominado *Resposta*, o modelo retornará um campo denominado *XF-Response* contendo os valores preditos para tal campo.

1. Especifique uma origem de dados. Ela pode ser uma origem de dados analítica que registre o resultado que você deseja prever.
2. Especifique um campo de destino que registre o resultado que você deseja prever. Por exemplo, o campo de destino pode indicar quais clientes foram perdidos, quais responderam às ofertas antigas, quais fizeram reclamações fraudulentas, entre outros.
Todos os outros campos serão automaticamente incluídos como entradas que podem ser úteis na predição desse valor.
3. Especifique configurações opcionais conforme desejado. Consulte o tópico “Configurações de modelo opcionais” na página 63 para obter mais informações.
4. Se desejar, clique no ícone Visão Geral dos Dados para ter uma visão geral dos dados que serão usados para a construção do modelo atual. Consulte o tópico “Visão geral de dados” na página 43 para obter mais informações.
5. Clique em **Modelo de Construção**.
É possível fechar o navegador ou trabalhar em outras tarefas durante a construção do modelo. O modelo pode ser acessado a partir da Galeria após a construção ser concluída. Consulte o tópico “Galeria” na página 26 para obter mais informações. Se o modelo estiver demorando mais do que o esperado na construção, clique no botão **Parar** para reverter para a versão salva anteriormente.
Quando a construção de modelo for concluída, os resultados serão exibidos. É possível escolher a visualização **Resultados de modelos combinados** (gráficos de desempenho para o modelo combinado que foi construído atualmente) ou **Resultados de modelos individuais** (detalhes de alto nível dos modelos individuais que compõem o modelo combinado).
Durante a visualização de *resultados do modelo individual*, o nome de cada modelo é um link que é aberto em um novo diálogo contendo informações sobre o modelo, como o campo de resposta e os campos de entrada. Se o modelo selecionado for um *modelo de árvore*, um visualizador de árvore também será exibido.
6. Opcionalmente, use as variáveis Avaliar e Testar para ver como o modelo se desempenha em seus dados de amostra. Consulte o tópico “Avaliando modelos” na página 67 para obter mais informações.

Resultados da modelagem preditiva

Quando um modelo preditivo é concluído, a seção de resultados permite se alternar entre a visualização **Resultados de Modelos Combinados** e **Resultados de Modelos Individuais**. Os resultados de modelo combinados incluem gráficos de importância de distribuição e variável para o modelo combinado que foi construído, fornecendo uma avaliação visual rápida do desempenho do modelo e da importância de diferentes preditores na obtenção do resultado final. Os resultados de modelos individuais listam informações sobre todos os modelos individuais que compõem o modelo combinado, permitindo clicar em cada modelo individual para visualizar seus detalhes.

Gráficos de distribuição

Os gráficos de distribuição de modelo exibem valores observados no eixo horizontal com valores preditos no eixo vertical ou sobreposição. Isso permite ver com que frequência cada valor ou resposta foi predito e com que frequência predição estava correto, fornecendo uma avaliação visual rápida da precisão do modelo.

- Para respostas flag ou configuradas, a distribuição é mostrada usando um gráfico de barras, com uma barra para cada categoria do campo de destino. A altura de cada barra indica quantas vezes tal resultado é predito pelo modelo. As regiões coloridas dentro de cada barra indicam quantas vezes o resultado correspondente foi de fato observado nos dados de teste. Se o modelo for razoavelmente exato, a maior área dentro de cada barra deverá corresponder à resposta correspondente.

- Para respostas do intervalo, a distribuição é representada por um gráfico de dispersão categorizado. Novamente, os valores preditos serão representados em gráficos no eixo horizontal e os valores observados serão representados em gráficos no eixo vertical. O ideal é que os pontos sejam agrupados ao longo de uma linha diagonal. Quanto mais dispersos os pontos (resultados) e mais distantes de uma distribuição linear, menos exato será o modelo.

Gráficos de importância do preditor

Normalmente você quer concentrar seus esforços de modelagem nos campos preditores mais importantes e considera descartar ou ignorar os que menos importam. O gráfico de importância do preditor lhe ajuda indicando a importância relativa de cada preditor na estimativa do modelo. Como os valores são relativos, a soma dos valores para todos os preditores na exibição é 1,0. A importância do preditor não tem relação com a precisão do modelo. Ela tem relação apenas com a importância de cada preditor na criação de uma predição, e não se a predição é ou não precisa.

O cálculo da importância do preditor pode levar muito mais tempo do que a construção de modelo, principalmente quando se usa conjuntos de dados maiores.

Se a opção para particionar dados automaticamente foi selecionada quando o modelo foi construído, a partição de teste será automaticamente selecionada na página Avaliar.

Se o particionamento não foi ativado, você precisará selecionar uma origem de dados para avaliar o modelo. A seleção de uma origem de dados além da usada para a construção do modelo é recomendada. Observe que o particionamento fica ativado por padrão, portanto, isso não seria necessário, a menos que você tenha mudado a configuração.

Resultados de modelos individuais

Resultados de modelos individuais mostram informações sobre cada modelo individual que compõe um modelo combinado. O nome de cada modelo é uma ligação que abre um novo diálogo contendo detalhes completos sobre esse modelo, como o campo de destino e os campos de entrada. Se o modelo selecionado for *modelo de árvore*, um visualizador de árvore especial também será exibido. O visualizador de árvore é interativo e fornece a funcionalidade a seguir:

- Aumentar zoom e diminuir zoom na árvore
- Mostrar os itens de árvore como tabelas, gráficos ou tabelas e gráficos combinados
- Alterar a orientação da árvore para de cima para baixo, da esquerda para a direita ou da direita para a esquerda
- Usar a *janela do mapa de árvore* e os ícones de seta para navegar por árvores grandes com mais facilidade

Construindo um modelo de regra preditivo

Modelos de regra preditivos permitem identificar segmentos que mostrem uma probabilidade maior ou menor com relação a um resultado específico. Por exemplo, você pode procurar clientes com menor probabilidade de migrarem para o concorrente ou com maior probabilidade de responderem sim para uma determinada oferta. O modelo resultante lembra uma lista de regras de seleção, com um atributo **Incluir** ou **Excluir** especificando como cada segmento é escorado.

Os modelos de regra preditivos lembram regras de seleção, mas diferem no modo como elas são escoradas. No caso de uma regra de seleção, registros excluídos são removidos, resultando em menos registros na saída. No caso de um modelo, nenhum registro é removido, mas campos adicionais são incluídos, indicando os valores de predição e propensão para cada registro. Os segmentos incluídos são escorados como *True* (ou *False*, ou qualquer resultado que você estiver tentando prever). Segmentos excluídos serão escorados como *NULL*, como quaisquer registros restantes, pois eles não correspondem a nenhuma regra.

Cada registro também recebe um escore de propensão indicando a probabilidade da resposta alvo, possibilitando identificar os clientes com mais chance de responder. Dessa forma, modelos de regra preditivos são mais adequados para gerar listas de distribuição, além de serem amplamente usados no gerenciamento de relacionamento com o cliente, incluindo central de atendimento ou aplicativos de marketing.

- A modelagem de regra preditiva está disponível somente para variáveis resposta categórica.
- É possível visualizar, modificar e reorganizar segmentos e escolher quais incluir ou excluir para os propósitos de escoragem. Por exemplo, é possível escolher excluir um grupo de clientes das ofertas futuras e incluir outros e ver imediatamente como isso afeta a taxa de ocorrência geral.
- Assim como acontece com regras de seleção, é possível incluir ou editar os segmentos manualmente conforme necessário. No entanto, geralmente é mais fácil deixar o modelo localizá-los para você, e inúmeras opções são fornecidas para lhe ajudar a ajustar o resultado.
- Regras são aplicadas na ordem listada, com a primeira regra de correspondência determinando o resultado para um determinado registro. Usadas independentemente, as regras ou condições podem se sobrepor, mas a ordem das regras resolve a ambiguidade. Se nenhuma regra corresponder, o registro será designado para as regras restantes.

Para obter um modelo de regra preditivo

1. Especifique uma origem de dados. Pode ser qualquer origem analítica que registre o resultado que você deseja prever.
2. Especifique um campo de destino que registre o resultado que você deseja prever.
3. Especifique a resposta ou valor específico que deseja localizar. Por exemplo, se o campo de destino fosse denominado *Resposta*, seria possível escolher *True* ou *False* para localizar os clientes que responderam--ou não -de acordo.
Por exemplo, a resposta pode indicar que os clientes migraram para o concorrente, responderam às ofertas antigas, fizeram reclamações fraudulentas, etc. Todos os outros campos serão automaticamente incluídos como entradas que podem ser úteis na previsão desse valor.
4. Especifique configurações opcionais conforme desejado. Consulte o tópico “Configurações de modelo opcionais” na página 63 para obter mais informações.
5. Se desejar, clique no ícone Visão Geral dos Dados para ter uma visão geral dos dados que serão usados para a construção do modelo atual. Consulte o tópico “Visão geral de dados” na página 43 para obter mais informações.
6. Clique em **Iniciar Construção**.
7. Especifique se deseja localizar segmentos com probabilidade alta ou baixa com relação à resposta especificada. Por exemplo, você poderia procurar grupos ou clientes com maior, ou menor, probabilidade de migrarem para o concorrente e incluí-los de acordo.
8. Especifique o número máximo de segmentos a ser localizado. Normalmente esse número é mantido pequeno em nome da simplicidade e do desempenho.
9. Clique em **Aumentar Modelo** para obter uma lista de uma ou mais regras de segmento que identifiquem grupos com maior ou menor probabilidade.
10. Clique nos ícones **Incluir/Excluir** para especificar como as regras devem ser usadas e como segmentos relevantes são escorados.
Segmentos incluídos recebem um escore correspondente ao valor da resposta (por exemplo *response=true*). Segmentos excluídos são escorados como *NULL*, mas não são removidos do arquivo.
11. Opcionalmente, é possível incluir ou editar segmentos manualmente, conforme desejado. Consulte o tópico “Definindo regras de segmento” na página 51 para obter mais informações.
12. Opcionalmente, use as variáveis Avaliar e Testar para ver como o modelo se desempenha em seus dados de amostra.

Dicas:

- Use as setas para cima/para baixo para mover as regras para cima ou para baixo. Cada registro é escorado com base na primeira ocorrência, portanto, a ordem das regras mudará o resultado. Por exemplo, um registro correspondente a várias regras será escorado com base na primeira que se aplicar.
- Altere uma ou mais opções e clique em **Aumentar Modelo** novamente para localizar regras adicionais. Por exemplo, você pode diminuir o limite para o intervalo de confiança ou segmento mínimo para identificar segmentos adicionais.
- Um clique em **Aumentar Modelo** não limpa a tabela ou reconstrói o modelo do zero, somente localizar mais regras de segmento para incluir na tabela.

Construindo um modelo de armazenamento em cluster

Os modelos de armazenamento em cluster permitem categorizar registros em um certo número de clusters. Isso pode ajudar a identificar grupos naturais em seus dados.

Os modelos de armazenamento em cluster estão focados na identificação de grupos de registros semelhantes e rotulagem de registros de acordo com o grupo ao qual pertencem. Isso é feito sem o benefício de conhecimento de informações a priori sobre os grupos e suas características. De fato, talvez você não saiba exatamente quantos grupos procurar. Isso é o que distingue os modelos de armazenamento em cluster de outras técnicas de aprendizado por máquina - não há saída predefinida ou campo de destino para o modelo prever. Esses modelos costumam ser chamados de modelos de **aprendizado sem supervisão**, já que não há nenhum padrão externo por meio do qual julgar o desempenho de classificação do modelo. Não há respostas *certas* ou *erradas* para esses modelos. Seu valor é determinado pela capacidade de captura de agrupamentos interessantes nos dados e do fornecimento de descrições úteis para esses agrupamentos.

Os métodos de armazenamento em cluster são baseados na medição de distâncias entre registros e entre clusters. Registros são designados a clusters de uma maneira que tende a minimizar a distância entre registros pertencentes ao mesmo cluster.

Os modelos de armazenamento em cluster geralmente são usados para criar clusters ou segmentos que depois são usados como entradas em análises subsequentes. Um exemplo comum disso são os segmentos de mercado usados pelos comerciantes para particionar seu mercado geral em subgrupos homogêneos. Cada segmento tem características especiais que afetam o sucesso dos esforços de marketing direcionados a ele. Se você estiver usando mineração de dados para otimizar sua estratégia de marketing, geralmente é possível melhorar seu modelo significativamente identificando os segmentos apropriados e usando as informações do segmento em seus modelos preditivos.

Para obter um modelo de armazenamento em cluster

1. Especifique uma origem de dados. Pode ser qualquer origem analítica que registre o resultado que você deseja prever.
2. Especifique configurações opcionais conforme desejado. Consulte o tópico “Configurações de modelo opcionais” na página 63 para obter mais informações.
3. Se desejar, clique no ícone Visão Geral dos Dados para ter uma visão geral dos dados que serão usados para a construção do modelo atual. Consulte o tópico “Visão geral de dados” na página 43 para obter mais informações.
4. Clique em **Localizar Clusters**.
5. Opcionalmente, é possível incluir clusters manuais. Consulte o tópico “Usando clusters manuais” para obter mais informações.
6. Opcionalmente, use as variáveis Avaliar e Testar para ver como o modelo se desempenha em seus dados de amostra.

Usando clusters manuais

Se desejar, você pode definir clusters manualmente, depois usar o armazenamento em cluster automatizado para localizar clusters de registros não identificados por quaisquer clusters definidos

manualmente. Crie clusters manuais via um conjunto de regras de alocação baseada em segmento "primeira ocorrência". As regras criadas definirão clusters e o restante será usado para modelagem de cluster automática.

Observe que os clusters manuais não estão incluídos nos resultados nos gráficos de sumarização.

1. Na seção **Cluster Manual**, clique no ícone **Criar uma nova regra** e inclua uma ou mais regras para definir clusters manuais.
2. Se desejar, altere os nomes de rótulos dos clusters.
3. Clique no ícone **Registros em clusters manuais** para executar uma contagem de registros do número total de registros incluídos nos clusters manuais. A coluna **Contagem** será atualizada para cada cluster manual definido.

Construindo um modelo de associação

Modelos de associação permitem prever quais itens têm maior probabilidade de aparecerem juntos e prever a intensidade do relacionamento entre eles.

Regras de associação associam uma determinada conclusão (a compra de um certo produto, por exemplo) a um conjunto de condições (a compra de vários outros produtos, por exemplo). Por exemplo, a regra

```
beer <= cannedveg & frozenmeal (173, 17.0%, 0.84)
```

diz que geralmente *beer* ocorre quando *cannedveg* e *frozenmeal* ocorrem juntos. A regra é 84% confiável e aplica-se a 17% dos dados ou 173 registros. Os algoritmos de regra de associação localizam automaticamente as associações que não puderam ser localizadas manualmente usando técnicas de visualização, como o nó da web .

A vantagem dos algoritmos da regra de associação sobre os algoritmos de árvore de decisão mais padrão (Árvores C5.0 e C&R Trees) é que as associações podem existir entre *quaisquer* atributos. Um algoritmo de árvore de decisão construirá regras com somente uma conclusão, enquanto algoritmos de associação tentam localizar várias regras, sendo que cada uma pode ter uma conclusão diferente.

A desvantagem dos algoritmos de associação é que eles estão tentando localizar padrões dentro de um espaço de procura potencialmente muito grande e, dessa forma, podem requerer muito mais tempo para execução do que um algoritmo de árvore de decisão. Os algoritmos usam um método **gerar e testar** para localizar regras--regras simples e geradas inicialmente, que são validadas com relação ao conjunto de dados. As boas regras são armazenadas, e todas as regras, sujeitas a várias restrições, são especializadas. **Especialização** é o processo de incluir condições em uma regra. Essas novas regras são validadas com relação aos dados e o processo armazena iterativamente as melhores e mais interessantes de todas as regras localizadas. Geralmente o usuário fornece algum limite para o possível número de antecedentes permitido em uma regra, e várias técnicas baseadas em teoria da informação e esquemas de indexação eficientes são usadas para reduzir espaço de procura potencialmente grande.

Para obter um modelo de associação

1. Selecione uma origem de dados. Pode ser qualquer origem analítica que registre o resultado que você deseja prever.
2. Especifique um formato de dados. Consulte o tópico "Dados tabulares versus transacionais" na página 63 para obter mais informações.
3. Especifique configurações opcionais conforme desejado. Consulte o tópico "Configurações de modelo opcionais" na página 63 para obter mais informações.
4. Se desejar, clique no ícone Visão Geral dos Dados para ter uma visão geral dos dados que serão usados para a construção do modelo atual. Consulte o tópico "Visão geral de dados" na página 43 para obter mais informações.
5. Clique em **Modelo de Construção**.

- Opcionalmente, use as variáveis de Teste para ver o desempenho do modelo em seus dados de amostra. Observe que a Avaliação não está disponível para o tipo de modelo de associação.

Com outros tipos de modelo, é possível escolher entre diferentes técnicas de modelagem. Para modelos de associação, o tipo de modelo *a priori* é usado.

Dados tabulares versus transacionais

Durante a construção de um modelo de associação, formatos de dados **Tabulares** ou **Transacionais** estão disponíveis. Dados tabulares são selecionados por padrão. Se você escolher **Transacional**, selecione campos para **ID** e **Conteúdo** nos menus suspensos que aparecem.

Formato transacional

Os dados transacionais têm um registro separado para cada transação ou item. Se um cliente fizesse várias compras, por exemplo, cada uma seria um registro separado, com itens associados vinculados por um ID de cliente. Às vezes isso é conhecido como formato **bobina de papel**.

Cliente	Compra
1	geleia
2	leite
3	geleia
3	pão
4	geleia
4	pão
4	leite

Dados tabulares

Dados tabulares (também conhecidos como dados de **cesta** ou **tabela de confiança**) têm os itens representados por flags, em que cada campo de flag representa a presença ou ausência de um item específico. Cada registro representa um conjunto completo de itens associados. Campos de flag podem ser categóricos ou numéricos, embora certos modelos possam ter requisitos mais específicos.

Cliente	Geleia	Pão	Leite
1	T	F	F
2	F	F	T
3	T	T	F
4	T	T	T

Configurações de modelo opcionais

Durante a construção de modelos, é possível escolher entre as seguintes configurações opcionais. Observe que dependendo do tipo de modelo e da resposta selecionada, talvez você não veja todas essas configurações. Se você alterar alguma dessas configurações para um modelo existente, o modelo deverá ser reconstruído para as mudanças se aplicarem.

Limpar e preparar automaticamente os dados para a construção de modelo confiável. Identifica e repara problemas de dados para tornar a modelagem mais rápida, mais previsível e mais confiável. Exibe os campos que são problemáticos ou que provavelmente não serão úteis, por exemplo, para manipular valores extremos ou omissos, derivar novos atributos quando apropriado e melhorar o desempenho por meio de técnicas inteligentes de triagem e amostra. A primeira vez que um modelo é construído com uma nova origem de dados, a análise é feita para identificar problemas e correções, o que pode deixar o

desempenho mais lento somente na primeira passagem. Em execuções subsequentes, as correções são aplicadas, mas a análise não é repetida, a menos que a origem de dados mude. Essa configuração pode ser desativada para alguns modelos, incluindo aqueles com configurações de preparação de dados customizados especificadas por um usuário especialista.

Nota: As configurações de preparação de dados automáticos só se aplicam ao criar um novo modelo e a categorização não é executada.

Particionar automaticamente os dados para permitir avaliação de modelo na origem de dados da construção para avaliação e teste. A seleção dessa opção divide os dados em amostras ou subconjuntos separados ou para treinamento e teste do modelo. Ao construir o modelo em um subconjunto e testá-lo em outro, você pode ter uma ideia de como o modelo será generalizado para outros conjuntos de dados. É possível também especificar um valor de porcentagem para o *tamanho da partição de treinamento* e *tamanho da partição de teste* gerados aleatoriamente.

O tamanho mínimo da partição de treinamento é 1, e o máximo de 100. O tamanho mínimo da partição de teste é 0, e o máximo é 100. O tamanho mínimo da partição de validação é 0, e o máximo é 100. A porcentagem total deve ser 100 ou menos para o modelo ser construído com êxito.

Essas opções de particionamento estão disponíveis para todos os tipos de modelo, exceto a modelagem de associação.

Escolher técnicas de modelo para construção de modelo. Se desejar, clique em **Selecionar** para escolher quais técnicas de modelagem serão incluídas na construção do modelo. Com base na resposta selecionada, todas as técnicas de modelagem disponíveis serão listadas. As técnicas de modelagem disponíveis (e aquelas selecionadas por padrão) variam dependendo do destino selecionado. Por exemplo, o algoritmo da *Lista de decisão* aparecerá apenas para destinos *flag*. Esse recurso está disponível apenas para modelos preditivos.

Para obter detalhes completos sobre as técnicas de modelagem, consulte o *Guia de algoritmos do IBM SPSS Modeler* e outras documentações enviadas com o produto IBM SPSS Modeler disponíveis na Web.

Número máximo de modelos a serem combinados. Permite que você configure o número máximo de modelos a serem retidos e combinados. Se você configurar essa opção como 1, um único modelo será construído e retido. Se houver menos modelos construídos do que o valor inserido aqui, então, todos os modelos construídos serão retidos e combinados. Essa opção está disponível apenas para modelos preditivos.

Crítérios de lucro (usado para ranqueamento de modelos com respostas binárias). Permite a você definir critérios de lucro antes de construir o modelo. Os valores definidos aqui serão definidos como o padrão para avaliar o modelo. Essa opção está disponível apenas para modelos preditivos com uma resposta *flag* selecionada.

Especificar entradas a serem usadas. Permite que você selecione os campos que deseja usar. Tipicamente essas seriam aquelas que têm algum relacionamento prático com o fato que você está tentando prever, como idade ou renda. Se você tiver um conjunto de dados grande, limitar o número de campos é uma forma de simplificar o modelo. Se os dados contiverem campos, tais como, ID do cliente ou número de contato, eles normalmente não seriam úteis na modelagem e não devem ser selecionados. Campos que duplicam outros dados também podem ser excluídos.

Clicar em um campo de entrada vinculado (uma expressão) abre a o visualizador de expressão para aquela expressão. Para editar uma expressão, consulte a guia *Dados*. Consulte o tópico “Construtor de Expressões” na página 47 para obter mais informações.

Especificar seleções a serem usadas. Especifica quais registros serão incluídos ou excluídos durante a modelagem. Você pode procurar regras existentes ou crie novas conforme apropriado. Consulte o tópico “Definindo regras de seleção” na página 53 para obter mais informações.

Além disso, se as seleções globais foram definidas, elas são exibidas aqui, e é possível especifica se elas devem ser aplicadas durante a modelagem. É possível optar por usar todas ou nenhuma das regras de seleção, não é possível escolher um subconjunto.

Opções de segmentação

Durante a construção de modelos de regra preditivos, inúmeras opções de segmentação ficam disponíveis. Elas estão disponíveis além daquelas com outros tipos de modelos.

Tamanho mínimo do segmento. Estas configurações ditam o tamanho mínimo do segmento, sendo que os dois valores maiores têm precedência. Por exemplo, se o valor da porcentagem for igual a um número maior que o valor absoluto, a configuração da porcentagem terá precedência.

- **Porcentagem do segmento anterior.** Especifica o tamanho mínimo do grupo como uma porcentagem de registros. A configuração mínima permitida é 0; a configuração máxima permitida é 99,9.
- **Valor absoluto.** Especifica o tamanho mínimo grupo como um número absoluto de registros. A configuração mínima permitida é 1; não há configuração máxima.

Número máximo de atributos. Especifica o número máximo de condições por regra de segmento. A configuração mínima permitida é 1; não há configuração máxima.

Permitir reutilização do atributo. Quando ativado, cada ciclo pode considerar todos os atributos, mesmo aqueles que foram usados em ciclos anteriores. As condições para um segmento são baseadas em ciclos, em que cada ciclo inclui uma nova condição. O número de ciclos é definido usando a configuração **Número máximo de atributos**.

Intervalo de confiança para novas condições. Especifica o nível de confiança para testar a significância do segmento. Essa configuração desempenha um papel significativo no número de segmentos retornados (se houver), bem como o número de condições por regra de segmento. Quanto maior o valor, menor o conjunto de resultados retornado. A configuração mínima permitida é 50; a configuração máximo permitida é 99,9.

Opções de cluster automático

Durante a construção de modelos de armazenamento em cluster, inúmeras opções de cluster automático ficam disponíveis. Elas estão disponíveis além daquelas com outros tipos de modelos.

Campo de avaliação. Especifica o campo a ser usado para avaliação. Se um campo de avaliação for selecionado, o método de ranqueamento no nó de construção do cluster automático do IBM SPSS Modeler será configurado para *Importância de avaliação*. Se nenhum campo de avaliação for selecionado, o método de ranqueamento será deixado como o padrão. Observe que um campo analítico pode ser selecionado como campo de avaliação aqui.

Configurar um intervalo desejado para um número de clusters localizado. Especifica o intervalo mínimo e máximo desejado para o número de clusters localizado pelo modelo de cluster automático.

Opções de escoragem e regra de associação

Durante a construção de modelos de associação, inúmeras opções de escoragem e regra de associação ficam disponíveis. Além delas, há outras opções disponíveis com outros tipos de modelos.

Opções de Regra de Associação.

- **Cobertura (%)**. Especifique a porcentagem de cobertura mínima para controlar até que ponto as regras serão aplicáveis de um modo geral. Por exemplo, a configuração de um valor de 80% significa que todas as regras se aplicarão a pelo menos 80% dos casos. Observe que valores mais altos gerarão menos regras ou, possivelmente, nenhuma.
- **Confiança (%)**. Especifique a porcentagem mínima de confiança.
- **Número máximo de condições prévias**. Especifique o número máximo de condições prévias.
- **Os dados já são ordenados por ID (somente formato transacional)**. Selecione esta opção para desativar a ordem que ocorre por padrão no início da construção do modelo para assegurar que os IDs sejam contíguos. A seleção dessa opção reduzirá o tempo de processamento. Essa opção não terá nenhum impacto se o formato dos dados for tabular.

Opções de Escoragem.

- **Número máximo de predições**. Especifique o número máximo de predições incluídas. Essa configuração é usada junto com o menu suspenso de critério de regra para produzir as "principais" predições, em que "principal" indica o nível mais alto de confiança, suporte, elevação, entre outros.
- **Critério de regra**. Selecione a medida usada para determinar a intensidade das regras. Regras são ordenadas pela intensidade de critérios selecionados aqui para retornar as principais predições para um conjunto de itens. Os critérios disponíveis são:
 - **Confiança** exibe a razão do suporte de regra para o suporte de condição da regra. Isso indica a proporção de registros com a(s) condição(ões) da regra especificada(s) para a(s) qual(is) a(s) predição(ões) de regra é(são) verdadeira(s). Por exemplo, se 50% dos dados de treinamento contiverem pão (indicando suporte de condição da regra), mas somente 20% contiverem pão e queijo (indicando suporte de regra), a predição para a regra pão -> queijo seria Suporte de Regra / Suporte de Condição da Regra ou, nesse caso, 40%.
 - **Cobertura (%)** exibe o suporte de condição da regra -- ou seja, a proporção de IDs para a qual condições de regra são verdadeiras com base nos dados de treinamento. Por exemplo, se 50% dos dados de treinamento incluir a compra de pão, a regra pão -> queijo terá um suporte de condição da regra de 50%.
 - **Suporte de Regra (%)** exibe a proporção de registros para a qual a regra inteira, condições da regra e predições da regra são verdadeiras. Por exemplo, se 20% dos dados de treinamento contiverem pão e queijo, o suporte de regra para a regra pão -> queijo será 20%.
 - **Elevação** exibe a razão de confiança para a regra para a probabilidade anterior de se ter uma predição de regra. Por exemplo, se 10% da população inteira comprar pão, uma regra que prevê se as pessoas comprarão pão com 20% de confiança terá uma elevação de $20/10 = 2$. Se outra regra lhe disser que as pessoas comprarão pão com 11% de confiança, a regra terá uma elevação próxima de 1, o que significa que ter a(s) condição(ões) da regra não faz muita diferença na probabilidade de se ter a predição da regra. Em geral, regras com uma elevação diferente de 1 serão mais interessantes do que as regras com elevação próxima de 1.
 - **Implementabilidade (%)** é uma medida de qual porcentagem dos dados de treinamento satisfaz condição(ões) de regra, mas não a predição de regra. Em termos de compra de produtos, basicamente isso significa qual porcentagem do total a base de clientes possui (ou comprou) a(s) condição(ões), mas ainda não comprou a(s) predição(ões) de regra. A estatística de implementabilidade é definida como $((\text{Rule Condition Support in \# of Records} - \text{Rule Support in \# of Records}) / \text{Number of Records}) * 100$, em que Rule Condition Support significa o número de registros para os quais as condições da regra são verdadeiras e Rule Support significa o número de registros para os quais ambas as condições da regra e a predição são verdadeiras.
- **Verificar transação inteira**. Selecione entre as opções a seguir:
 - **Assegurar que predições não estejam presentes** assegura que predições de regra também não estejam presentes. Por exemplo, se o propósito da escoragem for fazer uma recomendação de produto para mobília de casa, é improvável que uma cesta que já contém uma mesa de jantar compre outra. Nesse caso, você deve selecionar essa opção. Por outro lado, se os produtos forem

percebíveis ou descartáveis (como queijo, leite em pó ou lenço de papel), as regras nas quais o subsequente já está presente na cesta podem ter algum valor. No segundo caso, a opção mais útil pode ser **Não verificar predições**.

- **Assegurar que predições estejam presentes** assegura que as predições de regra também estejam presentes na cesta. Essa abordagem é útil quando se está tentando ganhar insight em clientes ou transações existentes. Por exemplo, talvez você queira identificar regras com elevação mais alta e depois explorar quais clientes se ajustam a essas regras.
- **Não verificar predições** inclui todas as regras durante a escoragem, independentemente da presença ou ausência de predições de regra na cesta.
- **Permitir predições de repetição.** Selecione esta opção para incluir diversas regras com a mesma predição de regra durante a escoragem. Por exemplo, a seleção desta opção permitiria que as seguintes regras fossem escoradas:
pão & queijo -> vinho
queijo & fruta -> vinho
Desative essa opção para excluir predições de repetição durante a escoragem.
- **Ignorar itens não correspondentes.** Selecione esta opção para ignorar a presença de itens adicionais no conjunto de itens. Por exemplo, quando essa opção é selecionada para uma cesta contendo [barraca & saco de dormir & chaleira], a regra barraca & saco de dormir -> forno a gás se aplica, apesar do item extra (chaleira) presente na cesta.

Avaliando modelos

Após a construção ou abertura de um modelo concluído, clique em **Avaliar** para obter gráficos como distribuição, distribuição de erro, matriz de coincidência, ganhos, elevações, resposta, lucro e ROI e para simular o lucro máximo possível. Os tipos de gráfico e outras opções disponíveis variam dependendo do tipo de modelo selecionado. O diálogo Avaliar só estará disponível se o modelo selecionado estiver completo. Observe que a avaliação não estará disponível para o tipo de modelo Associação.

1. No painel superior, selecione um conjunto de dados de avaliação. Assim como com a construção de modelo, esse deve ser um conjunto de dados analítico para o qual as respostas reais são conhecidas. (Caso contrário, não haverá como comparar as respostas reais com aquelas previstas pelo modelo.)
2. Se desejar modificar os modelos para incluir na avaliação, clique em **Modificar Seleção de Modelo**. Após fazer seleções, também é possível clicar no botão **Copiar seleções de modelo atuais de volta na tabela de resultados do modelo individual**, se desejado. Essas opções estarão disponíveis somente para o tipo de modelo Preditivo.



Figura 13. Copiar seleções de modelo atuais de volta na tabela de resultados do modelo individual

3. Selecione os gráficos que deseja construir.
4. Se estiver simulando um lucro, selecione a caixa **Simular lucro máximo** e insira a população na qual basear o cálculo do lucro, o orçamento e o valor mínimo. Consulte o tópico “Simulando lucro máximo” na página 69 para obter mais informações.
5. Se estiver solicitando um gráfico ROI ou de lucro, ou se estiver simulando um lucro, especifique os valores de custo e renda. Esses valores indicam o custo para cada oferta e a quantidade de renda esperada de cada resposta favorável, respectivamente.
6. Selecione o valor de valor considerado uma "ocorrência". Esse deve ser um dos valores para o campo de destino usado para a construção do modelo, como *response = true* para indicar uma resposta positiva ou favorável.
7. Se estiver solicitando um gráfico de desempenho, também é possível selecionar a opção **Incluir avaliações de modelo individual para gráficos de desempenho**.

8. Se desejar, selecione **Incluir resultados tabulares**. Isso requer mais tempo para retornados os resultados da avaliação.
9. Clique em **Executar** para visualizar os resultados.

Gráficos de distribuição (todas as respostas)

Gráficos de distribuição mostram valores observados versus esperados e são exibidos por padrão que o modelo é concluído. A caixa de diálogo Avaliar permite obter gráficos de distribuição para um conjunto de dados além do de teste. Para respostas de intervalo, o gráfico de distribuição é o único tipo de gráfico disponível.

É possível escolher entre os gráficos adicionais a seguir:

Gráficos de distribuição de erro

Os gráficos de distribuição de erro exibem um histograma dos erros nos resultados do modelo, em que erro é a diferença entre o valor predito e o valor real. Um erro positivo suscita quando o valor predito é maior que o valor real e um erro negativo suscita quando ele é menor. Esse gráfico está disponível somente para variáveis resposta contínua com o tipo de modelo Preditivo.

Gráficos de matriz de coincidência

Uma matriz de coincidência cria um gráfico das contagens de resultados *preditos* versus *reais*. Para uma resposta numérica, o intervalo é dividido pelo número de categorias especificado.

Gráfico de ganhos

Ganhos são usados para mostrar o ganho ou "elevação" que podem ser obtidos usando o modelo e são definidos como a proporção de acertos em cada incremento relativo ao número total de ocorrências disponíveis. A linha diagonal simples cria um gráfico da resposta esperada para a amostra inteira, caso nenhum modelo tenha sido usado. Nesse caso, a taxa de resposta seria constante, já a probabilidade de uma pessoa responder é igual à de outra. Para dobrar seu lucro, você precisaria perguntar para o dobro de pessoas. A linha curva indica quanto você pode melhorar esse resultado almejando aqueles com maior probabilidade de responder. Quanto mais íngreme a curva, maior o ganho.

Gráficos de ganho

O gráfico de ganho mostra a porcentagem de registros em cada incremento, que são ocorrências com a porcentagem geral de ocorrências no conjunto de dados de treinamento.

Gráficos de resposta

O gráfico de resposta mostra a porcentagem de registros no incremento que são ocorrências.

Gráficos de lucro

Lucro é igual à renda para cada registro, menos o custo para o registro. Os lucros para um quantil são simplesmente a soma de lucros de todos os registros no quantil. Supõe-se que as rendas se apliquem somente a ocorrências, mas os custos se aplicam a todos os registros. Consulte o tópico "Simulando lucro máximo" na página 69 para obter mais informações.

gráficos ROI

ROI (retorno sobre investimento) é semelhante ao lucro, já que envolve a definição de rendas e custos. O ROI compara lucros e custos para cada intervalo.

Simulando lucro máximo

Com base nas taxas de resposta previstas pelo modelo, é possível simular o lucro máximo a ser obtido por uma determinada população. Para isso, você precisa especificar valores de renda e custo, bem como o tamanho da população.

Para prever o lucro máximo:

1. Especifique a população, ou o número total de registros, em que a predição deve ser baseada. *Nota:* Na escoragem de um modelo, isso é configurado automaticamente para o número total de registros na origem de dados ou subconjunto.
2. Especifique um valor para orçamento se desejar que o parâmetro de orçamento seja incluído na simulação de lucro e um valor mínimo.
3. Especifique o custo associado a cada registro. Por exemplo, poderia ser o custo de impressão e correspondência, por pessoa, da literatura promocional para uma campanha.
4. Especifique a renda esperada que uma resposta positiva geraria. Por exemplo, o custo de varejo de um item em uma campanha.
5. Selecione a resposta em que está interessado.
6. Clique em **Executar**.

O gráfico de resultados previstos mostra o melhor resultado para lucro máximo e identifica a porcentagem de registros que você precisaria almejar para consegui-lo, com base na propensão de resposta, conforme identificado pelo modelo.

Abaixo do gráfico, é possível mover uma régua de controle de zero a 100% dos registros alvo para ver como isso afetaria seus lucros.

Quando os números forem aceitáveis, será possível escolher usar as configurações para escoragem de seus registros. Para isso, clique em **Usar para Escoragem**. A porcentagem de registros alvo é exibida no campo % **Principais** na área Opções de Propensão da guia Escorar.

Comparação de lucro

Para verificar se os registros escolhidos são uma boa seleção, é possível comparar a probabilidade de lucro a ser obtida a partir desses registros, um grupo aleatório da mesma porcentagem de registros e o conjunto inteiro de registros.

Os detalhes mostrados incluem o número de registros alvo, a probabilidade de taxa de resposta, custos, renda e lucro geral. É possível exportar essas informações em formato HTML para usar em outros aplicativos.

Gráfico de lucro

É possível exibir o lucro predito acumulativo e não acumulativo a partir dos registros selecionados como um gráfico.

Nota: Este gráfico é gerado quando você executa pela primeira vez a avaliação, e ele não é atualizado automaticamente com nenhuma mudança feita por meio da régua de controle na seção de simulação de lucro.

Testando o modelo

A caixa de diálogo Testar permite ver como um ou mais registros de amostra serão escorados. Resultados são específicos para o contexto em que o teste é executado e podem refletir os escores retornados por um modelo ou regra. A seguir estão alguns exemplos dos tipos de resultados que podem ser vistos, dependendo do seu tipo de aplicativo:

- Você pode consultar vários clientes específicos e ver quais ofertas são alocadas para cada um deles.
 - Você pode consultar vários registros de clientes e ver quais clientes têm maior probabilidade de responder a uma campanha de correio direto.
 - Você pode consultar várias reclamações de seguro específicas e ver quais ações são alocadas para cada uma delas.
 - Você pode consultar vários casos específicos e ver como cada regra é avaliada.
1. Na guia Modelagem, clique em **Testar**. O diálogo Testar é exibido.
 2. Selecione a origem de dados que deseja usar ou selecione **Especificar dados customizados** na seção Registros para inserir registros de teste manualmente.
 3. Especifique seleções para usar, se houver. Por exemplo, talvez você queira incluir ou excluir certos registros do teste. É possível procurar regras existentes ou criar novas regras conforme apropriado. Consulte o tópico “Definindo regras de seleção” na página 53 para obter mais informações.
 4. Se estiver usando uma origem de dados, especifique o número de registros que deseja retornar. É possível retornar os N primeiros registros, até um máximo de 100, lidos na parte superior dos dados ou em uma amostra aleatória de registros. Se você testar um modelo de associação usando dados transacionais, dados de entrada serão amostrados no nível das transações, e não no nível do registro.
 5. Se estiver usando dados customizados, selecione **Especificar dados customizados** e clique em **Incluir registro customizado**. Todos os campos definidos no modelo de dados do projeto são listados, permitindo especificar um valor para cada. Clique em **Salvar** ao concluir.
 6. Em **Selecionar Campos para Exibir**, selecione os campos que você quer que apareçam na saída de teste.
 7. Clique em **Executar**. A sumarização do teste exibe uma lista de alocações bem-sucedidas para cada registro.
 8. Clique no ícone de informações no início de cada linha para obter detalhes da linha.

Usando escores de modelo nos aplicativos

Escorar um modelo significa aplicá-lo a alguns dados ou população para obter um resultado ou uma predição que possam ser usados como entrada para decisões. Por exemplo, você pode escorar todos os clientes em seu banco de dados para identificar clientes em potencial para uma campanha de correio direto ou escorar um cliente por vez conforme eles contatam sua central de atendimento para decidir quais produtos ou serviços oferecer para eles. Dependendo do aplicativo, os resultados da escoragem podem ser gravados em um arquivo simples ou tabela de base de dados ou usados como entradas para as regras de segmento, seleção e alocação que direcionam as decisões em um aplicativo.

Os escores de modelo podem ser usados das seguintes formas:

- Como entradas para regras. Consulte o tópico “Criando regras” na página 51 para obter mais informações.
- Para aplicativos que suportam escoragem de lote, escores de modelo podem ser gravados em um arquivo simples ou tabela de base de dados. Consulte o tópico “Escoragem de modelos em uma tabela de base de dados, arquivo, Analytic Server ou servidor IBM Cognos BI” na página 73 para obter mais informações.

Campos de saída do modelo

Por convenção, os nomes dos campos de saída de escoragem são baseados no nome do campo de destino, mas com a inclusão de um prefixo customizado. Por exemplo, se o campo de destino for denominado *Resposta*, os campos de saída terão nomes como *\$XF-Response* e *\$XFRP-Response*. Os campos específicos dependem do tipo de modelo e do nível de medição para o campo de destino selecionado, conforme detalhado abaixo. Consulte o tópico “Níveis de medição” na página 43 para obter mais informações.

Respostas flag. Modelos automatizados com respostas flag geram os seguintes campos:

- **\$XF-<target>**. Valor predito para cada registro, relatado em termos de valores "true" e "false" definidos para o campo.
- **\$XFRP-<target>**. Escore de propensão; indica a probabilidade do valor "true" para cada registro. Os escores de propensão são diferentes dos escores de confiança, que normalmente são relatados em termos de predição para tal registro. Por exemplo, uma predição "false" com confiança alta seria convertida em um baixa propensão, já que isso indicaria uma alta probabilidade de *não* responder.

Variáveis resposta categórica. Modelos automatizados com variáveis resposta categórica (nominal ou ordinal) geram os seguintes campos:

- **\$XS-<target>**. Valor predito para cada registro.
- **\$XSC-<target>**. Valor de confiança associado à predição.

Variáveis resposta contínua. Modelos automatizados com respostas numéricas contínuas geram os seguintes campos:

- **\$XR-<target>**. Valor predito para cada registro.
- **\$XRE-<target>**. Erro padrão para a predição.

Modelos de regra preditivos (interativos). Modelos de regra preditivos (somente variáveis resposta categórica ou flag) geram os seguintes campos:

- **\$D-<target>**. Resultado predito para cada registro, sendo que o valor "true" é definido para o campo usado para indicar a resposta desejada e null indica qualquer outro valor.
- **\$DP-<target>**. Probabilidade do resultado desejado para cada registro; corresponde à probabilidade exibida para o segmento na tabela de resultados durante a construção do modelo. Consulte o tópico “Construindo um modelo de regra preditivo” na página 59 para obter mais informações.
- **\$DI-<target>**. Número do índice usado para identificar o segmento no qual o registro se encaixa (exibido na primeira coluna da esquerda nos resultados da modelagem).
- **\$DRP-<target>**. Escore de propensão indicando a probabilidade da resposta desejada. Disponível somente para respostas flag.

Capítulo 7. Modelos de escoragem

Escoragem de modelos em uma tabela de base de dados, arquivo, Analytic Server ou servidor IBM Cognos BI

Escorar um modelo significa aplicá-lo à amostra de dados ou preenchimento de seu interesse. Por exemplo, você pode usar seus clientes existentes para gerar um modelo que preveja quais clientes têm maior probabilidade de responderem a uma campanha e depois usar o modelo para escorar registros para pessoas que não são seus clientes atualmente, mas que você pode querer incluir em um email promocional.

Os passos básicos na escoragem de um modelo são:

1. Selecionar a origem de dados e quaisquer subconjuntos desses dados para escorar.
2. Selecionar o destino para os dados escorados.
3. Selecionar as informações para emitir e como escorá-las.
4. Especifique quais registros escorar. No menu suspenso Escorar, selecione se deseja escorar **Todos os registros**, os **N% principais**, os **N principais registros**, a **Propensão mín-máx** (disponível somente para respostas *flag*) ou **Escore mínimo** (disponível somente para *variável resposta contínua*).
5. Visualize os resultados esperados e execute uma visão geral dos dados, se desejar.
6. Clique em **Escorar Agora** para executar uma escoragem com relação ao modelo usando suas configurações. Ao concluir, revise a seção Relatório Sumarização de Escoragem. Consulte o tópico “Escoragem do relatório sumarização” na página 80 para obter mais informações.

Escores são gravados em um ou mais campos que são incluídos nos dados. Por convenção, os nomes desses campos são baseados no campo de destino, mas com um prefixo incluído. Por exemplo, se o campo de destino for denominado *Resposta*, os campos de saída terão nomes como *\$XF-Response* e *\$XFRP-Response*. Consulte o tópico “Selecionando campos de saída” na página 78 para obter mais informações.

Nota: Quando os dados de origem são originados em um banco de dados, a execução retrocede para o banco de dados sempre que possível (em um processo conhecido como "retrocesso de SQL"). Em geral, o IBM Analytical Decision Management suporta o mesmo nível de retrocesso SQL que o IBM SPSS Modeler, mas com as seguintes limitações:

- Se o projeto incluir regras de seleção ou campos adicionais que fazem referência a modelos ou regras externas, essas referências não retrocederão para o banco de dados e isso pode inibir o grau de retrocesso executável para o fluxo de projeto como um todo.
- Em aplicativos diferentes do IBM SPSS Modeler Advantage, o retrocesso é ativado somente até as regras de seleção (seleções globais e seleções de membros na primeira dimensão). O grau de retrocesso pode ser aumentado se você escolher um campo Número Inteiro como ID de entidade no modelo de dados do projeto.

IBM SPSS Modeler Advantage

Para **IBM SPSS Modeler Advantage**, o fluxo salvo conterà duas ramificações: a ramificação de construção e a ramificação de escore.

O nó de origem da ramificação de **construção** será sempre baseado na origem de dados de modelo de dados do projeto. Ele incluirá campos que são marcados como operacionais ou não operacionais.

O nó de origem para a ramificação de **escore** será sempre baseado na origem de dados de modelo de dados do projeto. No entanto, ele incluirá somente os campos marcados como operacionais. Os campos

marcados como não operacionais não serão incluídos, pois eles não são necessários para escoragem. Por exemplo, o campo de destino de um modelo preditivo é normalmente sinalizado como não operacional e não é necessário como uma entrada de escoragem.

Por padrão, a ramificação de escore do fluxo salvo terá um nó terminal Tabela. Esse nó de tabela será usado para indicar a ramificação de escore nas configurações de implementação de fluxo. O Serviço de Escoragem IBM SPSS Collaboration and Deployment Services usará essas informações para determinar o nó terminal a ser usado ao executar a escoragem em tempo real.

Se configurado, a guia **Escore** terá uma seção para **Escoragem de Lote**. Em **Escoragem de Lote**, há uma seção para **Opções Escorar Agora**. Essas configurações somente serão aplicadas ao escorar dentro da interface com o usuário do IBM Analytical Decision Management clicando no botão **Escorar Agora**. Elas não serão refletidas no fluxo gerado, que é salvo. Como resultado, elas não serão aplicadas ao escorar por meio do cliente do IBM SPSS Modeler nem ao planejar uma tarefa usando o cliente do IBM SPSS Collaboration and Deployment Services Deployment Manager.

A escoragem de lote também tem controles para selecionar os campos de saída e de destino. Ao salvar o fluxo, um nó terminal adicional será criado na ramificação de escore. As configurações do campo de saída e de destino serão usadas para criar um nó terminal Exportação na ramificação de escoragem. Essas configurações serão aplicadas ao executar o fluxo a partir do cliente do IBM SPSS Collaboration and Deployment Services Deployment Manager ou ao planejar uma tarefa usando o cliente do IBM SPSS Collaboration and Deployment Services Deployment Manager.

Projetos não IBM SPSS Modeler Advantage

Para projetos **diferentes de IBM SPSS Modeler Advantage**, o fluxo salvo terá somente a ramificação de escore. Assim como a ramificação de escore do IBM SPSS Modeler Advantage, o nó de origem para a ramificação de escore será sempre baseado na origem de dados de modelo de dados do projeto. Novamente, ele incluirá somente os campos marcados como operacionais. Se as tabelas secundárias forem usadas, os nós de origem adicionais serão criados. Esses nós de origem secundários são definidos pelas origens de dados secundários que são associadas à origem de dados primários.

A guia **Implementar**, se configurada, poderá ter seções para **Escoragem em Tempo Real** e **Escoragem de Lote**. Essas configurações são independentes umas das outras.

Ao salvar o fluxo, as configurações de **Escoragem em Tempo Real** serão usadas para criar um nó terminal Tabela na ramificação de escoragem. Esse nó de tabela será usado para indicar a ramificação de escore nas configurações de implementação do fluxo salvo. O Serviço de Escoragem do IBM SPSS Collaboration and Deployment Services usará essas informações para determinar o nó terminal a ser usado ao executar a escoragem em tempo real.

Em **Escoragem de Lote**, há uma seção para **Opções Escorar Agora**. Essas configurações somente serão aplicadas ao escorar dentro da interface com o usuário do IBM Analytical Decision Management clicando no botão **Escorar Agora**. Elas não serão refletidas no fluxo gerado, que é salvo. Como resultado, elas não serão aplicadas ao escorar por meio do cliente do Serviço de Escoragem do IBM SPSS Collaboration and Deployment Services nem ao planejar uma tarefa usando o cliente do IBM SPSS Collaboration and Deployment Services Deployment Manager .

A escoragem de lote também tem controles para selecionar os campos de saída e de destino. Ao salvar o fluxo, essas configurações serão usadas para criar um nó terminal Exportação na ramificação de escoragem. Essas configurações serão aplicadas ao planejar uma tarefa usando o cliente do IBM SPSS Collaboration and Deployment Services Deployment Manager ou ao executar o fluxo a partir do cliente do IBM SPSS Modeler.

Todos os projetos

Ambos os projetos IBM SPSS Modeler Advantage e não IBM SPSS Modeler Advantage podem construir nós terminais distintos para escoragem de lote em tempo real. Ao planejar tarefas usando o cliente do IBM SPSS Collaboration and Deployment Services Deployment Manager, certifique-se de configurar o passo da tarefa, para que apenas o nó de exportação designado para escoragem de lote seja executado.

Selecionando dados e subconjuntos

Na lista **Origem de dados para escorar**, selecione a origem de dados necessária.

Subconjuntos de dados

Se não desejar escorar a origem de dados inteira, é possível selecionar um subconjunto dos dados usando uma regra existente ou criando uma nova regra que será usada somente para este modelo. Consulte o tópico “Definindo regras de seleção” na página 53 para obter mais informações.

Selecionando registros para serem escorados

Para modelos que retornam escores de propensão, é possível usar esses valores para selecionar os registros a serem escorados. Por exemplo, você poderia escorar os 30% dos clientes com maior propensão para responder e usar isso para produzir uma lista de distribuição. Os escores de propensão indicam a probabilidade do valor "true" para a resposta e estão disponíveis para modelos preditivos com respostas flag, bem como modelos de regra preditivos (interativos). (No caso de modelos de regra preditivos, a resposta selecionada será o valor "true" e todas as outras serão false. Para outros tipos de modelos, essas opções não estão disponíveis.)

Há três maneiras de se escolher a seleção de registros com base em propensão:

- **A porcentagem mais alta.** Por padrão, seleciona os dez principais percentuais. Essa opção permite simular seus resultados de escoragem com base no lucro. Consulte o tópico “Simulando lucro máximo” na página 69 para obter mais informações.
- **O número máximo de registros.** Por padrão, seleciona os 1000 principais registros quando disponíveis.
- **Especificando valores mínimo e máximo de propensão.** Por exemplo, todos os registros que têm uma propensão para responder entre 75% e 100%.

Selecionando o destino de escoragem

É possível escolher emitir dados de escoragem para um banco de dados, arquivo, Analytic Server ou servidor Cognos BI. Os tipos de arquivos suportados incluem texto (*.txt, *.csv), Excel, o formato *.sav usado pelo IBM SPSS Statistics, SAS e aqueles suportados pela Coleta de Dados. Consulte o tópico “Destino do arquivo” na página 76 para obter mais informações.

Destino do banco de dados

Quando você escolhe enviar uma saída de escoragem para um banco de dados, é solicitado que você insira detalhes adicionais.

1. Insira o **Nome do banco de dados** ou navegue para o desejado.
2. Especifique como escores devem ser armazenados. É possível escolher criar uma nova tabela, em cujo será solicitado que você insira um nome para a tabela. Alternativamente, é possível anexá-las a uma tabela existente ou sobrescrever uma origem de dados existente, em cujo caso será solicitado que você selecione a tabela ou origem de dados relevante.
3. Altere o tamanho da sequência de caracteres padrão, se desejar.

Ao anexar, ou sobrescrever, um banco de dados existente, é possível mapear os campos de escoragem para campos no destino existente. Consulte o tópico “Mapear campos de escoragem” na página 78 para obter mais informações.

Destino do arquivo

Escores podem ser emitidos para diversos tipos de arquivos diferentes. Para cada tipo, você navega para um arquivo existente ou cria um novo arquivo no qual dados de escoragem são incluídos.

Ao anexar a, ou sobrescrever, um arquivo existente, é possível mapear os campos de escoragem para campos no destino existente. Consulte o tópico “Mapear campos de escoragem” na página 78 para obter mais informações.

Arquivo Excel

Tipo. Selecione o tipo de arquivo Excel sendo exportado.

Incluir nomes de campo. Especifica se nomes de campos devem ser incluídos na primeira linha da linha.

Arquivo simples

Quando você escolhe enviar saída de escoragem para um arquivo simples, é possível inserir detalhes adicionais para controlar o conteúdo do arquivo.

Modo de gravação. Se **Sobrescrever** for selecionada, todos os dados existentes no arquivo especificado serão sobrescritos. Se **Anexar** for selecionada, a saída será incluída no final do arquivo existente, preservando quaisquer dados que ele contiver.

- **Incluir nomes de campo.** Se esta opção for selecionada, os nomes de campo serão gravados na primeira linha do arquivo de saída. Esta opção está disponível somente para o modo de gravação **Sobrescrever**.

Nova linha após cada registro. Se esta opção for selecionada, cada registro será gravado em uma nova linha no arquivo de saída.

Separador de campo. Especifica o caractere a ser inserido entre valores de campo no arquivo de texto gerado. As opções são **Vírgula**, **Tab**, **Espaço** e **Outros**. Se você selecionar **Outros**, insira o(s) caractere(s) delimitador(es) desejado(s) na caixa de texto.

Símbolo de aspas. Especifica o tipo de aspas para usar para valores de campos simbólicos. As opções são **Nenhum** (valores não são colocados entre aspas), **Simples (')**, **Duplas (")** e **Outros**. Se você selecionar **Outros**, insira o(s) caractere(s) de aspa desejado(s) na caixa de texto.

Codificação. Especifica o método de codificação de texto usado. É possível escolher entre o padrão do sistema ou UTF-8.

Símbolo decimal. Especifica como decimais devem ser representados nos dados.

- **Ponto (.)**. O caractere de ponto será usado como separador decimal.
- **Vírgula (,)**. O caractere de vírgula será usado como separador decimal.

Arquivo de dados do IBM SPSS Statistics

Exportar nomes de campo. Especifica um método de tratamento de rótulos e nomes de variável mediante exportação para um arquivo *.sav* do IBM SPSS Statistics. Observe que quaisquer nomes inválidos são automaticamente corrigidos substituindo caracteres inválidos.

- **Nomes e rótulos de variáveis.** Selecione para exportar nomes de campo e rótulos de campo. Nomes são exportados como nomes de variável do IBM SPSS Statistics, enquanto rótulos são exportados como rótulos de variáveis do IBM SPSS Statistics.
- **Nomes como rótulos de variáveis.** Selecione para usar os nomes de campo como rótulos de variáveis no IBM SPSS Statistics. Para evitar a possibilidade de criar nomes de IBM SPSS Statistics inválidos, selecione **Nomes e rótulos de variáveis**.

Arquivo de dados SAS

Tipo. Selecione o formato de arquivo SAS para ser criado. É possível escolher entre três formatos de arquivo SAS: SAS for Windows/OS2 (*.sd2), SAS for UNIX (*.ssd) ou SAS Versão 7/8/9 (*sas7bdat).

Exportar nomes de campo. Selecione opções para exportar rótulos e nomes de campo para usar com SAS.

- **Nomes e rótulos de variáveis.** Selecione para exportar nomes de campo e rótulos de campo. Nomes são exportados como nomes de variável SAS, enquanto rótulos são exportados como rótulos de variáveis SAS.
- **Nomes como rótulos de variáveis.** Selecione para usar os nomes de campo como rótulos de variáveis no SAS. Para evitar a possibilidade de criar nomes de SAS inválidos, selecione **Nomes e rótulos de variáveis**.

Destino do Analytic Server

Quando você escolhe enviar uma saída de escoragem para um Analytic Server, é solicitado que você insira detalhes adicionais.

1. Especifique a **Origem do banco de dados**.
2. Especifique como escores devem ser armazenados. É possível anexá-los a uma tabela existente ou sobrescrever uma origem de dados existente, em cujo caso será solicitada a seleção da tabela ou origem de dados relevante.
3. Selecione os campos de saída para incluir com a saída de escoragem.
4. Especifique quais registros escorar. No menu suspenso Escorar, selecione se deseja escorar **Todos os registros**, os **N% principais**, os **N principais registros**, a **Propensão mín-máx** (disponível somente para respostas *flag*) ou **Escore mínimo** (disponível somente para *variável resposta contínua*).

Destino do servidor Cognos BI

Quando você escolhe enviar saída de escoragem para um servidor Cognos BI, é solicitado que você forneça detalhes adicionais.

Nota: É possível exportar somente dados relacionais, e não dados OLAP.

Para exportar dados para o Cognos BI, deve-se especificar o seguinte:

- **Conexão Cognos.** A conexão com o servidor Cognos BI.
- **Conexão ODBC.** A conexão com o servidor de dados Cognos usada pelo servidor Cognos BI.

As conexões devem apontar para o mesmo banco de dados, e o nome de usuário e a senha para o servidor Cognos se conectar ao ODBC devem ser os mesmos que os detalhes ODBC.

Você exporta os dados reais para o servidor de dados e os metadados do pacote para o servidor Cognos BI.

1. Selecione o **Destino** do Cognos BI Server e clique em **Conectar**; será solicitado que você insira os detalhes da conexão do servidor.
 - a. Digite a **URL do Servidor** do servidor IBM Cognos a partir do qual importar ou exportar dados. Entre em contato com o administrador do sistema Cognos se você não estiver certo quanto a qual URL usar.
 - b. Selecione o **Modo** no qual irá se conectar. Para efetuar login como um usuário específico (por exemplo, como um administrador), selecione **Configurar Credenciais** e insira o namespace, nome de usuário e senha do IBM Cognos:
 - Insira o **Namespace** do provedor de autenticação de segurança do IBM Cognos usado para efetuar logon no servidor. O provedor de autenticação é usado para definir e manter usuários, grupos e papéis e para controlar o processo de autenticação.
 - Insira o **Nome de Usuário** do IBM Cognos com o qual efetuar logon no servidor.
 - Insira a **Senha** associada ao nome de usuário especificado.

OU

Navegue e selecione uma credencial IBM Cognos armazenada no Repositório do IBM SPSS Collaboration and Deployment Services.

- c. Como alternativa, para efetuar login sem credenciais do usuário, caso você não possa especificar os campos de namespace, nome de usuário e senha, selecione **Usar Conexão Anônima**.

Nota: Algumas conexões do servidor não permitem a conexão anônima.

2. Selecione a **Origem de Dados**.
3. Digite o caminho e o nome da **Pasta** no servidor Cognos BI onde o pacote de exportação deve ser criado. Alternativamente, é possível navegar para a desejada.
4. Digite o **Nome do Pacote** do pacote que deve conter os metadados exportados. Ele deve ser um novo pacote; não é possível exportar para um pacote existente. Consulte o tópico “Selecionando detalhes de objetos Cognos” na página 40 para obter mais informações.
5. Para selecionar o Destino ODBC, insira o **Nome do Banco de Dados** ou navegue para o desejado.
6. Especifique como escores devem ser armazenados. É possível escolher criar uma nova tabela, em cujo será solicitado que você insira um nome para a tabela. Alternativamente, é possível anexá-las a uma tabela existente ou sobrescrever uma origem de dados existente, em cujo caso será solicitado que você selecione a tabela ou origem de dados relevante.

Ao anexar, ou sobrescrever, um banco de dados existente, é possível mapear os campos de escoragem para campos no destino existente. Consulte o tópico “Mapear campos de escoragem” para obter mais informações.

Mapear campos de escoragem

Quando você anexa ou sobrescreve escores em um arquivo ou tabela de base de dados existente, talvez seja necessário mapear os campos de escoragem para campos na tabela ou arquivo existente. Observe que a saída e o campo mapeado associado devem ter o mesmo tipo de dados.

- Durante a gravação de escores em uma tabela de base de dados existente, todos os campos devem ser mapeados para campos existentes para a escoragem continuar, ou sobrescrevendo ou anexando os novos escores. Se for necessário, clique em **Mapear Campos** para especificar saídas para quaisquer campos restantes.
 - Durante a gravação em um arquivo existente, o mapeamento só é necessário se **Anexar** estiver selecionado sob as opções do arquivo de saída. Quando você sobrescreve um arquivo existente, o mapeamento não é necessário, já que o novo arquivo simplesmente substituirá o antigo.
1. Clique em **Mapear Campos**. Um diálogo é aberto e o sistema faz a melhor suposição de mapeamento de campos para você. Em qualquer ponto, é possível retornar para a melhor suposição do sistema clicando em **Reconfigurar**.
 2. Selecione entre os destinos de saída disponíveis e clique em **Mapear** para mapeá-los para os campos de escoragem necessários.
 3. Ao concluir, clique em **Salvar**.

Selecionando campos de saída

Na guia *Score*, selecione **Especificar campos de saída** para escolher os campos para incluir com a saída de escoragem. Além dos escores reais, normalmente você vai querer pelo menos um campo que lhe permita identificar cada registro, como um campo de ID de cliente. Opcionalmente, alguns ou todos os campos de entrada usados na modelagem também podem ser incluídos.

Campos de saída do modelo

Por convenção, os nomes dos campos de saída de escoragem são baseados no nome do campo de destino, mas com a inclusão de um prefixo customizado. Por exemplo, se o campo de destino for denominado *Resposta*, os campos de saída terão nomes como *\$XF-Response* e *\$XFRP-Response*. Os campos

específicos dependem do tipo de modelo e do nível de medição para o campo de destino selecionado, conforme detalhado abaixo. Consulte o tópico “Níveis de medição” na página 43 para obter mais informações.

Respostas flag. Modelos automatizados com respostas flag geram os seguintes campos:

- **\$XF-<target>**. Valor predito para cada registro, relatado em termos de valores "true" e "false" definidos para o campo.
- **\$XFRP-<target>**. Escore de propensão; indica a probabilidade do valor "true" para cada registro. Os escores de propensão são diferentes dos escores de confiança, que normalmente são relatados em termos de predição para tal registro. Por exemplo, uma predição "false" com confiança alta seria convertida em uma baixa propensão, já que isso indicaria uma alta probabilidade de *não* responder.

Variáveis resposta categórica. Modelos automatizados com variáveis resposta categórica (nominal ou ordinal) geram os seguintes campos:

- **\$XS-<target>**. Valor predito para cada registro.
- **\$XSC-<target>**. Valor de confiança associado à predição.

Variáveis resposta contínua. Modelos automatizados com respostas numéricas contínuas geram os seguintes campos:

- **\$XR-<target>**. Valor predito para cada registro.
- **\$XRE-<target>**. Erro padrão para a predição.

Modelos de regra preditivos (interativos). Modelos de regra preditivos (somente variáveis resposta categórica ou flag) geram os seguintes campos:

- **\$D-<target>**. Resultado predito para cada registro, sendo que o valor "true" é definido para o campo usado para indicar a resposta desejada e null indica qualquer outro valor.
- **\$DP-<target>**. Probabilidade do resultado desejado para cada registro; corresponde à probabilidade exibida para o segmento na tabela de resultados durante a construção do modelo. Consulte o tópico “Construindo um modelo de regra preditivo” na página 59 para obter mais informações.
- **\$DI-<target>**. Número do índice usado para identificar o segmento no qual o registro se encaixa (exibido na primeira coluna da esquerda nos resultados da modelagem).
- **\$DRP-<target>**. Escore de propensão indicando a probabilidade da resposta desejada. Disponível somente para respostas flag.

Modelos de escoragem

Após ter especificado seus detalhes de escoragem, é possível visualizar uma amostra dos escores de modelo. Isso é útil principalmente quando você tem um grande conjunto de dados, já que lhe permite verificar rapidamente se os resultados da escoragem aparecem conforme esperado, sem precisar esperar o conjunto de dados inteiro ser escorado.

Nota: Se os dados forem considerados inválidos ou incompletos, um aviso será exibido quando você tentar visualizá-los.

A visualização mostra dados como uma tabela que exibe registros por linha e variáveis por colunas.

Escoragem

Para estabelecer quantos registros serão escorados, clique no ícone de atualização. Isso fornece uma contagem do número de registros que serão usados para a construção do modelo, levando em conta o número total de registros nas origens de dados e quaisquer configurações opcionais que você especificar.

Quando você clica em **Escorar Agora**, uma barra de progresso é exibida na parte inferior da guia. É possível parar a escoragem a qualquer momento enquanto ela for exibida.

Notas:

Se os dados forem considerados inválidos ou incompletos, um aviso será exibido quando você tentar escorá-los.

Durante a escoragem para um banco de dados, a execução retrocede para o banco de dados sempre que possível (o que geralmente é chamado de "retrocesso de SQL"). Observe que se o projeto incluir uma regra que contém uma referência a um modelo, a execução não retrocederá para o banco de dados.

Após os dados serem escorados com sucesso, uma resultado informa onde os resultados foram armazenados -- no banco de dados ou arquivo especificado -- e a seção Relatório Sumarização de Escoragem fornece uma sumarização. Consulte o tópico "Escoragem do relatório sumarização" para obter mais informações. Se você escolheu enviar a saída da escoragem para um arquivo simples, e se um administrador tiver ativado a funcionalidade, é possível fazer download de uma cópia local do arquivo, se desejado.

Escoragem do relatório sumarização

Após a escoragem de um projeto, a seção Escoragem do Relatório Sumarização fornece detalhes, como onde os dados foram escorados, tamanho do arquivo (se aplicável), nome da tabela e do banco de dados (se aplicável), nome do modelo ou nome do projeto, número de registros escorados; em alguns casos, um gráfico e/ou tabela resumindo o modelo. A seção Relatório Sumarização de Escoragem está disponível em IBM SPSS Modeler Advantage e na guia Implementar dos aplicativos que inclui a variável Escorar Agora.

Se você escolheu enviar a saída da escoragem para um arquivo simples, também é possível fazer download de uma cópia local do arquivo, se desejado. Os administradores podem usar IBM SPSS Collaboration and Deployment Services Deployment Manager baseada em navegador para ativar ou desativar a variável e para aumentar ou diminuir o limite de tamanho de arquivo máximo permitido. Se o botão **Download de Arquivo** estiver inativo, o tamanho do arquivo excede o limite de tamanho máximo configurado pelo administrador ou o administrador desativou a variável. A variável fica desativada por padrão .

Capítulo 8. Administrando aplicativos

Os administradores podem controlar a maneira como os usuários corporativos interagem com aplicativos IBM Analytical Decision Management. Esse nível de controle torna as tarefas do usuário corporativo mais fáceis para que eles possam usar o aplicativo de maneira mais rápida e fácil para ajudar na resolução de problemas de negócios. Ele também permite que os administradores restrinjam o acesso a certas funcionalidades.

Use IBM SPSS Collaboration and Deployment Services Deployment Manager para conceder autoridade administrativa a usuários ou grupos específicos conforme necessário. Somente os usuários com permissão para a ação *Administração do Decision Management* têm acesso às variáveis descritas nesta seção.

1. Abra o IBM SPSS Collaboration and Deployment Services Deployment Manager.
2. No menu Ferramentas, selecione **Administração do Servidor**.
3. Na guia Administração do Servidor, conecte seu servidor.
4. Em **Usuários e Grupos**, selecione **Repositório do Usuário Local**.
5. Clique em **Novo Grupo** para criar um novo grupo para os administradores e um novo grupo para os usuários corporativos. Por exemplo:
 - Administradores do Decision Management
 - Usuários do Decision Management

Também é possível criar novos usuários neste momento, se você desejar. Durante a criação de grupos, inclua os usuários apropriados em cada grupo. Por exemplo, os membros do grupo de administradores receberão a concessão de autoridade para administrar aplicativos IBM Analytical Decision Management em sua organização.

6. Acesse Papeis e clique em **Novo Papel** para criar um novo papel administrativo e um novo papel de usuário corporativo. Por exemplo:
 - Administrador do IBM Analytical Decision Management
 - Usuário do IBM Analytical Decision Management

Durante a criação de novos papeis, designe a eles pelo menos as ações a seguir.

Tabela 7. Ações do papel.

Papel	Ações Necessárias
Administrador do Decision Management	Acesso a Conteúdos e Pastas
	Administração do Decision Management
	Edição de Tarefa
	Execução Dinâmica de Relatório
	Modelo de Escore
	Envio de Trabalho
	Mostrar todas as Versões
Usuário do Decision Management	Acesso a Conteúdos e Pastas
	Edição de Tarefa
	Execução Dinâmica de Relatório
	Modelo de Escore
	Envio de Trabalho
	Mostrar mais recente OU Mostrar todas as Versões*

Nota: A ação **Mostrar mais recente ou **Mostrar todas as Versões** será necessária se os usuários precisarem de acesso às versões do projeto que eles criaram. Caso contrário, os projetos serão abertos no modo somente leitura se o usuário atual não tiver acesso à versão mais recente.*

7. Selecione o papel administrador do Decision Management e clique em **Editar Usuários e Grupos**. Inclua o grupo Administradores do Decision Management que você acabou de criar.
8. Selecione o papel usuário do Decision Management e clique em **Editar Usuários e Grupos**. Inclua o grupo Usuários do Decision Management que você acabou de criar.

Opções gerais

Certas opções gerais estão disponíveis para administradores em todas as guias da interface com o usuário do IBM Analytical Decision Management. Essas opções permitem que os administradores ocultem ou bloqueiem guias dos usuários corporativos.

- A parte superior de cada guia tem a opção **Ocultar guia <nome da guia> de não administradores**, em que **<nome da guia>** é o nome da guia. Isso oculta a guia inteira dos usuários corporativos. Por exemplo, se você não quiser que os usuários vejam a guia Dados para um aplicativo, efetue logon no aplicativo, acesse a guia Dados e selecione **Ocultar guia Dados de não administradores**.
- A parte superior de cada guia tem a opção **Bloquear todas as opções de <nome da guia>**, em que **<nome da guia>** é o nome da guia. Isso bloqueia todos os controles na guia para que ela seja somente leitura para o usuário corporativo. Por exemplo, se você quiser que os usuários corporativos possam visualizar as atuais configurações de dados, mas não alterá-las, é possível configurar essa opção na guia Dados de um aplicativo. Essa configuração também é bastante útil para situações nas quais você deseja configurar certas antecipadamente para o usuário corporativo para simplificar sua tarefa.
- Os administradores também podem bloquear e desbloquear configurações individuais em cada guia clicando no ícone cadeado quando ele estiver disponível.



Por exemplo, em aplicativos com uma guia Definir, os administradores podem bloquear a árvore de hierarquia da dimensão para que usuários corporativos não possam alterá-la. Um clique no ícone cadeado bloqueia ou desbloqueia seu controle associado. Os usuários corporativos verão o ícone cadeado trancado para os controles que o administrador bloqueia, mas nunca verão um ícone cadeado destrancado.

Itens individuais que são bloqueados também podem ser exibidos com um segundo plano esmaecido ou cores silenciadas para indicar que eles estão indisponíveis para seleção; uma mensagem de erro será exibida se um usuário tentar selecioná-los. Quando uma guia inteira está bloqueada, uma mensagem para esse efeito é exibida na parte superior da guia.

Controlando atalhos do aplicativo

A Página de Ativação de Aplicativos fornece uma maneira fácil para os usuários corporativos abrirem aplicativos e incluírem atalhos para versões rotuladas específicas dos aplicativos. Os administradores podem especificar quais atalhos de aplicativo aparecerão na página de ativação para todos os usuários. Após a criação de log com autoridade do administrador:

- Se você desejar, clique no ícone cadeado em um atalho do aplicativo para que os usuários corporativos não possam removê-lo de sua página de ativação. Observe que os usuários corporativos ainda podem incluir e remover outros atalhos de aplicativo. Os usuários verão o ícone cadeado trancado no lugar do ícone fechar para atalhos bloqueados. Consulte o tópico “Ativando aplicativos” na página 25 para obter mais informações.

Ocultando a página inicial

Cada aplicativo fornece um fluxo de trabalho passo a passo, conforme representado pelo gráfico na página inicial. Um clique no gráfico leva você até a seção da interface com o usuário. No entanto, talvez nem sempre os usuários corporativos precisem acessar cada passo do fluxo de trabalho de um aplicativo.

Os administradores podem ocultar a página inicial dos usuários corporativos. Por exemplo, em um determinado aplicativo, os usuários corporativos podem usar somente a guia Relatórios. Nesse caso, um administrador poderia ocultar a página inicial e todas as outras guias dos usuários corporativos. Portanto, quando o usuário corporativo efetuar logon, ele será levado diretamente para a guia Relatórios e nunca verá outras guias.

1. Na página inicial, selecione **Ocultar página inicial do aplicativo**.
2. No menu suspenso, especifique qual guia exibir quando os usuários abrirem o aplicativo pela primeira vez.

Observe que se você ocultar essa guia selecionada subsequentemente, a próxima guia não oculta no aplicativo será exibida.

Opções de bloqueio de dados

Dados

A guia Dados define os conjuntos de dados que o aplicativo usa para análise, simulação e teste, escoragem ou outro uso operacional. Essas origens de dados são definidas com relação à sua instalação de servidor; assim, todos os caminhos para arquivos de dados, origens ODBC e outros são relativos ao servidor.

Como administrador, talvez você queira configurar alguns ou todos os dados antes de os usuários corporativos começarem a usar seus aplicativos. Os administradores podem:

- Selecione **Bloquear Origem de Dados** ao criar uma nova origem de dados para bloqueá-la após sua criação.
- Bloqueie e desbloqueie origens de dados individuais clicando no ícone cadeado para uma origem de dados na tabela ou no diálogo Editor de Origem de Dados.
- Configure e bloqueie o modelo de dados usado no aplicativo selecionando um modelo de dados no menu suspenso **Modelo de Dados do Projeto** clicando no ícone cadeado.

Observe que os usuários corporativos podem copiar uma origem de dados bloqueada dos Conjuntos de Dados do Projeto para sua seção Meus Conjuntos de Dados, mas ela permanece bloqueada em ambos os conjuntos. Consulte o tópico Capítulo 4, “Gerenciando origens de dados”, na página 33 para obter mais informações.

- Se o modelo de dados do projeto for uma *Visualização de Dados*, use o menu suspenso **Permitir o uso de regras do ODM com base em dados hierárquicos** para controlar o uso de regras do IBM Operational Decision Management (ODM) que têm um modelo de dados hierárquicos. O uso de tais regras pode resultar em baixo desempenho. As opções a seguir estão disponíveis no menu suspenso:
 - **Não permitir** é recomendado para você obter o melhor desempenho. Proíbe que regras do ODM com estrutura de dados complexos sejam usadas em qualquer lugar no projeto. Será exibido um erro se um usuário selecionar uma regra complexa em qualquer ponto no aplicativo.
 - **Somente na guia Modelagem** permite que regras complexas sejam referenciadas durante a construção da guia Modelagem. Isso permite a inclusão de flexibilidade com degradação de desempenho. Essa opção aplica-se ao IBM SPSS Modeler Advantage.

Bloqueando opções de destino de escoragem

Para aplicativos configurados para usar a seção Escorar Agora (guia Implementar ou Escorar), há uma seção que permite que usuários corporativos especifiquem o destino para armazenar dados de escore (no banco de dados ou arquivo). Por exemplo, se você estiver armazenando escores em um banco de dados, especifique o nome do banco de dados e o método para armazenar os escores (como criar uma nova tabela no banco de dados).

Como administrador, é possível bloquear as opções de destino para escoragem. Por exemplo, talvez você queira que os escores sejam anexados somente a uma tabela existente em um certo de banco de dados -

em vez de dar aos usuários corporativos a opção de sobrescrever dados existentes, anexar novos dados a uma tabela existente, criar uma nova tabela ou armazenar dados em um arquivo.

- Para bloquear todas as opções de destino de escoragem, clique no ícone cadeado próximo do menu suspenso Destino na guia Implementar. Os usuários corporativos poderão ver as opções de destino, mas não editá-las.

Fazendo download dos metadados do projeto para uso na criação de regras externas

Regras criadas em um Business Rules Management System, como IBM Operational Decision Management, podem ser referenciadas e usadas em aplicativos IBM Analytical Decision Management, contanto que tenham sido desenvolvidas para suportar o mesmo modelo de dados usado no atual projeto IBM Analytical Decision Management e possam ser implementadas como um serviço da web para serem usadas pelo aplicativo. Para isso, um administrador do IBM Analytical Decision Management deve fazer o download de metadados relevantes. Há duas maneiras de se fazer o download dos metadados, e a complexidade do modelo de dados e a potencial complexidade da regra resultante podem diferir:

Fazendo download dos metadados do projeto no IBM Analytical Decision Management

Um administrador pode clicar no ícone **Fazer Download de Metadados** no IBM Analytical Decision Management para fazer o download de um arquivo .ZIP contendo metadados para o projeto atual. O arquivo transferido por download é salvo localmente no sistema usado por um administrador para acionar o download.



Figura 14. Ícone Download de metadados

Os metadados transferidos por download dessa maneira suportam a criação de regras com base no modelo de dados do projeto. O modelo de dados do projeto será simples (não hierárquico), portanto, quaisquer regras criadas com base nesses metadados só poderão utilizar um modelo de dados simples. Uma regra criada com base nesse modelo de dados simples pode ser referenciada em um projeto IBM Analytical Decision Management, independentemente do tipo de origem de dados do modelo de dados do projeto.

Os arquivos a seguir são incluídos no arquivo .ZIP:

- **Definição de esquema XML (*.XSD)**. Contém definições dos tipos de campo no modelo de dados para o atual projeto IBM Analytical Decision Management. Esse arquivo pode ser importado em uma ferramenta de desenvolvimento externa, como IBM Operational Decision Management Rule Designer, para desenvolver regras para serem usadas com o projeto atual.
- **Modelo Web Service Description Language (*.WSDL)**. Incluído para ajudar no desenvolvimento de serviços da web que usam modelo de dados. O arquivo .WSDL é fornecido como um modelo que pode ser usado para construir serviços da web responsáveis por mediar as mensagens entre o aplicativo IBM Analytical Decision Management e um serviço de regra externo. Os detalhes sobre a mediação de serviço da web podem ser fornecidas manualmente por um integrador.

Fazendo download de metadados do modelo de dados para uma visualização de dados analíticos do IBM SPSS Collaboration and Deployment Services

No IBM SPSS Collaboration and Deployment Services Deployment Manager, é possível exportar metadados para uma determinada visualização de dados analíticos.

Metadados transferidos por download dessa maneira suportam a criação de regras com base no modelo de dados lógicos de uma visualização de dados analíticos. Um modelo de dados lógicos da visualização de dados analíticos pode ser hierárquico. Uma origem Visualização de Dados do IBM Analytical Decision Management suporta a geração de uma representação XML dos dados hierárquicos para cada registro, que podem ser passados para regras externas durante a execução. As regras criadas com base nesses metadados podem ser complexas e utilizar o modelo de dados hierárquicos.

No IBM Analytical Decision Management, uma regra externa baseada em um modelo de dados hierárquicos só pode ser referenciada em um projeto no qual o modelo de dados do projeto usa uma origem Visualização de Dados. Para assegurar a compatibilidade, a origem Visualização de Dados e a regra externa devem ser baseadas no mesmo modelo de dados lógicos da visualização de dados analíticos. Somente a origem Visualização de Dados pode gerar os dados hierárquicos com base na instância da visualização de dados analíticos à qual fazem referência. Em outras palavras, para usar regras externas, o modelo de dados do projeto no projeto IBM Analytical Decision Management precisa corresponder ao usado para criar a regra, e a visualização de dados analíticos torna isso possível.

Para fazer o download de um arquivo .ZIP com metadados para o modelo de dados lógicos da visualização de dados analíticos:

1. No IBM SPSS Collaboration and Deployment Services Deployment Manager, abra uma visualização de dados analíticos a partir do repositório e acesse a guia **Lógico**.
2. Clique com o botão direito em uma tabela e exporte os metadados XOM.
3. Selecione uma localização para salvar o arquivo e clique em **Salvar**.

Os arquivos a seguir estão incluídos no arquivo .ZIP:

Definições de esquema XML (*.XSD). Duas ou mais definições de esquema XML estão incluídas. Uma contém uma definição que pode ser útil para um tipo de resposta de regra. O(s) outro(s) esquema(s) contém(êm) definições das tabelas, incluindo detalhes do relacionamento hierárquico e tipos de campo no modelo de dados lógicos da visualização de dados analíticos.

É possível importar esses arquivos de esquema em uma ferramenta de desenvolvimento externa, como IBM Operational Decision Management Rule Designer, para desenvolver regras para serem usadas com um projeto, cujo modelo de dados do projeto usa a origem Visualização de Dados com base na mesma instância da visualização de dados analíticos.

Para obter informações sobre como criar regras externas para usar com IBM Analytical Decision Management, consulte *Application Designers Guide*.

Observe que os administradores podem restringir o uso dessas regras externas complexas porque elas podem comprometer o desempenho. Para obter detalhes, consulte “Opções de bloqueio de dados” na página 83.

Editando a definição de esquema XML

O arquivo .XSD transferido por download a partir do IBM Analytical Decision Management contém dois tipos de dados complexos: um representando o modelo de dados do projeto IBM Analytical Decision Management, que deve ser usado para o parâmetro de entrada dos projetos de regra do IBM Operational Decision Management; e um representando um tipo de retorno de regras, que deve ser usado para o parâmetro de saída dos projetos de regras do IBM Operational Decision Management.

Os metadados da visualização de dados analíticos do IBM SPSS Collaboration and Deployment Services Deployment Manager contém dois ou mais arquivos .XSD: um relacionado ao modelo de dados lógicos de uma visualização de dados analíticos, que deve ser usado para o parâmetro de entrada dos projetos de regras do IBM Operational Decision Management; e um ou mais contendo um tipo complexo representando um tipo de retorno de regras, que deve ser usado para o parâmetro de saída dos projetos de regras do IBM Operational Decision Management.

Usuários especialistas podem escolher editar a definição de esquema .XSD gerada para customizá-la para a regra ou regras específicas que eles desejam criar. Há circunstâncias nas quais o usuário deseja editar a entrada e/ou os tipos relacionados a resposta definidos na definição de esquema .XSD relacionada antes da criação de suas regras. Por exemplo, o usuário especialista pode precisar editar o esquema relacionado ao parâmetro de entrada para permitir que as regras criadas por ele manipulem o potencial dos campos de entrada com valores nulos. O usuário especialista pode fazer isso indicando via a definição de esquema .XSD quais campos/elementos de parâmetro de entrada são obrigatórios e quais são opcionais.

Editando o tipo de parâmetro de entrada

O número de ocorrências possíveis de cada elemento relacionado a campo não é configurado explicitamente no esquema .XSD (ou seja, os atributos `minOccurs` e `maxOccurs` não estão sendo configurados e o valor padrão para ambos é 1). Dessa forma, se o esquema não for editado antes de ser usado para gerar regras do IBM Operational Decision Management, todos os campos no tipo de parâmetro de entrada serão requeridos pelas regras geradas.

Editando o tipo de parâmetro de entrada relacionado ao modelo de dados do projeto

Se o esquema definindo o tipo de parâmetro de entrada foi gerado por meio do download dos metadados do projeto de um IBM Analytical Decision Management, o modelo de dados dos parâmetros de entrada será simples. Nesse caso, o IBM Analytical Decision Management criará diretamente a carga útil requerida pelo serviço de regra em execução. Isso dá ao usuário especialista mais liberdade durante a edição do tipo de parâmetro de entrada enquanto mantém a compatibilidade (ou seja, o usuário especialista pode remover campos/elementos não requeridos para criação de regra ou indicar que campos/elementos são opcionais).

Editando o tipo de parâmetro de entrada relacionado ao modelo de dados lógicos da visualização de dados analítica

Se o esquema definindo o tipo de parâmetro de entrada foi gerado por meio da exportação de um modelo de dados lógicos da visualização de dados analítica do cliente IBM SPSS Collaboration and Deployment Services Deployment Manager, o modelo de dados do parâmetro de entrada pode ser hierárquico. Se ele for hierárquico, o IBM Analytical Decision Management *não* criará diretamente a carga útil requerida pelo serviço de regra em execução. A principal parte da carga útil será gerada pela origem Visualização de Dados Analítica. Isso limita como o usuário especialista pode editar o tipo de parâmetro de entrada e manter a compatibilidade (por exemplo, o usuário especialista só pode indicar quais campos/elementos são opcionais).

Informações gerais

Antes de chamar uma regra criada em um Business Rules Management System, como IBM Operational Decision Management, IBM Analytical Decision Management pergunta para o serviço da regra se seu modelo de dados de parâmetro de entrada é hierárquico ou simples. Se o parâmetro de entrada for simples, ele também verifica quais entradas são requeridas pelo serviço de regra.

Se o modelo de dados do parâmetro de entrada for simples, e o registro para score contiver um valor nulo para qualquer um dos campos obrigatórios do serviço de regra, o IBM Analytical Decision Management não chamará a regra e os campos de saída/resposta da regra serão preenchidos com valores padrão. O valor depende do tipo de campo saída/resposta da seguinte forma:

Tabela 8. Valores para tipos de campo de resposta.

Tipo de campo de resposta	Valor
xsd:string	""
xsd:boolean	""
xsd:int	0

Tabela 8. Valores para tipos de campo de resposta (continuação).

Tipo de campo de resposta	Valor
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

Isso pode afetar os resultados dos registros de escoragem via um serviço de regra, mas é preciso tomar cuidado na edição do parâmetro de entrada no esquema *.XSD*.

Se o modelo de dados do parâmetro de entrada for hierárquico, e o registro a ser escorado contiver um valor nulo para quaisquer campos obrigatórios do serviço de regra, o IBM Analytical Decision Management chamará a regra, mas o campo ou campos com um valor nulo serão omitidos da carga útil criada pelo nó de origem Visualização de Dados. Se o campo omitido era um campo obrigatório, isso fará com que a chamada da regra falhe, portanto, tenha cuidado ao editar o parâmetro de entrada no esquema *.XSD*.

A edição de esquema requer conhecimento dos serviços de regra que serão criados. Se um determinado campo de entrada não for requerido pela regra, e o modelo de dados do parâmetro de entrada da regra for **simples**, o elemento relacionado ao campo pode ser **removido** do parâmetro de entrada no esquema ou **marcado como opcional**. Se um determinado campo não for requerido pela regra, e o parâmetro de entrada da regra for **hierárquico**, o elemento relacionado ao campo **não poderá ser removido**; mas ele pode ser **marcado como opcional**. Isso acontece porque o IBM Analytical Decision Management não gera diretamente a carga útil para regras baseadas em um modelo de dados hierárquico.

Quando você remove um campo de parâmetro de entrada, o valor do campo é ignorado quando a regra é chamada e a regra não pode se referir ao campo. Quando você torna um campo de parâmetro de entrada opcional, um valor nulo é permitido e a regra deve ser gravada defensivamente para verificar isso. Por exemplo, se o modelo de dados do parâmetro de entrada da regra for simples e você souber que o campo Age não será usado pela regra, é possível remover o elemento relacionado para indicar que ele não é um campo obrigatório. No exemplo de esquema *.XSD* a seguir, se o campo Age não for obrigatório, a seção `<xsd:element name="Age">` inteira poderá ser removida do 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:sequence>
</xsd:complexType>
<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>

```

Ou, se o modelo de dados do parâmetro de entrada da regra for simples ou hierárquico, você pode tornar um elemento de campo de entrada opcional se ele não for requerido pela(s) regra(s)--configurando o valor do atributo `minOccurs` do elemento para 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>

```

Exemplo de modelo de dados hierárquico:

```

<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>

```

As regras também podem ser gravadas defensivamente para manipular valores nulos, em cujo caso o campo pode ser marcado como opcional no esquema. A seguir está um exemplo de uma regra do IBM Operational Decision Management chamada *currentOfferIsReplacePhone* que foi gravada para lidar com 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 ;

```

Editando o tipo de parâmetro de resposta

Observe que é possível criar e utilizar regras do IBM Operational Decision Management que retornam uma variedade de tipos para o IBM Analytical Decision Management (*boolean* para usar em Seleções, *numeric* para usar em regras agregadas, por exemplo, e *string*). No entanto, para criar qualquer coisa diferente de uma regra de resposta de dados *string*, a definição de esquema XML (.XSD) que contém o tipo de resposta deve ser editada manualmente.

Além disso, um projeto de regra pode conter várias regras, cada uma retornando diferentes tipos. É necessário manter diversas versões de .XSD para suportar esta (por exemplo, uma versão na qual a variável de resposta de dados foi configurada para numeric, outra para string, outra para boolean, etc.).

O arquivo .XSD transferido por download a partir do IBM Analytical Decision Management contém dois tipos de dados complexos: um representando o modelo de dados do projeto IBM Analytical Decision Management, que deve ser usado para o parâmetro de entrada dos projetos de regra do IBM Operational Decision Management, e um representando um tipo de retorno de regras, que deve ser usado para o parâmetro de saída dos projetos de regra do IBM Operational Decision Management.

O arquivo .XSD relacionado à resposta transferido por download a partir do IBM SPSS Collaboration and Deployment Services Deployment Manager contém a definição de tipo de resposta e conterá somente o tipo de dados complexo.

Em ambos os casos acima, o tipo de retorno (cujo nome sempre tem um sufixo "response") define um elemento value que é padronizado para *string*. Se você deseja criar um projeto de regras com um tipo diferente de valor de retorno, é preciso editar manualmente o .XSD transferido por download antes de importar o projeto de regras do IBM Operational Decision Management para que o elemento value do tipo de retorno seja definido conforme desejado. A seguir estão alguns exemplos:

```

<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: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. Obtendo ajuda

Três níveis de assistência ao usuário são fornecidos:

- Ajuda customizada é fornecida para cada aplicativo criado previamente empacotado com IBM Analytical Decision Management. Para ativar a ajuda, clique no ícone de ajuda em qualquer tela ou diálogo:



- Texto de treinamento, que é exibido em uma pequena janela pop-up, fornecendo detalhes sobre o atual problema de aplicativo ou negócio. O texto de treinamento fornece uma camada adicional de ajuda contextual customizada que pode ser facilmente configurada para um aplicativo ou site específico. O texto de treinamento aparece em toda a interface com o usuário quando você clica no pequeno ícone de texto de treinamento:



- Dicas de ferramenta padrão, que são fornecidas para controles individuais dentro de uma página ou diálogo. Para exibir uma dica de ferramenta, passe o mouse sobre o botão, ligação ou outro controle.

Documentação disponível

A documentação do IBM Analytical Decision Management está disponível em dois formatos. Ao acessar ajuda no aplicativo, o IBM Knowledge Center online é ativado, com o tópico da ajuda relevante exibido. Se você tiver problema em ativar a ajuda, entre em contato com seu administrador no caso de não ter autoridade para acessar a URL externa do Knowledge Center.

O Knowledge Center é uma coleção online e centralizada de toda a documentação. Ele está sempre disponível, mesmo quando o produto não está em execução. É recomendado marcá-lo como favorito em seu navegador. Ele também contém a documentação para outros produtos IBM, como IBM SPSS Modeler, IBM SPSS Collaboration and Deployment Services e Estatísticas.

Em momentos em que você está sem uma conexão de Internet ou quando o Knowledge Center pode estar temporariamente indisponível para manutenção, toda a documentação também fica disponível em Portable Document Format (*.pdf). Os arquivos PDF estão disponíveis com os arquivos de instalação transferidos por download e também são instalados no diretório de instalação o IBM SPSS Collaboration and Deployment Services (por exemplo, *C:\Program Files\IBM\SPSS\Deployment\8.0\Server\documentation\DecisionManagement*).

A documentação em PDF também está disponível na Web, em <http://www-01.ibm.com/support/docview.wss?uid=swg27047202>, e as notas sobre a liberação estão disponíveis em <http://www-01.ibm.com/support/docview.wss?uid=swg27047203>.

Apêndice A. Compartilhando fluxos

Modelos e projetos criados no IBM Analytical Decision Management são salvos como arquivos de fluxo e podem ser modificados e usados no IBM SPSS Modeler se disponíveis. Por exemplo, talvez você queira criar um fluxo de projeto ou modelo usando IBM SPSS Modeler Advantage ou outro aplicativo IBM Analytical Decision Management e depois customizar o fluxo no IBM SPSS Modeler antes de implementar por meio do IBM Analytical Decision Management. Alternativamente, fluxos criados no IBM SPSS Modeler podem ser abertos no IBM Analytical Decision Management, contanto que as configurações de implementação tenham sido especificadas conforme descrito neste apêndice.

Modelos que aderem ao formato descrito neste apêndice podem ser visualizados, editados, escorados e atualizados usando IBM SPSS Modeler Advantage. Outros modelos também podem ser escorados no IBM SPSS Modeler Advantage, mas podem não ser exibidos corretamente e nem atualizados.

Abrindo fluxos no IBM Analytical Decision Management

Durante a abertura de fluxos no IBM Analytical Decision Management ou IBM SPSS Modeler Advantage:

- Se o tipo de implementação não for especificado (nenhum), o fluxo poderá ser aberto, mas somente a guia Dados estará disponível.
- Se o tipo de implementação for **Somente Escoragem**, a guia Escorar também pode estar disponível, contanto que a ramificação de escoragem comece com um único nó de origem de dados de um tipo suportado pelo IBM SPSS Modeler Advantage ou IBM Analytical Decision Management. Em outras palavras, todos os tipos de origem suportados pelo IBM SPSS Modeler também são suportados pelo IBM SPSS Modeler Advantage e IBM Analytical Decision Management, contanto que haja somente um nó de origem no fluxo (e não conjuntos de dados mesclados).
- Se o tipo de implementação for **Atualização de Modelo**, a guia Modelagem também pode estar disponível, contanto que a ramificação de modelagem comece com um único nó de origem de dados de um tipo suportado pelo IBM SPSS Modeler Advantage ou IBM Analytical Decision Management e que o nó de modelagem selecionado seja um nó de modelagem automatizado (suportado para modelagem preditiva), um nó de lista de decisão (suportado para modelagem de regra preditiva), um nó de armazenamento em cluster (suportado para modelagem de cluster) ou um nó de regra de associação (suportado para modelagem de associação). Além disso, um nó Tipo deve estar presente para definir o campo de destino, bem como outras informações usadas em modelagem.

Especificando configurações de implementação no IBM SPSS Modeler

No IBM SPSS Modeler, especifique configurações de implementação na caixa de diálogo Propriedades do Fluxo, como a seguir:

1. Nos menus do IBM SPSS Modeler, escolha:
Ferramentas > Propriedades do Fluxo > Implementação
2. Na lista **Tipo de Implementação**, escolha **Somente Escoragem** ou **Atualização de Modelo** conforme apropriado.
3. Selecione **Nó Escoragem**, **Nó Modelagem** e **Nó Modelo** conforme necessário.
4. Selecione **Verificar** para confirmar se todas as opções necessárias foram especificadas ou selecione **Armazenar** para armazenar o fluxo no Repositório do IBM SPSS Collaboration and Deployment Services.

Para fluxos criados no IBM SPSS Modeler Advantage ou IBM Analytical Decision Management, as informações de implementação são configuradas automaticamente.

Dicas para usar fluxos do IBM SPSS Modeler no IBM Analytical Decision Management

- Talvez seja mais fácil criar fluxos usando IBM SPSS Modeler Advantage ou IBM Analytical Decision Management e depois modificar no IBM SPSS Modeler em vez de construir do zero usando IBM SPSS Modeler. Os fluxos criados dessa maneira serão estruturados de um modo que permita que eles fiquem totalmente funcionais no IBM Analytical Decision Management.
- Para todos os fluxos criados no IBM SPSS Modeler Advantage, são incluídos supernós em certas localizações dentro da ramificação que podem ser customizados por um usuário especialista. É altamente recomendado que quaisquer customizações ou adições no fluxo sejam incluídas nesses supernós para assegurar que elas sejam preservadas no fluxo e reconhecidas pelo IBM SPSS Modeler Advantage.
- Todos os nós de regras locais contendo um recebimento de dados imediato de regra de expressão de cada nó de origem terão seus campos incluídos no diálogo Incluir/Editar Campos Adicionais. Observe que as regras locais só podem ser criadas dentro do IBM Analytical Decision Management ou IBM SPSS Modeler Advantage, e não pelo IBM SPSS Modeler.
- Se houver um envio de dados do nó Partição do nó de modelagem (e recebimento de dados do nó Tipo), **Particionar dados automaticamente para ativar a avaliação de modelo na construção da origem de dados para avaliação e teste** será ativada no IBM SPSS Modeler Advantage e o estado da caixa de seleção será configurado, dependendo se o nó de partição estiver ou não ativado.
- Se houver um recebimento de dados de nó Automated Data Preparation (ADP) do nó Tipo, **Limpar e preparar automaticamente os dados para construção de modelo confiável** será ativada e o estado da caixa de seleção será configurado, dependendo se o nó ADP estiver ou não ativado.
- Se os nós Regra ou Seleção forem um envio de dados presente do nó de modelagem, eles serão usados para preencher a seção **Especificar Seleções para Usar**.
- Se os parâmetros de fluxo tiverem sido definidos (guia Parâmetros na caixa de diálogo Propriedades do Fluxo), uma ligação Parâmetros será exibida nas guias Modelagem e Score no IBM SPSS Modeler Advantage. O usuário pode clicar nessa ligação para especificar esses valores.

Ativando IBM SPSS Modeler Advantage a partir do IBM SPSS Modeler

Na tela inicial do IBM SPSS Modeler, é possível escolher criar um modelo com IBM SPSS Modeler Advantage.

Apêndice B. Acessibilidade

Os recursos de acessibilidade ajudam os usuários com deficiência física, como mobilidade restrita ou deficiência visual, a utilizarem os produtos de tecnologia da informação com êxito. A IBM esforça-se para fornecer produtos com acesso utilizável para todos, independentemente da idade ou capacidade. Este produto usa teclas de navegação padrão do navegador da Web.

Como o IBM Analytical Decision Management é um aplicativo baseado em navegador, as opções de acessibilidade para usuários com deficiência, como aqueles que são deficientes visuais, são controladas a partir das configurações do navegador da Web. Para obter detalhes sobre as ações na interface com o usuário, incluindo os atalhos de teclado, leitores de tela e assim por diante, consulte a documentação do navegador da Web.

Software do fornecedor

O IBM Analytical Decision Management pode requerer o uso de alguns softwares de fornecedor que não são cobertos pelo contrato de licença da IBM. A IBM não representa nenhum recurso de acessibilidade desses produtos. Entre em contato com o fornecedor para obter informações de acessibilidade sobre seus produtos.

IBM e Acessibilidade

Consulte o IBM Human Ability and Accessibility Center para obter informações adicionais sobre o compromisso da IBM com a acessibilidade.

Avisos

Estas informações foram desenvolvidas para produtos e serviços oferecidos nos EUA. Esta material pode estar disponível na IBM em outros idiomas. No entanto, pode ser solicitado que você tenha uma cópia do produto ou da versão de produto nesse idioma para acessá-lo.

O IBM pode não oferecer os produtos, serviços ou recursos discutidos neste documento em outros países. Consulte um representante IBM local para obter informações sobre produtos e serviços disponíveis atualmente em sua área. Toda referência a um produto, programa ou serviço IBM não tem a intenção de declarar ou implica em que apenas esse produto, programa ou serviço IBM possa ser usado. Qualquer produto, programa ou serviço funcionalmente equivalente, que não infrinja nenhum direito de propriedade intelectual da IBM, poderá ser usado em substituição. Entretanto, a avaliação e verificação da operação de qualquer produto, programa ou serviço não IBM são de responsabilidade do usuário.

A IBM pode ter patentes ou solicitações de patentes pendentes relativas a assuntos tratados nesta publicação. O fornecimento desta publicação não garante ao Cliente nenhum direito sobre tais patentes. O Cliente pode enviar consultas sobre licença, por escrito, para:

Gerência de Relações Comerciais e Industriais da IBM Brasil
Av. Pasteur, 138/146
Av. Pasteur, 138-146
Armonk, NY 10504-1785
CEP 22290-240

Para pedidos de licença relacionados a informações de Conjunto de Caracteres de Byte Duplo (DBCS), entre em contato com o Departamento de Propriedade Intelectual da IBM em seu país ou envie pedidos de licença, por escrito, para:

Intellectual Property Licensing
Lei de Propriedade Legal e Intelectual
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Japan

A INTERNATIONAL BUSINESS MACHINES CORPORATION FORNECE ESTA PUBLICAÇÃO "NO ESTADO EM QUE SE ENCONTRA", SEM GARANTIA DE NENHUM TIPO, SEJA EXPRESSA OU IMPLÍCITA, INCLUINDO, MAS NÃO SE LIMITANDO ÀS GARANTIAS IMPLÍCITAS DE NÃO-VIOLAÇÃO, MERCADO OU ADEQUAÇÃO A UM DETERMINADO PROPÓSITO. Algumas jurisdições não permitem a isenção de responsabilidade de garantias diretas ou indiretas em algumas transações; portanto, esta declaração pode não se aplicar a você.

Esta publicação pode incluir imprecisões técnicas ou erros tipográficos. Periodicamente são feitas alterações nas informações aqui contidas; tais alterações serão incorporadas em novas edições da publicação. A IBM pode, a qualquer momento, aperfeiçoar e/ou alterar os produtos e/ou programas descritos nesta publicação, sem aviso prévio.

Referências nestas informações a Web sites não IBM são fornecidas apenas por conveniência e não representam de forma alguma um endosso a esses Web sites. Os materiais contidos nesses websites não fazem parte dos materiais desse produto IBM e a utilização desses Websites é de inteira responsabilidade do Cliente.

A IBM pode utilizar ou distribuir as informações fornecidas da forma que julgar apropriada sem incorrer em qualquer obrigação para com o Cliente.

Licenciados deste programa que desejam obter informações sobre este assunto com objetivo de permitir: (i) a troca de informações entre programas criados independentemente e outros programas (incluindo este) e (ii) a utilização mútua das informações trocadas, devem entrar em contato com:

Gerência de Relações Comerciais e Industriais da IBM Brasil
Av. Pasteur, 138/146
Av. Pasteur, 138-146
Armonk, NY 10504-1785
CEP 22290-240

Tais informações podem estar disponíveis, sujeitas a termos e condições apropriadas, incluindo em alguns casos o pagamento de uma taxa.

O programa licenciado descrito nesta publicação e todo o material licenciado disponível são fornecidos pela IBM sob os termos do Contrato com o Cliente IBM, do Contrato de Licença de Programa Internacional IBM ou de qualquer outro contrato equivalente.

Os exemplos de clientes e dados de desempenho citados são apresentados com propósitos meramente ilustrativos. Os resultados de desempenho reais podem variar, dependendo de configurações e condições operacionais específicas.

As informações relativas a produtos não IBM foram obtidas junto aos fornecedores dos respectivos produtos, de seus anúncios publicados ou de outras fontes disponíveis publicamente. A IBM não testou estes produtos e não pode confirmar a precisão de seu desempenho, compatibilidade nem qualquer outra reivindicação relacionada a produtos não IBM. Dúvidas sobre os recursos de produtos não IBM devem ser encaminhadas diretamente a seus fornecedores.

Declarações relacionadas aos objetivos e intenções futuras da IBM estão sujeitas a alterações ou cancelamento sem aviso prévio e representam apenas metas e objetivos.

Estas informações contêm exemplos de dados e relatórios utilizados em operações comerciais diárias. Para ilustrá-los da forma mais completa possível, os exemplos podem incluir nomes de indivíduos, empresas, marcas e produtos. Todos esses nomes são fictícios e qualquer semelhança com pessoas ou empresas reais é mera coincidência.

Marcas Comerciais

IBM, o logotipo IBM e ibm.com são marcas ou marcas registradas da International Business Machines Corp., registradas em vários países no mundo todo. Outros nomes de empresas, produtos e serviços podem ser marcas registradas ou marcas da IBM ou de outras empresas. Uma lista atual de marcas registradas da IBM está disponível na web em "Copyright and trademark information" em www.ibm.com/legal/copytrade.shtml.

Adobe, o logotipo Adobe, PostScript e o logotipo PostScript são marcas ou marcas registradas da Adobe Systems Incorporated nos Estados Unidos e/ou em outros países.

Intel, logotipo Intel, Intel Inside, logotipo Intel Inside, Intel Centrino, logotipo Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium e Pentium são marcas comerciais ou marcas registradas da Intel Corporation ou de suas subsidiárias nos Estados Unidos e em outros países.

Linux é uma marca registrada de Linus Torvalds nos Estados Unidos e/ou em outros países.

Microsoft, Windows, Windows NT e o logotipo Windows são marcas registradas da Microsoft Corporation nos Estados Unidos e/ou em outros países.

UNIX é uma marca registrada da The Open Group nos Estados Unidos e em outros países.

Java e todas as marcas registradas e logotipos baseados em Java são marcas ou marcas registradas da Oracle e/ou de suas afiliadas.

Termos e condições da documentação do produto

As permissões para uso destas publicações são concedidas sujeitas aos seguintes termos e condições.

Aplicabilidade

Estes termos e condições estão em adição a quaisquer termos de uso para o website IBM.

Uso Pessoal

O Cliente pode reproduzir essas publicações para seu uso pessoal, não comercial, desde que todos os avisos do proprietário sejam preservados. O Cliente não deve distribuir, exibir ou criar trabalhos derivativos destas publicações ou de qualquer parte delas, sem o consentimento expresso da IBM.

Uso Comercial

É possível reproduzir, distribuir e exibir essas publicações unicamente dentro de sua empresa, contanto que todos os avisos do proprietário sejam preservados. O Cliente não poderá criar trabalhos derivativos destas publicações ou reproduzir, distribuir ou exibir estas publicações ou qualquer parte delas fora de sua empresa, sem o consentimento expresso da IBM.

Direitos

Exceto conforme expressamente concedido nesta permissão, nenhuma outra permissão, licença ou direitos serão concedidos, sejam expressos ou implícitos, para as publicações ou qualquer informação, dados, software ou outra propriedade intelectual contida nela.

A IBM reserva-se no direito de retirar as permissões aqui concedidas sempre que, a seu critério, o uso das publicações for prejudicial aos seus interesses ou, conforme determinado pela IBM, as instruções acima não estiverem sendo seguidas corretamente.

O Cliente não pode fazer download, exportar ou reexportar estas informações, exceto se elas estiverem totalmente em conformidade com todas as leis e regulamentações aplicáveis, incluindo todas as leis e regulamentações de exportação dos Estados Unidos.

A IBM NÃO FAZ QUALQUER TIPO DE GARANTIA QUANTO AO CONTEÚDO DESTAS PUBLICAÇÕES. AS PUBLICAÇÕES SÃO FORNECIDAS "NO ESTADO EM QUE SE ENCONTRAM" E SEM GARANTIA DE NENHUM TIPO, SEJA EXPRESSA OU IMPLÍCITA, INCLUINDO, MAS A ELAS NÃO SE LIMITANDO, GARANTIAS IMPLÍCITAS DE COMERCIALIZAÇÃO, NÃO INFRAÇÃO E ADEQUAÇÃO A UM DETERMINADO PROPÓSITO.

Índice Remissivo

A

- abrindo arquivos 28
- acessibilidade 95
- administração
 - atalhos do aplicativo 82
 - opções de dados 83
 - opções de destino de escoragem 83
 - opções gerais 82
 - página inicial 82
 - visão geral 81
- ajuda 91
- ajuda online 91
- alterar idioma 25
- alterar idioma do navegador 25
- anotações
 - incluindo em regras 55
- aplicativos
 - ativação 11
 - testando 70
- aplicativos criados previamente 10
- aplicativos de amostra 10
- arquivos de amostra 9
- arquivos de dados 10
- arquivos de dados baseados em texto 36
- arquivos de dados do Excel 36
- arquivos de dados do IBM SPSS Statistics 36
- arquivos para aplicativos de amostra 9, 10
- artefatos de demonstração 9
- artefatos para aplicativos de amostra 9
- atalhos
 - aplicativo 25
- atalhos do aplicativo 25, 82
- avaliação de modelos 67
- avaliando modelos 15
 - lista de modelos interativos 18
 - modelos da regra preditiva 18
 - modelos de armazenamento em cluster 19, 21

B

- bloqueando opções de destino de escoragem 83
- bloqueando projetos 30
- booleana 49

C

- campo de destino
 - construindo modelos 13, 57
- campos
 - derivando 47
 - em expressões 48
 - mapeando 45
- campos de entrada
 - definindo 42
 - mapeando 45
- campos de escoragem de mapa 78

- campos de mapa 78
- campos de saída em modelos 78
- cluster manual 61
- compartilhando fluxos entre IBM Analytical Decision Management e IBM SPSS Modeler 93
- construindo modelos
 - campo de destino 13
- construtor de expressões 45, 47, 49
- controlando
 - opções gerais 82
- controlando atalhos do aplicativo 82
- criando campos usando regras de segmento 45

D

- dados de amostra 10
- dados do Analytic Server 42
- dados do Hadoop 42
- dados do HDFS 42
- dados para aplicativos de amostra 10
- data da simulação 27
- derivar campos 47
- descrição 31
- destino de escore em modelos 75
- destino do Analytic Server nos modelos 77
- destino do arquivo em modelos 76
- destino do banco de dados nos modelos 75
- destino do servidor Cognos BI em modelos 77
- dicas de ferramenta 91
- documentação disponível 91

E

- editando campos adicionais 45
- editando campos derivados 45
- editando valores de campo 44
- editor de expressão 47
- exemplo de mala direta 9
- exemplo de modelagem de associação 21
- exemplo de modelagem do cluster 19
- exportando regras 54
- expressões que retornam booleano 49

F

- fazendo upload de arquivos de fluxo 28
- fazendo upload de fluxos 28
 - fluxos
 - fazendo upload 28
- formatos de data 27
- formatos de horário 27
- formatos de moeda 27
- funções
 - em expressões 47

G

- galeria
 - tarefas de longa execução 26
 - trabalho não concluído 26
- gráfico de distribuição 58
- gráfico de importância de variável 58
- gráfico de importância do preditor 58
- gráficos de coincidência 67
- gráficos de distribuição 67
- gráficos de distribuição de erro 67
- gráficos de ganho 67
- gráficos de ganhos 15, 67
- gráficos de lucro 15, 67
- gráficos de resposta 67
- gráficos em modelos 58
- gráficos ROI 67
- guia Derivado 45

I

- IBM SPSS Collaboration and Deployment Services Repository
 - salvando em 28, 29, 30
- IBM SPSS Modeler Advantage
 - exemplo 9
- ícones, IBM Cognos BI 41
- idioma 25
- incluindo campos 45
- incluindo campos adicionais 45
- incluindo campos usando regras de segmento 45
- incluindo expressões 45
- incluindo saídas de modelo 45
- incluir/editar campos adicionais 45
- incluir/editar campos derivados 45
- itens não concluídos
 - galeria 26

L

- lucro
 - maximizando 9
 - simulando 15

M

- mapeando campos 45
- máximo de membros para campos nominais 31
- minhas origens de dados 33
- modelo de dados do projeto 33
- modelos
 - armazenamento em cluster 61
 - cluster manual 61
 - associação 62
 - avaliando 15, 67
 - campo de destino 13, 57
 - campos de escoragem de mapa 78
 - configurações opcionais 63, 65

- modelos (*continuação*)
 - configurando preferências 27
 - construção com tempo restrito 27
 - construindo 57
 - edifício 13
 - editor de tarefas 19
 - escoragem 73, 79, 80
 - gráficos 58
 - lista interativa 59
 - preditivos 57
 - regra preditiva 59
 - relatório sumarização de
 - escoragem 80
 - resultados 58
 - seleção de campo de saída 78
 - seleção de dados 75
 - seleção de destino de escore 75
 - seleção de destino do Analytic Server 77
 - seleção de destino do arquivo 76
 - seleção de destino do banco de dados 75
 - seleção de destino do servidor Cognos BI 77
 - seleção de registro 75
 - simulando lucro 69
 - testando 19, 70
 - usando em regras 51
- modelos de associação 62, 65
 - dados tabulares 63
 - dados tabulares versus transacionais 63
 - dados transacionais 63
- modelos de clusterização 61, 65
- modelos de contagem 19
- modelos de escoragem 73, 79
 - campos de mapa 78
 - Destino do Analytic Server 77
 - destino do arquivo 76
 - destino do banco de dados 75
 - Destino do servidor Cognos BI 77
 - selecionando campos de saída 78
 - selecionando dados 75
 - selecionando destino de escoragem 75
- modelos de lista interativa 59, 65
- modelos de regra preditiva 59, 65
- modelos preditivos 57

N

- níveis de medição 43
- nó de origem do IBM Cognos BI
 - ícones 41
- novas variáveis 1, 2, 3

O

- o que há de novo 1, 2, 3
- obtendo ajuda 91
 - documentação disponível 91
- ocultando a página inicial 82
- opções de bloqueio de dados 83
- opções de dados 83
- opções de destino de escoragem 83
- opções gerais 82

- operadores
 - em expressões 47
- Operational Decision Management
 - regras 55
- origens de banco de dados 38
- origens de dados
 - Analytic Server 42
 - banco de dados 38
 - baseado em texto 36
 - campos de entrada 42
 - construtor de expressões 47, 49
 - criando campos usando regras de segmento 45
 - definindo 33, 35
 - derivando campos 45
 - editando valores de campo 44
 - Excel 36
 - guia Derivado 45
 - Hadoop 42
 - HDFS 42
 - IBM Cognos BI 40
 - IBM SPSS Statistics 36
 - incluindo campos 45
 - incluindo expressões 45
 - incluindo saídas de modelo 45
 - mapeando campos 45
 - níveis de medição 43
 - origens do arquivo 36
 - visão geral de dados 43
 - visualização de dados 39
 - visualizando 43
- origens de dados do projeto 33
- origens do Cognos BI 40
- origens do IBM Cognos BI 40

P

- página de ativação 11, 25, 82
 - customizando 25
 - incluindo aplicativos 25
- página de ativação de aplicativo 25, 82
 - customizando 25
 - incluindo aplicativos 25
- página inicial 26, 82
- página inicial do aplicativo 26
- palavras-chaves 28, 31
- parâmetros
 - em expressões 48
- preferências 27
- preferências do usuário 27
- propriedades
 - descrição 31
 - máximo de membros para campos nominais 31
 - palavras-chaves 31
 - projeto 31
 - repositório 31

R

- regras 51
 - anotações 55
 - compartilhada 54, 55
 - criando 51
 - exportando 54
 - externas 55

- regras (*continuação*)
 - IBM Operational Decision Management 55
 - lógica de seleção 53
 - reutilizando 54, 55
 - segmento 51
 - seleção 53
 - usando modelos com 51
- regras compartilhadas 54
- regras de exclusão 53
- regras de inclusão 53
- regras de negócios 51
- regras de segmento
 - usando modelos com 51
- regras de seleção 53
- regras do IBM Operational Decision Management
 - usando em aplicativos 55
- regras externas
 - usando em aplicativos 55
- relatório sumarização de escoragem 80
- rotulando versões 30
- rótulos de versão 30

S

- salvando arquivos 28, 29, 30
- seleção de dados em modelos 75
- seleção de registro nos modelos 75
- simulação
 - data padrão 27
- simulação de lucro em modelos 69

T

- testando modelos 19
- testando um aplicativo 70
- testando um modelo 70
- texto de treinamento 91
- tipos de dados 43

U

- URL
 - aplicativo 25
- URL do aplicativo 25

V

- valores globais
 - em expressões 48
- verificando expressões 48
- visão geral
 - administração 81
- visão geral de administração 81
- visão geral de dados 43
- visualização de dados 39
- visualizando dados 43

