

*Guia do Usuário do IBM SPSS Modeler
17.1*

IBM

Observação

Antes de usar estas informações e o produto por elas suportadas, leia as informações em “Avisos” na página 213.

Informações do produto

Essa edição se aplica à versão 17, liberação 1, modificação 0 do IBM(r) SPSS(r) Modeler e a todas as liberações e modificações subsequentes até que seja indicado de outra forma em novas edições.

Índice

Capítulo 1. Sobre o IBM SPSS Modeler 1

Produtos IBM SPSS Modeler	1
IBM SPSS Modeler	1
IBM SPSS Modeler Server	1
IBM SPSS Modeler Administration Console	2
IBM SPSS Modeler Batch	2
IBM SPSS Modeler Solution Publisher	2
IBM SPSS Modeler Server Adapters for IBM SPSS Collaboration and Deployment Services	2
Edições do IBM SPSS Modeler	2
Documentação do IBM SPSS Modeler	3
Documentação do SPSS Modeler Professional	3
Documentação do SPSS Modeler Premium	4
Exemplos de Aplicativos	5
Pasta Demos	5

Capítulo 2. Novas variáveis no IBM SPSS Modeler 17.1 7

Capítulo 3. Visão geral do IBM SPSS Modeler 9

Introdução	9
Iniciando o IBM SPSS Modeler	9
Ativando a partir da linha de comandos	9
Conectando a IBM SPSS Modeler Server.	10
Alterando o diretório temporário	12
Iniciando várias sessões do IBM SPSS Modeler	12
Visão rápida da interface do IBM SPSS Modeler	13
Tela de fluxo do IBM SPSS Modeler	13
Paleta de Nós	14
Gerenciadores do IBM SPSS Modeler	15
Projetos do IBM SPSS Modeler	17
Barra de ferramentas do IBM SPSS Modeler	17
Customizando a barra de ferramentas	18
Customizando a janela do IBM SPSS Modeler	19
Alterando o tamanho do ícone para um fluxo	20
Usando o mouse no IBM SPSS Modeler	20
Usando teclas de atalho	20
Imprimindo	21
Automatizando IBM SPSS Modeler	22

Capítulo 4. Entendendo a mineração de dados 23

Visão geral de mineração de dados	23
Avaliando dados	24
Uma estratégia para mineração de dados	25
O modelo de processo CRISP-DM	26
Tipos de modelos	27
Exemplos de mineração de dados	32

Capítulo 5. Construindo fluxos 33

Visão geral da construção de fluxo	33
Construindo fluxos de dados	33
Trabalhando com nós	33

Trabalhando com fluxos	39
Descrições de fluxo	52
Executando fluxos	53
Trabalhando com modelos	54
Incluindo comentários e anotações em nós e fluxos	54
Salvando fluxos de dados	59
Carregando arquivos	61
Mapeando fluxos de dados	62
Dicas e atalhos	64

Capítulo 6. Trabalhando com saída 67

Visualizador	67
Mostrando e ocultando resultados	67
Movendo, excluindo e copiando saída	68
Alterando o alinhamento inicial	68
Alterando o alinhamento dos itens de saída	68
Estrutura de tópicos do Visualizador	68
Incluindo itens no Visualizador	70
Localizando e substituindo informações no Visualizador	70
Copiando saída em outros aplicativos	71
Saída interativa	72
Exportar saída	73
Opções de HTML	74
Opções de relatório da web	75
Opções do Word/RTF	75
Opções do Excel	76
Opções do PowerPoint	77
Opções de PDF	77
Opções de texto	78
Opções somente de gráficos	79
Opções de formato gráfico	79
Impressão do Visualizador	80
Para imprimir saída e gráficos	80
Visualização prévia da impressão	80
Atributos de página: cabeçalhos e rodapés	81
Atributos de Página: Opções.	81
Salvando saída	82
Para salvar um documento do Visualizador	82
Tabelas dinâmicas	83
Tabelas dinâmicas	83
Manipulando uma tabela dinâmica	83
Trabalhando com camadas	86
Mostrando e ocultando itens.	86
TableLooks	87
Propriedades de tabela	88
Propriedades de célula	90
Notas de rodapé e legendas	91
Larguras de células de dados	92
Alterando a largura da coluna	93
Exibindo bordas ocultas em uma tabela dinâmica	93
Selecionando linhas, colunas e células em uma tabela dinâmica	93
Imprimindo tabelas dinâmicas	93

Criando um gráfico a partir de uma tabela dinâmica	94
Tabelas anteriores	94
Opções	94
Opções	94
Opções gerais	95
Opções do Visualizador	95
Opções de tabela dinâmica	95
Opções de saída	96

**Capítulo 7. Manipulando valores
omissos 97**

Visão geral de valores omissos	97
Manipulando valores omissos	98
Manipulando registros com valores omissos	98
Manipulando campos com valores omissos	98
Tratamento de registros com valores omissos do sistema	99
Imputando ou preenchendo valores omissos	101
Funções do CLEM para valores omissos	102

**Capítulo 8. Construindo expressões
do CLEM 103**

Sobre o CLEM	103
CLEMExemplos	103
Valores e tipos de dados	105
Expressões e condições	105
Parâmetros Stream, Session e SuperNode	106
Trabalhando com sequências de caracteres	107
Manipulando valores omissos e em branco	108
Trabalhando com números	108
Trabalhando com datas e horas	108
Resumindo diversos campos	109
Trabalhando com dados de múltiplas respostas	110
O construtor de expressões	111
Acessando o construtor de expressões	111
Criando expressões	111
Selecionando funções	111
Selecionando campos, parâmetros e variáveis globais	114
Visualizando ou selecionando valores	115
Verificando expressões do CLEM	115
Localizar e substituir	116

**Capítulo 9. Referência de linguagem
do CLEM 119**

Visão geral de referência do CLEM	119
Tipos de dados do CLEM	119
Números Inteiros	119
Reais	120
Caracteres	120
Sequência de caracteres	120
Listas	120
Campos	121
Datas	121
Hora	122
Operadores do CLEM	122
Referência de função	124
Convenções em descrições de função	125
Funções de informações	126

Funções de conversão	126
Funções de comparação	128
Funções lógicas	130
Funções numéricas	130
Funções trigonométricas	131
Funções de probabilidade	132
Funções espaciais	132
Operações de número inteiro bit a bit	133
Funções aleatórias	134
Funções de sequência de caracteres	135
Funções SoundEx	140
Funções de data e hora	140
Funções de sequência	145
Funções globais	150
Funções manipulando valores em branco e nulos	150
Campos Especiais	151

**Capítulo 10. Usando IBM SPSS
Modeler com um repositório 153**

Sobre o IBM SPSS Collaboration and Deployment Services Repository	153
Armazenando e implementando objetos do repositório	154
Conectando ao repositório	155
Inserindo credenciais para o repositório	155
Navegação pelas credenciais do repositório	155
Navegando no conteúdo do repositório	155
Armazenando objetos no repositório	156
Configurando propriedades do objeto	156
Armazenando fluxos	158
Armazenando projetos	158
Armazenando nós	159
Armazenando objetos de saída	159
Armazenando modelos e paletas de modelos	159
Recuperando objetos do repositório	160
Escolhendo um objeto para recuperar	160
Selecionando versão de um objeto	161
Procurando objetos no repositório	161
Modificando objetos do repositório	162
Criando, renomeando e excluindo pastas	162
Bloqueando e desbloqueando objetos do repositório	163
Excluindo objetos do repositório	163
Gerenciando propriedades dos objetos do repositório	163
Visualizando propriedades da pasta	163
Visualizando e editando propriedades do objeto	164
Gerenciando rótulos de versão do objeto	165
Implementando fluxos	166
Opções de Implementação do Fluxo	166
A ramificação de escoragem	168

**Capítulo 11. Exportando para
aplicativos externos 173**

Sobre Exportando para aplicativos externos	173
Abrindo um fluxo no IBM SPSS Modeler Advantage	173
Importando e exportando modelos como PMML	174
Tipos de modelo que suportam PMML	175

Capítulo 12. Projetos e relatórios . . .	177
Introdução aos Projetos	177
Visualização CRISP-DM	177
Visualização Classes	178
Construindo um projeto	178
Criando um novo projeto	178
Incluindo em um projeto	178
Transferindo projetos para o IBM SPSS Collaboration and Deployment Services Repository	179
Configurando propriedades do projeto	180
Anotando um projeto.	180
Propriedades do Objeto	181
Encerrando um projeto	181
Gerando um relatório	181
Salvando e exportando relatórios gerados	183
Capítulo 13. Customizando o IBM SPSS Modeler	185
Customizando opções do IBM SPSS Modeler	185
Configurando opções do IBM SPSS Modeler	185
Opções de Sistema	185
Configurando diretórios padrão	186
Configurando opções do usuário	186
Customizando a Paleta de Nós	190
Customizando o Gerenciador de Paleta.	190
Alterando uma visualização da guia paleta	192
Gerenciamento do nó CEMI	192
Capítulo 14. Considerações de desempenho para fluxos e nós.	193
Ordem dos nós	193

Caches de nó	194
Desempenho: Nós de Processamento	195
Desempenho: Nós de Modelagem	196
Desempenho: expressões do CLEM	197

Capítulo 15. Acessibilidade no IBM SPSS Modeler	199
Visão geral de acessibilidade no IBM SPSS Modeler	199
Tipos de suporte à acessibilidade.	199
Acessibilidade para deficiência visual	199
Acessibilidade para usuários cegos	200
Teclado de acessibilidade	200
Usando um leitor de tela	207
Dicas de uso	208
Interferência de outro software	209
JAWS e Java.	209
Usando gráficos no IBM SPSS Modeler.	209

Capítulo 16. Suporte a Unicode.	211
Suporte Unicode no IBM SPSS Modeler	211

Avisos	213
Marcas comerciais	214

Índice Remissivo	217
-----------------------------------	------------

Capítulo 1. Sobre o IBM SPSS Modeler

O IBM® SPSS Modeler é um conjunto de ferramentas de mineração de dados que permite desenvolver rapidamente modelos preditivos usando o conhecimento de negócios, e implementá-los em operações de negócios para melhorar a tomada de decisão. Projetado em torno do modelo CRISP-DM padrão de mercado, o IBM SPSS Modeler suporta todo o processo de mineração de dados, a partir dos dados para melhores resultados de negócios.

O IBM SPSS Modeler oferece uma variedade de métodos de modelagem a partir do aprendizado de máquina, inteligência artificial e estatísticas. Os métodos disponíveis na paleta Modelagem permitem derivar informações novas a partir dos dados, e desenvolver modelos preditivos. Cada método possui determinadas intensidades e é mais bem adequado para tipos de problemas específicos.

O SPSS Modeler pode ser comprado como um produto independente, ou usado como um cliente na combinação com o SPSS Modeler Server. Várias opções adicionais também estão disponíveis, conforme resumidas nas seções a seguir. Para obter mais informações, consulte <http://www.ibm.com/software/analytics/spss/products/modeler/>.

Produtos IBM SPSS Modeler

A família de produtos IBM SPSS Modeler e o software associado abrangem o seguinte.

- IBM SPSS Modeler
- IBM SPSS Modeler Server
- IBM SPSS Modeler Administration Console
- IBM SPSS Modeler Batch
- IBM SPSS Modeler Solution Publisher
- Adaptadores IBM SPSS Modeler Server para IBM SPSS Collaboration and Deployment Services

IBM SPSS Modeler

SPSS Modeler é uma versão funcionalmente completa do produto que você instala e executa em seu computador pessoal. É possível executar o SPSS Modeler no modo local como um produto independente ou usá-lo no modo distribuído com IBM SPSS Modeler Server para melhorar o desempenho em conjuntos de dados grandes.

Com o SPSS Modeler, é possível construir modelos preditivos exatos de maneira rápida e intuitiva, sem programação. Usando a interface visual exclusiva, é possível visualizar facilmente o processo de mineração de dados. Com o suporte da análise avançada integrada ao produto, é possível descobrir tendências e padrões ocultos anteriormente em seus dados. É possível modelar resultados e entender os fatores que os influenciam, permitindo que você aproveite as vantagens das oportunidades de negócios e diminua os riscos.

SPSS Modeler está disponível em duas edições: SPSS Modeler Professional e SPSS Modeler Premium. Consulte o tópico “Edições do IBM SPSS Modeler” na página 2 para obter mais informações.

IBM SPSS Modeler Server

SPSS Modeler usa uma arquitetura de cliente/servidor para distribuir solicitações para operações cheias de recursos para poderosos softwares de servidor, resultando em desempenho mais rápido em conjuntos de dados maiores.

SPSS Modeler Server é um produto licenciado separadamente que é executado de forma contínua no modo de análise distribuído em um host do servidor com uma ou mais instalações do IBM SPSS Modeler. Dessa maneira, o SPSS Modeler Server fornece desempenho superior em conjuntos de dados grandes, pois operações com uso intensivo de memória podem ser executadas no servidor sem fazer download dos dados no computador cliente. IBM SPSS Modeler Server também fornece suporte para otimização de SQL e capacidades de modelagem dentro da base de dados, entregando mais benefícios para o desempenho e a automação.

IBM SPSS Modeler Administration Console

O Modeler Administration Console é um aplicativo gráfico para gerenciar muitas das opções de configuração do SPSS Modeler Server, que também são configuráveis por meio de um arquivo de opções. O aplicativo fornece uma interface com o usuário do console para monitorar e configurar suas instalações do SPSS Modeler Server e está disponível gratuitamente para os atuais clientes do SPSS Modeler Server. O aplicativo pode ser instalado somente em computadores Windows; no entanto, ele pode administrar um servidor instalado em qualquer plataforma suportada.

IBM SPSS Modeler Batch

Embora geralmente a mineração de dados seja um processo interativo, também é possível executar o SPSS Modeler a partir de uma linha de comandos, sem a necessidade de uma interface gráfica com o usuário. Por exemplo, você pode ter tarefas repetidas ou de longa execução que deseja executar sem intervenção do usuário. SPSS Modeler Batch é uma versão especial do produto que fornece suporte para capacidades de análise completa do SPSS Modeler sem acessar a interface com o usuário regular. SPSS Modeler Server é necessário para usar o SPSS Modeler Batch.

IBM SPSS Modeler Solution Publisher

SPSS Modeler Solution Publisher é uma ferramenta que permite criar uma versão do pacote de um fluxo do SPSS Modeler que pode ser executado por um mecanismo de tempo de execução externo ou integrado a um aplicativo externo. Dessa maneira, é possível publicar e implementar fluxos completos do SPSS Modeler para uso em ambientes que não têm o SPSS Modeler instalado. SPSS Modeler Solution Publisher é distribuído como parte do serviço IBM SPSS Collaboration and Deployment Services - Scoring, para o qual uma licença separada é necessária. Com essa licença, você recebe o SPSS Modeler Solution Publisher Runtime, que permite executar os fluxos publicados.

Para obter mais informações sobre SPSS Modeler Solution Publisher, consulte a documentação do IBM SPSS Collaboration and Deployment Services. O IBM SPSS Collaboration and Deployment Services Knowledge Center contém seções chamadas "IBM SPSS Modeler Solution Publisher" e "IBM SPSS Analytics Toolkit."

IBM SPSS Modeler Server Adapters for IBM SPSS Collaboration and Deployment Services

Inúmeros adaptadores para o IBM SPSS Collaboration and Deployment Services estão disponíveis para permitir que o SPSS Modeler e o SPSS Modeler Server interajam com um repositório do IBM SPSS Collaboration and Deployment Services. Dessa forma, um fluxo do SPSS Modeler implementado no repositório pode ser compartilhado por diversos usuários ou acessado a partir do aplicativo thin client IBM SPSS Modeler Advantage. Você instala o adaptador no sistema que hospeda o repositório.

Edições do IBM SPSS Modeler

SPSS Modeler está disponível nas seguintes edições.

SPSS Modeler Professional

SPSS Modeler Professional fornece todas as ferramentas necessárias para você trabalhar com a maioria dos tipos de dados estruturados, como comportamentos e interações controlados em sistemas CRM,

demográficos, comportamento de compra e dados de vendas.

SPSS Modeler Premium

SPSS Modeler Premium é um produto licenciado separadamente que estende o SPSS Modeler Professional para trabalhar com dados especificados, como aqueles usados para análise de entidade ou rede social, e com dados de texto não estruturados. SPSS Modeler Premium abrange os componentes a seguir.

O **IBM SPSS Modeler Entity Analytics** inclui uma dimensão extra à análise preditiva do IBM SPSS Modeler. Considerando que a análise preditiva tenta prever comportamento futuro de dados passados, a análise de entidade foca melhorar a coerência e a consistência dos dados atuais resolvendo conflitos de identidade nos registros em si. Uma identidade pode ser de um indivíduo, de uma organização, de um objeto ou de qualquer outra entidade para a qual possa existir ambiguidade. A resolução de identidade pode ser vital em diversos campos, incluindo gerenciamento de relacionamento com o cliente, detecção de fraude, medidas contra lavagem de dinheiro e segurança nacional e internacional.

IBM SPSS Modeler Social Network Analysis transforma informações sobre relacionamentos em campos que caracterizam o comportamento social dos indivíduos e grupos. Usando dados que descrevem os relacionamentos subjacentes das redes sociais, o IBM SPSS Modeler Social Network Analysis identifica os líderes sociais que influenciam o comportamento dos outros na rede. Além disso, é possível determinar quais pessoas são mais afetadas por outros participantes da rede. Combinando esses resultados com outras medidas, é possível criar perfis abrangentes de indivíduos nos quais basear modelos preditivos. Os modelos que incluem essas informações sociais serão executados melhor do que os modelos que não incluem.

IBM SPSS Modeler Text Analytics usa tecnologias de linguística avançada e processamento de linguagem natural (NLP) para processar rapidamente uma grande variedade de dados de texto não estruturados, extrair e organizar conceitos chave e agrupar esses conceitos em categorias. Categorias e conceitos extraídos podem ser combinados com dados estruturados existentes, como demográficos, e aplicados à modelagem usando o conjunto completo de ferramentas de mineração de dados do IBM SPSS Modeler para gerar decisões melhores e mais focadas.

Documentação do IBM SPSS Modeler

A documentação em formato de ajuda online está disponível no menu Ajuda do SPSS Modeler. Isso inclui documentação para SPSS Modeler, SPSS Modeler Server, bem como o Guia de Aplicativos (também chamado de Tutorial) e outros materiais de apoio.

A documentação completa para cada produto (incluindo instruções de instalação) está disponível em formato PDF sob a pasta *Documentation* em cada produto DVD. Os documentos de instalação também podem ser transferidos por download a partir da web em <http://www.ibm.com/support/docview.wss?uid=swg27046116>.

A documentação em ambos os formatos também está disponível no SPSS Modeler Knowledge Center em http://www-01.ibm.com/support/knowledgecenter/SS3RA7_17.1.0.

Documentação do SPSS Modeler Professional

O conjunto de documentações do SPSS Modeler Professional (excluindo instruções de instalação) é o seguinte.

- **IBM SPSS Modeler User's Guide.** Introdução geral ao uso do SPSS Modeler, incluindo como construir fluxos de dados, manipular valores omissos, construir expressões do CLEM, trabalhar com projetos e relatórios e empacotar fluxos para implementação no IBM SPSS Collaboration and Deployment Services, Predictive Applications ou IBM SPSS Modeler Advantage.

- **IBM SPSS Modeler Source, Process, and Output Nodes.** Descrições de todos os nós usados para ler, processar e emitir dados em diferentes formatos. Efetivamente, isso significa todos os nós além dos de modelagem.
- **IBM SPSS Modeler Modeling Nodes.** Descrições de todos os nós usados para criar modelos de mineração de dados. O IBM SPSS Modeler oferece uma variedade de métodos de modelagem a partir do aprendizado de máquina, inteligência artificial e estatísticas.
- **IBM SPSS Modeler Algorithms Guide.** Descrições das bases matemáticas dos métodos de modelagem usados no IBM SPSS Modeler. Este guia está disponível somente em formato PDF.
- **IBM SPSS Modeler Applications Guide.** Os exemplos neste guia fornecem introduções sintetizadas e direcionadas para técnicas e métodos de modelagem específicos. Uma versão online deste guia também está disponível no menu Ajuda. Consulte o tópico “Exemplos de Aplicativos” na página 5 para obter mais informações.
- **IBM SPSS Modeler Python Scripting and Automation.** Informações sobre como automatizar o sistema por meio de script Python, incluindo as propriedades que podem ser usadas para manipular nós e fluxos.
- **IBM SPSS Modeler Deployment Guide.** Informações sobre como executar fluxos e cenários do IBM SPSS Modeler como passos no processamento de tarefas sob IBM SPSS Collaboration and Deployment Services Deployment Manager.
- **IBM SPSS Modeler CLEF Developer's Guide.** CLEF fornece a capacidade de integrar programas de terceiros, como rotinas de processamento de dados ou algoritmos de modelagem como nós, no IBM SPSS Modeler.
- **IBM SPSS Modeler In-Database Mining Guide.** Informações sobre como usar o poder do seu banco de dados para melhorar o desempenho e ampliar o intervalo de capacidades analíticas por meio de algoritmos de terceiros.
- **IBM SPSS Modeler Server Administration and Performance Guide.** Informações sobre como configurar e administrar o IBM SPSS Modeler Server.
- **IBM SPSS Modeler Administration Console User Guide.** Informações sobre a instalação e o uso da interface com o usuário do console para monitorar e configurar o IBM SPSS Modeler Server. O console é implementado como um plug-in no aplicativo Deployment Manager.
- **Guia do IBM SPSS Modeler CRISP-DM.** Guia passo a passo para o uso da metodologia CRISP-DM para mineração de dados com SPSS Modeler.
- **IBM SPSS Modeler Batch User's Guide.** Guia completo para o uso do IBM SPSS Modeler no modo em lote, incluindo detalhes da execução do modo em lote e argumentos de linha de comandos. Este guia está disponível somente em formato PDF.

Documentação do SPSS Modeler Premium

O conjunto de documentações do SPSS Modeler Premium (excluindo instruções de instalação) é o seguinte.

- **IBM SPSS Modeler Entity Analytics User Guide.** Informações sobre o uso de análise de entidade com SPSS Modeler, cobrindo instalação e configuração de repositório, nós de análise de entidade e tarefas administrativas.
- **IBM SPSS Modeler Social Network Analysis User Guide.** Um guia para executar análise de rede social com SPSS Modeler, incluindo análise de grupo e análise de difusão.
- **SPSS Modeler Text Analytics User's Guide.** Informações sobre o uso de analítica de texto com SPSS Modeler, cobrindo os nós de mineração de texto, ambiente de trabalho interativo, modelos e outros recursos.

Exemplos de Aplicativos

Enquanto as ferramentas de mineração de dados no SPSS Modeler podem ajudar a resolver uma ampla variedade de negócios e problemas organizacionais, os exemplos de aplicativos fornecem introduções breves e destinadas aos métodos e técnicas de modelagem específicos. Os conjuntos de dados usados aqui são muito menores do que os enormes armazenamentos de dados gerenciados por alguns mineradores de dados, mas os conceitos e métodos envolvidos devem ser escaláveis para aplicativos reais.

É possível acessar os exemplos clicando em **Exemplos de Aplicativos** no menu de Ajuda no SPSS Modeler. Os arquivos de dados e os fluxos de amostra são instalados na pasta *Demos* no diretório de instalação do produto. Consulte o tópico “Pasta Demos” para obter mais informações.

Exemplos de modelagem da base de dados. Consulte os exemplos no *Guia de Mineração dentro do Banco de Dados do IBM SPSS Modeler*.

Exemplos de script. Consulte os exemplos no *Guia de Script e Automação do IBM SPSS Modeler*.

Pasta Demos

Os arquivos de dados e fluxos de amostras usados com exemplos de aplicativo são instalados na pasta *Demos* sob o diretório de instalação do produto. Essa pasta também pode ser acessada a partir do grupo do programa IBM SPSS Modeler no menu Iniciar do Windows ou clicando em *Demos* na lista de diretórios recentes na caixa de diálogo Arquivo Aberto.

Capítulo 2. Novas variáveis no IBM SPSS Modeler 17.1

O IBM SPSS Modeler inclui as variáveis a seguir nesta liberação.

Modelando

Nó Linear do AS. Um novo nó, Linear do AS, está disponível. Esse novo nó é semelhante ao nó Linear existente, mas é executado quando conectado ao IBM SPSS Analytic Server para aumentar seu poder de processamento.

Nó LSVM. Um novo nó, LSVM, está disponível. Esse novo nó é semelhante ao nó SVM existente, mas é linear e melhor na manipulação de um número grande de registros. Ele é executado apenas quando conectado ao IBM SPSS Analytic Server para aumentar a sua energia de processamento.

Nó Árvore do AS. Um novo nó de árvore de decisão Árvore do AS está disponível. Esse novo nó é semelhante ao nó CHAID existente, mas é executado quando conectado ao IBM SPSS Analytic Server para aumentar seu poder de processamento.

Nó de Árvores Aleatórias. Um novo nó de Árvores Aleatórias está disponível. Esse novo nó é semelhante ao nó CHAID existente, mas é executado quando conectado ao IBM SPSS Analytic Server para aumentar seu poder de processamento.

Nó GLE. Um novo nó, GLE, está disponível. Esse novo nó é semelhante aos nós GenLin e GLMM existentes, mas é executado quando conectado ao IBM SPSS Analytic Server para aumentar seu poder de processamento.

Nós importação e exportação do TM1. Novas versões desses dois nós foram criadas para melhorar a velocidade da conexão entre o SPSS Modeler e o servidor TM1. Se você tiver quaisquer fluxos existentes que usam uma das versões anteriores desses nós, os fluxos ainda funcionarão, já que as versões mais antigas ainda são suportadas.

Melhorias de usabilidade do IBM SPSS Analytic Server. No cliente do SPSS Modeler, há uma nova opção Servidor Analítico para filtrar a Paleta de Nós para exibir apenas nós que sejam executados no Analytic Server.

Suporte de Python Spark para o Construtor de Diálogo Personalizado. Anteriormente, disponível para uso apenas com R, o Construtor de Diálogo Customizado agora também suporta o Python for Spark, que fornece acesso ao Apache Spark e a sua biblioteca de aprendizado de máquina (mllib). Por exemplo, um analista pode usar o Construtor de Diálogo Personalizado para criar um nó de modelagem que explora um algoritmo a partir do mllib e, em seguida, compartilhar esse nó com os outros. Observe as limitações a seguir:

- Não há nós independentes para a criação de script Python como os nós do R.
- Nós de Python criados a partir do Construtor de Diálogo Customizado dependem do ambiente do Spark e serão executados apenas com relação ao IBM SPSS Analytic Server (os dados são gerados a partir de um nó de origem de IBM SPSS Analytic Server e não são extraídos para o IBM SPSS Modeler Server Para obter mais informações sobre o IBM SPSS Analytic Server, consulte o IBM Knowledge Center.
- Os scripts Python devem usar a API Spark, porque os dados serão apresentados nro formato de um Spark DataFrame.

Suporte do PowerLinux. Red Hat Enterprise Linux (RHEL) ou Canonical Ubuntu em sistemas IBM POWER são agora ambientes suportados. Para obter mais informações sobre os requisitos do sistema,

consulte <http://www-01.ibm.com/software/analytics/spss/products/modeler/requirements.html>.

Capítulo 3. Visão geral do IBM SPSS Modeler

Introdução

Como um aplicativo de mineração de dados, o IBM SPSS Modeler oferece uma abordagem estratégica para localizar relacionamentos úteis em grandes conjuntos de dados. Em contraste com métodos estatísticos mais tradicionais, você não precisará necessariamente saber o que está procurando ao iniciar. Você pode explorar os dados, ajustando diferentes modelos e investigando diferentes relacionamentos até localizar informações úteis.

Iniciando o IBM SPSS Modeler

Para iniciar o aplicativo, clique em:

Iniciar > [Todos] Programas > IBM SPSS Modeler 17.1 > IBM SPSS Modeler 17.1

A janela principal é exibida após alguns segundos.

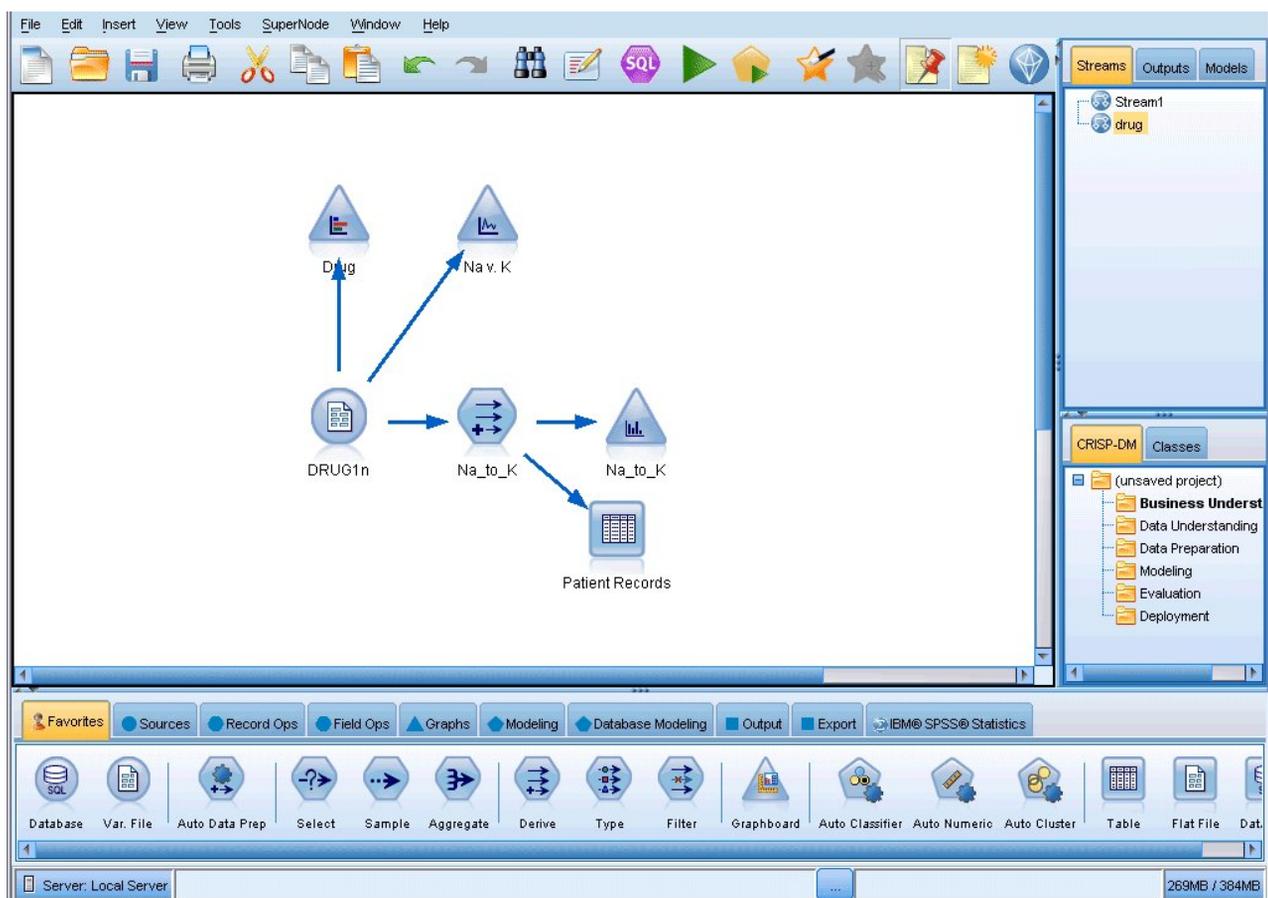


Figura 1. Janela do aplicativo principal do IBM SPSS Modeler

Ativando a partir da linha de comandos

É possível usar a linha de comandos do seu sistema operacional para ativar o IBM SPSS Modeler da seguinte forma:

1. Em um computador no qual o IBM SPSS Modeler está instalado, abra um DOS, um prompt de comandos ou uma janela.
2. Para ativar a interface do IBM SPSS Modeler no modo interativo, digite o comando `modelerclient` seguido pelos argumentos necessários; por exemplo:

```
modelerclient -stream report.str -execute
```

Os argumentos (flags) disponíveis permitem se conectar a um servidor, carregar fluxos, executar scripts ou especificar outros parâmetros conforme necessário.

Conectando a IBM SPSS Modeler Server

IBM SPSS Modeler pode ser executado como um aplicativo independente, ou como um cliente conectado diretamente ao IBM SPSS Modeler Server, ou a um IBM SPSS Modeler Server ou cluster de servidores por meio do plug-in Coordinator of Processes do IBM SPSS Collaboration and Deployment Services. O status da conexão atual é exibido na parte inferior esquerda da janela do IBM SPSS Modeler.

Sempre que você desejar se conectar a um servidor, é possível inserir manualmente o nome do servidor ao qual deseja se conectar ou selecionar um nome que você definiu anteriormente. No entanto, se você tiver o IBM SPSS Collaboration and Deployment Services, é possível procurar em uma lista de servidores ou clusters de servidores na caixa de diálogo Login do Servidor. A capacidade de navegar pelos serviços do Statistics em execução em uma rede é disponibilizada por meio do Coordinator of Processes.

Para se conectar a um servidor

1. No menu Ferramentas, clique em **Login do Servidor**. A caixa de diálogo Login do Servidor é aberta. Alternativamente, dê um clique duplo na área de status da conexão da janela IBM SPSS Modeler.
2. Usando a caixa de diálogo, especifique opções para se conectar ao computador servidor local ou selecione uma conexão da tabela.
 - Clique em **Incluir** ou **Editar** para incluir ou editar uma conexão. Consulte o tópico “Incluindo e editando a conexão do IBM SPSS Modeler Server” na página 11 para obter mais informações.
 - Clique em **Procurar** para acessar um servidor ou cluster de servidores no Coordinator of Processes. Consulte o tópico “Procurando servidores no IBM SPSS Collaboration and Deployment Services” na página 11 para obter mais informações.

Tabela de servidores. Essa tabela contém o conjunto de conexões de servidor definidas. A tabela exibe a conexão padrão, o nome do servidor, a descrição e o número da porta. É possível incluir manualmente uma nova conexão, bem como selecionar ou procurar uma conexão existente. Para configurar um determinado servidor como conexão padrão, marque a caixa de seleção na coluna Padrão na tabela para a conexão.

Caminho de dados padrão. Especifique um caminho usado para dados no computador servidor. Clique no botão de reticências (...) para navegar para o local necessário.

Configurar Credenciais. Deixe essa caixa desmarcada para ativar a variável **conexão única**, que tenta efetuar seu login no servidor usando detalhes de nome de usuário e senha do computador local. Se uma conexão única não for possível ou se você marcar essa caixa para desativar a conexão única (por exemplo, para efetuar login em uma conta do administrador), os campos a seguir serão ativados para você inserir suas credenciais.

ID do Usuário. Insira o nome de usuário com o qual deseja efetuar logon no servidor.

Senha. Insira a senha associada ao nome de usuário especificado.

Domínio. Especifique o domínio usado para efetuar logon no servidor. Um nome de domínio só é necessário quando o computador servidor está em um domínio do Windows diferente daquele do computador cliente.

3. Clique em **OK** para concluir a conexão.

Para se desconectar de um servidor

1. No menu Ferramentas, clique em **Login do Servidor**. A caixa de diálogo Login do Servidor é aberta. Alternativamente, dê um clique duplo na área de status da conexão da janela IBM SPSS Modeler.
2. Na caixa de diálogo, selecione o Servidor Local e clique em **OK**.

Incluindo e editando a conexão do IBM SPSS Modeler Server

É possível editar ou incluir manualmente uma conexão do servidor na caixa de diálogo Login do Servidor. Clicando em Incluir, é possível acessar uma caixa de diálogo Incluir/Editar Servidor vazia na qual é possível inserir detalhes da conexão do servidor. Com a seleção de uma conexão existente e um clique em Editar na caixa de diálogo Login do Servidor, a caixa de diálogo Incluir/Editar é aberta com os detalhes para essa conexão, de modo que seja possível fazer quaisquer mudanças.

Nota: Não é possível editar uma conexão do servidor que foi incluída do IBM SPSS Collaboration and Deployment Services, já que o nome, a porta e outros detalhes são definidos no IBM SPSS Collaboration and Deployment Services. A melhor prática determina que as mesmas portas devem ser usadas para comunicação com IBM SPSS Collaboration and Deployment Services e SPSS Modeler Client. Isso pode ser configurado como `max_server_port` e `min_server_port` no arquivo `options.cfg`.

Para incluir conexões do servidor

1. No menu Ferramentas, clique em **Login do Servidor**. A caixa de diálogo Login do Servidor é aberta.
2. Nesta caixa de diálogo, clique em **Incluir**. A caixa de diálogo Incluir/Editar Servidor de Login do Servidor é aberta.
3. Insira os detalhes de conexão do servidor e clique em **OK** para salvar a conexão e retornar para a caixa de diálogo Login do Servidor.
 - **Servidor.** Especifique um servidor disponível ou selecione um na lista. O computador servidor pode ser identificado por um nome alfanumérico (por exemplo, *myserver*) ou um endereço IP designado ao computador servidor (por exemplo, 202.123.456.78).
 - **Porta.** Forneça o número da porta no qual o servidor está atendendo. Se o padrão não funcionar, peça ao administrador do sistema o número da porta correto.
 - **Descrição.** Insira uma descrição opcional para essa conexão do servidor.
 - **Assegurar conexão segura (usar SSL).** Especifica se uma conexão SSL (**Secure Sockets Layer**) deve ser usada. SSL é um protocolo comumente usado para proteger dados enviados por uma rede. Para usar essa variável, o SSL deve ser ativado no servidor hospedando o IBM SPSS Modeler Server. Se for necessário, entre em contato com seu administrador local para obter detalhes.

Para editar conexões do servidor

1. No menu Ferramentas, clique em **Login do Servidor**. A caixa de diálogo Login do Servidor é aberta.
2. Nessa caixa de diálogo, selecione a conexão que deseja editar e clique em **Editar**. A caixa de diálogo Incluir/Editar Servidor de Login do Servidor é aberta.
3. Altere os detalhes de conexão do servidor e clique em **OK** para salvar as mudanças e retornar para a caixa de diálogo Login do Servidor.

Procurando servidores no IBM SPSS Collaboration and Deployment Services

Em vez de inserir uma conexão do servidor manualmente, é possível selecionar um servidor ou cluster de servidores disponíveis na rede por meio do Coordinator of Processes, disponível no IBM SPSS Collaboration and Deployment Services. Um cluster de servidores é um grupo de servidores a partir do qual o Coordinator of Processes determina o servidor mais adequado para responder a uma solicitação de processamento.

Embora seja possível incluir servidores manualmente na caixa de diálogo Login do Servidor, a procura pelos servidores disponíveis permite se conectar aos servidores sem a necessidade de saber o nome do servidor e o número da porta corretos. Essas informações são fornecidas automaticamente. Entretanto, você ainda precisará das informações de logon corretas, como nome de usuário, domínio e senha.

Nota: Se você não tiver acesso à capacidade Coordinator of Processes, ainda será possível inserir manualmente o nome do servidor ao qual deseja se conectar ou selecionar um nome que você definiu anteriormente. Consulte o tópico “Incluindo e editando a conexão do IBM SPSS Modeler Server” na página 11 para obter mais informações.

Para procurar servidores e clusters

1. No menu Ferramentas, clique em **Login do Servidor**. A caixa de diálogo Login do Servidor é aberta.
2. Nesta caixa de diálogo, clique em **Procurar** para abrir a caixa de diálogo Procurar por Servidores. Se você não tiver efetuado logon no IBM SPSS Collaboration and Deployment Services quando tentou navegar no Coordinator of Processes, será solicitado que faça isso agora.
3. Selecione o servidor ou cluster de servidores da lista.
4. Clique em **OK** para fechar a caixa de diálogo e incluir essa conexão na tabela na caixa de diálogo Login do Servidor.

Alterando o diretório temporário

Algumas operações executadas pelo IBM SPSS Modeler Server podem requerer que arquivos temporários sejam criados. Por padrão, IBM SPSS Modeler usa o diretório temporário do sistema para criar arquivos temporários. É possível alterar a localização do diretório temporário usando os passos a seguir.

1. Crie um novo diretório chamado *spss* e um subdiretório chamado *servertemp*.
2. Edite *options.cfg*, localizado no diretório */config* do diretório de instalação do IBM SPSS Modeler. Edite o parâmetro *temp_directory* nesse arquivo para ler: *temp_directory, "C:/spss/servertemp"*.
3. Depois disso, deve-se reiniciar o serviço do IBM SPSS Modeler Server. Isso pode ser feito clicando na guia **Serviços** do Painel de Controle do Windows. Apenas pare o serviço e inicie-o para ativar as mudanças feitas. A reinicialização da máquina também reiniciará o serviço.

Todos os arquivos temporários agora serão gravados nesse novo diretório.

Nota: O erro mais comum quando você tenta fazer isso é usar o tipo errado de barra; as barras normais são usadas.

Iniciando várias sessões do IBM SPSS Modeler

Se você precisar ativar mais de uma sessão do IBM SPSS Modeler por vez, deve-se fazer algumas mudanças nas configurações do IBM SPSS Modeler e do Windows. Por exemplo, talvez seja necessário fazer isso se você tiver duas licenças de servidor separadas e desejar executar dois fluxos com relação a dois servidores diferentes a partir do mesmo computador cliente.

Para ativar várias sessões do IBM SPSS Modeler:

1. Clique em:
Iniciar > [Todos] Programas > IBM SPSS Modeler 17.1
2. No atalho do IBM SPSS Modeler 17.1 (o que tem o ícone), clique com o botão direito e selecione **Propriedades**.
3. Na caixa de texto **Resposta**, inclua **-noshare** no final da sequência de caracteres.
4. No Windows Explorer, selecione:
Ferramentas > Opções de Pasta...
5. Na guia Tipos de Arquivo, selecione a opção Fluxo do IBM SPSS Modeler e clique em **Avançado**.
6. Na caixa de diálogo Editar Tipo de Arquivo, selecione Abrir com IBM SPSS Modeler e clique em **Editar**.
7. Na caixa de texto **Aplicativo Usado para Executar Ação**, inclua **-noshare** antes do argumento **-stream**.

Visão rápida da interface do IBM SPSS Modeler

Em cada ponto no processo de mineração de dados, a interface do IBM SPSS Modeler fácil de usar convida seus conhecimentos de negócios específicos. Os algoritmos de modelagem, como predição, classificação, segmentação e detecção de associação, asseguram modelos poderosos e exatos. Os resultados do modelo podem ser facilmente implementados e lidos nos bancos de dados, IBM SPSS Statistics e em uma grande variedade de outros aplicativos.

O trabalho com o IBM SPSS Modeler é um processo de três passos de trabalho com dados.

- Primeiro, você lê dados no IBM SPSS Modeler.
- Depois, você executa os dados por meio de uma série de manipulações.
- Por fim, você envia os dados para um destino.

Essa sequência de operações é conhecida como **fluxo de dados**, pois os dados fluem registro por registro, desde a origem, passando por cada manipulação e, por fim, até seu destino--um modelo ou tipo de saída de dados.

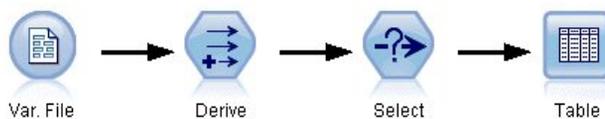


Figura 2. Um fluxo simples

Tela de fluxo do IBM SPSS Modeler

A tela de fluxo é a maior área da janela do IBM SPSS Modeler e é onde você construirá e manipulará fluxos de dados.

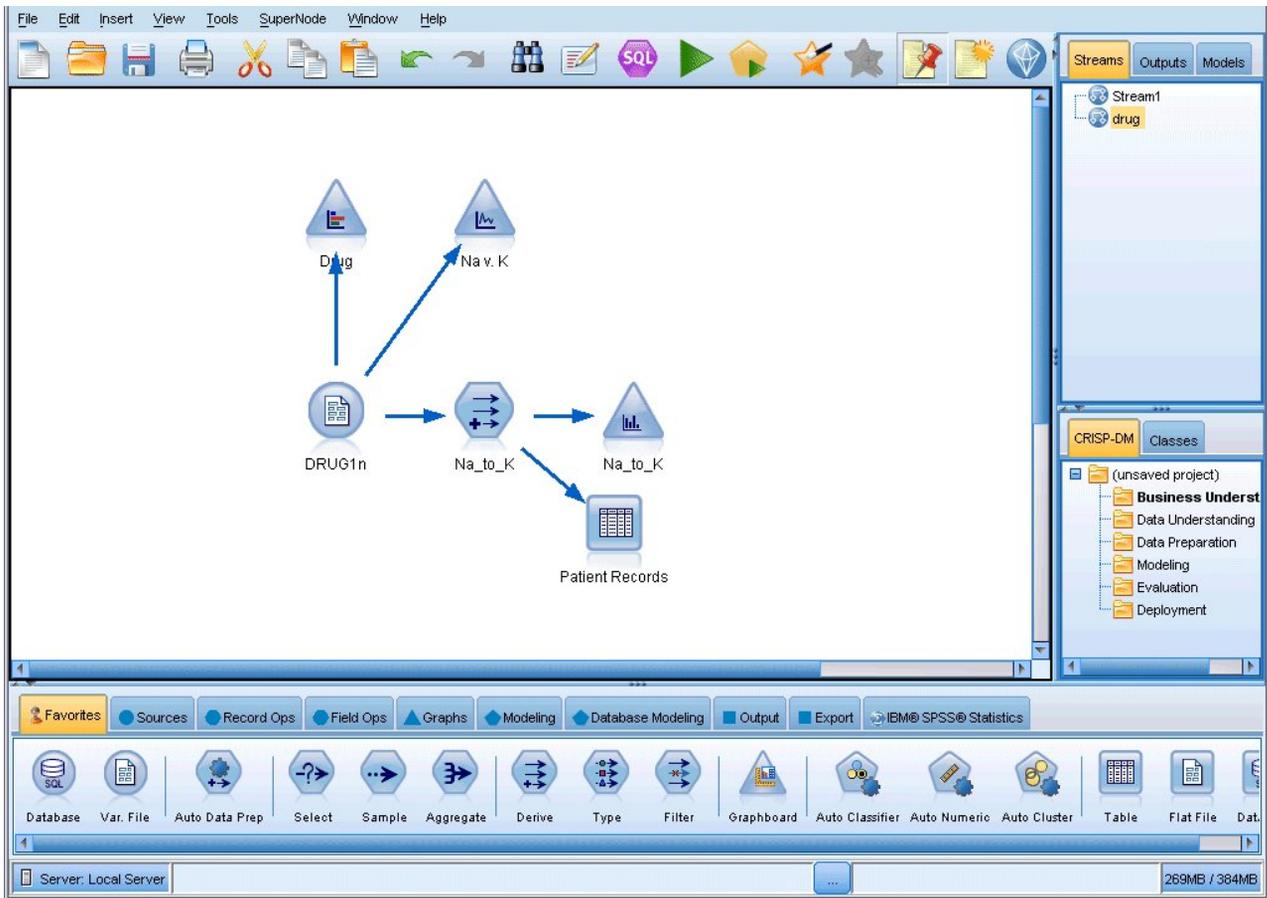


Figura 3. Área de trabalho do IBM SPSS Modeler (visualização padrão)

Fluxos são criados desenhando diagramas de operações de dados relevantes para seus negócios na tela principal da interface. Cada operação é representada por um ícone ou **nó**, e os nós são vinculados em um **fluxo** representando o fluxo de dados em cada operação.

É possível trabalhar com vários fluxos de uma vez no IBM SPSS Modeler, ou na mesma tela de fluxo ou abrindo uma nova. Durante uma sessão, fluxos são armazenados no gerenciador de Fluxos, no lado superior direito da janela do IBM SPSS Modeler.

Paleta de Nós

A maioria dos dados e das ferramentas de modelagem no IBM SPSS Modeler reside na **Paleta de Nós** na parte inferior da janela abaixo da tela de fluxo.

Por exemplo, a guia da paleta Operações de Registro contém nós que podem ser usados para executar operações nos **registros** de dados, como selecionar, mesclar e anexar.

Para incluir nós na tela, dê um clique duplo nos ícones da Paleta de Nós ou arraste e solte-os na tela. Depois conecte-os para criar um **fluxo**, representando o fluxo de dados.

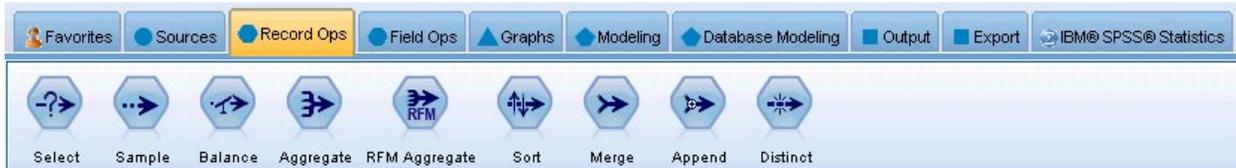


Figura 4. Guia Operações de Registro na paleta de nós

Cada guia da paleta contém uma coleção de nós relacionados usados para diferentes fases de operações de fluxo, como:

- **Origens.** Nós podem ser carregados no IBM SPSS Modeler.
- **Operações de Registro.** Nós executam operações em **registros** de dados, como selecionar, mesclar e anexar.
- **Operações de Campo.** Nós executam operações em **campos** de dados, como filtragem, derivação de novos campos e determinação do nível de medição para os campos especificados.
- **Gráficos.** Nós exibem dados graficamente antes e depois da modelagem. Gráficos incluem plots, histogramas, nós da web e gráficos de avaliação.
- **Modelagem.** Nós usam os algoritmos de modelagem disponíveis no IBM SPSS Modeler, como redes neurais, árvores de decisão, algoritmos de armazenamento em cluster e sequenciamento de dados.
- **Modelagem de Banco de Dados.** Nós usam os algoritmos de modelagem disponíveis no Microsoft SQL Server, IBM DB2 e bancos de dados Oracle e Netezza.
- **Saída.** Nós produzem uma variedade de saída para os dados, gráficos e resultados de modelo que podem ser visualizados no IBM SPSS Modeler.
- **Exportação.** Nós produzem uma variedade de saídas que podem ser visualizadas em aplicativos externos, como IBM SPSS Data Collection ou Excel.
- **IBM SPSS Statistics.** Os nós importam dados do, ou exportam dados para o, IBM SPSS Statistics, bem como procedimentos do IBM SPSS Statistics em execução.

Conforme você se familiariza com IBM SPSS Modeler, é possível customizar o conteúdo da paleta para seu próprio uso.

No lado esquerdo da Paleta de Nós, é possível filtrar os nós que são exibidos ao selecionar o Analytic Server, a Classificação, a Associação ou a Segmentação. Consulte Visão Geral de Nós de Modelagem para obter mais informações.

Localizada abaixo da Paleta de Nós, uma área de janela de relatório fornece feedback sobre o progresso de várias operações, como quando os dados estão sendo lidos no fluxo de dados. Também localizada abaixo da Paleta de Nós, uma área de janela de status fornece informações sobre o que o aplicativo está fazendo no momento, bem como indicações de quando um feedback do usuário é necessário.

Gerenciadores do IBM SPSS Modeler

Na parte superior direita da janela está a área de janela de gerenciadores. Ela possui três guias, que são usadas para gerenciar fluxos, saída e modelos.

É possível usar a guia Fluxos para abrir, renomear, salvar e excluir os fluxos criados em uma sessão.



Figura 5. Guia Fluxos



Figura 6. Guia Saídas

A guia Saídas contém uma variedade de arquivos, como gráficos e tabelas, produzidos por operações de fluxo no IBM SPSS Modeler. É possível exibir, salvar, renomear e fechar as tabelas, gráficos e relatórios listados nessa guia.

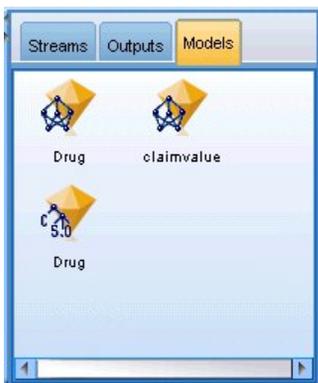


Figura 7. Guia Modelos contendo nuggets de modelo

A guia Modelos é a mais poderosa das guias de gerenciadores. Essa guia contém todos os **nuggets** de modelo, que contêm os modelos gerados no IBM SPSS Modeler, para a sessão atual. Esses modelos podem ser navegados diretamente a partir da guia Modelos ou incluídos no fluxo na tela.

Projetos do IBM SPSS Modeler

No lado inferior direito da janela está a área de janela do projeto, usada para criar e gerenciar **projetos** de mineração de dados (grupos de arquivos relacionados a uma tarefa de mineração de dados). Há duas maneiras de se visualizar projetos que você cria no IBM SPSS Modeler—na visualização Classes e na visualização CRISP-DM.

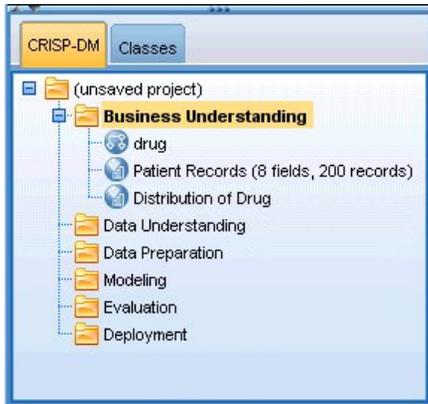


Figura 8. Visualização CRISP-DM

A guia CRISP-DM fornece uma maneira de organizar projetos de acordo com o processo padrão de vários segmentos de mercados para mineração de dados, uma metodologia não proprietária comprovada pela indústria. Tanto para mineradores de dados iniciantes quanto para os experientes, o uso da ferramenta CRISP-DM ajudará você a organizar e comunicar melhor seus esforços.

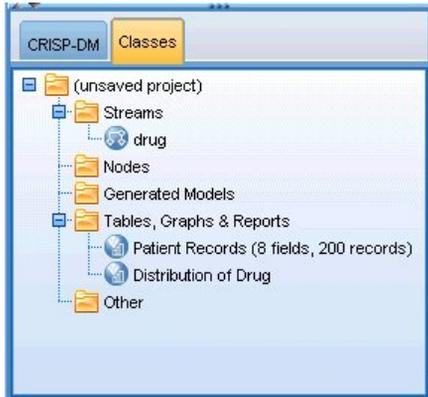


Figura 9. Visualização Classes

A guia Classes fornece uma maneira de organizar seu trabalho no IBM SPSS Modeler categoricamente—por tipos de objetos criados. Essa visualização é útil durante a realização de inventário de dados, fluxos e modelos.

Barra de ferramentas do IBM SPSS Modeler

Na parte superior da janela do IBM SPSS Modeler, você localizará uma barra de ferramentas de ícones que fornecem inúmeras funções úteis. A seguir estão os botões da barra de ferramentas e suas funções.



Criar novo fluxo



Abrir fluxo

	Salvar fluxo		Imprimir fluxo atual
	Cortar e mover para a área de transferência		Copiar para área de transferência
	Colar seleção		Desfazer última ação
	Refazer		Procurar nós
	Edite propriedades do fluxo		Visualizar geração de SQL
	Executar fluxo atual		Executar seleção de fluxo
	Parar fluxo (Ativo somente enquanto o fluxo está em execução)		Adicionar Supernó
	Aumentar zoom (Somente SuperNodes)		Diminuir zoom (Somente SuperNodes)
	Nenhuma marcação no fluxo		Inserir comentário
	Ocultar marcação de fluxo (se houver)		Mostrar marcação de fluxo oculta
	Abrir fluxo no IBM SPSS Modeler Advantage		

Marcação de fluxo consiste em comentários do fluxo, ligações de modelo e indicações de ramificação de escoragem.

Ligações de modelo são descritas no guia *Nós de Modelagem do IBM SPSS*.

Customizando a barra de ferramentas

É possível alterar vários aspectos da barra de ferramentas, como:

- Se ela é exibida
- Se os ícones têm dicas de ferramentas disponíveis
- Se ela usa ícones grandes ou pequenos

Para ativar e desativar a exibição da barra de ferramentas:

1. No menu principal, clique em:
Visualizar > Barra de Ferramentas > Exibir

Para alterar as configurações de dica de ferramenta ou tamanho do ícone:

1. No menu principal, clique em:
Visualizar > Barra de Ferramentas > Customizar

Clique em **Mostrar Dicas de Ferramenta** ou **Botões Grandes** conforme necessário.

Customizando a janela do IBM SPSS Modeler

Usando as divisórias entre várias partes da interface do IBM SPSS Modeler, é possível redimensionar ou fechar ferramentas para atender às suas preferências. Por exemplo, se você estiver trabalhando com um grande fluxo, é possível usar as pequenas setas localizadas em cada divisória para fechar a paleta do nó, a área de janela de gerenciadores e a área de janela de projeto. Isso maximiza a tela de fluxo, fornecendo um espaço de trabalho suficiente para fluxos grandes ou para vários fluxos.

Alternativamente, no menu Visualizar, clique em **Paleta de Nós**, **Gerenciadores** ou **Projeto** para ativar ou desativar a exibição desses itens.

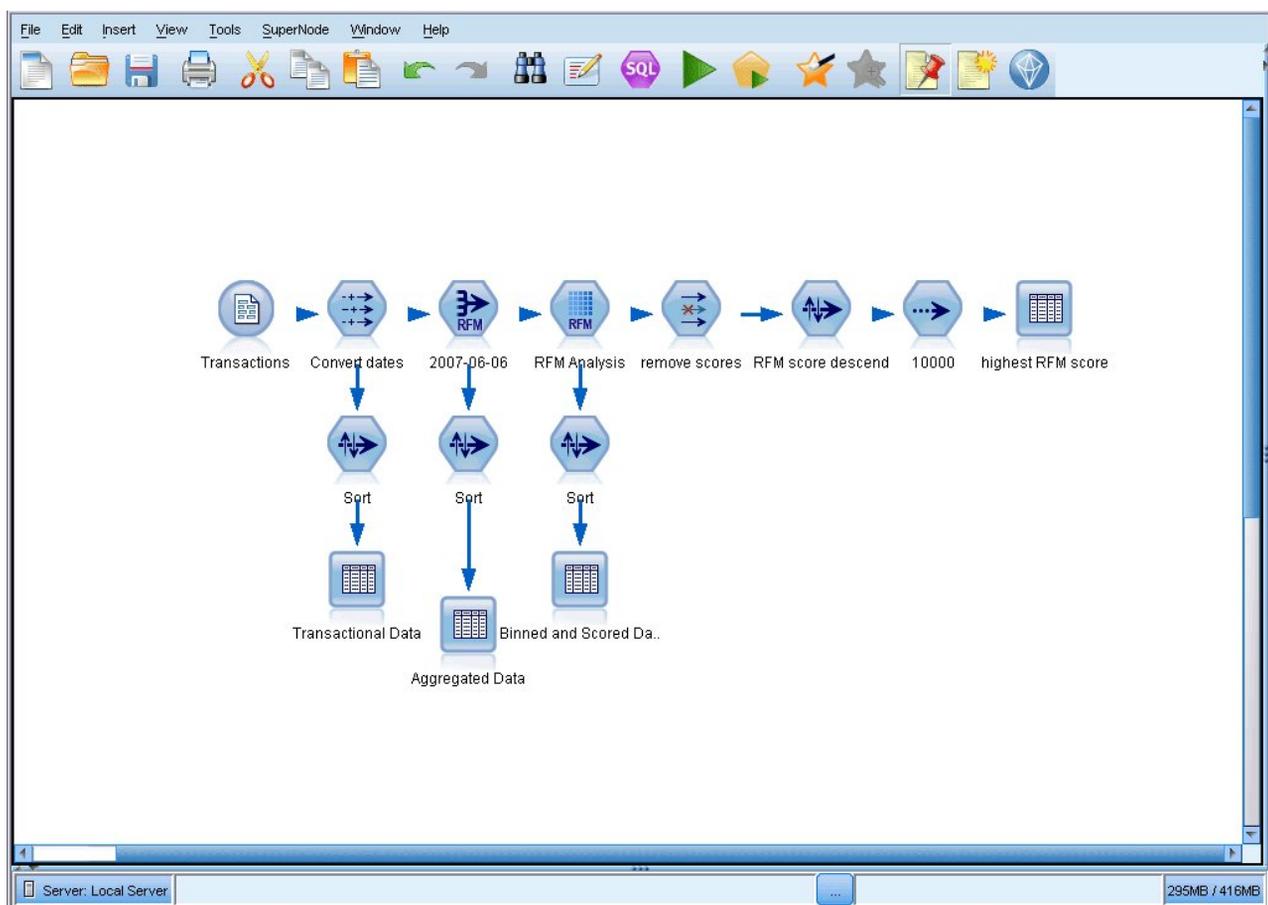


Figura 10. Tela de fluxo maximizada

Como uma alternativa para o fechamento da paleta de nós e para as áreas de janela de gerenciador e projeto, é possível usar a tela de fluxo como uma página rolável se movendo vertical e horizontalmente com as barras de rolagem ao lado e na parte inferior da janela do IBM SPSS Modeler.

Também é possível controlar a exibição da marcação da tela, o que consiste em comentários do fluxo, ligações de modelo e indicações de ramificação de escoragem. Para ativar ou desativar essa exibição, clique em:

Visualizar > Marcação de Fluxo

Alterando o tamanho do ícone para um fluxo

É possível alterar o tamanho dos ícones de fluxo das seguintes formas.

- Por meio de uma configuração de propriedades de fluxo
- Por meio de um menu pop-up no fluxo
- Usando o teclado

É possível escalar a visualização de fluxo inteira para um de vários tamanhos entre 8% e 200% do tamanho do ícone padrão.

Para escalar o fluxo inteiro (método propriedades do fluxo)

1. No menu principal, escolha **Ferramentas > Propriedades do Fluxo > Opções > Layout**.
2. Escolha o tamanho desejado no menu Tamanho do Ícone.
3. Clique em **Aplicar** para ver o resultado.
4. Clique em **OK** para salvar a mudança.

Para escalar o fluxo inteiro (método menu)

1. Clique com o botão direito no segundo plano do fluxo na tela.
2. Escolha **Tamanho do Ícone** e selecione o tamanho desejado.

Para escalar o fluxo inteiro (método teclado)

1. Pressione **Ctrl + [-]** no teclado principal para diminuir zoom para o próximo menor tamanho.
2. Pressione **Ctrl + Shift + [+]** no teclado principal para aumentar zoom para o próximo maior tamanho.

Essa variável é particularmente útil para se ter uma visualização geral de um fluxo complexo. Também é possível usá-la para minimizar o número de páginas necessárias para se imprimir um fluxo.

Usando o mouse no IBM SPSS Modeler

Os usos mais comuns do mouse no IBM SPSS Modeler incluem o seguinte:

- **Clique único.** Use o botão direito ou esquerdo do mouse para selecionar opções dos menus, abrir menus pop-up e acessar vários outros controles e opções padrão. Clique e segure o botão para mover e arrastar nós.
- **Clique duplo.** Dê um clique duplo usando o botão esquerdo do mouse para colocar nós na tela de fluxo e editar nós existentes.
- **Clique do meio.** Clique com o botão do meio do mouse e arraste o cursor para conectar nós na tela de fluxo. Dê um clique duplo no botão do meio do mouse para desconectar um nó. Se você não tiver um mouse com três botões, é possível simular essa variável pressionando a tecla **Alt** ao clicar e arrastar o mouse.

Usando teclas de atalho

Muitas operações de programação visual no IBM SPSS Modeler têm teclas de atalho associadas a elas. Por exemplo, é possível excluir um nó clicando nele e pressionando a tecla **Delete** em seu teclado. Da mesma forma, é possível salvar rapidamente um fluxo pressionando a tecla **S** enquanto você mantém pressionada a tecla **Ctrl**. Comandos de controle como esse são indicados por uma combinação de **Ctrl** e outra tecla -- por exemplo, **Ctrl+S**.

Há inúmeras teclas de atalho usadas em operações padrão do Windows, como Ctrl+X para cortar. Esses atalhos são suportados no IBM SPSS Modeler com os seguintes atalhos específicos do aplicativo.

Nota: Na maioria dos casos, teclas de atalho antigas usadas no IBM SPSS Modeler entram em conflito com teclas de atalho padrão do Windows. Esses atalhos antigos são suportados com a adição da tecla Alt. Por exemplo, Ctrl+Alt+C pode ser usado para ativar e desativar o cache.

Tabela 1. Teclas de atalho suportadas

Tecla de Atalho	Função
Ctrl+A	Selecionar todos
Ctrl+X	Cortar
Ctrl+N	Novo fluxo
Ctrl+O	Abrir fluxo
Ctrl+P	Imprimir
Ctrl+C	Copiar
Ctrl+V	Colar
Ctrl+Z	Desfazer
Ctrl+Q	Selecionar todos os nós de recebimentos de dados do nó selecionado
Ctrl+W	Cancelar todos os nós de recebimento de dados (alterna-se com Ctrl+Q)
Ctrl+E	Executar a partir do nó selecionado
Ctrl+S	Salvar fluxo atual
Alt+Teclas de Seta	Mova os nós selecionados na tela de fluxo na direção da tecla usada
Shift+F10	Abre o menu pop-up para o nó selecionado

Tabela 2. Atalhos suportados para antigas teclas de atalho

Tecla de Atalho	Função
Ctrl+Alt+D	Duplicar nó
Ctrl+Alt+L	Carregar nó
Ctrl+Alt+R	Renomear nó
Ctrl+Alt+U	Criar nó de Entrada do Usuário
Ctrl+Alt+C	Ativar/desativar cache
Ctrl+Alt+F	Limpar cache
Ctrl+Alt+X	Expandir Supernó
Ctrl+Alt+Z	Aumentar zoom/diminuir aumentar zoom
Excluir	Excluir nó ou conexão

Imprimindo

Os objetos a seguir podem ser impressos no IBM SPSS Modeler:

- Diagramas de fluxo
- Gráficos
- Tabelas
- Relatórios (no nó Relatório e Relatórios do Projeto)
- Scripts (das caixas de diálogo de propriedades do fluxo, Script Independente ou Script SuperNode)
- Modelos (navegadores Modelo, guias de caixa de diálogo com foco atual, visualizadores de árvore)

- Anotações (usando a guia Anotações para saída)

Para imprimir um objeto:

- Para imprimir sem visualizar, clique no botão Imprimir na barra de ferramentas.
- Para configurar a página antes da impressão, selecione **Configuração de Página** no menu Arquivo.
- Para visualizar antes de imprimir, selecione **Visualização de Impressão** no menu Arquivo.
- Para visualizar a caixa de diálogo de impressão padrão com opções para selecionar impressoras e especificar opções de aparência, selecione **Imprimir** no menu Arquivo.

Automatizando IBM SPSS Modeler

Como a mineração de dados avançada pode ser um processo complexo e, às vezes, longo, o IBM SPSS Modeler inclui vários tipos de codificação e o suporte de automação.

- **Control Language for Expression Manipulation (CLEM)** é uma linguagem para analisar e manipular os dados que fluem ao longo dos fluxos do IBM SPSS Modeler. Os mineradores de dados usam CLEM extensivamente em operações de fluxo para executar tarefas tão simples quanto derivar lucros dos dados de custo e renda ou tão complexas quanto transformar dados de log da web em um conjunto de campos e registros com informações utilizáveis.
- **Script** é uma ferramenta poderosa para automatizar processos na interface com o usuário. Scripts podem executar os mesmos tipos de ações que os usuários executam com um mouse ou teclado. Também é possível especificar saída e manipular modelos gerados.

Capítulo 4. Entendendo a mineração de dados

Visão geral de mineração de dados

Por meio de uma variedade de técnicas, **mineração de dados** identifica nuggets de informações em corpos de dados. A mineração de dados extrai informações de uma maneira que elas possam ser usadas em áreas como suporte à decisão, predição, previsões e estimação. Geralmente os dados são volumosos, mas têm baixo valor e pouca utilidade direta em sua forma bruta. São as informações ocultas nos dados que têm valor.

Na mineração de dados, o sucesso vem da combinação do seu conhecimento (ou do especialista) em dados com técnicas de análise ativa avançadas, nas quais o computador identifica as variáveis e os relacionamentos subjacentes nos dados. O processo de mineração de dados gera modelos a partir de dados históricos que, posteriormente, são usados para predições, detecção de padrões e mais. A técnica para a construção desses modelos é chamada **aprendizado por máquina** ou **modelagem**.

Técnicas de modelagem

IBM SPSS Modeler inclui inúmeras tecnologias de modelagem e aprendizado por máquina, que podem ser grosseiramente agrupadas de acordo com os tipos de problemas que pretendem resolver.

- Métodos de modelagem preditiva incluem árvores de decisão, redes neurais e modelos estatísticos.
- Os modelos de armazenamento em cluster focam na identificação de grupos de registros semelhantes e no rótulo desses registros de acordo com o grupo ao qual eles pertencem. Os métodos de armazenamento em cluster incluem Kohonen, k -médias e TwoStep.
- As regras de associação associam uma determinada conclusão (como a compra de um certo produto) a um conjunto de condições (a compra de vários outros produtos).
- Os modelos de rastreamento podem ser usados para rastrear dados para localizar campos e registros com maior probabilidade de serem interessantes para modelagem e identificar valores discrepantes que podem não se ajustar aos padrões conhecidos. Os métodos disponíveis incluem seleção de variável e detecção de anomalias.

Manipulação de dados e descoberta

IBM SPSS Modeler também inclui vários recursos que lhe permitem aplicar seu conhecimento aos dados:

- **Manipulação de dados.** Constrói novos itens de dados derivados dos existentes e divide os dados em subconjuntos significativos. Dados de várias origens podem ser mesclados e filtrados.
- **Navegação e visualização.** Exibe aspectos dos dados usando o nó de Auditoria de Dados para executar uma auditoria inicial, incluindo gráficos e estatísticas. A visualização avançada inclui gráficos interativos, que podem ser exportados para inclusão em relatórios do projeto.
- **Estatísticas.** Confirma relacionamentos suspeitos entre as variáveis nos dados. As estatísticas do IBM SPSS Statistics também podem ser usadas dentro do IBM SPSS Modeler.
- **Testes de hipótese.** Constrói modelos de como os dados se comportam e verifica esses modelos.

Normalmente você usará essas variáveis para identificar um conjunto promissor de atributos nos dados. Esses atributos podem ser alimentados para as técnicas de modelagem, que tentarão identificar relacionamentos e regras subjacentes.

Aplicativos típicos

Aplicativos típicos de técnicas de mineração de dados incluem o seguinte:

Correio direto. Determina quais grupos demográficos têm a taxa de resposta mais alta. Use essas informações para maximizar a resposta para futuras correspondências.

Escoragem de crédito. Use um histórico de crédito individual para tomar decisões sobre crédito.

Recursos humanos. Entenda antigas práticas de contratação e crie regras de decisão para aperfeiçoar seu processo de contratação.

Pesquisa médica. Crie regras de decisão que sugiram procedimentos apropriados com base em evidências médicas.

Análise de mercado. Determine quais variáveis, como geografia, preço e características do cliente, estão associadas às vendas.

Controle de qualidade. Analise dados de manufatura do produto e identifique variáveis determinando defeitos do produto.

Estudos de políticas. Use dados de pesquisas de opinião para formular políticas aplicando regras de decisão para selecionar as variáveis mais importantes.

Assistência médica. Os dados clínicos e as pesquisas de opinião do usuário podem ser combinados para descobrir variáveis que contribuam com a saúde.

Terminologia

Os termos **atributo**, **campo** e **variável** se referem a um único item de dados comum para todos os casos sendo considerados. Uma coleção de valores de atributo que se refere a um caso específico é chamada **registro**, **exemplo** ou **caso**.

Avaliando dados

Não é provável que uma mineração de dados seja produtiva, a menos que os dados que você deseja usar atendam a determinados critérios. As seções a seguir apresentam alguns dos aspectos dos dados e sua aplicação que deve ser considerada.

Assegure que os dados estejam disponíveis

Isso pode parecer óbvio, mas saiba que, embora os dados possam estar disponíveis, talvez eles não estejam em um formato que possa ser usado facilmente. IBM SPSS Modeler pode importar dados de bancos de dados (por meio de ODBC) ou de arquivos. Os dados, no entanto, podem ser retidos de alguma outra forma em uma máquina que não pode ser acessada diretamente. Eles precisarão ser transferidos por download ou submetidos a um dump de uma maneira adequada antes de poderem ser usados. Eles podem ser dispersados para diferentes bancos de dados e origens e precisam colaborar. Eles podem até não estar online. Se existir somente no papel, a entrada de dados será necessária para que seja possível iniciar a mineração de dados.

Verifique se os dados cobrem os atributos relevantes

O objetivo da mineração de dados é identificar atributos relevantes, portanto, a inclusão dessa verificação pode parecer estranha a princípio. No entanto, é muito útil examinar quais dados estão disponíveis e tentar identificar os prováveis fatores relevantes que não estão registrados. Na tentativa de prever as vendas de sorvetes, por exemplo, talvez você tenha inúmeras informações sobre estabelecimentos varejistas ou histórico de vendas, mas não tenha informações sobre o tempo e a temperatura, o que provavelmente tem um papel significativo. Atributos omissos não significam necessariamente que a mineração de dados não produzirá resultados úteis, mas podem limitar a precisão das previsões resultantes.

Uma maneira rápida de avaliar a situação é executar uma auditoria abrangente de seus dados. Antes de continuar, considere anexar um nó de Auditoria de Dados a sua origem de dados e executá-lo para gerar um relatório completo.

Tome cuidado com dados ruidosos

Dados costumam conter erros ou podem conter julgamentos subjetivos e, portanto, variáveis. Esses fenômenos são chamados coletivamente de **ruído**. Às vezes um ruído nos dados é normal. Pode haver regras subjacentes, mas não em 100% dos casos.

Normalmente, quanto mais ruído há nos dados, mais difícil fica a obtenção de resultados precisos. Porém, os métodos de aprendizado por máquina no IBM SPSS Modeler podem manipular dados ruidosos e vêm sendo usados com sucesso em conjuntos de dados contendo quase 50% de ruído.

Assegure que haja dados suficientes

Na mineração de dados, não é necessariamente o tamanho do conjunto de dados que importa. A representatividade do conjunto de dados é muito mais significativa, com sua cobertura de possíveis resultados e combinações de variáveis.

Normalmente, quanto mais atributos são considerados, mais registros serão necessários para fornecer uma cobertura representante.

Se os dados forem representantes e houver regras gerais subjacentes, talvez uma amostra de dados de alguns milhares (ou centenas) de registros forneçam resultados igualmente bons como milhões, e você terá esses resultados mais rapidamente.

Busque especialistas em dados

Em muitos casos, você estará trabalhando com seus próprios dados e, portanto, estará totalmente familiarizado com seu conteúdo e significado. No entanto, se você estiver trabalhando com dados de outro departamento de sua organização ou de seu cliente, talvez seja desejável que você tenha acesso aos especialistas que conhecem os dados. Eles podem guiá-lo na identificação de atributos relevantes e ajudá-lo a interpretar os resultados da mineração de dados, distinguindo nuggets verdadeiros de informações "baratas" ou artefatos causados por anomalias nos conjuntos de dados.

Uma estratégia para mineração de dados

Assim como acontece com a maioria dos esforços de negócios, a mineração de dados é muito mais efetiva se for feita de uma maneira sistemática e planejada. Mesmo com ferramentas de mineração de dados de ponta, como IBM SPSS Modeler, a maioria do trabalho na mineração de dados requer que um analista de negócios instruído mantenha o processo atualizado. Para guiar seu planejamento, responda as questões a seguir:

- Qual problema substantivo você deseja resolver?
- Quais origens de dados estão disponíveis e quais partes dos dados são relevantes para o problema atual?
- Que tipo de pré-processamento e limpeza de dados você precisa fazer antes de iniciar a mineração de dados?
- Qual(is) técnica(s) de mineração de dados você usará?
- Como você avaliará os resultados da análise de mineração de dados?
- Como você irá tirar o máximo proveito das informações obtidas da mineração de dados?

O processo de mineração de dados típico pode se tornar complicado muito rapidamente. Há muito o que acompanhar -- problemas de negócios complexos, várias origens de dados, qualidade de dados variada nas origens de dados, matriz de técnicas de mineração de dados, diferentes maneiras de medir o sucesso da mineração de dados, entre outras coisas.

Para você se manter atualizado, é útil ter um modelo de processo definido explicitamente para mineração de dados. O modelo de processo ajuda você a responder as questões listadas anteriormente nesta seção e se certificar de que os pontos importantes serão abordados. Isso serve como um roteiro de mineração de dados para que você não se perca ao se aprofundar nas complexidades de seus dados.

O processo de mineração de dados sugerido para uso com SPSS Modeler é o processo padrão de vários segmentos de mercados para mineração de dados (CRISP-DM). Como o nome já diz, esse modelo é projetado como um modelo geral que pode ser aplicado a uma grande variedade de problemas de mercado e negócios.

O modelo de processo CRISP-DM

O modelo de processo CRISP-DM geral inclui seis fases que abordam os principais problemas da mineração de dados. As seis fases se ajustam em um processo cíclico projetado para incorporar mineração de dados a suas práticas de negócios maiores.

As seis fases incluem:

- **Entendimento de negócios.** Talvez esta seja a fase mais importante da mineração de dados. Entendimento de negócios inclui determinar objetivos de negócios, avaliar a situação, determinar objetivos de mineração de dados e produzir um plano de projeto.
- **Entendimento de dados.** Dados fornecem os "materiais brutos" da mineração de dados. Essa fase aborda a necessidade de entender quais são seus recursos de dados e as características desses recursos. Isso inclui coletar dados iniciais, descrever dados, explorar dados e verificar a qualidade dos dados. O nó de Auditoria de Dados disponível na paleta de nós de Saída é uma ferramenta indispensável para o entendimento dos dados.
- **Preparação de dados.** Após catalogar seus recursos de dados, você precisará preparar seus dados para mineração. As preparações incluem selecionar, limpar, construir, integrar e formatar dados.
- **Modelagem.** Esta é, sem dúvida, a parte mais chamativa da mineração de dados, na qual métodos de análise sofisticados são usados para extrair informações dos dados. Essa fase envolve a seleção de técnicas de modelagem, geração de designs de teste e construção e avaliação de modelo.
- **Avaliação.** Após ter escolhido seus modelos, você estará pronto para avaliar como os resultados da mineração de dados podem ajudá-lo a atingir seus objetivos de negócios. Elementos dessa fase incluem avaliar resultados, revisar o processo de mineração de dados e determinar os próximos passos.
- **Implementação.** Agora que você investiu todo seu esforço, é hora de colher os benefícios. Essa fase foca na integração do seu conhecimento em processos de negócios do dia a dia para resolver seus problemas de negócios originais. Essa fase inclui implementação de plano, monitoramento e manutenção, produção de relatório final e revisão de projeto.

Há alguns pontos chave nesse modelo de processo. Primeiro, embora haja uma tendência geral para o processo fluir pelos passos na ordem esboçada nos parágrafos anteriores, também há inúmeros locais nos quais as fases influenciam umas às outras de maneira não linear. Por exemplo, geralmente a preparação de dados precede a modelagem. No entanto, geralmente as decisões tomadas e as informações reunidas durante a fase de modelagem podem levá-lo a repensar partes da fase de preparação de dados, o que pode então apresentar novos problemas de modelagem. As duas fases realimentam uma à outra até que ambas tenham sido resolvidas adequadamente. Da mesma forma, a fase de avaliação pode levá-lo a reavaliar seu entendimento de negócios original e você pode concluir que esteve tentando responder à questão errada. Nesse ponto, é possível revisar seu entendimento de negócios e continuar com o restante do processo novamente com uma resposta melhor em mente.

O segundo ponto chave é a natureza iterativa da mineração de dados. Raramente, ou nunca, você irá simplesmente planejar um projeto de mineração de dados, concluí-lo e depois empacotar seus dados e ir para casa. A mineração de dados para abordar demandas de clientes é um esforço constante. O conhecimento ganho com um ciclo de mineração de dados quase invariavelmente levará a novas questões, novos problemas e novas oportunidades para identificar e atender às necessidades do cliente. Essas novas questões, problemas e oportunidades geralmente podem ser abordados minerando os dados novamente. Esse processo de mineração e identificação de novas oportunidades deve fazer parte da forma como você pensa sobre negócios e da base de sua estratégia geral de negócios.

Essa introdução fornece apenas uma breve visão geral do modelo de processo CRISP-DM. Para obter detalhes completos sobre o modelo, consulte os seguintes recursos:

- O *Guia do CRISP-DM*, que pode ser acessado com outra documentação na pasta `\Documentation` no disco de instalação.
- Sistema de Ajuda CRISP-DM, disponível no menu Iniciar ou clicando em **Ajuda do CRISP-DM** no menu Ajuda no IBM SPSS Modeler.

Tipos de modelos

O IBM SPSS Modeler oferece uma variedade de métodos de modelagem a partir do aprendizado de máquina, inteligência artificial e estatísticas. Os métodos disponíveis na paleta Modelagem permitem derivar informações novas a partir dos dados, e desenvolver modelos preditivos. Cada método possui determinadas intensidades e é mais bem adequado para tipos de problemas específicos.

O *Guia de Aplicativos do IBM SPSS Modeler* fornece exemplos para vários desses métodos, com uma introdução geral ao processo de modelagem. Esse guia está disponível como um tutorial online e também em formato PDF. Consulte o tópico “Exemplos de Aplicativos” na página 5 para obter informações adicionais.

Os métodos de modelagem são divididos em quatro categorias:

- Analytic Server
- Classificação
- Associação
- Segmentação

Modelos de Classificação

Modelos de classificação usam os valores de um ou mais campos de **entrada** para prever o valor de um ou mais campos de saída ou **resposta**. Alguns exemplos dessas técnicas são: árvores de decisão (algoritmo Árvore C e R, algoritmo de árvore estatística eficiente, rápido e imparcial, algoritmo Detector de Interação Automático Chi-quadrado e C5.0), regressão (algoritmos linear, de logística, linear generalizado e algoritmo de regressão de Cox), redes neurais, Support Vector Machines e redes bayesianas.

Os modelos de classificação ajudam as organizações a prever um resultado conhecido, como se um cliente comprará ou sairá ou se uma transação se ajusta a um padrão de fraude conhecido. As técnicas de modelagem incluem aprendizado por máquina, indução de regra, identificação de subgrupo, métodos estatísticos e geração de modelos múltiplos.

Nós de classificação



O nó Previsor Categórico Automático cria e compara inúmeros modelos diferentes para resultados binários (sim ou não, perda de clientes ou não, e assim por diante), permitindo escolher a melhor abordagem para uma determinada análise. Vários algoritmos de modelagem são suportados, possibilitando a seleção dos métodos que você deseja usar, as opções específicas para cada e os critérios para comparar os resultados. O nó gera um conjunto de modelos com base nas opções especificadas e ranqueia os melhores candidatos de acordo com os critérios especificados.



O nó Previsor Contínuo Automático estima e compara modelos para resultados de intervalos numéricos contínuos usando vários métodos diferentes. O nó trabalha da mesma maneira que o nó Previsor Categórico Automático, permitindo escolher os algoritmos para usar e experimentar com várias combinações de opções em uma única passagem de modelagem. Os algoritmos suportados incluem redes neurais, Árvore C e R, algoritmo Detector de Interação Automático Chi-quadrado, regressão linear, regressão linear generalizada e Support Vector Machines (SVM). Os modelos podem ser comparados com base em correlação, erro relativo ou número de variáveis usadas.



O nó Classificação e Regressão (C e R) gera uma árvore de decisão que permite prever ou classificar futuras observações. O método usa particionamento recursivo para dividir os registros de treinamento em segmentos, minimizando as impurezas de cada passo, em que um nó na árvore é considerado “puro” se 100% dos casos no nó estiverem dentro de uma categoria específica do campo de resposta. Os campos de entrada e resposta podem ser intervalos numéricos ou categóricos (nominal, ordinal ou flags); todas as divisões são binárias (somente dois subgrupos).



O nó QUEST fornece um método de classificação binário para construir árvores de decisão, projetadas para reduzir o tempo de processamento necessário para grandes análises de C e R, enquanto também reduz a tendência localizada nos métodos de árvore de classificação para favorecer entradas que permitam mais divisões. Campos de entrada podem ser intervalos numéricos (contínuos), mas o campo de resposta deve ser categórico. Todas as divisões são binárias.



O nó CHAID gera árvores de decisão usando estatísticas qui-quadrado para identificar as divisões ideais. Ao contrário dos nós Árvore C e R e QUEST, o CHAID pode gerar árvores não binárias, o que significa que algumas divisões têm mais de duas ramificações. Os campos de entrada e de resposta podem ser intervalos numéricos (contínuos) ou categóricos. CHAID exaustivo é uma modificação de CHAID que executa uma tarefa mais completa de exame de todas as divisões possíveis, mas que leva mais tempo para ser calculada.



O nó C5.0 constrói uma árvore de decisão ou um conjunto de regras. O modelo trabalha dividindo a amostra com base no campo que fornece o máximo de informações obtidas em cada nível. O campo de resposta deve ser categórico. Divisões múltiplas em mais de dois subgrupos são permitidas.



O nó Lista de Decisão identifica subgrupos, ou segmentos, que mostram uma probabilidade maior ou menor de um determinado resultado binário relativo à população geral. Por exemplo, você pode examinar os clientes com menor probabilidade de migrarem para o concorrente ou com maior probabilidade de responderem a uma campanha de forma favorável. É possível incorporar seu conhecimento de negócios no modelo incluindo seus próprios segmentos customizados e visualizando modelos alternativos lado a lado para comparar os resultados. Os modelos de Lista de Decisão consistem em uma lista de regras na qual cada regra tem uma condição e um resultado. Regras são aplicadas em ordem, e a primeira regra correspondente determina o resultado.



Os modelos de regressão linear preveem uma variável resposta contínua baseada em relacionamentos lineares entre a resposta e um ou mais preditores.



O nó PCA/Factor fornece técnicas poderosas de redução de dados para reduzir a complexidade de seus dados. A análise de componentes principais (PCA) localiza combinações lineares dos campos de entrada que realizam as melhores tarefas de captura de variância no conjunto de campos inteiro, no qual os componentes são ortogonais (perpendiculares) uns aos outros. A análise fatorial tenta identificar os fatores subjacentes que explicam o padrão de correlações dentro de um conjunto observado de campos. Para ambas as abordagens, o objetivo é localizar um pequeno número de campos derivados que resumizam efetivamente as informações no conjunto de campos original.



O nó de Seleção de Variável exhibe campos de entrada para remoção com base em um conjunto de critérios (como porcentagem de valores omissos); depois ele ranqueia a importância das entradas restantes relativas a uma resposta especificada. Por exemplo, dado um conjunto de dados com centenas de entradas potenciais, quais têm maior probabilidade de serem úteis na modelagem de resultados de pacientes?



A análise discriminante faz suposições mais rigorosas que a regressão logística, mas pode ser uma alternativa de valor ou um suplemento para uma análise de regressão logística quando essas suposições são atendidas.



A regressão logística é uma técnica estatística para ordenar registros com base em valores de campos de entrada. Ela é análoga à regressão linear, mas usa um campo de variável resposta categórica em vez de um intervalo numérico.



O Modelo Linear Generalizado expande o modelo linear geral para que a variável dependente esteja linearmente relacionada aos fatores e às covariáveis por meio de uma função de ligação especificada. Além disso, o modelo permite que a variável dependente tenha uma distribuição não normal. Ele cobre a funcionalidade de um amplo número de modelos estatísticos, incluindo regressão linear, regressão logística, modelos log-linear para dados de contagem e modelos de sobrevivência censurados por intervalo.



Um modelo linear generalizado misto (GLMM) estende o modelo linear para que a resposta possa ter uma distribuição não normal, esteja linearmente relacionada aos fatores e às covariáveis via uma função de ligação especificada e para que as observações possam ser correlacionadas. Modelos lineares generalizados mistos cobrem uma grande variedade de modelos, de regressão linear simples a modelos multinível complexos para dados longitudinais não normais.



O nó de Regressão de Cox permite construir um modelo de sobrevivência para dados de sobrevivência na presença de registros censurados. O modelo produz uma função de sobrevivência que prevê a probabilidade de o evento de interesse ter ocorrido em um determinado momento (t) para determinados valores das variáveis de entrada.



O nó Support Vector Machine (SVM) permite ordenar dados em dois grupos sem super ajuste. SVM trabalha bem com conjuntos de dados grandes, como aqueles com um número muito grande de campos de entrada.



O nó de Rede Bayesiana permite construir um modelo de probabilidade combinando evidências observadas e registradas com conhecimento do mundo real para estabelecer a probabilidade de ocorrências. O nó foca nas redes Tree Augmented Naïve Bayes (TAN) e Markov Blanket que são usadas principalmente para classificação.



O nó Self-Learning Response Model (SLRM) permite construir um modelo no qual um único novo caso, ou um pequeno número de novos casos, pode ser usado para estimar novamente o modelo sem precisar treinar o modelo outra vez usando todos os dados.



O nó Séries Temporais estima modelos de suavização exponencial, Média Móvel Integrada AutoRegressiva (ARIMA) univariada e ARIMA multivariada (ou função de transferência) para dados de séries temporais e produz previsões do desempenho futuro. Um nó Séries Temporais deve ser sempre precedido por um nó Intervalos de Tempo.



O nó k -Nearest Neighbor (KNN) associa um novo caso à categoria ou valor dos objetos k mais próximos dele no espaço do preditor, em que k é um número inteiro. Casos semelhantes ficam próximos uns dos outros e os casos diferentes ficam distantes uns dos outros.



O nó Spatio-Temporal Prediction (STP) usa dados que contêm dados de localização, campos de entrada para previsão (preditores), um campo de tempo e um campo de resposta. Cada localização tem inúmeras linhas nos dados que representam os valores de cada preditor em cada momento da medição. Após os dados serem analisados, eles podem ser usados para prever valores de resposta em qualquer localização dentro dos dados de forma usados na análise.

Modelos da Associação

Modelos de associação localizam padrões em seus dados nos quais uma ou mais entidades (como eventos, compras ou atributos) são associadas a uma ou mais entidades. Os modelos constroem conjuntos de regras que definem esses relacionamentos. Aqui os campos dentro dos dados podem agir como entradas e respostas. Você poderia localizar essas associações manualmente, mas os algoritmos de regra de associação fazem isso com muito mais rapidez e podem explorar padrões mais complexos. Os modelos A Priori e Carma são exemplos do uso desses algoritmos. Um outro tipo de modelo de associação é um modelo de detecção de sequência, que localiza padrões sequenciais em dados estruturados por tempo.

Modelos de associação são mais úteis na previsão de resultados múltiplos—por exemplo, clientes que compraram o produto X também compraram os produtos Y e Z. Os modelos de associação associam uma determinada conclusão (como decisão de comprar algo) a um conjunto de condições. A vantagem dos algoritmos de regra de associação sobre os algoritmos de árvore de decisão mais padrão (C5.0 and C&RT) é que associações podem existir entre qualquer um dos 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.

Nós de associação



O nó A Priori extrai um conjunto de regras dos dados, removendo as regras com o conteúdo de informações mais alto. A Priori oferece cinco métodos diferentes de seleção de regras e usa um esquema de indexação sofisticado para processar conjuntos de dados grandes com eficiência. Para grandes problemas, o A Priori geralmente é mais rápido para treinar; ele não tem um limite arbitrário no número de regras que podem ser retidas e pode manipular regras com até 32 condições prévias. O A Priori requer que os campos de entrada e saída sejam todos categóricos, mas entrega melhor desempenho por ser otimizado para esse tipo de dado.



O modelo CARMA extrai um conjunto de regras dos dados sem requerer a especificação de campos de entrada ou saída. Em contraste com o A Priori o nó CARMA oferece configurações de construção para suporte de regra (suporte para antecedente e consequente) em vez de apenas suporte antecedente. Isso significa que as regras geradas podem ser usadas para uma variedade maior de aplicativos—por exemplo, para localizar uma lista de produtos ou serviços (antecedentes) cujo consequente é o item que você deseja promover nesta temporada de férias.



O nó de Sequência descobre regras de associação em dados sequenciais ou orientados por tempo. Uma sequência é uma lista de conjuntos de itens que tende a ocorrer em uma ordem previsível. Por exemplo, um cliente que compra um aparelho de barbear e uma loção pós-barba pode comprar um creme de barbear na próxima compra. O nó de Sequência é baseado no algoritmo das regras de associação CARMA, que usa um método eficiente de dois passos para localizar sequências.



O nó de Regras de Associação é semelhante ao nó A Priori; no entanto, ao contrário do A Priori, o nó de Regras de Associação pode processar dados de lista. Além disso, o nó de Regras de Associação pode ser usado com IBM SPSS Analytic Server para processar Big Data e aproveitar a vantagem do processamento paralelo mais rápido.

Modelos de Segmentação

Modelos de segmentação dividem os dados em segmentos, ou clusters, de registros com padrões semelhantes de campos de entrada. Como eles estão interessados somente nos campos de entrada, os modelos de segmentação não têm um conceito de campos de saída ou resposta. Os exemplos de modelos de segmentação são redes Kohonen, armazenamento em cluster de k-médias, clusterização em dois passos e detecção de anomalias.

Segmentação de modelos (também conhecidos como "modelos de armazenamento em cluster") são úteis nos casos em que o resultado específico é desconhecido (por exemplo, durante a identificação de novos padrões de fraude ou identificação de grupos de interesse em sua base de cliente). Os modelos de armazenamento em cluster focam na identificação de grupos de registros semelhantes e na rotulagem dos registros de acordo com o grupo ao qual eles pertencem. Isso é feito sem o benefício do conhecimento prévio sobre os grupos e suas características, e isso distingue os modelos de armazenamento em cluster de outras técnicas de modelagem porque não há saída predefinida ou campo de resposta para o modelo prever. Não há respostas certas ou erradas para esses modelos. Seu valor é determinado pela sua capacidade de capturar agrupamentos interessantes nos dados e fornecer descrições úteis desses agrupamentos. Os modelos de armazenamento em cluster geralmente são usados para criar clusters ou segmentos que são usados como entradas em análises subsequentes (por exemplo, segmentando possíveis clientes em subgrupos homogêneos).

Nós de segmentação



O nó Cluster Automático estima e compara modelos de armazenamento em cluster, que identificam grupos de registros com características semelhantes. O nó trabalha da mesma maneira que outros nós de modelagem automatizados, permitindo experimentar várias combinações de opções em uma única passagem de modelagem. Modelos podem ser comparados usando medidas básicas com as quais você tentar filtrar e ranquear a utilidade dos modelos de cluster e fornecer uma medida baseada na importância de determinados campos.



O nó K-médias armazena em cluster os dados configurados em grupos distintos (ou clusters). O método define um número fixo de clusters, designa registros aos clusters iterativamente e ajusta os centros de cluster até que os refinamentos adicionais não possam mais melhorar o modelo. Em vez de tentar prever um resultado, o *k*-médias usa um processo conhecido como aprendizado não supervisionado para descobrir padrões no conjunto de campos de entrada.



O nó Kohonen gera um tipo de rede neural que pode ser usado para armazenar o conjunto de dados em cluster em grupos distintos. Quando a rede estiver totalmente treinada, os registros semelhantes deverão estar próximos no mapa de saída, enquanto os registros diferentes deverão ficar mais afastados. É possível consultar o número de observações capturadas por cada unidade no nugget do modelo para identificar as unidades fortes. Isso lhe dará uma noção do número apropriado de clusters.



O nó TwoStep usa um método de clusterização em dois passos. O primeiro passo faz uma simples passagem pelos dados para compactar os dados de entrada brutos em um conjunto gerenciável de subclusters. O segundo passo usa um método de armazenamento em cluster hierárquico para mesclar progressivamente os subclusters em clusters cada vez maiores. TwoStep tem a vantagem de estimar automaticamente o número ideal de clusters para dados de treinamento. Ele pode manipular tipos mistos de campos e conjuntos grandes de dados de forma eficiente.



O nó de Detecção de Anomalias identifica casos incomuns, ou valores discrepantes, que não estão em conformidade com os padrões de dados “normais”. Com esse nó, é possível identificar valores discrepantes, mesmo que eles não se ajustem a nenhum padrão conhecido anteriormente e mesmo que você não esteja certo do que está procurando exatamente.

Modelos de Mineração dentro da Base de Dados

IBM SPSS Modeler suporta integração com ferramentas de modelagem e mineração de dados que estão disponíveis a partir dos fornecedores de base de dados, incluindo Oracle Data Miner, IBM DB2 InfoSphere Warehouse e Microsoft Analysis Services. É possível construir, escorar e armazenar modelos dentro do banco de dados-tudo isso de dentro do aplicativo IBM SPSS Modeler. Para obter detalhes completos, consulte o *Guia de Mineração dentro da Base de Dados do IBM SPSS Modeler*, disponível no produto DVD.

IBM SPSS Statistics Modelos

Se você tiver uma cópia do IBM SPSS Statistics instalada e licenciada em seu computador, é possível acessar e executar certas rotinas do IBM SPSS Statistics de dentro do IBM SPSS Modeler para construir e escorar modelos.

Informações Adicionais

A documentação detalhada dos algoritmos de modelagem também está disponível. Para obter mais informações, consulte o *Guia de Algoritmos do IBM SPSS Modeler*, disponível no produto DVD.

Exemplos de mineração de dados

A melhor maneira de se aprender sobre mineração de dados na prática é começar com um exemplo. Inúmeros exemplos de aplicativo estão disponíveis no *Guia de Aplicativos do IBM SPSS Modeler*, que fornece introduções direcionadas sintetizadas para técnicos e métodos de modelagem específicos. Consulte o tópico “Exemplos de Aplicativos” na página 5 para obter mais informações.

Capítulo 5. Construindo fluxos

Visão geral da construção de fluxo

A mineração de dados usando IBM SPSS Modeler foca no processo de execução de dados por meio de uma série de nós chamada **fluxo**. Essa série de nós representa operações a serem executadas nos dados, enquanto ligações entre os nós indicam a direção do fluxo de dados. Normalmente você usa um fluxo de dados para ler dados no IBM SPSS Modeler, executá-lo por meio de uma série de manipulações e depois enviá-lo para um destino, como uma tabela ou um visualizador.

Por exemplo, suponha que você deseja abrir uma origem de dados, incluir um novo campo, selecionar registros com base em valores no novo campo e depois exibir os resultados em uma tabela. Nesse caso, seu fluxo de dados consistiria em quatro nós:



Um nó de Arquivo Variável, que você configura para ler os dados da origem de dados.



Um nó de Derivação, que você usa para incluir o campo novo calculado no conjunto de dados.



Um nó de Seleção, que você usa para configurar critérios de seleção para excluir registros do fluxo de dados.



Um nó de Tabela, que você usa para exibir os resultados de suas manipulações na tela.

Construindo fluxos de dados

A interface do IBM SPSS Modeler exclusiva permite minar seus dados visualmente trabalhando com diagramas de fluxos de dados. No nível mais básico, é possível construir um fluxo de dados usando os passos a seguir:

- Inclua nós na tela de fluxo.
- Conecte os nós para formar um fluxo.
- Especifique as opções de nó ou de fluxo.
- Execute o fluxo.

Esta seção contém informações mais detalhadas sobre como trabalhar com nós para criar fluxos de dados mais complexos. Discute também as opções e configurações para nós e fluxos. Para exemplos passo a passo de construção de fluxo usando os dados enviados com o IBM SPSS Modeler (na pasta *Demos* da instalação do programa, consulte “Exemplos de Aplicativos” na página 5.

Trabalhando com nós

Nós são usados no IBM SPSS Modeler para ajudar a explorar dados. Vários nós na área de trabalho representam diferentes objetos e ações. A paleta na parte inferior da janela do IBM SPSS Modeler contém todos os nós possíveis usados na construção do fluxo.

Há inúmeros tipos de nós. **Nós de origem** reúnem dados no fluxo e estão localizados na guia Origens da paleta de nós. **Nós de processo** executam operações em campos e registros de dados individuais e podem ser localizados nas guias Operações de Registro e Operações de Campo da paleta. **Nós de saída** produzem uma variedade de saída de para dados, gráficos e resultados de modelo, além de serem incluídos nas guias Gráficos, Saída e Exportação da paleta de nós. **Nós de modelagem** usam algoritmos estatísticos para criar nuggets de modelos e estão localizados na guia Modelagem e (se ativada) na guia Modelagem da Base de Dados da paleta de nós. Consulte o tópico “Paleta de Nós” na página 14 para obter mais informações.

Você conecta os nós para formar fluxos que, quando executados, permitem visualizar relacionamentos e tirar conclusões. Fluxos são como scripts--é possível salvá-los e reutilizá-los com diferentes arquivos de dados.

Um nó executável que processa dados de fluxo é conhecido como **nó terminal**. Um nó de saída ou modelagem será um nó terminal se estiver localizado no final do fluxo ou ramificação do fluxo. Não é possível conectar nós adicionais a um nó terminal.

Nota: É possível customizar a paleta Nós. Consulte o tópico “Customizando a Paleta de Nós” na página 190 para obter mais informações.

Incluindo nós em um fluxo

Há várias maneiras de incluir nós em um fluxo a partir da paleta de nós:

- Dê um clique duplo em um nó na paleta. *Nota:* Um clique duplo em um nó o conecta automaticamente ao fluxo atual. Consulte o tópico “Conectando nós em um fluxo” para obter mais informações.
- Arraste e solte um nó da paleta para a tela de fluxo.
- Clique em um nó na paleta e, em seguida, clique na tela de fluxo.
- Selecione uma opção apropriada no menu Inserir do IBM SPSS Modeler.

Após ter incluído um nó na tela de fluxo, dê um clique duplo no nó para exibir sua caixa de diálogo. As opções disponíveis dependem do tipo de nó sendo incluído. Para obter informações sobre controles específicos dentro da caixa de diálogo, clique em seu botão **Ajuda**.

Removendo Nós

Para remover um nó do fluxo de dados, clique nele e pressione a tecla Delete ou clique com o botão direito e selecione **Excluir** no menu.

Conectando nós em um fluxo

Nós incluídos na tela de fluxo não formam um fluxo de dados até que eles tenham sido conectados. As conexões entre os nós indicam a direção dos dados conforme eles fluem de uma operação para a seguinte. Há inúmeras maneiras de se conectar nós para formar um fluxo: dando um clique duplo, usando o botão do meio do mouse ou manualmente.

Para incluir e conectar nós dando um clique duplo

A maneira mais simples de formar um fluxo é dando um clique duplo nos nós na paleta. Esse método conecta automaticamente o novo nó ao nó selecionado na tela de fluxo. Por exemplo, se a tela contiver um nó de Banco de Dados, é possível selecionar esse nó e depois dar um clique duplo no próximo nó da paleta, como um nó de Derivação. Essa ação conecta automaticamente o nó de Derivação ao nó de Banco de Dados existente. É possível repetir esse processo até atingir um nó terminal, como um nó de Histograma ou Tabela, em cujo ponto quaisquer novos nós serão conectados ao último nó de envio de dados não terminal.

Para conectar nós usando o botão do meio do mouse

Na tela de fluxo, é possível clicar e arrastar de um nó para outro usando o botão do meio do mouse. (Se seu mouse não tiver um botão do meio, é possível simular isso pressionando a tecla Alt enquanto você arrasta com o mouse de um nó para outro.)

Para conectar nós manualmente

Se você não tiver um botão do meio do mouse e preferir conectar nós manualmente, é possível usar o menu pop-up para um nó para conectá-lo a outro nó já na tela.

1. Clique com o botão direito no nó do qual deseja iniciar a conexão. Isso abre o menu do nó.
2. No menu, clique em **Conectar**.
3. Um ícone de conexão é exibido no nó inicial e cursor. Clique em um segundo nó na tela para conectar os dois nós.

Durante a conexão dos nós, há várias diretrizes para serem seguidas. Você receberá uma mensagem de erro se tentar estabelecer algum dos tipos de conexão a seguir:

- Uma conexão levando a um nó de origem
- Uma conexão a partir de um nó terminal
- Um nó com mais do que seu número máximo de conexões de entrada
- Conexão de dois nós já conectados
- Circularidade (dados retornam para um nó do qual já fluíram)

Efetuando bypass nos nós em um fluxo

Quando você efetua bypass em um nó no fluxo de dados, todas as suas conexões de entrada e saída são substituídas pelas conexões que levaram diretamente dos seus nós de entrada para seus nós de saída. Se o nó não tiver conexões de entrada e saída, todas as suas conexões serão excluídas em vez de roteadas novamente.

Por exemplo, você pode ter um fluxo que deriva um novo campo, filtra campos e depois explora os resultados em um histograma e em uma tabela. Se também deseja visualizar o mesmo gráfico e tabela para os dados *antes* de os campos serem filtrados, é possível incluir novos nós de Histograma e Tabela ou efetuar bypass no nó de Filtro. Quando você efetua bypass no nó de Filtro, as conexões com o gráfico e a tabela passam diretamente do nó de Derivação. O nó de Filtro é desconectado do fluxo.

Para efetuar bypass em um nó

1. Na tela de fluxo, use o botão do meio do mouse para dar um clique duplo no nó no qual deseja efetuar bypass. Como alternativa, é possível usar Alt+clique duplo.

Nota: É possível desfazer essa ação clicando em **Desfazer** no menu Editar ou pressionando Ctrl+Z.

Desativando nós em um fluxo

Os nós de processamento com uma única entrada dentro dos fluxos podem ser desativados, cujo resultado é o nó ser ignorado durante a execução do fluxo. Isso evita que você precise remover ou efetuar bypass do nó e significa que é possível deixá-lo conectado aos nós restantes. Ainda é possível abrir e editar as configurações do nó; no entanto, nenhuma mudança entrará em vigor até você ativar o nó novamente.

Por exemplo, você pode ter um fluxo que filtre vários campos e depois construa modelos com o conjunto de dados reduzido. Se também deseja construir os mesmos modelos *sem* os campos serem filtrados, para ver se eles melhoram os resultados do modelo, é possível desativar o nó de Filtro. Quando você desativa o nó de Filtro, as conexões com o nó de modelagem passam diretamente pelo nó de Derivação para o nó Tipo.

Para desativar um nó

1. Na tela de fluxo, clique com o botão direito no nó que deseja desativar.

2. Clique em **Desativar Nó** no menu pop-up.

Como alternativa, é possível clicar em **Nó > Desativar Nó** no menu Editar. Quando quiser incluir o nó de volta no fluxo, clique em **Ativar Nó** da mesma maneira.

Nota: É possível desfazer essa ação clicando em **Desfazer** no menu Editar ou pressionando Ctrl+Z.

Incluindo nós em conexões existentes

É possível incluir um novo nó entre dois nós conectados arrastando a seta que conecta os dois nós.

1. Com o botão do meio do mouse, clique e arraste a seta de conexão na qual deseja inserir o nó. Alternativamente, é possível manter a tecla ALT pressionada enquanto você clica e arrasta para simular o botão do meio do mouse.
2. Arraste a conexão para o nó que deseja incluir e solte o botão do mouse.

Nota: É possível remover novas conexões do nó e restaurar o original **efetuando bypass** no nó.

Excluindo conexões entre nós

Para excluir a conexão entre dois nós:

1. Clique com o botão direito na seta de conexão.
2. No menu, clique em **Excluir Conexão**.

Para excluir todas as conexões para e de um nó, execute um dos seguintes procedimentos:

- Selecione o nó e pressione F3.
- Selecione o nó e, no menu principal, clique em:

Editar > Nó > Desconectar

Configurando opções para nós

Após você ter criado e conectado nós, há várias opções para customizar nós. Clique com o botão direito em um nó e selecione uma das opções de menu.

- Clique em **Editar** para abrir a caixa de diálogo para o nó selecionado.
- Clique em **Conectar** para conectar manualmente um nó ao outro.
- Clique em **Desconectar** para excluir todos as ligações para e do nó.
- Clique em **Renomear e Anotar** para abrir a guia Anotações da caixa de diálogo de edição.
- Clique em **Novo Comentário** para incluir um comentário relacionado ao nó. Consulte o tópico “Incluindo comentários e anotações em nós e fluxos” na página 54 para obter mais informações.
- Clique em **Desativar Nó** para “ocultar” o nó durante o processamento. Para deixar o nó visível novamente para processamento, clique em **Ativar Nó**. Consulte o tópico “Desativando nós em um fluxo” na página 35 para obter mais informações.
- Clique em **Cortar** ou **Excluir** para remover o(s) nó(s) selecionado(s) da tela de fluxo. *Nota:* Um clique em **Cortar** permite colar nós, enquanto **Excluir** não.
- Clique em **Copiar Nó** para fazer uma cópia do nó sem conexões. Isso pode ser incluído em um fluxo novo ou existente.
- Clique em **Carregar Nó** para abrir um nó salvo anteriormente e carregar suas opções no nó atualmente selecionado. *Nota:* Os nós devem ter tipos idênticos.
- Clique em **Recuperar Nó** para recuperar um nó de um IBM SPSS Collaboration and Deployment Services Repository conectado.
- Clique em **Salvar Nó** para salvar os detalhes do nó em um arquivo. É possível carregar detalhes do nó somente em outro nó do mesmo tipo.
- Clique em **Armazenar Nó** para armazenar o nó selecionado em um IBM SPSS Collaboration and Deployment Services Repository conectado.

- Clique em **Cache** para expandir o menu com opções para armazenar em cache o nó selecionado.
- Clique em **Mapeamento de Dados** para expandir o menu com opções para mapear dados para uma nova origem ou especificar campos obrigatórios.
- Clique em **Criar SuperNode** para expandir o menu com opções para criar um SuperNode no fluxo atual.
- Clique em **Gerar Nó de Entrada do Usuário** para substituir o nó selecionado. Os exemplos gerados por esse nó terão os mesmos campos que o nó atual.
- Clique em **Executar Daqui** para executar o recebimento de dados de todos os nós terminais a partir do nó selecionado.

Armazenando opções em cache para nós

Para otimizar a execução do fluxo, é possível configurar um **cache** em um nó não terminal. Quando você configura um cache em um nó, o cache é preenchido com os dados que são passados pelo nó na próxima vez que você executa o fluxo de dados. Daí em diante, os dados são lidos do cache (que é armazenado em disco em um diretório temporário) e não da origem de dados.

O armazenamento em cache é mais útil após uma operação demorada, como uma ordenação, mesclagem ou agregação. Por exemplo, suponha que você tenha um nó de origem configurado para ler dados de vendas de um banco de dados e um nó agregado que resuma as vendas por localização. É possível configurar um cache no nó agregado em vez de no nó de origem, pois você deseja que o cache armazene os nós agregados, e não o conjunto de dados inteiro.

Nota: O armazenamento em cache em nós de origem, que simplesmente armazena uma cópia dos dados originais como eles são lidos no IBM SPSS Modeler, não melhora o desempenho na maioria das circunstâncias.

Nós com o armazenamento em cache ativado são exibidos com um ícone de documento pequeno no canto superior direito. Quando os dados são armazenados em cache no nó, o ícone de documento fica verde.

Para ativar um cache

1. Na tela de fluxo, clique com o botão direito no nó e clique em **Cache** no menu.
2. No submenu de armazenamento em cache, clique em **Ativar**.
3. É possível desativar o cache clicando com o botão direito no nó e clicando em **Desativar** no submenu de armazenamento em cache.

Nós de armazenamento em cache em um banco de dados

Para fluxos executados em um banco de dados, dados podem ser armazenados em cache no midstream para uma tabela temporária no banco de dados em vez de para o sistema de arquivo. Quando combinado com otimização de SQL, isso pode resultar em ganhos significativos no desempenho. Por exemplo, a saída de um fluxo que mescla várias tabelas para criar uma visualização de mineração de dados pode ser armazenada em cache e reutilizada conforme necessário. Ao gerar SQL automaticamente para todos os nós de recebimento de dados, o desempenho pode ser melhorado.

Durante o uso do armazenamento em cache do banco de dados com sequências de caracteres com mais de 255 caracteres, assegure que haja um envio de dados de nó Tipo a partir do nó de armazenamento em cache e de que os valores de campo sejam lidos, ou configure o comprimento da sequência de caracteres por meio do parâmetro `default_sql_string_length` no arquivo `options.cfg`. Isso assegura que a coluna correspondente na tabela temporária seja configurada para a largura correta para acomodar as sequências de caracteres.

Para aproveitar a vantagem do armazenamento em cache do banco de dados, a otimização de SQL e o armazenamento em cache do banco de dados devem estar ativados. Observe que as configurações de

otimização do Servidor substituem as do Cliente. Consulte o tópico “Configurando opções de otimização para fluxos” na página 42 para obter mais informações.

Com o armazenamento em cache do banco de dados ativado, basta clicar com o botão direito em qualquer nó não terminal para armazenar dados em cache nesse ponto e, automaticamente, o cache será criado diretamente no banco de dados na próxima vez que o fluxo for executado. Se o armazenamento em cache do banco de dados ou a otimização de SQL não estiverem ativados, o cache será gravado no sistema de arquivos.

Nota: Os bancos de dados a seguir suportam tabelas temporárias para o propósito de armazenamento em cache: DB2, Netezza, Oracle, SQL Server e Teradata. Outros bancos de dados usarão uma tabela normal para armazenamento em cache do banco de dados. O código SQL pode ser customizado para bancos de dados específicos - entre em contato com os Serviços para obter assistência.

Para limpar um cache

Um ícone de documento em branco em um nó indica que seu cache está vazio. Quando o cache estiver cheio, o ícone de documento ficará verde sólido. Se desejar substituir o conteúdo do cache, primeiro você deve limpar o cache e depois executar o fluxo de dados novamente para voltar a preenchê-lo.

1. Na tela de fluxo, clique com o botão direito no nó e clique em **Cache** no menu.
2. No submenu de armazenamento em cache, clique em **Limpeza**.

Para salvar um cache

É possível salvar o conteúdo de um cache como um arquivo de dados do IBM SPSS Statistics (*.sav). Então é possível recarregar o arquivo como um cache ou configurar um nó que use o arquivo de cache como sua origem de dados. Também é possível carregar um cache que você salvou de outro projeto.

1. Na tela de fluxo, clique com o botão direito no nó e clique em **Cache** no menu.
2. No submenu de armazenamento em cache, clique em **Salvar Cache**.
3. Na caixa de diálogo Salvar Cache, navegue para a localização onde deseja salvar o arquivo de cache.
4. Insira um nome na caixa de texto Nome do Arquivo.
5. Certifique-se de que *.sav esteja selecionado na lista Arquivos do Tipo e clique em **Salvar**.

Para carregar um cache

Se você tiver salvo um arquivo de cache antes de removê-lo do nó, é possível recarregá-lo.

1. Na tela de fluxo, clique com o botão direito no nó e clique em **Cache** no menu.
2. No submenu de armazenamento em cache, clique em **Carregar Cache**.
3. Na caixa de diálogo Carregar Cache, navegue para a localização onde deseja salvar o arquivo de cache, selecione-o e clique em **Carregar**.

Visualizando dados em nós

Para assegurar que os dados estejam sendo alterados da maneira esperada conforme você constrói um fluxo, seria possível executar seus dados por meio de um nó de Tabela em cada passo significativo. Para evitar a necessidade disso, é possível gerar uma visualização a partir de cada nó que exibe uma amostra dos dados que serão criados, reduzindo assim o tempo levado para a construção de cada nó.

Para o envio de dados de nós de um nugget do modelo, a visualização mostra os campos de entrada; para um nugget do modelo ou recebimento de dados de nós do nugget (exceto nós terminais), a visualização mostra campos de entrada e gerados.

O número padrão de linhas exibido é 10; no entanto, é possível alterar isso nas propriedades do fluxo. Consulte o tópico “Configurando opções gerais para fluxos” na página 40 para obter mais informações.

No menu Gerar, é possível criar vários tipos de nós.

Bloqueando nós

Para evitar que outros usuários aditem as configurações de um ou mais nós em um fluxo, é possível conter o nó ou nós em um tipo especial de nó chamado SuperNode e depois bloquear o SuperNode aplicando proteção de senha.

Trabalhando com fluxos

Após ter conectado os nós de origem, processo e terminal na tela de fluxo, você criou um fluxo. Como uma coleção de nós, fluxos podem ser salvos, anotados e incluídos em projetos. Também é possível configurar inúmeras opções para fluxos, como otimização, configurações de data e hora, parâmetros e scripts. Essas propriedades serão discutidas nos tópicos a seguir.

No IBM SPSS Modeler, é possível usar e modificar mais de um fluxo de dados na mesma sessão do IBM SPSS Modeler. O lado direito da janela principal contém a área de janela de gerenciadores, que lhe ajuda a navegar nos fluxos, nas saídas e nos modelos atualmente abertos. Se não for possível ver a área de janela de gerenciadores, clique em **Gerenciadores** no menu Visualizar e, em seguida, clique na guia **Fluxos**.

Nessa guia, é possível:

- Acessar fluxos.
- Salvar fluxos.
- Salvar fluxos no projeto atual.
- Fechar fluxos.
- Abrir novos fluxos.
- Armazenar e recuperar fluxos a partir de um repositório do IBM SPSS Collaboration and Deployment Services (se disponível em seu site). Consulte o tópico “Sobre o IBM SPSS Collaboration and Deployment Services Repository” na página 153 para obter mais informações.

Clique com o botão direito no fluxo na guia Fluxos para acessar essas opções.

Configurando opções para fluxos

É possível especificar inúmeras opções para aplicar ao fluxo atual. Também é possível salvar essas opções como padrões para aplicar a todos os fluxos. As opções são as seguintes.

- **Geral.** Opções variadas, como símbolos e codificação de texto, para usar no fluxo. Consulte o tópico “Configurando opções gerais para fluxos” na página 40 para obter mais informações.
- **Data/Hora.** Opções relacionadas ao formato das expressões de data e hora. Consulte o tópico “Configurando opções de data e hora para fluxos” na página 41 para obter mais informações.
- **Formatos de Número.** Opções controlando o formato de expressões numéricas. Consulte o tópico “Configurando opções de formato numérico para fluxos” na página 42 para obter mais informações.
- **Otimização.** Opções para otimizar o desempenho do fluxo. Consulte o tópico “Configurando opções de otimização para fluxos” na página 42 para obter mais informações.
- **Criação de Log e Status.** Opções controlando criação de log SQL e status de registro. Consulte o tópico “Configurando opções de status de registro e criação de log SQL para fluxos” na página 44 para obter mais informações.
- **Layout.** Opções relacionadas ao layout do fluxo na tela. Consulte o tópico “Configurando opções de layout para fluxos” na página 45 para obter mais informações.
- **Servidor Analítico.** Opções relacionadas ao uso de Analytic Server com SPSS Modeler. Consulte o tópico “Propriedades de fluxo do Analytic Server” na página 45 para obter mais informações.
- **Geoespacial.** Opções relacionadas à formatação de dados geoespaciais para uso no fluxo. Consulte o tópico “Configurando opções geoespaciais para fluxos” na página 46 para obter mais informações.

Para configurar opções de fluxo

1. No menu Arquivo, clique em **Propriedades do Fluxo** (ou selecione o fluxo na guia Fluxos na área de janela dos gerenciadores, clique com o botão direito e clique em **Propriedades do Fluxo** no menu pop-up).
2. Clique na guia **Opções**.

Alternativamente, no menu Ferramentas, clique em:

Propriedades do Fluxo > Opções

Configurando opções gerais para fluxos: As opções gerais são um conjunto de opções diversas que se aplicam a vários aspectos do fluxo atual.

A seção **Configurações Básicas** inclui as opções básicas a seguir:

- **Símbolo Decimal.** Selecione uma vírgula (,) ou um ponto final (.) como um separador decimal.
- **Símbolo de Agrupamento.** Para formatos de exibição de número, selecione o símbolo usado para agrupar valores (por exemplo, a vírgula em 3.000,00). As opções incluem nenhum, ponto final, vírgula, espaço e definido por código de idioma (em cujo caso o padrão para o código de idioma atual é usado).
- **Codificação.** Especifique o método padrão de fluxo para codificação de texto. (*Nota:* Aplica-se a Var. Somente nó de origem Arquivo e nó de exportação Arquivo Simples. Nenhum outro nó usa essa configuração; a maioria dos arquivos de dados tem informações de codificação integradas.) É possível escolher o padrão do sistema ou UTF-8. O padrão do sistema é especificado no Painel de Controle do Windows ou, se estiver em execução no modo distribuído, no computador servidor. Consulte o tópico “Suporte Unicode no IBM SPSS Modeler” na página 211 para obter mais informações.
- **Avaliação de Conjunto de Regras.** Determina como modelos de conjunto de regras são avaliados. Por padrão, conjuntos de regras usam **Voto** para combinar previsões de regras individuais e determinar a previsão final. Para assegurar que os conjuntos de regras usem a primeira regra de ocorrência por padrão, selecione **Primeira Ocorrência**. Observe que essa opção não se aplica aos modelos Lista de Decisão, que sempre usam a primeira ocorrência conforme definido pelo algoritmo.

Número máximo de linhas para mostrar na Visualização de Dados. Especifique o número de linhas a ser mostrado quando uma visualização de dados for solicitada para um nó. Consulte o tópico “Visualizando dados em nós” na página 38 para obter mais informações.

Máximo de membros para campos nominais. Selecione para especificar um número máximo de membros para campos nominais (configurados) após o qual o tipo de dado do campo se torna **Sem Tipo**. Essa opção é útil quando se trabalha com grandes campos nominais. *Nota:* Quando o nível de medição de um campo é configurado para **Sem Tipo**, sua função é automaticamente configurada para **Nenhum**. Isso significa que os campos não estão disponíveis para modelagem.

Limitar tamanho da configuração para modelagem Kohonen e k-médias. Selecione para especificar um número máximo de membros para campos nominais usados em redes Kohonen e modelagem K-médias. O tamanho de configuração padrão é 20; depois disso, o campo é ignorado e um aviso é levantado, fornecendo informações sobre o campo em questão.

Observe que, para compatibilidade, essa opção também se aplica ao antigo nó Rede Neural que foi substituído na versão 14 do IBM SPSS Modeler; alguns fluxos anteriores ainda podem conter esse nó.

Atualizar nós de origem na execução. Selecione para atualizar automaticamente todos os nós de origem durante a execução do fluxo atual. Essa ação é análoga a um clique no botão **Atualizar** em um nó de origem, exceto que atualiza automaticamente todos os nós de origem (exceto nós Entrada do Usuário) para o fluxo atual.

Nota: A seleção dessa opção limpa os caches dos nós de recebimento de dados, mesmo se os dados não tiverem mudado. Entretanto, a limpeza ocorre somente uma vez por execução do fluxo, o que significa que ainda é possível usar os caches de recebimento de dados como armazenamento temporário para uma única execução. Por exemplo, digamos que você tenha configurado um midstream de cache após uma operação de derivação complexa e tenha um recebimento de dados com vários gráficos e relatórios anexados desse nó Derivar. Durante a execução do fluxo, o cache no nó Derivar será limpo e preenchido novamente, mas somente para o primeiro gráfico ou relatório. Os nós terminais subsequentes lerão dados do cache do nó Derivar.

Rótulos de valor e campo de exibição na saída. Rótulos de valor e campo de exibição em tabelas, gráficos e outra saída. Se os rótulos não existirem, os nomes de campo e os valores de dados serão exibidos. Os rótulos ficam desativados por padrão; no entanto, é possível alternar os rótulos individualmente em qualquer lugar no IBM SPSS Modeler. Também é possível escolher a exibição de rótulos na janela de saída usando um botão de alternância disponível na barra de ferramentas.



Figura 11. Ícone da barra de ferramentas usado para alternar rótulos de valor e campo

Exibir tempos de execução. Exibe tempos de execução individuais para nós de fluxo na guia Tempos de Execução após o fluxo ser executado. Consulte o tópico “Visualizando tempos de execução de nós” na página 47 para obter mais informações.

A seção **Criação de Nó Automática** inclui as seguintes opções para criar nós automaticamente em fluxos individuais. Essas opções controlam se deve-se ou não inserir os nuggets de modelagem na tela de fluxo durante a geração de novos nuggets. Por padrão, essas opções se aplicam somente aos fluxos criados na versão 16 ou posterior. No IBM SPSS Modeler Server 16 ou posterior, se você abrir um fluxo criado na versão 15 ou anterior e executar um nó de modelagem, o nugget não será colocado na tela de fluxo como nas liberações anteriores. Se você criar um novo fluxo usando IBM SPSS Modeler Server 16 ou posterior e executar um nó de modelagem, o nugget gerado será colocado na tela de fluxo. Isso acontece conforme projetado, pois, por exemplo, a opção **Criar nós de aplicação de modelo para nova saída do modelo** provavelmente quebraria os fluxos pré-16 executados em lote, no IBM SPSS Collaboration and Deployment Services e em outros ambientes nos quais a interface com o usuário do cliente IBM SPSS Modeler Server não está presente.

- **Criar nós de aplicação de modelo para nova saída do modelo.** Cria automaticamente nós de aplicação de modelo para a nova saída do modelo. Se você selecionar essa opção, também é possível escolher em **Criar ligações de atualização de modelo** se você irá configurar as ligações como ativadas, desativadas ou não irá criá-las.

Quando um aplicador de novo modelo ou nó de origem é criado, as opções de ligação nos menus suspensos controlam se as ligações de atualização entre o nó do construtor e o novo nó são criadas e, se sim, em que modo elas estão. Se as ligações forem criadas, é provável que você as queira ativadas, mas essas opções fornecem controle completo ao usuário.

- **Criar nós de origem a partir de construtores de origem.** Cria automaticamente nós de origem a partir de construtores de origem. Semelhante à opção anterior, se você selecionar essa opção, também é possível escolher no menu suspenso **Criar ligações de atualização de origem** se você irá configurar ligações de atualização como ativadas, desativadas ou não irá criá-las.

Salvar Como Padrão. As opções especificadas se aplicam somente ao fluxo atual. Clique nesse botão para configurar essas opções como padrão para todos os fluxos.

Configurando opções de data e hora para fluxos: Essas opções especificam o formato a ser usado para várias expressões de data e hora no fluxo atual.

Importar data/hora como Selecione entre usar armazenamento de data/hora para campos de data/hora ou importá-los como variáveis de sequência.

Formato de data Selecione um formato de data para ser usado para campos de armazenamento de data ou quando sequências forem interpretadas como datas pelas funções de data do CLEM.

Formato de hora Selecione um formato de hora para ser usado para campos de armazenamento de hora ou quando sequências forem interpretadas como hora pelas funções de hora do CLEM.

Dias/minutos de rolagem Para formatos de hora, selecione se diferenças de tempo negativo são interpretadas como referência a um dia ou hora anterior.

Linha de base de data (1° Jan) Selecione os anos de linha de base (sempre 1° de janeiro) para serem usados pelas funções de data do CLEM que trabalham com data única.

Início de datas de 2 dígitos Especifica o ano de corte para incluir dígitos de século para anos que estão denotados com apenas 2 dígitos. Por exemplo, especificando 1930 como ano de corte significa que 11/05/02 é no ano 2002. A mesma configuração usará o século 20 para datas após o ano 30; assim, 05/11/73 será assumido como 1973.

Fuso horário Seleciona como o fuso horário é escolhido para ser usado com a expressão `datetime_now` de CLEM.

- Se **Servidor** for selecionado, o fuso horário dependerá dos seguintes itens:
 - Se o fluxo atual usa uma origem de dados de Analytic Server, a expressão `datetime_now` usará a hora do Analytic Server; por padrão, o servidor usa a Hora Universal Coordenada.
 - Se o fluxo atual usa um nó de origem de Banco de Dados, os bancos de dados suportados usarão o SQL pushback e a expressão `datetime_now` usará a hora do bando de dados.
 - Para todos os outros fluxos, o fuso horário usa a hora do SPSS Modeler Server.
- Se **Cliente do Modelador** for selecionado, o fuso horário refletirá os detalhes de fuso horário da máquina em que o SPSS Modeler está instalado.
- Alternativamente, é possível selecionar qualquer valor de Hora Universal Coordenada para o fuso horário.

Salvar como Padrão. As opções especificadas são aplicadas apenas ao fluxo atual. Clique nesse botão para configurar essas opções como padrão para todos os fluxos.

Configurando opções de formato numérico para fluxos: Estas opções especificam o formato a ser usado para várias expressões numéricas no fluxo atual.

Formato de exibição numérico. É possível escolher entre os formatos de exibição padrão (####.###), científico (#.###E+##) ou moeda (\$###.##).

Casas decimais (padrão, científico, moeda). Para formatos de exibição numéricos, especifica o número de casas decimais a serem usadas durante a exibição ou impressão de números reais. Essa opção é especificada separadamente para cada formato de exibição.

Cálculos em. Selecione **Radianos** ou **Graus** como unidade de medida a ser usada em expressões trigonométricas do CLEM. Consulte o tópico “Funções trigonométricas” na página 131 para obter mais informações.

Salvar Como Padrão. As opções especificadas se aplicam somente ao fluxo atual. Clique nesse botão para configurar essas opções como padrão para todos os fluxos.

Configurando opções de otimização para fluxos: É possível usar configurações de Otimização para otimizar o desempenho do fluxo. Observe que as configurações de desempenho e otimização no IBM SPSS Modeler Server (se usado) substituem quaisquer configurações equivalentes no cliente. Se essas configurações estiverem desativadas no servidor, o cliente não poderá ativá-las. Mas, se elas estiverem ativadas no servidor, o cliente pode optar por desativá-las.

Nota: A modelagem da base de dados e a otimização de SQL requerem que a conectividade do IBM SPSS Modeler Server esteja ativada no computador do IBM SPSS Modeler. Com essa configuração ativada, é possível acessar os algoritmos de banco de dados, realizar SQL pushback diretamente do IBM SPSS Modeler e acessar IBM SPSS Modeler Server. Para verificar o atual status da licença, escolha o seguinte no menu do IBM SPSS Modeler.

Ajuda > Sobre > Detalhes Adicionais

Se a conectividade estiver ativada, você verá a opção **Ativação do Servidor** na guia Status da Licença.

Consulte o tópico “Conectando a IBM SPSS Modeler Server” na página 10 para obter mais informações.

Nota: Se a otimização e o SQL pushback são suportados depende do tipo de banco de dados em uso. Para obter as informações mais recentes sobre quais bancos de dados e drivers ODBC são suportados e testados para uso com IBM SPSS Modeler 17.1, consulte o site de Suporte corporativo em <http://www.ibm.com/support>.

Ativar regravação de fluxo. Selecione esta opção para ativar a regravação de fluxo no IBM SPSS Modeler. Há quatro tipos de regravação disponíveis, e é possível selecionar um ou mais deles. A regravação de fluxo reordena os nós em um fluxo por trás dos cenários para uma operação mais eficiente, sem alterar a semântica do fluxo.

- **Otimizar geração de SQL.** Esta opção permite que os nós sejam reordenados dentro do fluxo, de modo que mais operações possam retroceder usando geração de SQL para execução no banco de dados. Ao localizar um nó que não pode ser renderizado em SQL, o otimizador aguardará para ver se há algum nó de recebimento de dados que pode ser renderizado em SQL e movido com segurança para a frente do nó problemático sem afetar a semântica do fluxo. Não só o banco de dados pode executar operações com mais eficiência que o IBM SPSS Modeler, mas esses retrocessos agem para reduzir o tamanho do conjunto de dados que é retornado para o IBM SPSS Modeler para processamento. Isso, por sua vez, pode reduzir o tráfego de rede e acelerar as operações do fluxo. Observe que a caixa de seleção **Gerar SQL** deve ser selecionada para que a otimização de SQL tenha algum efeito.
- **Otimizar expressão do CLEM.** Esta opção permite que o otimizador procure expressões do CLEM que possam ser pré-processadas antes de o fluxo ser executado para aumentar a velocidade do processamento. Como um exemplo simples, se você tivesse uma expressão como $\log(\text{salário})$, o otimizador calcularia o valor do salário real e o passaria para processamento. Isso pode ser usado para melhorar o SQL pushback e o desempenho do IBM SPSS Modeler Server.
- **Otimizar execução de sintaxe.** Esse método de regravação de fluxo aumenta a eficiência das operações que incorporam mais de um nó contendo sintaxe do IBM SPSS Statistics. A otimização é obtida combinando os comandos de sintaxe em uma única operação, em vez de executar cada um como uma operação separada.
- **Otimizar outra execução.** Este método de regravação de fluxo aumenta a eficiência das operações que não podem ser delegadas ao banco de dados. A otimização é obtida reduzindo a quantidade de dados no fluxo o mais cedo possível. Enquanto mantém a integridade dos dados, o fluxo é regravado para enviar as operações por push para mais perto da origem de dados, reduzindo assim o recebimento de dados para as operações mais caras, como junções.

Ativar processamento paralelo. Durante a execução em um computador com vários processadores, esta opção permite que o sistema equilibre a carga nesses processadores, o que pode resultar em um desempenho mais rápido. O uso de diversos nós ou o uso dos seguintes nós individuais pode ser beneficiado pelo processamento paralelo: C5.0, Mesclagem (por chave), Ordem, Categorização (métodos de ranqueamento e ladrilho) e Agregado (usando um ou mais campos-chave).

Gerar SQL. Selecione esta opção para ativar geração de SQL, permitindo que operações de fluxo retrocedam para o banco de dados usando código SQL para gerar processos de execução, o que pode melhorar o desempenho. Para melhorar ainda mais o desempenho, **Otimizar Geração de SQL** também

pode ser selecionada para maximizar o número de operações que retrocedem para o banco de dados. Quando operações para um nó retrocederem para o banco de dados, o nó será destacado em púrpura quando o fluxo for executado.

- **Armazenamento em cache do banco de dados.** Para que fluxos que geram SQL sejam executados no banco de dados, dados podem ser armazenados em cache no midstream para uma tabela temporária no banco de dados em vez de para o sistema de arquivo. Quando combinado com otimização de SQL, isso pode resultar em ganhos significativos no desempenho. Por exemplo, a saída de um fluxo que mescla várias tabelas para criar uma visualização de mineração de dados pode ser armazenada em cache e reutilizada conforme necessário. Com o armazenamento em cache do banco de dados ativado, basta clicar com o botão direito em qualquer nó não terminal para armazenar dados em cache nesse ponto e, automaticamente, o cache será criado diretamente no banco de dados na próxima vez que o fluxo for executado. Isso permite que o SQL seja gerado para nós de recebimento de dados, melhorando ainda mais o desempenho. Alternativamente, essa opção pode ser desativada se for necessário, como quando políticas ou permissões impedem que dados sejam gravados no banco de dados. Se o armazenamento em cache do banco de dados ou a otimização de SQL não estiverem ativados, o cache será gravado no sistema de arquivos. Consulte o tópico “Armazenando opções em cache para nós” na página 37 para obter mais informações.
- **Usar conversão afrouxada.** Esta opção permite a conversão de dados de sequências de caracteres em números, ou de números em sequências de caracteres, se eles estiverem armazenados em um formato adequado. Por exemplo, se os dados forem mantidos no banco de dados como uma sequência de caracteres, mas de fato contiverem um número significativo, os dados poderão ser convertidos para uso quando ocorrer um retrocesso.

Nota: Devido a pequenas diferenças na implementação de SQL, os fluxos executados em um banco de dados podem retornar resultados um pouco diferentes daqueles retornados durante a execução no IBM SPSS Modeler. Por motivos semelhantes, essas diferenças também podem variar, dependendo do fornecedor do banco de dados.

Salvar Como Padrão. As opções especificadas se aplicam somente ao fluxo atual. Clique nesse botão para configurar essas opções como padrão para todos os fluxos.

Configurando opções de status de registro e criação de log SQL para fluxos: Essas configurações incluem várias opções controlando a exibição de instruções SQL geradas pelo fluxo, além da exibição do número de registros processados pelo fluxo.

Exibir SQL no log de mensagens durante a execução de fluxo. Especifica se a SQL gerada durante a execução do fluxo é passada para o log de mensagens.

Exibir detalhes de geração de SQL no log de mensagens durante a preparação do fluxo. Durante a visualização do fluxo, especifica se uma visualização da SQL que seria gerada é passada para o log de mensagens.

Exibir SQL. Especifica se alguma SQL que é exibida no log deve conter funções de SQL nativo ou funções de ODBC padrão na forma {fn FUNC(...)}, conforme gerado por SPSS Modeler. O último conta com a funcionalidade do driver ODBC que pode não estar implementada.

Reformatar SQL para capacidade de leitura melhorada. Especifica se a SQL exibida no log deve ser formatada para capacidade de leitura.

Mostrar status para registros. Especifica quando os registros devem ser relatados conforme eles chegam aos nós terminais. Especifique um número que seja usado para atualizar o status a cada *N* registros.

Salvar Como Padrão. As opções especificadas se aplicam somente ao fluxo atual. Clique nesse botão para configurar essas opções como padrão para todos os fluxos.

Configurando opções de layout para fluxos: Estas configurações fornecem inúmeras opções relacionadas à exibição e ao uso da tela de fluxo.

Largura mínima da tela de fluxo. Especifique a largura mínima da tela de fluxo em pixels.

Altura mínima da tela de fluxo. Especifique a altura mínima da tela de fluxo em pixels.

Taxa de rolagem do fluxo. Especifique a taxa de rolagem para a tela de fluxo para controlar a rapidez com que a área de janela da tela de fluxo rola quando um nó está sendo arrastado de um local para outro na tela. Os números mais altos especificam uma taxa de rolagem mais rápida.

Nome máximo do ícone. Especifique um limite nos caracteres para os nomes dos nós na tela de fluxo.

Tamanho do ícone. Selecione uma opção para escalar a visualização de fluxo inteira para um de vários tamanhos entre 8% e 200% do tamanho do ícone padrão.

Tamanho da célula de grade. Selecione um tamanho de célula de grade da lista. Esse número é usado para alinhar nós na tela de fluxo usando uma grade invisível. O tamanho da célula de grade padrão é 0,25.

Ajustar à grade. Selecione para alinhar ícones para um padrão de grade invisível (selecionado por padrão).

Localização do ícone gerado. Escolha onde na tela colocar ícones para nós gerados a partir dos nuggets de modelo. O padrão é superior esquerdo.

Salvar Como Padrão. As opções especificadas se aplicam somente ao fluxo atual. Clique nesse botão para configurar essas opções como padrão para todos os fluxos.

Propriedades de fluxo do Analytic Server: Estas configurações fornecem diversas opções para trabalhar com o Analytic Server.

Número máximo de registros a processar fora do Analytic Server

Especifique o número máximo de registros a ser importado no servidor SPSS Modeler a partir de uma origem de dados do Analytic Server.

Notificação quando um nó não pode ser processado no Analytic Server

Esta configuração determina o que acontece quando um fluxo que seria enviado para o Analytic Server contém um nó que não pode ser processado no Analytic Server. Especifique se um aviso será emitido e continue processando o fluxo, ou lance um erro e pare o processamento.

Configurações de Armazenamento de Modelo de Divisão

Armazene os modelos de divisão por referência no Analytic Server quando o tamanho do modelo (MB) for excedido

Nuggets do modelo geralmente são armazenados como parte do fluxo. Modelos de divisão com muitas divisões podem produzir nuggets grandes, e mover o nugget para frente e para trás entre o fluxo e o Analytic Server pode causar impacto no desempenho. Como uma solução, quando um modelo de divisão exceder o tamanho especificado, ele será armazenado no Analytic Server, e o nugget no SPSS Modeler conterá uma referência ao modelo.

Pasta padrão para armazenar modelos por referência no Analytic Server quando a execução for concluída

Especifique o caminho padrão no qual deseja armazenar modelos de divisão no Analytic Server. O caminho deve começar com um nome do projeto válido do Analytic Server.

Pasta para armazenar modelos promovidos

Especifique o caminho padrão no qual deseja armazenar modelos "promovidos". Um modelo promovido não é limpo quando a sessão do SPSS Modeler termina.

Configurando opções geoespaciais para fluxos: Qualquer campo geoespacial, independentemente de ser uma forma, uma coordenada ou um valor de eixo único (como x ou y, ou latitude e longitude) tem um sistema de coordenadas associado. Esse sistema de coordenadas configura atributos como o ponto de origem (0,0) e as unidades associadas aos valores.

Há inúmeros sistemas de coordenadas e dois tipos: Geográfico e Projetado. Todas as funções espaciais no SPSS Modeler podem ser usadas somente com um sistema de coordenadas projetadas.

Devido à natureza dos sistemas de coordenadas, a mesclagem ou anexação de dados de duas origens de dados geoespaciais separadas requer que as origens usem o mesmo sistema de coordenadas. Por isso, você deve especificar uma configuração de coordenadas para quaisquer dados geoespaciais usados no fluxo.

Os dados são reprojeto automaticamente para usar o sistema de coordenadas de fluxo escolhido nas seguintes situações:

- Para funções espaciais (como area, closeto, within), o parâmetro passado para a função é automaticamente reprojeto; no entanto, os dados de linha originais são mantidos inalterados.
- Durante o uso dos nós de construção ou escoragem (nugget) no Spatio-Temporal Prediction (STP), o campo de localização é automaticamente reprojeto. Durante a escoragem, a localização que é revelada do nugget é a localização original.
- Durante o uso do nó Visualização de Mapa.

Sistema de coordenadas do fluxo. Disponível somente se você marcar a caixa de seleção. Clique em **Mudança** para exibir uma lista de Sistemas de Coordenadas Projetadas disponíveis e selecionar o desejado para usar para o fluxo atual.

Salvar Como Padrão. O sistema de coordenadas selecionado se aplica somente ao fluxo atual. Para selecionar o sistema como padrão para todos os fluxos, clique nesse botão.

Selecionando sistemas de coordenadas geoespaciais: Todas as funções espaciais no SPSS Modeler podem ser usadas somente com um sistema de coordenadas projetadas.

A caixa de diálogo Selecionar Sistema de Coordenadas de Fluxo contém uma lista de todos os sistemas de coordenadas projetadas que podem ser selecionados para quaisquer dados geoespaciais usados em um fluxo.

As informações a seguir são listadas para cada sistema de coordenadas.

- **WKID** O Well Known ID que é exclusivo para cada sistema de coordenadas.
- **Nome** O nome do sistema de coordenadas.
- **Unidades** A unidade de medição associada ao sistema de coordenadas.

Além da lista de todos os sistemas de coordenadas, a caixa de diálogo tem um controle de **Filtragem**. Se você souber todo ou parte do nome do sistema de coordenadas requerido, digite-o no campo **Nome** na parte inferior da caixa de diálogo. A lista de sistemas de coordenadas da qual é possível escolher é filtrada automaticamente para mostrar somente os sistemas com nomes que contêm o texto que você inseriu.

Visualizando mensagens de operação de fluxo

Mensagens referentes a operações de fluxo, como execução, otimização e tempo decorrido para construção e avaliação de modelo, podem ser facilmente visualizadas usando a guia Mensagens na caixa de diálogo de propriedades do fluxo. Mensagens de erro também são relatadas nessa tabela.

Para visualizar mensagens de fluxo

1. No menu Arquivo, clique em **Propriedades do Fluxo** (ou selecione o fluxo na guia Fluxos na área de janela dos gerenciadores, clique com o botão direito e clique em **Propriedades do Fluxo** no menu pop-up).
2. Clique na guia **Mensagens**.

Alternativamente, no menu Ferramentas, clique em:

Propriedades do Fluxo > Mensagens

Além das mensagens referentes às operações de fluxo, mensagens de erro são relatadas aqui. Quando o fluxo em execução é finalizado devido a um erro, essa caixa de diálogo é aberta para a guia Mensagens com a mensagem de erro visível. Além disso, o nó com erros é destacado em vermelho na tela de fluxo.

Se a otimização de SQL e as opções de criação de log estiverem ativadas na caixa de diálogo Opções do Usuário, as informações no SQL gerado também serão exibidas. Consulte o tópico “Configurando opções de otimização para fluxos” na página 42 para obter mais informações.

É possível salvar mensagens relatadas aqui para um fluxo clicando em **Salvar Mensagens** na lista suspensa do botão Salvar (à esquerda, abaixo da guia Mensagens). Também é possível limpar todas as mensagens para um determinado fluxo clicando em **Limpar todas as Mensagens** na lista do botão Salvar.

Observe que tempo de CPU é a quantidade de tempo que o processo do servidor está utilizando CPU. Tempo decorrido é o tempo total entre o início da execução e o fim da execução, portanto, ele também inclui coisas como transferência de arquivos e renderização de saídas. O tempo de CPU pode ser mais que o tempo decorrido quando um fluxo está alavancando várias CPUs (execução paralela). Quando um fluxo for totalmente enviado de volta para o banco de dados sendo usado como uma origem de dados, o tempo de CPU será zero.

Visualizando tempos de execução de nós

Na guia Mensagens, é possível escolher exibir Tempos de Execução, onde é possível ver os tempos de execução individuais para todos os nós no fluxo que são executados no IBM SPSS Modeler Server. Observe que os itens podem não ser exatos para fluxos executados em outras áreas, como R ou Analytic Server.

Nota: Para essa variável funcionar, a caixa de seleção **Exibir Tempos de Execução** deve ser selecionada na configuração **Geral** da guia **Opções**.

Na tabela de tempos de execução de nó, as colunas são as seguintes. Clique em um título da coluna para ordenar as entradas em ordem crescente e decrescente (por exemplo, para ver quais nós têm os tempos de execução mais longos).

Nó Terminal. O identificador da ramificação à qual o nó pertence. O identificador é o nome do nó terminal no final da ramificação.

Rótulo do Nó. O nome do nó ao qual o tempo de execução se refere.

ID do Nó. O identificador exclusivo do nó ao qual o tempo de execução se refere. Esse identificador é gerado pelo sistema quando o nó é criado.

Tempo(s) de Execução. O tempo em segundos levado para a execução desse nó.

Configurando parâmetros de sessão e fluxo

Parâmetros podem ser definidos para uso em expressões do CLEM e em scripts. Na verdade, eles são variáveis definidas pelo usuário que são salvas e persistem no fluxo, sessão ou SuperNode atuais e que

podem ser acessadas a partir da interface com o usuário, bem como por meio de scripts. Se você salvar um fluxo, por exemplo, quaisquer parâmetros configurados para esse fluxo também serão salvos. (Isso os distingue das variáveis do script local, que podem ser usadas somente no script no qual elas são declaradas.) Geralmente os parâmetros são usados em scripts para controlar o comportamento do script, fornecendo informações sobre campos e valores que não precisam ser codificados permanentemente no script.

O escopo do parâmetro depende de onde ele está configurado:

- Parâmetros de fluxo podem ser configurados em um script de fluxo ou na caixa de diálogo de propriedades do fluxo e estão disponíveis para todos os nós no fluxo. Eles são exibidos na lista Parâmetros no construtor de expressões.
- Parâmetros de sessão podem ser configurados em um script independente ou na caixa de diálogo de parâmetros de sessão. Eles estão disponíveis para todos os fluxos usados na sessão atual (todos os fluxos listados na guia Fluxos na área de janela de gerenciadores).

Parâmetros também podem ser configurados para SuperNodes, em cujo caso ficam visíveis somente para os nós contidos nesse SuperNode.

Para Configurar Parâmetros de Sessão e Fluxo por Meio da Interface com o Usuário

1. Para configurar parâmetros de fluxo, no menu principal, clique em:
Ferramentas > Propriedades do Fluxo > Parâmetros
2. Para configurar parâmetros de sessão, clique em **Configurar Parâmetros da Sessão** no menu Ferramentas.

Prompt? Marque esta caixa se desejar que o usuário receba um prompt no tempo de execução para inserir um valor para esse parâmetro.

Nome. Nomes de parâmetros são listados aqui. É possível criar um novo parâmetro inserindo um nome nesse campo. Por exemplo, para criar um parâmetro para temperatura mínima, você poderia digitar minvalue. Não inclua o prefixo \$P- que denota um parâmetro nas expressões do CLEM. Esse nome também é usado para exibição no construtor de expressões do CLEM.

Nome longo. Lista o nome descritivo para cada parâmetro criado.

Armazenamento. Selecione um tipo de armazenamento da lista. Armazenamento indica como valores de dados são armazenados no parâmetro. Por exemplo, quando você trabalha com valores contendo zeros à esquerda que deseja preservar (como 008), é necessário selecionar **Sequência de Caracteres** como o tipo de armazenamento. Caso contrário, os zeros serão removidos do valor. Os tipos de armazenamento disponíveis são sequência de caracteres, número inteiro, real, hora, data e registro de data e hora. Para parâmetros de data, observe que os valores devem ser especificados usando uma notação padrão ISO, conforme mostrado no parágrafo seguinte.

Valor. Lista o valor atual de cada parâmetro. Ajuste o parâmetro conforme necessário. Observe que para parâmetros de data, os valores devem ser especificados na notação padrão ISO (ou seja, YYYY-MM-DD). Datas especificadas em outros formatos não são aceitas.

Tipo (opcional). Se você pretende implementar o fluxo em um aplicativo externo, selecione um nível de medição da lista. Caso contrário, é recomendável deixar a coluna *Tipo* no estado em que se encontra. Se desejar especificar restrições de valor para o parâmetro, como limites mínimo e máximo para um intervalo numérico, selecione **Especificar** na lista.

Observe que as opções nome longo, armazenamento e tipo podem ser configuradas para parâmetros somente por meio da interface com o usuário. Essas opções não podem ser configuradas usando scripts.

Clique nas setas à direita para mover o parâmetro selecionado mais para cima ou para baixo na lista de parâmetros disponíveis. Use o botão Excluir (marcado com um X) para remover o parâmetro selecionado.

Especificando prompts de tempo de execução para valores de parâmetro

Se você tiver fluxos nos quais talvez seja necessário inserir valores diferentes para o mesmo parâmetro em ocasiões diferentes, é possível especificar prompts de tempo de execução para um ou mais valores de parâmetro de fluxo ou sessão.

Parâmetros. (Opcional) Insira um valor para o parâmetro ou deixe o valor padrão, se houver algum.

Desative esses prompts. Selecione essa caixa se não quiser que esses prompts sejam exibidos quando você executar o fluxo. É possível fazer com que eles sejam exibidos novamente marcando a caixa de seleção **Prompt?** nas propriedades do fluxo ou na caixa de diálogo de propriedades do fluxo onde os parâmetros foram definidos. Consulte o tópico “Configurando parâmetros de sessão e fluxo” na página 47 para obter mais informações.

Especificando restrições de valor para um tipo de parâmetro

É possível criar restrições de valor para um parâmetro disponível durante a implementação do fluxo em um aplicativo externo que lê fluxos de modelagem de dados. Essa caixa de diálogo permite especificar valores disponíveis para um usuário externo executando o fluxo. Dependendo do tipo de dados, as restrições de valor variam dinamicamente na caixa de diálogo. As opções mostradas aqui são idênticas às opções disponíveis para valores do nó Tipo.

Tipo. Exibe o nível de medição atualmente selecionado. É possível alterar esse valor para refletir a maneira como você pretende usar o parâmetro no IBM SPSS Modeler.

Armazenamento. Exibe o tipo de armazenamento se ele for conhecido. Os tipos de armazenamento não são afetados pelo nível de medição (contínuo, nominal ou flag) que você escolher para trabalhar no IBM SPSS Modeler. É possível alterar o tipo de armazenamento na guia Parâmetros principal.

A metade inferior da caixa de diálogo muda dinamicamente, dependendo do nível de medição selecionado no campo **Tipo**.

Nível de Medição Contínuos

Inferior. Especifique um limite inferior para os valores de parâmetro.

Superior. Especifique um limite superior para os valores de parâmetro.

Rótulos. É possível especificar rótulos para qualquer valor de um campo de intervalo. Clique no botão **Rótulos** para abrir uma caixa de diálogo separada para especificar rótulos de valores.

Níveis de Medição Nominais

Valores. Esta opção permite especificar valores para um parâmetro que serão usados como um campo nominal. Valores não serão forçados no fluxo do IBM SPSS Modeler, mas serão usados em uma lista suspensa para aplicativos de implementação externos. Usando os botões de seta e exclusão, é possível modificar valores existentes, bem como reordenar ou excluir valores.

Nível de Medição Flag

True. Especifique um valor de flag para o parâmetro quando a condição for atendida.

False. Especifique um valor de flag para o parâmetro quando a condição não for atendida.

Rótulos. É possível especificar rótulos para os valores de um campo de flag.

Opções de Implementação do Fluxo

A guia Implementação da caixa de diálogo de propriedades do fluxo permite especificar opções para implementar o fluxo como um cenário dentro do IBM SPSS Collaboration and Deployment Services com os propósitos de atualização de modelo, planejamento de tarefa automatizado ou uso adicional pelo IBM Analytical Decision Management. Todos os fluxos requerem uma ramificação de escoragem designada antes de serem implementados; as opções e os requisitos adicionais dependem do tipo de implementação. Consulte o tópico “Armazenando e implementando objetos do repositório” na página 154 para obter mais informações.

Loop da execução de fluxos

Usando a guia Execução na caixa de diálogo de propriedades do fluxo, é possível configurar condições de loop para automatizar tarefas repetitivas no fluxo atual.

Após a configuração dessas condições, é possível usá-las como uma introdução aos scripts, já que elas preenchem a janela de scripts com scripts básicos para seu fluxo que, depois, pode ser modificado - talvez para você usar como base para a construção de scripts melhores. Consulte o tópico “Funções globais” na página 150 para obter mais informações.

Para configurar um loop para um fluxo

1. No menu Arquivo, clique em **Propriedades do Fluxo** (ou selecione o fluxo na guia Fluxos na área de janela dos gerenciadores, clique com o botão direito e clique em **Propriedades do Fluxo** no menu pop-up).
2. Clique na guia **Execução**.
3. Selecione o modo de execução **Loop / Execução Condicional**.
4. Clique na guia **Loop**.

Alternativamente, no menu Ferramentas, clique em:

Propriedades do Fluxo > Execução

Como uma alternativa adicional, clique com o botão direito no nó e, no menu de contexto, clique em:

Loop / Execução Condicional > Editar Configurações de Loop

Iteração. Não é possível editar esse valor de número da linha, mas é possível incluir, excluir ou mover uma iteração para cima ou para baixo usando os botões à direita da tabela.

Cabeçalhos da tabela. Refletem a chave de iteração e quaisquer variáveis de iteração criadas durante a configuração do loop.

Visualizar valores globais para fluxos

Usando a guia Globais na caixa de diálogo de propriedades do fluxo, é possível visualizar os valores globais configurados para o fluxo atual. Valores globais são criados usando um nó Configurar Globais para determinar estatísticas como média, soma ou desvio padrão para os campos selecionados.

Após o nó Configurar Globais ser executado, esses valores serão disponibilizados para uma variedade de usos em operações de fluxo. Consulte o tópico “Funções globais” na página 150 para obter mais informações.

Para visualizar valores globais para um fluxo

1. No menu Arquivo, clique em **Propriedades do Fluxo** (ou selecione o fluxo na guia Fluxos na área de janela dos gerenciadores, clique com o botão direito e clique em **Propriedades do Fluxo** no menu pop-up).
2. Clique na guia **Globais**.

Alternativamente, no menu Ferramentas, clique em:

Propriedades do Fluxo > Globais

Globais disponíveis. Globais disponíveis são listados nesta tabela. Não é possível editar valores globais aqui, mas é possível limpar todos os valores globais para um fluxo usando o botão Limpar Todos os Valores à direita da tabela.

Procurando nós em um fluxo

É possível procurar por nós em um fluxo especificando inúmeros critérios de procura, como nome do nó, categoria e identificador. Essa variável pode ser útil principalmente para fluxos complexos contendo um grande número de nós.

Para procurar nós em um fluxo

1. No menu Arquivo, clique em **Propriedades do Fluxo** (ou selecione o fluxo na guia Fluxos na área de janela dos gerenciadores, clique com o botão direito e clique em **Propriedades do Fluxo** no menu pop-up).
2. Clique na guia **Procura**.

Alternativamente, no menu Ferramentas, clique em:

Propriedades do Fluxo > Procurar

É possível especificar mais de uma opção para limitar a procura, exceto que a procura por ID do nó (usando o campo **ID é igual a**) exclui as outras opções.

Rótulo do nó contém. Marque esta caixa e insira todo ou parte do rótulo de um nó para procurar um nó específico. As procuras não fazem distinção entre maiúsculas e minúsculas, e várias palavras são tratadas como uma única parte de texto.

Categoria do nó. Marque esta caixa e clique em uma categoria na lista para procurar por um tipo de nó específico. **Nó de Processo** significa um nó da guia Operações de Registro ou Operações de Campo da paleta de nós; **Aplicar Nó de Modelo** se refere a um nugget de modelo.

Palavras-chave incluem. Marque esta caixa e insira uma ou mais palavras-chave completas para procurar campos com esse texto no campo Palavras-chave na guia Anotações da caixa de diálogo do nó. O texto da palavra-chave inserido deve ser uma correspondência exata. Separe palavras-chave múltiplas com pontos e vírgulas para procurar alternativas (por exemplo, a inserção de `proton;neutron` localizará todos os nós com qualquer uma dessas palavras-chave. Consulte o tópico “Anotações” na página 59 para obter mais informações.

Anotação contém. Marque esta caixa e insira uma ou mais palavras para procurar por nós que contenham esse texto na área de texto principal na guia Anotações da caixa de diálogo do nó. As procuras não fazem distinção entre maiúsculas e minúsculas, e várias palavras são tratadas como uma única parte de texto. Consulte o tópico “Anotações” na página 59 para obter mais informações.

Gera campo chamado. Marque esta caixa e insira o nome de um campo gerado (por exemplo, `$C-Drug`). É possível usar essa opção para procurar nós de modelagem que gerem um campo específico. Insira somente um nome de campo, que deve ser uma correspondência exata.

ID é igual a. Marque esta caixa e insira um ID de nó para procurar um nó específico com esse identificador (a seleção dessa opção desativa todas as opções precedentes. IDs de nó são designados pelo sistema quando o nó é criado e podem ser usados para referenciar o nó para os propósitos de script ou automação. Insira somente um ID de nó, que deve ser uma correspondência exata. Consulte o tópico “Anotações” na página 59 para obter mais informações.

Procura em SuperNodes. Esta caixa é marcada por padrão, o que significa que a procura é executada nos nós dentro e fora de SuperNodes. Limpe essa caixa se desejar executar a procura somente em nós fora de SuperNodes, no nível superior do fluxo.

Localizar. Quando você tiver especificado todas as opções desejadas, clique neste botão para iniciar a procura.

Nós que correspondem às opções especificadas são listados na parte inferior da caixa de diálogo. Selecione um nó na lista para destacá-lo na tela de fluxo.

Renomeando fluxos

Usando a guia Anotações na caixa de diálogo de propriedades de fluxo, é possível incluir anotações descritivas para um fluxo e criar um nome customizado para o fluxo. Essas opções são úteis principalmente durante a geração de relatórios para fluxos incluídos na área de janela do projeto. Consulte o tópico “Anotações” na página 59 para obter mais informações.

Descrições de fluxo

Para cada fluxo criado, o IBM SPSS Modeler produz uma descrição do fluxo contendo informações sobre o conteúdo do fluxo. Isso pode ser útil se você estiver tentando ver o que um fluxo faz, mas não tiver o IBM SPSS Modeler instalado, por exemplo, ao acessar um fluxo por meio do IBM SPSS Collaboration and Deployment Services.

A descrição do fluxo é exibida em forma de um documento HTML consistindo em inúmeras seções.

Informações Gerais do Fluxo

Esta seção contém o nome do fluxo, com os detalhes de quando o fluxo foi criado e salvo pela última vez.

Descrição e Comentários

Esta seção inclui quaisquer:

- Anotações de fluxo (consulte “Anotações” na página 59)
- Comentários não conectados a nós específicos
- Comentários conectados a nós nas ramificações de modelagem e escoragem do fluxo

Informações de pontuação

Esta seção contém informações sob vários títulos relacionados à ramificação de escoragem do fluxo.

- **Comentários.** Inclui comentários vinculados aos nós na ramificação de escoragem.
- **Entradas.** Lista os campos de entrada com seus tipos de armazenamento (por exemplo, sequência de caracteres, número inteiro, real, entre outros).
- **Saídas.** Lista os campos de saída, incluindo os campos adicionais geradas pelo nó de modelagem, juntamente com seus tipos de armazenamento.
- **Parâmetros.** Lista quaisquer parâmetros relativos à ramificação de escoragem do fluxo e que podem ser visualizados ou editados cada vez que o modelo é escorado. Esses parâmetros são identificados quando você clica no botão **Parâmetros de Escoragem** na guia **Implementação** da caixa de diálogo propriedades do fluxo.
- **Nó de Modelo.** Mostra o nome e o tipo do modelo (por exemplo, Rede Neural, Árvore C&R e assim por diante). Esse é o nugget do modelo selecionado para o campo **Nó de Modelo** na guia **Implementação** da caixa de diálogo propriedades do fluxo.
- **Detalhes do Modelo.** Mostra detalhes do nugget do modelo identificado sob o título anterior. Onde for possível, os gráficos de avaliação e importância do preditor para o modelo serão incluídos.

Informações de modelagem

Contém informações relativas à ramificação modelagem do fluxo.

- **Comentários.** Lista quaisquer comentários ou anotações que estão conectados aos nós na ramificação de modelagem.
- **Entradas.** Lista os campos de entrada com seu papel na ramificação de modelagem (em forma de valor do papel do campo, por exemplo, Entrada, Resposta, Divisão, entre outros).
- **Parâmetros.** Lista quaisquer parâmetros relativos à ramificação de modelagem do fluxo e que podem ser visualizados ou editados cada vez que o modelo é atualizado. Esses parâmetros são identificados quando você clica no botão **Parâmetros de Construção de Modelo** na guia **Implementação** da caixa de diálogo propriedades do fluxo.
- **Nó de Modelagem.** Mostra o nome e o tipo do nó de modelagem usado para gerar ou atualizar o modelo.

Visualizando descrições de fluxos

É possível visualizar o conteúdo de uma descrição de fluxo em um navegador da web clicando em uma opção na caixa de diálogo de propriedades do fluxo. O conteúdo da descrição depende das opções especificadas na guia Implementação da caixa de diálogo. Consulte o tópico “Opções de Implementação do Fluxo” na página 166 para obter mais informações.

Para visualizar uma descrição do fluxo:

1. No menu principal do IBM SPSS Modeler, clique em:
Ferramentas > Propriedades do Fluxo > Implementação
2. Configure o tipo de implementação, o nó de escoragem designado e quaisquer parâmetros de escoragem.
3. Se o tipo de implementação for Atualização de Modelo, opcionalmente, você poderá selecionar:
 - Nó de modelagem e quaisquer parâmetros de construção de modelo
 - Nugget de modelo na ramificação de escoragem do fluxo
4. Clique no botão **Visualizar Descrição do Fluxo**.

Exportando descrições de fluxo

É possível exportar conteúdo da descrição do fluxo para um arquivo HTML.

Para exportar uma descrição de fluxo:

1. No menu principal, clique em:
Arquivo > Exportar Descrição do Fluxo
2. Insira um nome para o arquivo HTML e clique em **Salvar**.

Executando fluxos

Após você ter especificado as opções necessárias para fluxos e ter conectado os nós necessários, é possível executar o fluxo executando os dados por meio dos nós no fluxo. Há várias maneiras de se executar um fluxo dentro do IBM SPSS Modeler. É possível:

- Clique em **Executar** no menu Ferramentas.
- Clique em um dos botões **Executar...** na barra de ferramentas. Esses botões permitem executar o fluxo inteiro ou simplesmente o nó terminal selecionado. Consulte o tópico “Barra de ferramentas do IBM SPSS Modeler” na página 17 para obter mais informações.
- Execute um único fluxo de dados clicando com o botão direito em um nó terminal e clicando em **Executar** no menu pop-up.
- Execute parte de um fluxo de dados clicando com o botão direito em qualquer nó não terminal e clicando em **Executar Daqui** no menu pop-up. Isso faz com que somente aquelas operações após o nó selecionado sejam executadas.

Para parar a execução de um fluxo em andamento, é possível clicar no botão Parar vermelho na barra de ferramentas ou clicar em **Parar Execução** no menu Ferramentas.

Se algum fluxo levar mais tempo do que três segundos para ser executado, a caixa de diálogo Feedback de Execução será exibida para indicar o progresso.

Alguns nós têm exibições adicionais fornecendo informações adicionais sobre a execução do fluxo. Elas são exibidas selecionando a linha correspondente na caixa de diálogo. A primeira linha é selecionada automaticamente.

Trabalhando com modelos

Se um fluxo incluir um nó de modelagem (ou seja, da guia Modelagem ou Modelagem da Base de Dados da paleta de nós), um **nugget de modelo** será criado quando o fluxo for executado. Um nugget do modelo é um contêiner para um **modelo**, ou seja, o conjunto de regras, fórmulas ou equações que permite gerar previsões com relação a seus dados de origem e que está no centro da análise preditiva.



Figura 12. Nugget do modelo

Quando você executa um nó de modelagem com sucesso, um nugget do modelo correspondente é colocado na tela de fluxo, onde é representado por um ícone em forma de diamante dourado (por isso o nome "nugget"). É possível abrir o nugget e navegar por seu conteúdo para visualizar detalhes sobre o modelo. Para visualizar as previsões, você anexa e executa um ou mais nós terminais, cuja saída apresenta as previsões em formato legível.

Um fluxo de modelagem típico consiste em duas ramificações. A **ramificação de modelagem** contém o nó de modelagem, com os nós de processamento e de origem que o precedem. A **ramificação de escoragem** é criada quando você executa o nó de modelagem e contém o nugget do modelo e o nó ou nós terminais que podem ser usados para visualizar as previsões.

Para obter mais informações, consulte o guia *Nós de Modelagem do IBM SPSS Modeler*.

Incluindo comentários e anotações em nós e fluxos

Você pode precisar descrever um fluxo para outras pessoas em sua organização. Para ajudar com isso, é possível anexar comentários explicativos aos fluxos, nós e nuggets de modelo.

Outras pessoas poderão então visualizar esses comentários na tela ou é possível imprimir uma imagem do fluxo que inclui os comentários.

É possível listar todos os comentários para um fluxo ou SuperNode, alterar a ordem dos comentários na lista, editar o texto do comentário e alterar as cores do primeiro plano e do segundo plano de um comentário. Consulte o tópico "Listando comentários do fluxo" na página 58 para obter mais informações.

Também é possível incluir notas em forma de anotações de texto em fluxos, nós e nuggets por meio da guia Anotações de uma caixa de diálogo de propriedades do fluxo, caixa de diálogo do nó ou janela do nugget de modelo. Essas notas ficam visíveis somente quando a guia Anotações está aberta, exceto que as anotações de fluxo também podem ser mostradas como comentários na tela. Consulte o tópico "Anotações" na página 59 para obter mais informações.

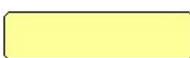
Comentários

Comentários têm forma de caixas de texto nas quais é possível inserir qualquer quantidade de texto, e é possível incluir quantos comentários você quiser. Um comentário pode ser independente (não anexado a nenhum objeto de fluxo) ou ser conectado a um ou mais nós ou nuggets de modelo no fluxo. Os comentários independentes normalmente são usados para descrever o propósito geral do fluxo; os comentários conectados descrevem o nó ou nugget ao qual estão conectados. Nós e nuggets podem ter mais de um comentário conectado, e o fluxo pode ter qualquer número de comentários independentes.

Nota: Também é possível mostrar anotações de fluxo como comentários na tela, embora elas não possam ser anexadas a nós ou nuggets. Consulte o tópico “Convertendo anotações em comentários” na página 58 para obter mais informações.

A aparência da caixa de texto muda para indicar o atual modo do comentário (ou anotação mostrada como um comentário), como mostra a tabela a seguir.

Tabela 3. Modos de caixa de texto de comentário e anotação

Caixa de texto do comentário	Caixa de texto da anotação	Modo	Indica	Obtido ao...
		Editar	Comentário é aberto para edição.	Criar um novo comentário ou anotação ou dar um clique duplo em um existente.
		Última seleção	Comentário pode ser movido, redimensionado ou excluído.	Clicar no segundo plano do fluxo após a edição ou clicar em um comentário ou anotação existente.
		Visualizar	Edição está concluída.	Clicar em outro nó, comentário ou anotação após a edição.

Quando você cria um novo comentário independente, ele é exibido inicialmente no canto superior esquerdo da tela de fluxo.

Se você estiver anexando um comentário a um nó ou nugget, o comentário será exibido inicialmente acima do objeto de fluxo ao qual está anexado.

A caixa de texto tem a cor branca para mostrar que um texto pode ser inserido. Quando você tiver inserido o texto, clique fora da caixa de texto. O segundo plano do comentário muda para amarelo para mostrar que a entrada de texto foi concluída. O comentário permanece selecionado, o que permite que ele possa ser movido, redimensionado ou excluído.

Quando você clicar novamente, as bordas mudarão para linhas sólidas para mostrar que a edição foi concluída.

Um clique duplo em um comentário muda a caixa de texto para o modo de edição--o segundo plano muda para branco e o texto do comentário pode ser editado.

Também é possível anexar comentários a SuperNodes.

Operações envolvendo comentários: É possível executar inúmeras operações nos comentários. É possível:

- Incluir um comentário independente
- Anexar um comentário a um nó ou nugget
- Editar um comentário

- Redimensionar um comentário
- Mover um comentário
- Desconectar um comentário
- Excluir um comentário
- Mostrar ou ocultar todos os comentários para um fluxo

Para incluir um comentário independente

1. Assegure que nada esteja selecionado no fluxo.
2. Execute uma das seguintes ações:
 - No menu principal, clique em:
Inserir > Novo Comentário
 - Clique com o botão direito no segundo plano do fluxo e clique em **Novo Comentário** no menu pop-up.
 - Clique no botão **Novo Comentário** na barra de ferramentas.
3. Insira o texto do comentário (ou cole-o da área de transferência).
4. Clique em um nó no fluxo para salvar o comentário.

Para anexar um comentário a um nó ou nugget

1. Selecione um ou mais nós ou nuggets na tela de fluxo.
2. Execute uma das seguintes ações:
 - No menu principal, clique em:
Inserir > Novo Comentário
 - Clique com o botão direito no segundo plano do fluxo e clique em **Novo Comentário** no menu pop-up.
 - Clique no botão **Novo Comentário** na barra de ferramentas.
3. Insira um texto de comentário.
4. Clique em outro nó no fluxo para salvar o comentário.
Alternativamente, é possível:
5. Inserir um comentário independente (consulte a seção anterior).
6. Execute uma das seguintes ações:
 - Selecione o comentário, pressione F2 e, em seguida, selecione o nó ou nugget.
 - Selecione o nó ou nugget, pressione F2 e, em seguida, selecione o comentário.
 - (Somente mouses com três botões) Mova o ponteiro do mouse sobre o comentário, mantenha o botão do meio pressionado, arraste o ponteiro do mouse sobre o nó ou nugget e solte o botão.

Para anexar um comentário a um nó ou nugget adicional

Se um comentário já estiver anexado a um nó ou nugget, ou se ele estiver atualmente no nível do fluxo e você desejar anexá-lo a um nó ou nugget adicional, execute o seguinte:

- Selecione o comentário, pressione F2 e, em seguida, selecione o nó ou nugget.
- Selecione o nó ou nugget, pressione F2 e, em seguida, selecione o comentário.
- (Somente mouses com três botões) Mova o ponteiro do mouse sobre o comentário, mantenha o botão do meio pressionado, arraste o ponteiro do mouse sobre o nó ou nugget e solte o botão.

Para editar um comentário existente

1. Execute uma das seguintes ações:
 - Dê um clique duplo na caixa de texto do comentário.
 - Selecione a caixa de texto e pressione Enter.

- Clique com o botão direito na caixa de texto para exibir seu menu e clique em Editar.
2. Edite o texto do comentário. É possível usar teclas de atalho do Windows padrão durante a edição, por exemplo, Ctrl+C, para copiar texto. Outras opções durante a edição são listadas no menu pop-up para o comentário.
 3. Clique fora da caixa de texto uma vez para exibir os controles de redimensionamento e depois novamente para concluir o comentário.

Para redimensionar uma caixa de texto de comentário

1. Selecione o comentário para exibir os controles de redimensionamento.
2. Clique e arraste um controle para redimensionar a caixa.
3. Clique fora da caixa de texto para salvar a mudança.

Para mover um comentário existente

Se desejar mover um comentário, mas não seus objetos anexados (se houver algum), execute o seguinte:

- Mova o ponteiro do mouse sobre o comentário, mantenha pressionado o botão esquerdo do mouse e arraste o comentário para a nova posição.
- Selecione o comentário, mantenha pressionada a tecla Alt e mova o comentário usando as teclas de seta.

Se desejar mover um comentário com quaisquer nós ou nuggets aos quais o comentário está anexado:

1. Selecione todos os objetos que deseja mover.
2. Execute uma das seguintes ações:
 - Mova o ponteiro do mouse sobre um dos objetos, mantenha pressionado o botão esquerdo do mouse e arraste os objetos para a nova posição.
 - Selecione um dos objetos, mantenha pressionada a tecla Alt e mova os objetos usando as teclas de seta.

Para desconectar um comentário de um nó ou nugget

1. Selecione um ou mais comentários para serem desconectados.
2. Execute uma das seguintes ações:
 - Pressione F3.
 - Clique com o botão direito em um comentário selecionado e clique em Desconectar em seu menu.

Para excluir um comentário

1. Selecione um ou mais comentários para serem excluídos.
2. Execute uma das seguintes ações:
 - Pressione a tecla Delete.
 - Clique com o botão direito em um comentário selecionado e clique em Excluir em seu menu.

Se o comentário estava anexado a um nó ou nugget, a linha de conexão será excluída também.

Se o comentário era originalmente um fluxo ou uma anotação SuperNode que foi convertida para um comentário independente, o comentário será excluído da tela, mas seu texto ficará retido na guia Anotações para o fluxo ou SuperNode.

Para mostrar ou ocultar comentários para um fluxo

1. Execute uma das seguintes ações:
 - No menu principal, clique em:
Visualizar > Comentários
 - Clique no botão **Mostrar/Ocultar Comentários** na barra de ferramentas.

Listando comentários do fluxo: É possível visualizar uma lista de todos os comentários que foram feitos para um determinado fluxo ou SuperNode.

Nessa lista, é possível

- Alterar a ordem dos comentários
- Editar o texto do comentário
- Alterar as cores do primeiro plano e do plano de fundo de um comentário

Listando Comentários

Para listar os comentários feitos para um fluxo, execute uma das opções a seguir:

- No menu principal, clique em:
Ferramentas > Propriedades do Fluxo > Comentários
- Clique com o botão direito em um fluxo na área de janela de gerenciadores e clique em **Propriedades do Fluxo**, depois em **Comentários**.
- Clique com o botão direito no segundo plano de um fluxo na tela e clique em **Propriedades do Fluxo**, depois em **Comentários**.

Texto. O texto do comentário. Dê um clique duplo no texto para alterar o campo para uma caixa de texto editável.

Links. O nome do nó ao qual o comentário é anexado. Se esse campo estiver vazio, o comentário se aplicará ao fluxo.

Botões de posicionamento. Movem um comentário selecionado para cima ou para baixo na lista.

Cores do comentário. Para alterar a cor do plano de fundo ou do primeiro plano de um comentário, selecione o comentário, marque a caixa de seleção **Cores Customizadas** e, em seguida, selecione uma cor da lista **Segundo Plano** ou **Primeiro Plano** (ou ambas). Clique em **Aplicar** e, em seguida, clique no segundo plano do fluxo para ver o efeito da mudança. Clique em **OK** para salvar a mudança.

Convertendo anotações em comentários: Anotações feitas em fluxos ou SuperNodes podem ser convertidas em comentários.

No caso dos fluxos, a anotação é convertida em um comentário independente (ou seja, não é anexado a nenhum outro nó) na tela de fluxo.

Quando uma anotação SuperNode é convertida em um comentário, o comentário não é anexado ao SuperNode na tela de fluxo, mas fica visível quando você aumenta o zoom no SuperNode.

Para converter uma anotação de fluxo em um comentário

1. Clique em **Propriedades do Fluxo** no menu Ferramentas. (Alternativamente, é possível clicar com o botão direito em um fluxo na área de janela dos gerenciadores e clicar em **Propriedades do Fluxo**.)
2. Clique na guia **Anotações**.
3. Marque a caixa de seleção **Mostrar Anotação como Comentário**.
4. Clique em **OK**.

Para converter uma anotação SuperNode em um comentário

1. Dê um clique duplo no ícone SuperNode na tela.
2. Clique na guia **Anotações**.
3. Marque a caixa de seleção **Mostrar Anotação como Comentário**.
4. Clique em **OK**.

Anotações

Nós, fluxos e modelos podem ser anotados de inúmeras formas. É possível incluir anotações descritivas e especificar um nome customizado. Essas opções são úteis principalmente durante a geração de relatórios para fluxos incluídos na área de janela do projeto. Para nós e nuggets de modelo, também é possível incluir um texto de dica de ferramenta para ajudar a distinguir entre nós semelhantes na tela de fluxo.

Incluindo Anotações

A edição de um nó ou nugget de modelo abre uma caixa de diálogo em aba contendo uma guia Anotações usada para configurar uma variedade de opções de anotação. Também é possível abrir a guia Anotações diretamente.

1. Para anotar um nó ou nugget, clique com o botão direito no nó ou nugget na tela de fluxo e clique em **Renomear e Anotar**. A caixa de diálogo de edição abre com a guia Anotações visível.
2. Para anotar um fluxo, clique em **Propriedades do Fluxo** no menu Ferramentas. (Alternativamente, é possível clicar com o botão direito em um fluxo na área de janela dos gerenciadores e clicar em **Propriedades do Fluxo**.) Clique na guia Anotações.

Nome. Selecione **Customizado** para ajustar o nome gerado automaticamente ou para criar um nome exclusivo para o nó conforme exibido na tela de fluxo.

Texto de dica de ferramenta. (Somente para nós e nuggets de modelo) Insira o texto usado como uma dica de ferramenta na tela de fluxo. Isso é particularmente útil quando se trabalha com um grande número de nós semelhantes.

Palavras-chave. Especifique palavras-chave para serem usadas em relatórios de projeto e durante a procura por nós em um fluxo ou durante o rastreamento de objetos armazenados no repositório (consulte “Sobre o IBM SPSS Collaboration and Deployment Services Repository” na página 153). Várias palavras-chave podem ser separadas por pontos e vírgulas--por exemplo, receita; tipo de agricultura; valor da indenização. Espaços em branco no início e no final de cada palavra-chave são aparados--por exemplo, receita ; tipo de agricultura produzirá os mesmos resultados que receita;tipo de agricultura. (No entanto, espaços em branco dentro de palavras-chave não são aparados. Por exemplo, tipo de agricultura com um espaço e tipo de agricultura com dois espaços não são a mesma coisa.)

A área de texto principal pode ser usada para inserir anotações longas referentes às operações do nó ou decisões tomadas no nó. Por exemplo, quando você está compartilhando e reutilizando fluxos, é útil fazer anotações sobre decisões, como descartar um campo com inúmeros espaços em branco usando um nó de Filtro. A anotação do nó armazena essas informações com o nó. Também é possível escolher incluir essas anotações em um relatório do projeto criado a partir da área de janela do projeto. Consulte o tópico “Introdução aos Projetos” na página 177 para obter mais informações.

Mostrar anotação como comentário. (Somente para anotações de fluxo e SuperNode) Marque esta caixa para converter a anotação em um comentário independente que ficará visível na tela de fluxo. Consulte o tópico “Incluindo comentários e anotações em nós e fluxos” na página 54 para obter mais informações.

ID. Exibe um ID exclusivo que pode ser usado para referenciar o nó para o propósito de script ou automação. Esse valor é gerado automaticamente quando o nó é criado e não mudará. Observe também que, para evitar confusão com a letra "O", zeros não são usados em IDs de nós. Use o botão de cópia à direita para copiar e colar o ID nos scripts ou em onde for necessário.

Salvando fluxos de dados

Após você ter criado um fluxo, é possível salvá-lo para reutilização futura.

Para salvar um fluxo

1. No menu Arquivo, clique em **Salvar Fluxo** ou **Salvar Fluxo como**.

2. Na caixa de diálogo Salvar, navegue para a pasta em que deseja salvar o arquivo de fluxo.
3. Insira um nome para o fluxo na caixa de texto Nome do Arquivo.
4. Selecione **Incluir no Projeto** se quiser incluir o fluxo salvo no projeto atual.

Um clique em **Salvar** armazena o fluxo com a extensão **.str* no diretório especificado.

Arquivos de backup automático. Cada vez que um fluxo é salvo, a versão salva anteriormente do arquivo é automaticamente preservada como um backup, com um hífen anexado ao nome do arquivo (por exemplo, *mystream.str-*). Para restaurar a versão submetida a backup, simplesmente exclua o hífen e reabra o arquivo.

Salvando estados

Além de fluxos, é possível salvar **estados**, que incluem o diagrama de fluxo atualmente exibido e quaisquer nuggets de modelo que você criou (listados na guia Modelos na área de janela de gerenciadores).

Para salvar um estado

1. No menu Arquivo, clique em:
Estado > Salvar Estado ou Salvar Estado Como
2. Na caixa de diálogo Salvar, navegue para a pasta em que deseja salvar o arquivo de estado.

Um clique em **Salvar** armazena o estado com a extensão **.cst* no diretório especificado.

Salvando nós

Também é possível salvar um nó individual clicando nele com o botão direito na tela de fluxo e clicando em **Salvar Nó** no menu pop-up. Use a extensão do arquivo **.nod*.

Salvando vários objetos de fluxo

Ao sair do IBM SPSS Modeler com vários objetos não salvos, como fluxos, projetos ou nuggets de modelo, será solicitado que você salve antes de fechar completamente o software. Se escolher salvar itens, uma caixa de diálogo exibirá opções para salvar cada objeto.

1. Basta marcar as caixas de seleção para os objetos que deseja salvar.
2. Clique em **OK** para salvar cada objeto no local necessário.

Será aberto um prompt com a caixa de diálogo Salvar padrão para cada objeto. Após você concluir o salvamento, o aplicativo será fechado como instruído originalmente.

Salvando saída

Tabelas, gráficos e relatórios gerados a partir dos nós de saída do IBM SPSS Modeler podem ser salvos em formato de objeto de saída (**.cou*).

1. Durante a visualização da saída que você deseja salvar, nos menus da janela de saída, clique em:
Arquivo > Salvar
2. Especifique um nome e uma localização para o arquivo de saída.
3. Opcionalmente, selecione **Incluir Arquivo no Projeto** na caixa de diálogo Salvar para incluir o arquivo no projeto atual. Consulte o tópico “Introdução aos Projetos” na página 177 para obter mais informações.

Alternativamente, é possível clicar com o botão direito em qualquer objeto de saída listado na área de janela de gerenciadores e selecionar **Salvar** no menu pop-up.

Criptografando e descriptografando informações

Quando você salva um fluxo, nó, projeto, arquivo de saída ou nugget do modelo, é possível criptografá-los para evitar seu uso desautorizado. Para isso, você seleciona uma opção extra ao salvar e

inclui uma senha no item sendo salvo. Essa criptografia pode ser configurada para qualquer um dos itens salvos e inclui segurança extra; isso não será o mesmo que a criptografia SSL usada se você estiver passando arquivos entre o IBM SPSS Modeler e o IBM SPSS Modeler Server.

Quando você tenta abrir um item criptografado, é solicitado que você insira uma senha. Após a inserção da senha correta, o item é decriptografado automaticamente e aberto como usual.

Para criptografar um item

1. Na caixa de diálogo Salvar, para o item ser criptografado, clique em **Opções**. A caixa de diálogo Opções de Criptografia é aberta.
2. Selecione **Criptografar este Arquivo**.
3. Opcionalmente, para maior segurança, selecione **Máscara de Senha**. Isso exhibe qualquer coisa que você inserir como uma série de pontos.
4. Insira a senha. *Aviso:* Se você esqueceu a senha, o arquivo ou modelo não poderá ser aberto.
5. Se você selecionou **Máscara de Senha**, insira a senha novamente para confirmar que ela foi inserida corretamente.
6. Clique em **OK** para retornar à caixa de diálogo Salvar.

Nota: Se você salvar uma cópia de qualquer item protegido por criptografia, o novo item será salvo automaticamente em um formato criptografado usando a senha original, a menos que você altere as configurações na caixa de diálogo Opções de Criptografia.

Carregando arquivos

É possível recarregar inúmeros objetos salvos no IBM SPSS Modeler:

- Fluxos (.str)
- Estados (.cst)
- Modelos (.gm)
- Paleta de modelos (.gen)
- Nós (.nod)
- Saída (.cou)
- Projetos (.cpi)

Abrindo novos arquivos

Fluxos podem ser carregados diretamente no menu Arquivo.

1. No menu Arquivo, clique em **Abrir Fluxo**.

Todos os outros tipos de arquivo podem ser abertos usando os itens de submenu disponíveis no menu Arquivo. Por exemplo, para carregar um modelo, no menu Arquivo, clique em:

Modelos > Abrir Modelo ou Carregar Paleta de Modelos

Abrindo arquivos usados recentemente

Para o carregamento rápido de arquivos usados recentemente, é possível usar as opções para parte inferior do menu Arquivo.

Selecione **Fluxos Recentes**, **Projetos Recentes** ou **Estados Recentes** para expandir uma lista de arquivos usados recentemente.

Mapeando fluxos de dados

Usando a ferramenta de mapeamento, é possível conectar uma nova origem de dados a um fluxo pré-existente. A ferramenta de mapeamento não só irá configurar a conexão, mas também ajudará a especificar como os campos na nova origem substituirão aqueles no fluxo existente. Em vez de recriar um fluxo de dados inteiro para uma nova origem de dados, é possível simplesmente se conectar a um fluxo existente.

A ferramenta de mapeamento de dados permite unir dois fragmentos de fluxo e se certificar de que todos os nomes de campos (essenciais) sejam correspondentes. Essencialmente, os dados de mapeamento resultam simplesmente na criação de um novo nó de Filtro, o que faz a correspondência dos campos apropriados por meio da renomeação.

Há duas maneiras equivalentes de se mapear dados:

Selecionar o nó de substituição. Esse método começa com o nó a ser substituído. Primeiro, você clica com o botão direito no nó a ser substituído; em seguida, usando a opção **Mapeamento de Dados > Selecionar Nó de Substituição** no menu pop-up, selecione o nó pelo qual irá substituí-lo.

Mapear para. Esse método começa com o nó a ser introduzido no fluxo. Primeiro, você clica com o botão direito no nó a ser introduzido; em seguida, usando a opção **Mapeamento de Dados > Mapear para** no menu pop-up, selecione o nó com o qual deseja fazer a junção. Esse método é útil principalmente para o mapeamento para um nó terminal. *Nota:* Não é possível mapear para os nós de Mesclagem ou Anexação. Em vez disso, é necessário simplesmente conectar o fluxo ao nó de Mesclagem de maneira normal.

O mapeamento de dados é fortemente integrado à construção do fluxo. Se tentar se conectar a um nó que já tem uma conexão, você terá a opção de substituir a conexão ou mapeamento para esse nó.

Mapeando dados para um modelo

Para substituir a origem de dados para um fluxo de modelo por um novo nó de origem trazendo seus próprios dados para o IBM SPSS Modeler, é necessário usar a opção **Selecionar Nó de Substituição** no menu pop-up Mapeamento de Dados. Essa opção está disponível para todos os nós, exceto nós de Mesclagem, Agregado e todos os nós terminais. O uso da ferramenta de mapeamento de dados para executar essa ação ajuda a assegurar que os campos sejam adequadamente correspondentes entre as operações de fluxo existentes e a nova origem de dados. Os passos a seguir fornecem uma visão geral do processo de mapeamento de dados.

Passo 1: Especificar campos essenciais no nó de origem original. Para que operações de fluxo sejam executadas adequadamente, os campos essenciais devem ser especificados. Consulte o tópico “Especificando campos essenciais” na página 63 para obter mais informações.

Passo 2: Incluir uma nova origem de dados na tela de fluxo. Usando um dos nós de origem, traga novos dados de substituição.

Passo 3: Substituir o nó de origem do modelo. Usando a opção Mapeamento de Dados no menu pop-up para o nó de origem do modelo, clique em **Selecionar Nó de Substituição** e, em seguida, selecione o nó de origem para os dados de substituição.

Passo 4: Verificar campos mapeados. Na caixa de diálogo que é aberta, verifique se o software está mapeando os campos de maneira adequada da origem de dados de substituição para o fluxo. Quaisquer campos essenciais não mapeados serão exibidos em vermelho. Esses campos são usados em operações de fluxo e devem ser substituídos por um campo semelhante na nova origem de dados para que operações de recebimento de dados funcionem adequadamente. Consulte o tópico “Examinando campos mapeados” na página 63 para obter mais informações.

Após o uso da caixa de diálogo para assegurar que todos os campos essenciais sejam mapeados adequadamente, a origem de dados antiga é desconectada e a nova origem de dados é conectada ao fluxo

usando um nó de Filtro chamado *Mapear*. Esse nó de Filtro direciona o mapeamento real dos campos no fluxo. Um nó de Filtro *Remover Mapeamento* também é incluído na tela de fluxo. O nó de filtro *Remover Mapeamento* pode ser usado para inverter o mapeamento de nome de campo incluindo-o no fluxo. Isso desfará os campos mapeados, mas observe que você terá que editar quaisquer nós terminais de recebimento de dados para selecionar novamente os campos e as sobreposições.

Mapeamento entre fluxos

Assim como a conexão de nós, esse método de mapeamento de dados não requer a configuração de campos essenciais antecipadamente. Com esse método, você simplesmente se conecta de um fluxo a outro usando **Mapear para** no menu pop-up Mapeamento de Dados. Esse tipo de mapeamento de dados é útil para mapear para nós terminais e para copiar e colar entre fluxos. *Nota:* Usando a opção **Mapear para**, não é possível mapear para os nós de Mesclagem, Anexação e todos os tipos de nós de origem.

Para mapear dados entre fluxos

1. Clique com o botão direito no nó que deseja usar para se conectar ao novo fluxo.
2. No menu, clique em:
Mapeamento de Dados > Mapear para
3. Use o cursor para selecionar um nó de destino no fluxo de destino.
4. Na caixa de diálogo que é aberta, assegure que os campos estejam correspondidos de forma adequada e clique em **OK**.

Especificando campos essenciais

Durante o mapeamento para um fluxo existente, normalmente campos essenciais são especificados pelo autor do fluxo. Esses campos essenciais indicam se um determinado campo é usado nas operações de recebimento de dados. Por exemplo, o fluxo existente pode construir um modelo que usa um campo chamado *Perda de Cliente*. Nesse fluxo, *Perda de Cliente* é um campo essencial porque não seria possível construir o modelo sem ele. Da mesma forma, os campos usados em nós de manipulação, como nó de Derivação, são necessários para derivar o novo campo. A configuração desse tipo de campo explicitamente ajuda a assegurar que os campos adequados no nó de origem sejam mapeados para eles. Se campos obrigatórios não forem mapeados, você receberá uma mensagem de erro. Se você decidir que certas manipulações ou nós de saída são desnecessários, é possível excluir os nós do fluxo e remover os campos apropriados da lista Campos Essenciais.

Para configurar campos essenciais

1. Clique com o botão direito no nó de origem no fluxo de modelo que será substituído.
2. No menu, clique em:
Mapeamento de Dados > Especificar Campos Essenciais
3. Usando o Seletor de Campo, é possível incluir ou remover campos da lista. Para abrir o Seletor de Campo, clique no ícone à direita da lista de campos.

Examinando campos mapeados

Após ter selecionado o ponto em que uma origem de dados ou um fluxo de dados será mapeado para outro, uma caixa de diálogo será exibida para você selecionar campos para mapeamento ou para assegurar que o mapeamento padrão do sistema esteja correto. Se os campos essenciais tiverem sido configurados para o fluxo ou origem de dados e não forem correspondentes, eles serão exibidos em vermelho. Quaisquer campos não mapeados a partir da origem de dados passarão pelo nó de Filtro não alterado, mas observe que é possível mapear um campo não essencial também.

Original. Lista todos os campos no modelo ou fluxo existente - todos os campos que estão presentes além do recebimento de dados. Os campos da nova origem de dados serão mapeados para esses campos.

Mapeado. Lista os campos selecionados para mapeamento para campos modelo. Esses são os campos cujos nomes podem precisar de mudança para corresponderem aos campos originais usados nas operações de fluxo. Clique na célula de tabela para um campo para ativar uma lista de campos disponíveis.

Se você não tiver certeza de quais campos mapear, talvez seja útil examinar os dados de origem mais de perto antes do mapeamento. Por exemplo, é possível usar a guia Tipos no nó de origem para revisar um resumo dos dados de origem.

Dicas e atalhos

Trabalhe de maneira rápida e fácil se familiarizando com os atalhos e dicas a seguir:

- **Construir fluxos rapidamente dando um clique duplo.** Basta dar um clique duplo em um nó na paleta para incluí-lo e conectá-lo ao fluxo atual.
- **Usar combinações de teclas para selecionar nós de recebimento de dados.** Pressione Ctrl+Q e Ctrl+W para alternar a seleção de todos os nós de recebimento de dados.
- **Usar teclas de atalho para conectar e desconectar nós.** Quando um nó for selecionado na tela, pressione F2 para iniciar uma conexão, pressione Tab para se mover para o nó necessário e pressione Shift+Barra de Espaço para concluir a conexão. Pressione F3 para desconectar todas as entradas e saídas do nó selecionado.
- **Customizar a guia Paleta de Nós com seus nós favoritos.** No menu Ferramentas, clique em **Gerenciar Paletas** para abrir uma caixa de diálogo para incluir, remover ou mover os nós mostrados na Paleta de Nós.
- **Renomear nós e incluir Dicas de ferramenta.** Cada caixa de diálogo inclui uma guia Anotações na qual é possível especificar um nome customizado para nós na tela, bem como incluir Dicas de ferramenta para ajudar a organizar seu fluxo. Também é possível incluir anotações detalhadas para controlar o progresso, salvar detalhes do processo e denotar quaisquer decisões de negócios necessárias ou arquivadas.
- **Inserir valores automaticamente em uma expressão do CLEM.** Usando o construtor de expressões, acessível em várias caixas de diálogo (como aquelas para os nós de Derivação e Filtro), é possível inserir valores de campo automaticamente em uma expressão do CLEM. Clique no botão de valores no construtor de expressões para escolher entre os valores de campo existentes.



Figura 13. Botão de valores

- **Navegar para arquivos rapidamente.** Durante a navegação em arquivos em uma caixa de diálogo Abrir, use a lista Arquivos (clique no botão de diamante amarelo na parte superior da caixa de diálogo e, em seguida, no campo Verificar) para acessar os diretórios usados anteriormente, bem como diretórios padrão do IBM SPSS Modeler. Use os botões Avançar e Voltar para rolar pelos diretórios acessados.
- **Minimizar desordem na janela de saída.** É possível fechar e excluir uma saída rapidamente usando o botão X vermelho no canto superior direito de todas as janelas de saída. Isso permite que você mantenha somente os resultados mais promissores e interessantes na guia Saída da área de janela de gerenciadores.

Um intervalo completo de atalhos de teclado está disponível para o software. Consulte o tópico “Teclado de acessibilidade” na página 200 para obter mais informações.

Você sabia que é possível...

- Arrastar e selecionar um grupo de nós na tela de fluxo usando seu mouse.
- Copiar e colar nós de um fluxo em outro.

- Acessar Ajuda de cada caixa de diálogo e janela de saída.
- Obter Ajuda sobre CRISP-DM o processo padrão de vários segmentos de mercados para mineração de dados. (No menu Ajuda, clique em **Ajuda do CRISP-DM**.)

Capítulo 6. Trabalhando com saída

Quando você executa alguns fluxos, os resultados ficam disponíveis no Visualizador via a guia **Modelo** ou guia **Avançado** dos nós do nugget do modelo. No Visualizador, é possível navegar facilmente para a saída que você deseja ver. Também é possível manipular a saída e criar um documento que contenha exatamente a saída desejada. Algumas saídas de gráfico também usam o Visualizador.

O Visualizador é usado para a seguinte saída no IBM SPSS Modeler:

- Nuggets do modelo TCM
- Nuggets do modelo STP
- Nuggets do modelo TwoStep-AS Cluster
- Nuggets do modelo GSAR
- Nó de gráfico Visualização de Mapa

Visualizador

Resultados são exibidos no Visualizador. É possível usar o Visualizador para:

- Navegar nos resultados
- Mostrar ou ocultar tabelas e gráficos selecionados
- Alterar a ordem de exibição dos resultados movendo itens selecionados
- Mover itens entre o Visualizador e outros aplicativos

O Visualizador é dividido em duas áreas de janela:

- A área de janela esquerda contém uma visualização da estrutura de tópicos do conteúdo.
- A área de janela direita contém tabelas estatísticas, gráficos e saída de texto.

É possível clicar em um item na estrutura de tópicos para ir diretamente à tabela ou gráfico correspondente. É possível clicar e arrastar a borda direita da área de janela da estrutura de tópicos para alterar sua largura.

Mostrando e ocultando resultados

No Visualizador, é possível mostrar e ocultar seletivamente tabelas individuais ou resultados de um procedimento inteiro. Esse processo é útil quando você deseja diminuir a quantidade de saída visível na área de janela de conteúdo.

Para ocultar tabelas e gráficos

1. Dê um clique duplo no ícone de livro do item na área de janela da estrutura de tópicos do Visualizador.
ou
2. Clique no item para selecioná-lo.
3. Nos menus, escolha:
Visualizar > Ocultar
ou
4. Clique no ícone de livro fechado (Ocultar) na barra de ferramentas Estrutura de Tópicos.

O ícone de livro aberto (Mostrar) se torna o ícone ativo, indicando que agora o item está oculto.

Para ocultar resultados do procedimento

1. Clique na caixa à esquerda do nome do procedimento na área de janela da estrutura de tópicos.

Isso oculta todos os resultados do procedimento e reduz a visualização da estrutura de tópicos.

Movendo, excluindo e copiando saída

É possível reorganizar os resultados copiando, movendo ou excluindo um item ou um grupo de itens.

Para mover saída no Visualizador

1. Selecione os itens na área de janela de conteúdo ou estrutura de tópicos.
2. Arraste e solte os itens selecionados em uma localização diferente.

Para excluir saída no Visualizador

1. Selecione os itens na área de janela de conteúdo ou estrutura de tópicos.
2. Pressione a tecla **Delete**.
- ou
3. Nos menus, escolha:

Editar > Excluir

Alterando o alinhamento inicial

Por padrão, todos os resultados são inicialmente alinhados à esquerda. Para alterar o alinhamento inicial de novos itens de saída:

1. Nos menus, escolha:
Editar > Opções
2. Clique na guia **Visualizador**.
3. No grupo Estados de Saída Iniciais, selecione o tipo de item (por exemplo, tabela dinâmica, gráfico, saída de texto).
4. Selecione a opção de alinhamento desejada.

Alterando o alinhamento dos itens de saída

1. Na área de janela de conteúdo ou estrutura de tópicos, selecione os itens que deseja alinhar.
2. Nos menus, escolha:
Formatar > Alinhar à Esquerda
ou
Formatar > Centralizar
ou
Formatar > Alinhar à Direita

Estrutura de tópicos do Visualizador

A área de janela da estrutura de tópicos fornece um índice do documento do Visualizador. É possível usar a área de janela da estrutura de tópicos para navegar pelos resultados e controlar a exibição. A maioria das ações na área de janela da estrutura de tópicos tem um efeito correspondente na área de janela de conteúdo.

- A seleção de um item na área de janela da estrutura de tópicos exibe o item correspondente na área de janela de conteúdo.
- A movimentação de um item na área de janela da estrutura de tópicos move o item correspondente na área de janela de conteúdo.

- A redução da visualização da estrutura de tópicos oculta os resultados de todos os itens nos níveis reduzidos.

Controlando a exibição da estrutura de tópicos. Para controlar a exibição da estrutura de tópicos, é possível:

- Expandir e reduzir a visualização da estrutura de tópicos
- Alterar o nível da estrutura de tópicos para os itens selecionados
- Alterar o tamanho dos itens na exibição da estrutura de tópicos
- Alterar a fonte usada na exibição da estrutura de tópicos

Para reduzir ou expandir a visualização da estrutura de tópicos

1. Clique na caixa à esquerda do item da estrutura de tópicos que deseja reduzir ou expandir.
ou
2. Clique no item na estrutura de tópicos.
3. Nos menus, escolha:
Visualizar > Reduzir
ou

Visualizar > Expandir

Para alterar o nível da estrutura de tópicos

1. Clique no item na área de janela da estrutura de tópicos.
2. Clique na seta para a esquerda na barra de ferramentas Estrutura de Tópicos para promover o item (mover o item para a esquerda).
ou
Clique na seta para a direita na barra de ferramentas Estrutura de Tópicos para rebaixar o item (mover o item para a direita).
ou
3. Nos menus, escolha:
Editar > Estrutura de Tópicos > Promover
ou
Editar > Estrutura de Tópicos > Rebaixar

A mudança do nível da estrutura de tópicos é útil principalmente após você mover itens no nível da estrutura de tópicos. A movimentação dos itens pode alterar o nível da estrutura de tópicos dos itens, e é possível usar os botões de seta para a esquerda e para a direita na barra de ferramentas Estrutura de Tópicos para restaurar o nível de estrutura de tópicos original.

Para alterar o tamanho dos itens da estrutura de tópicos

1. Nos menus, escolha:
Visualizar > Tamanho da Estrutura de Tópicos
2. Selecione o tamanho da estrutura de tópicos (**Pequeno**, **Médio** ou **Grande**).

Para alterar a fonte na estrutura de tópicos

1. Nos menus, escolha:
Visualizar > Fonte da Estrutura de Tópicos...
2. Selecione uma fonte.

Incluindo itens no Visualizador

No Visualizador, é possível incluir itens como títulos, texto novo, gráficos ou materiais de outros aplicativos.

Para incluir um título ou texto

Itens de texto que não estão conectados a uma tabela ou gráfico podem ser incluídos no Visualizador.

1. Clique na tabela, gráfico ou outro objeto que precederá o título ou texto.
2. Nos menus, escolha:

Inserir > Novo Título

ou

Inserir > Novo Texto

3. Dê um clique duplo no novo objeto.
4. Insira o texto.

Para incluir um arquivo de texto

1. Na área de janela da estrutura de tópicos ou área de janela de conteúdo do Visualizador, clique na tabela, gráfico ou outro objeto que precederá o texto.
2. Nos menus, escolha:
Inserir > Arquivo de Texto...
3. Selecione um arquivo de texto.

Para editar o texto, dê um clique duplo nele.

Colando objetos no Visualizador

Objetos de outros aplicativos podem ser colados no Visualizador. É possível usar Colar Após ou Colar Especial. Qualquer tipo de colagem coloca o novo objeto após o objeto atualmente selecionado no Visualizador. Use Colar Especial quando quiser escolher o formato do objeto colado.

Localizando e substituindo informações no Visualizador

1. Para localizar e substituir informações no Visualizador, nos menus, escolha:

Editar > Localizar

ou

Editar > Substituir

É possível usar Localizar e Substituir para:

- Procurar o documento inteiro ou apenas os itens selecionados.
- Procurar acima ou abaixo da localização atual.
- Procurar em ambas as áreas de janela ou restringir a procura à área de janela da estrutura de tópicos ou conteúdo.
- Procurar itens ocultos. Isso inclui quaisquer itens ocultos na área de janela de conteúdo (por exemplo, tabelas Notas, que ficam ocultas por padrão) e linhas e colunas ocultas em tabelas dinâmicas.
- Restringir os critérios de procura a correspondências com distinção entre maiúsculas e minúsculas.
- Restringir os critérios de procura em tabelas dinâmicas a correspondências do conteúdo da célula inteira.
- Restringir os critérios de procura em tabelas dinâmicas somente a marcadores de nota de rodapé. Essa opção não estará disponível se a seleção no Visualizador incluir qualquer coisa além de tabelas dinâmicas.

Itens Ocultos e Camadas da Tabela Dinâmica

- As camadas abaixo da camada atualmente visível de uma tabela dinâmica multidimensional não são consideradas ocultas e serão incluídas na área de procura, mesmo quando itens ocultos não estiverem incluídos na procura.
- Itens ocultos incluem aqueles na área de janela de conteúdo (itens com ícones de livro fechado na área de janela da estrutura de tópicos ou incluídos em blocos reduzidos da área de janela da estrutura de tópicos) e linhas e colunas em tabelas dinâmicas ocultas por padrão (por exemplo, linhas e colunas vazias ficam ocultas por padrão) ou ocultas manualmente por meio da edição da tabela e ocultando linhas ou colunas seletivamente. Itens ocultos só serão incluídos na procura se você selecionar **Incluir Itens Ocultos** explicitamente.
- Em ambos os casos, o elemento oculto ou não visível que contém o valor ou texto da procura será exibido quando localizado, mas o item será retornado para seu estado original posteriormente.

Localizando um intervalo de valores em tabelas dinâmicas

Para localizar valores que estejam dentro do intervalo especificado de valores em tabelas dinâmicas:

1. Ative uma tabela dinâmica ou selecione uma ou mais tabelas dinâmicas no Visualizador. Certifique-se de que somente tabelas dinâmicas estejam selecionadas. Se algum outro objeto estiver selecionado, a opção Intervalo não estará disponível.
2. Nos menus, escolha:
Editar > Localizar
3. Clique na guia **Intervalo**.
4. Selecione o tipo de intervalo: Entre, Maior que ou igual a ou Menor que ou igual a.
5. Selecione o valor ou os valores que definem o intervalo.
 - Se o valor contiver caracteres não numéricos, ambos os valores serão tratados como sequências de caracteres.
 - Se ambos os valores forem números, somente valores numéricos serão procurados.
 - Não é possível usar a guia Intervalo para substituir valores.

Essa variável não está disponível para tabelas anteriores. Consulte o tópico “Tabelas anteriores” na página 94 para obter mais informações.

Copiando saída em outros aplicativos

Objetos de saída podem ser copiados e colados em outros aplicativos, como programa de processamento de texto ou planilha. É possível colar saída em vários formatos. Dependendo do aplicativo de destino e do(s) objeto(s) de saída selecionado(s), alguns ou todos os formatos podem estar disponíveis:

Metarquivo. Formato de metarquivo WMF e EMF. Esses formatos estão disponíveis somente em sistemas operacionais Windows.

RTF (formato Rich Text). Vários objetos selecionados, saída de texto e tabelas dinâmicas podem ser copiados e colados em formato RTF. Para tabelas dinâmicas, na maioria dos aplicativos, isso significa que as tabelas são coladas como tabelas que depois podem ser editadas em outro aplicativo. Tabelas dinâmicas muito grandes para a largura do documento são quebradas, diminuídas para se ajustarem à largura do documento ou deixadas inalteradas, dependendo das configurações de opções da tabela dinâmica. Consulte o tópico “Opções de tabela dinâmica” na página 95 para obter mais informações.

Nota: O Microsoft Word pode não exibir tabelas extremamente grandes adequadamente.

Imagem. Formatos de imagem JPG e PNG.

BIFF. Tabelas dinâmicas e saída de texto podem ser coladas em uma planilha em formato BIFF. Números em tabelas dinâmicas retêm precisão numérica. Esse formato está disponível somente em sistemas operacionais Windows.

Texto. Tabelas dinâmicas e saída de texto podem ser copiadas e coladas como texto. Esse processo pode ser útil para aplicativos como email, em que o aplicativo pode aceitar ou transmitir somente texto.

Se o aplicativo de destino suportar diversos formatos disponíveis, ele pode ter um item de menu Colar Especial que permita selecionar o formato ou pode exibir automaticamente uma lista de formatos disponíveis.

Copiando e Colando Diversos Objetos de Saída

As limitações a seguir se aplicam durante a colagem de vários objetos de saída em outros aplicativos:

- **Formato RTF.** Na maioria dos aplicativos, tabelas dinâmicas são coladas como tabelas que podem ser editadas nesse aplicativo. Gráficos, árvores e visualizações de modelo são colados como imagens.
- **Formatos de metarquivo e imagem.** Todos os objetos de saída selecionados são colados como um único objeto no outro aplicativo.
- **Formato BIFF.** Gráficos, árvores e visualizações de modelo são excluídos.

Copiar Especial

Durante a cópia e a colagem de grandes quantidades de saída, principalmente tabelas dinâmicas muito grandes, é possível melhorar a velocidade da operação usando Editar > Copiar Especial para limitar o número de formatos copiados na área de transferência.

Também é possível salvar os formatos selecionados como o conjunto padrão de formatos para copiar na área de transferência. Essa configuração persiste nas sessões.

Saída interativa

Objetos de saída interativa contêm diversos objetos de saída relacionados. A seleção em um objeto pode alterar o que é exibido ou destacado no outro objeto. Por exemplo, a seleção de uma linha em uma tabela pode destacar uma área em um mapa ou exibir um gráfico para uma categoria diferente.

Objetos de saída interativa não suportam a edição de variáveis, como alterar texto, cores, fontes ou bordas da tabela. Os objetos individuais podem ser copiados do objeto interativo no Visualizador. Tabelas copiadas da saída interativa podem ser editadas no editor de tabela dinâmica.

Copiando objetos da saída interativa

Arquivo>Copiar no Visualizador copia objetos de saída individuais na janela Visualizador.

- As opções disponíveis dependem do conteúdo da saída interativa.
- **Gráfico e Mapa** criam objetos de gráfico.
- **Tabela** cria uma tabela dinâmica que pode ser editada no editor de tabela dinâmica.
- **Captura Instantânea** cria uma imagem da visualização atual.
- **Modelo** cria uma cópia do atual objeto de saída interativa.

Editar>Copiar Objeto copia objetos de saída individuais na área de transferência.

- A colagem do objeto copiado no Visualizador equivale a **Arquivo>Copiar no Visualizador**.
- A colagem do objeto em outro aplicativo cola o objeto como uma imagem.

Zoom e panorama

Para mapas, é possível usar **Visualizar>Zoom** para aplicar zoom na visualização do mapa. Na visualização de mapa com zoom, é possível usar **Visualizar>Panoramizar** para mover a visualização.

Configurações de impressão

Arquivo>Configurações de Impressão controla como objetos interativos são impressos.

- **Imprimir somente visualização visível.** Imprime somente a visualização atualmente exibida. Essa opção é a configuração padrão.
- **Imprimir todas as visualizações.** Imprime todas as visualizações contidas na saída interativa.
- A opção selecionada também determina a ação padrão para exportar o objeto de saída.

Exportar saída

Exportar Saída salva saída do Visualizador em formatos HTML, de texto, Word/RTF, Excel, PowerPoint (requer PowerPoint 97 ou posterior) e PDF. Gráficos também podem ser exportados em inúmeros formatos gráficos diferentes.

Nota: A exportação para o PowerPoint está disponível somente em sistemas operacionais Windows.

Para exportar saída

1. Torne o Visualizador a janela ativa (clique em qualquer lugar na janela).
2. Clique no botão **Exportar** na barra de ferramentas ou clique com o botão direito na janela de saída e selecione **Exportar**.
3. Insira um nome de arquivo (ou prefixo para gráficos) e selecione um formato de exportação.

Objetos para Exportar. É possível exportar todos os objetos no Visualizador, todos os objetos visíveis ou somente os objetos selecionados.

Tipo de Documento. As opções disponíveis são:

- **Word/RTF.** Tabelas dinâmicas são exportadas como tabelas do Word com todos os atributos de formatação intactos (por exemplo, bordas da célula, estilos de fonte e cores do plano de fundo). A saída de texto é exportada como RTF formatado. Gráficos, diagramas de árvore e visualizações de modelo são incluídos em formato PNG. Observe que o Microsoft Word pode não exibir as tabelas extremamente grandes de forma adequada.
- **Excel.** Linhas, colunas e células de tabelas dinâmicas são exportadas como linhas, colunas e células do Excel, com todos os atributos de formatação intactos (por exemplo, bordas de célula, estilos de fonte e cores de plano de fundo). A saída de texto é exportada com todos os atributos de fonte intactos. Cada linha na saída de texto é uma linha no arquivo do Excel, com o conteúdo inteiro da linha em uma única célula. Gráficos, diagramas de árvore e visualizações de modelo são incluídos em formato PNG. A saída pode ser exportada como *Excel 97-2004* ou *Excel 2007 e superior*.
- **HTML.** Tabelas dinâmicas são exportadas como tabelas HTML. A saída de texto é exportada como HTML pré-formatado. Gráficos, diagramas de árvore e visualizações de modelo são integrados no documento no formato gráfico selecionado. Um navegador compatível com HTML 5 é necessário para visualização de saída exportada em formato HTML.
- **Relatório da Web.** Um relatório da web é um documento interativo compatível com a maioria dos navegadores. Muitos dos recursos interativos das tabelas dinâmicas disponíveis no Visualizador também estão disponíveis em relatórios da web. Também é possível exportar um relatório da web como IBM Cognos Active Report.
- **Portable Document Format.** Toda saída é exportada como aparece na Visualização de Impressão, com todos os atributos de formatação intactos.

- **Arquivo do PowerPoint.** Tabelas dinâmicas são exportadas como tabelas do Word e integradas em slides separados no arquivo do PowerPoint, com um slide para cada tabela dinâmica. Todos os atributos de formatação da tabela dinâmica são retidos (por exemplo, bordas de célula, estilos de fonte e cores do plano de fundo). Gráficos, diagramas de árvore e visualizações de modelo são exportados em formato TIFF. A saída de texto não é incluída.

A exportação para o PowerPoint está disponível somente em sistemas operacionais Windows.

- **Texto.** Os formatos de saída de texto incluem texto simples, UTF-8 e UTF-16. Tabelas dinâmicas podem ser exportadas em formato separado por guias ou separado por espaços. Toda saída de texto é exportada em formato separado por espaços. Para gráficos, diagramas de árvore e visualizações de modelo, uma linha é inserida no arquivo de texto para cada gráfico, indicando o nome do arquivo de imagem.
- **Nenhum (Somente Gráfico).** Os formatos de exportação disponíveis incluem: EPS, JPEG, TIFF, PNG e BMP. Em sistemas operacionais Windows, o formato EMF (enhanced metafile) também está disponível.

Abrir a pasta contêiner. Abre a pasta que contém os arquivos que são criados pela exportação.

Opções de HTML

A exportação HTML requer um navegador que seja compatível com HTML 5. As opções a seguir estão disponíveis para exportar saída no formato HTML:

Camadas em tabelas dinâmicas. Por padrão, a inclusão ou exclusão de camadas de tabela dinâmica são controladas pelas propriedades de tabela para cada tabela dinâmica. É possível substituir essa configuração e incluir todas as camadas ou excluir todas, exceto a camada atualmente visível. Consulte o tópico “Propriedades da tabela: impressão” na página 90 para obter mais informações.

Exportar tabelas em camadas como interativas. Tabelas em camadas são exibidas como elas aparecem no Visualizador, e é possível alterar interativamente a camada exibida no navegador. Se essa opção não for selecionada, cada camada da tabela será exibida como uma tabela separada.

Tabelas como HTML. Isso controla as informações de estilo incluídas para tabelas dinâmicas exportadas.

- **Exportar com estilos e largura de coluna fixa.** Todas as informações de estilo da tabela dinâmica (estilos de fonte, cores de plano de fundo, etc.) e larguras de coluna são preservadas.
- **Exportar sem estilos.** Tabelas dinâmicas são convertidas em tabelas HTML padrão. Nenhum atributo de estilo é preservado. A largura da coluna é determinada automaticamente.

Incluir notas de rodapé e legendas. Controla a inclusão ou exclusão de notas de rodapé e legendas de todas as tabelas dinâmicas.

Visualizações de Modelos. Por padrão, a inclusão ou exclusão de visualizações de modelo são controladas pelas propriedades de modelo para cada modelo. É possível substituir essa configuração e incluir todas as visualizações ou excluir todas, exceto a que está atualmente visível. Consulte o tópico Propriedades do modelo para obter mais informações. (Nota: todas as visualizações de modelo, incluindo tabelas, são exportadas como gráficos.)

Nota: Para HTML, também é possível controlar o formato de arquivo de imagem para gráficos exportados. Consulte o tópico “Opções de formato gráfico” na página 79 para obter mais informações.

Para configurar opções de exportação HTML

1. Selecione **HTML** como o formato de exportação.
2. Clique em **Alterar Opções**.

Opções de relatório da web

Um relatório da web é um documento interativo compatível com a maioria dos navegadores. Muitos dos recursos interativos das tabelas dinâmicas disponíveis no Visualizador também estão disponíveis em relatórios da web.

Título do Relatório. O título que é exibido no cabeçalho do relatório. Por padrão, o nome do arquivo é usado. É possível especificar um título customizado para ser usado no lugar do nome do arquivo.

Formato. Há duas opções de formato do relatório:

- **Relatório da Web SPSS (HTML 5).** Esse formato requer um navegador que seja compatível com HTML 5.
- **Cognos Active Report (mht).** Esse formato requer um navegador que suporte arquivos em formato MHT ou aplicativo Cognos Active Report.

Excluir Objetos. É possível excluir tipos de objeto selecionados do relatório:

- **Texto.** Objetos de texto que não são logs. Essa opção inclui objetos de texto que contêm informações sobre o conjunto de dados ativo.
- **Logs.** Objetos de texto que contêm uma listagem da sintaxe de comando que foi executada. Itens de log também incluem mensagens de aviso e de erro que são encontradas por comandos que não produzem nenhuma saída do Visualizador.
- **Tabelas de Notas.** A saída dos procedimentos gráficos e estatísticos inclui uma tabela Notas. Essa tabela contém informações sobre o conjunto de dados que foi usado, valores omissos e a sintaxe de comando que foi usada para executar o procedimento.
- **Mensagens de Aviso e de Erro.** Mensagens de aviso e de erro de procedimentos gráficos e estatísticos.

Redefinir o estilo das tabelas e gráficos para corresponder ao relatório da web. Esta opção aplica o estilo de relatório da web padrão a todas as tabelas e gráficos. Essa configuração substitui quaisquer fontes, cores ou outros estilos na saída, conforme exibido no Visualizador. Não é possível modificar o estilo de relatório da web padrão.

Conexão do servidor da web. É possível incluir a localização da URL de um ou mais servidores de aplicativos que estão executando o Servidor de Aplicativos de Relatório da Web IBM SPSS Statistics. O servidor de aplicativos da web fornece variáveis para colocar tabelas sobre o eixo, editar gráficos e salvar relatórios da web modificados.

- Selecione **Usar** para cada servidor de aplicativos que deseja incluir no relatório da web.
- Se um relatório da web contiver uma especificação de URL, o relatório da web se conectará ao servidor de aplicativos para fornecer variáveis de edição adicionais.
- Se você especificar várias URLs, o relatório da web tentará se conectar a cada servidor na ordem em que eles estão especificados.

O Servidor de Aplicativos de Relatório da Web IBM SPSS Statistics pode ser transferido por download em <http://www.ibm.com/developerworks/spssdevcentral>.

Opções do Word/RTF

As opções a seguir estão disponíveis para exportar saída no formato do Word:

Camadas em tabelas dinâmicas. Por padrão, a inclusão ou exclusão de camadas de tabela dinâmica são controladas pelas propriedades de tabela para cada tabela dinâmica. É possível substituir essa configuração e incluir todas as camadas ou excluir todas, exceto a camada atualmente visível. Consulte o tópico “Propriedades da tabela: impressão” na página 90 para obter mais informações.

Tabelas Dinâmicas Grandes. Controla o tratamento de tabelas que são muito grandes para a largura do documento definida. Por padrão, a tabela é quebrada para se ajustar. A tabela é dividida em seções, e os

rótulos de linha são repetidos para cada seção da tabela. Como alternativa, é possível reduzir as tabelas grandes ou não fazer mudanças nas tabelas grandes para estendê-las além da largura do documento definida.

Preservar pontos de interrupção. Se você tiver definido pontos de interrupção, essas configurações serão preservadas nas tabelas do Word.

Incluir notas de rodapé e legendas. Controla a inclusão ou exclusão de notas de rodapé e legendas de todas as tabelas dinâmicas.

Visualizações de Modelos. Por padrão, a inclusão ou exclusão de visualizações de modelo são controladas pelas propriedades de modelo para cada modelo. É possível substituir essa configuração e incluir todas as visualizações ou excluir todas, exceto a que está atualmente visível. Consulte o tópico Propriedades do modelo para obter mais informações. (Nota: todas as visualizações de modelo, incluindo tabelas, são exportadas como gráficos.)

Configuração da Página para Exportar. Isso abre um diálogo onde é possível definir o tamanho da página e as margens para o documento exportado. A largura do documento usado para determinar o comportamento de agrupamento e redução é a largura da página menos as margens esquerda e direita.

Para configurar opções de exportação do Word

1. Selecione **Word/RTF** como formato de exportação.
2. Clique em **Alterar Opções**.

Opções do Excel

As opções a seguir estão disponíveis para exportar saída no formato Excel:

Criar uma planilha ou pasta de trabalho ou modificar uma planilha existente. Por padrão, uma nova pasta de trabalho será criada. Se um arquivo com o nome especificado já existir, ele será sobrescrito. Se você selecionar a opção para criar uma planilha, e uma planilha com o nome especificado já existir no arquivo especificado, ela será sobrescrita. Se você selecionar a opção para modificar uma planilha existente, também deve-se especificar o nome da planilha. (Isso é opcional para a criação de uma planilha.) Nomes de planilha não podem exceder 31 caracteres e não podem conter barras ou barras invertidas, colchetes, pontos de interrogação ou asteriscos.

Ao exportar para Excel 97-2004, se você modificar uma planilha existente, os gráficos, as visualizações de modelo e os diagramas de árvore não serão incluídos na saída exportada.

Localização na planilha. Controla a localização dentro da planilha para a saída exportada. Por padrão, a saída exportada será incluída após a última coluna que tenha qualquer conteúdo, começando na primeira linha, sem modificar qualquer conteúdo existente. Essa é uma boa opção para incluir novas colunas para uma planilha existente. A inclusão de saída exportada após a última linha é uma boa opção para incluir novas linhas em uma planilha existente. A inclusão de saída exportada começando em uma localização de célula específica sobrescreverá qualquer conteúdo existente na área onde a saída exportada será incluída.

Camadas em tabelas dinâmicas. Por padrão, a inclusão ou exclusão de camadas de tabela dinâmica são controladas pelas propriedades de tabela para cada tabela dinâmica. É possível substituir essa configuração e incluir todas as camadas ou excluir todas, exceto a camada atualmente visível. Consulte o tópico “Propriedades da tabela: impressão” na página 90 para obter mais informações.

Incluir notas de rodapé e legendas. Controla a inclusão ou exclusão de notas de rodapé e legendas de todas as tabelas dinâmicas.

Visualizações de Modelos. Por padrão, a inclusão ou exclusão de visualizações de modelo são controladas pelas propriedades de modelo para cada modelo. É possível substituir essa configuração e

incluir todas as visualizações ou excluir todas, exceto a que está atualmente visível. Consulte o tópico Propriedades do modelo para obter mais informações. (Nota: todas as visualizações de modelo, incluindo tabelas, são exportadas como gráficos.)

Para configurar opções de exportação de Excel

1. Selecione **Excel** como o formato de exportação.
2. Clique em **Alterar Opções**.

Opções do PowerPoint

As opções a seguir estão disponíveis para PowerPoint:

Camadas em tabelas dinâmicas. Por padrão, a inclusão ou exclusão de camadas de tabela dinâmica são controladas pelas propriedades de tabela para cada tabela dinâmica. É possível substituir essa configuração e incluir todas as camadas ou excluir todas, exceto a camada atualmente visível. Consulte o tópico “Propriedades da tabela: impressão” na página 90 para obter mais informações.

Tabelas Dinâmicas Grandes. Controla o tratamento de tabelas que são muito grandes para a largura do documento definida. Por padrão, a tabela é quebrada para se ajustar. A tabela é dividida em seções, e os rótulos de linha são repetidos para cada seção da tabela. Como alternativa, é possível reduzir as tabelas grandes ou não fazer mudanças nas tabelas grandes para estendê-las além da largura do documento definida.

Incluir notas de rodapé e legendas. Controla a inclusão ou exclusão de notas de rodapé e legendas de todas as tabelas dinâmicas.

Usar entradas de estrutura de tópicos do Visualizador como títulos de slide. Inclui um título em cada slide criado pela exportação. Cada slide contém um único item que é exportado do Visualizador. O título é formado a partir da entrada de estrutura de tópicos para o item na área de janela da estrutura de tópicos do Visualizador.

Visualizações de Modelos. Por padrão, a inclusão ou exclusão de visualizações de modelo são controladas pelas propriedades de modelo para cada modelo. É possível substituir essa configuração e incluir todas as visualizações ou excluir todas, exceto a que está atualmente visível. Consulte o tópico Propriedades do modelo para obter mais informações. (Nota: todas as visualizações de modelo, incluindo tabelas, são exportadas como gráficos.)

Configuração da Página para Exportar. Isso abre um diálogo onde é possível definir o tamanho da página e as margens para o documento exportado. A largura do documento usado para determinar o comportamento de agrupamento e redução é a largura da página menos as margens esquerda e direita.

Para configurar opções de exportação do PowerPoint

1. Selecione **PowerPoint** como o formato de exportação.
2. Clique em **Alterar Opções**.

Nota: A exportação para o PowerPoint está disponível somente em sistemas operacionais Windows.

Opções de PDF

As opções a seguir estão disponíveis para PDF:

Favoritos integrados. Essa opção inclui favoritos no documento PDF que correspondem às entradas de estrutura de tópicos do Visualizador. Como a área de janela da estrutura de tópicos do Visualizador, os favoritos facilitam a navegação em documentos com um grande número de objetos de saída.

Fontes integradas. A integração de fontes assegura que o documento PDF fique igual em todos os computadores. Caso contrário, se algumas fontes usadas no documento não estiverem disponíveis no computador sendo usado para visualizar (ou imprimir) o documento PDF, a substituição da fonte pode gerar resultados pouco satisfatórios.

Camadas em tabelas dinâmicas. Por padrão, a inclusão ou exclusão de camadas de tabela dinâmica são controladas pelas propriedades de tabela para cada tabela dinâmica. É possível substituir essa configuração e incluir todas as camadas ou excluir todas, exceto a camada atualmente visível. Consulte o tópico “Propriedades da tabela: impressão” na página 90 para obter mais informações.

Visualizações de Modelos. Por padrão, a inclusão ou exclusão de visualizações de modelo são controladas pelas propriedades de modelo para cada modelo. É possível substituir essa configuração e incluir todas as visualizações ou excluir todas, exceto a que está atualmente visível. Consulte o tópico Propriedades do modelo para obter mais informações. (Nota: todas as visualizações de modelo, incluindo tabelas, são exportadas como gráficos.)

Para configurar opções de exportação de PDF

1. Selecione **Portable Document Format** como o formato de exportação.
2. Clique em **Alterar Opções**.

Outras configurações que afetam a saída PDF

Configuração de Página/Atributos de Página. O tamanho da página, a orientação, as margens, o conteúdo e a exibição dos rodapés e dos cabeçalhos da página e o tamanho do gráfico impresso em documentos PDF são controlados pela configuração da página e pelas opções de atributo da página.

Propriedades da Tabela/TableLooks. O ajuste de escala de tabelas grandes e/ou longas e a impressão de camadas de tabela são controlados pelas propriedades de tabela de cada tabela. Essas propriedades também podem ser salvas em TableLooks.

Impressora Padrão/Atual. A resolução (DPI) do documento PDF é a configuração de resolução atual para a impressora padrão ou atualmente selecionada (que pode ser alterada usando Configuração de Página). A resolução máxima é 1200 DPI. Se a configuração da impressora for superior, a resolução do documento PDF será 1200 DPI.

Nota: Documentos de alta resolução podem gerar baixos resultados quando impressos em impressoras de resolução inferior.

Opções de texto

As opções a seguir estão disponíveis para exportação de texto:

Formato de Tabela Dinâmica. Tabelas dinâmicas podem ser exportadas em formato separado por guias ou separado por espaços. Para formato separado por espaço, também é possível controlar:

- **Largura da Coluna. Ajuste Automático** não quebra o conteúdo de nenhuma coluna, e cada coluna é tão grande quanto o maior rótulo ou valor nessa coluna. **Customizado** configura a largura máxima da coluna que é aplicada a todas as colunas na tabela, e os valores que excedem essa largura são quebrados para a próxima linha nessa coluna.
- **Caractere da Borda da Linha/Coluna.** Controla os caracteres usados para criar bordas de linha e coluna. Para suprimir a exibição das bordas da linha e coluna, insira espaços em branco para os valores.

Camadas em tabelas dinâmicas. Por padrão, a inclusão ou exclusão de camadas de tabela dinâmica são controladas pelas propriedades de tabela para cada tabela dinâmica. É possível substituir essa

configuração e incluir todas as camadas ou excluir todas, exceto a camada atualmente visível. Consulte o tópico “Propriedades da tabela: impressão” na página 90 para obter mais informações.

Incluir notas de rodapé e legendas. Controla a inclusão ou exclusão de notas de rodapé e legendas de todas as tabelas dinâmicas.

Visualizações de Modelos. Por padrão, a inclusão ou exclusão de visualizações de modelo são controladas pelas propriedades de modelo para cada modelo. É possível substituir essa configuração e incluir todas as visualizações ou excluir todas, exceto a que está atualmente visível. Consulte o tópico Propriedades do modelo para obter mais informações. (Nota: todas as visualizações de modelo, incluindo tabelas, são exportadas como gráficos.)

Para configurar opções de exportação de texto

1. Selecione **Texto** como formato de exportação.
2. Clique em **Alterar Opções**.

Opções somente de gráficos

As opções a seguir estão disponíveis para exportação somente de gráficos:

Visualizações de Modelos. Por padrão, a inclusão ou exclusão de visualizações de modelo são controladas pelas propriedades de modelo para cada modelo. É possível substituir essa configuração e incluir todas as visualizações ou excluir todas, exceto a que está atualmente visível. Consulte o tópico Propriedades do modelo para obter mais informações. (Nota: todas as visualizações de modelo, incluindo tabelas, são exportadas como gráficos.)

Opções de formato gráfico

Para documentos de texto e HTML e somente para exportar gráficos, é possível selecionar o formato gráfico e, para cada formato gráfico, controlar várias configurações opcionais.

Para selecionar o formato gráfico e as opções para gráficos exportados:

1. Selecione **HTML**, **Texto** ou **Nenhum (somente Gráficos)** como o tipo de documento.
2. Selecione o formato de arquivo gráfico na lista suspensa.
3. Clique em **Alterar Opções** para alterar as opções para o formato de arquivo gráfico selecionado.

Opções de exportação de gráfico JPEG

- **Tamanho da imagem.** Porcentagem do tamanho do gráfico original, até 200 por cento.
- **Converter em escala de tons.** Converte cores em tonalidades de cinza.

Opções de exportação de gráfico BMP

- **Tamanho da imagem.** Porcentagem do tamanho do gráfico original, até 200 por cento.
- **Compactar imagem para reduzir o tamanho do arquivo.** Uma técnica de compactação sem perdas que cria arquivos menores sem afetar a qualidade da imagem.

Opções de exportação de gráfico PNG

Tamanho da imagem. Porcentagem do tamanho do gráfico original, até 200 por cento.

Intensidade de Cor. Determina o número de cores no gráfico exportado. Um gráfico que é salvo sob qualquer intensidade terá um mínimo do número de cores realmente usadas e um máximo do número de cores permitidas por intensidade. Por exemplo, se o gráfico contiver três cores -- vermelho, branco e preto -- e você salvá-lo como 16 cores, o gráfico permanecerá como três cores.

- Se o número de cores no gráfico exceder o número de cores para essa intensidade, as cores ficarão confusas para replicar cores no gráfico.
- **Intensidade da tela atual** é o número de cores atualmente exibido no monitor do seu computador.

Opções de exportação de gráfico EMF e TIFF

Tamanho da imagem. Porcentagem do tamanho do gráfico original, até 200 por cento.

Nota: O formato EMF (enhanced metafile) está disponível somente em sistemas operacionais Windows.

Opções de exportação de gráfico EPS

Tamanho da imagem. É possível especificar o tamanho como uma porcentagem do tamanho da imagem original (até 200 por cento) ou especificar a largura de uma imagem em pixels (com a altura determinada pelo valor de largura e a proporção de aspecto). A imagem exportada é sempre proporcional à original.

Incluir imagem da visualização TIFF. Salva uma visualização com a imagem EPS em formato TIFF para exibição em aplicativos que não podem exibir imagens EPS na tela.

Fontes. Controla o tratamento das fontes em imagens EPS.

- **Usar Referências de Fonte.** Se as fontes que são usadas no gráfico estiverem disponíveis no dispositivo de saída, elas serão usadas. Caso contrário, o dispositivo de saída usará fontes alternativas.
- **Substituir Fontes por Curvas.** Transforma fontes em dados de curva PostScript. O texto em si não é mais editável como texto em aplicativos que podem editar gráficos EPS. Essa opção será útil se as fontes que são usadas no gráfico não estiverem disponíveis no dispositivo de saída.

Impressão do Visualizador

Há duas opções para imprimir o conteúdo da janela Visualizador:

Toda saída visível. Imprime somente os itens que estão sendo exibidos atualmente na área de janela de conteúdo. Itens ocultos (itens com um ícone de livro fechado na área de janela da estrutura de tópicos ou ocultos em camadas de estrutura de tópicos reduzidas) não são impressos.

Seleção. Imprime somente os itens atualmente selecionados nas áreas de janela de estrutura de tópicos e/ou conteúdo.

Para imprimir saída e gráficos

1. Torne o Visualizador a janela ativa (clique em qualquer lugar na janela).
2. Nos menus, escolha:
 Arquivo > Imprimir...
3. Selecione as configurações de impressão desejadas.
4. Clique em **OK** para imprimir.

Visualização prévia da impressão

A Visualização de Impressão mostra o que será impresso em cada página para documentos do Visualizador. Uma boa ideia é marcar Visualização de Impressão antes de realmente imprimir um documento do Visualizador, pois a Visualização de Impressão mostra os itens que podem não estar visíveis por meio da consulta da área de janela de conteúdo do Visualizador, incluindo:

- Quebras de páginas
- Camadas ocultas das tabelas dinâmicas
- Quebras em tabelas grandes
- Cabeçalhos e rodapés impressos em cada página

Se alguma saída estiver atualmente selecionada no Visualizador, a visualização exibirá somente a saída selecionada. Para ver uma visualização de todas as saídas, certifique-se de que nada esteja selecionado no Visualizador.

Atributos de página: cabeçalhos e rodapés

Cabeçalhos e rodapés são as informações que são impressas nas partes superior e inferior de cada página. É possível inserir qualquer texto que você deseja usar como cabeçalhos e rodapés. Também é possível usar a barra de ferramentas no meio da caixa de diálogo para inserir:

- Data e hora
- Números de página
- Nome do arquivo do Visualizador
- Rótulos de título da estrutura de tópicos
- Títulos e subtítulos da página
- **Tornar Padrão** usa as configurações especificadas aqui como configurações padrão para novos documentos do Visualizador. (Nota: isso torna as configurações atuais na guia Cabeçalho/Rodapé e na guia Opções as configurações padrão.)
- Rótulos de título de estrutura de tópicos indicam o título de estrutura de tópicos do primeiro, segundo, terceiro e/ou quarto nível para o primeiro item em cada página.
- Os títulos e subtítulos de página imprimem os títulos e subtítulos da página atual. Isso pode ser criado em Título da Nova Página no menu Inserção do Visualizador ou com os comandos TITLE e SUBTITLE. Se você não tiver especificado nenhum título ou subtítulo de página, essa configuração será ignorada.
Nota: Características de fonte para títulos e subtítulos de páginas novas são controladas na guia Visualizador da caixa de diálogo Opções (acessada escolhendo Opções no menu Editar). Características de fonte para títulos e subtítulos de páginas existentes podem ser alteradas editando os títulos no Visualizador.

Para ver como seus cabeçalhos e rodapés ficarão na página impressa, escolha Visualização de Impressão no menu Arquivo.

Para inserir cabeçalhos e rodapés da página

1. Torne o Visualizador a janela ativa (clique em qualquer lugar na janela).
2. Nos menus, escolha:
Arquivo > Cabeçalho e Rodapé...
3. Insira o cabeçalho e/ou o rodapé que você deseja que apareça em cada página.

Atributos de Página: Opções

Esta caixa de diálogo controla o tamanho do gráfico impresso, o espaço entre os itens de saída impressos e numeração de página.

- **Tamanho do Gráfico Impresso.** Controla o tamanho do gráfico impresso relativo ao tamanho da página definido. A proporção de aspecto do gráfico (proporção de largura para altura) não é afetado pelo tamanho do gráfico impresso. O tamanho da impressão geral de um gráfico é limitado por sua altura e largura. Quando as bordas externas de um gráfico atingem as bordas esquerda e direita da página, o tamanho do gráfico não pode aumentar mais para preencher mais a altura da página.
- **Espaço entre Itens.** Controla o espaço entre itens impressos. Cada tabela dinâmica, gráfico e objeto de texto é um item separado. Essa configuração não afeta a exibição de itens no Visualizador.
- **Número de Páginas Começando com.** Numera páginas sequencialmente, começando com o número especificado.
- **Tornar Padrão.** Esta opção usa as configurações que são especificadas aqui como padrão para os novos documentos do Visualizador. (Observe que essa opção torna as atuais configurações de Cabeçalho/Rodapé e configurações de Opções o padrão.)

Para alterar o tamanho do gráfico impresso, a numeração de página e o espaço entre itens impressos

1. Torne o Visualizador a janela ativa (clique em qualquer lugar na janela).
2. Nos menus, escolha:

Arquivo > Atributos de Página...

3. Clique na guia **Opções**.
4. Altere as configurações e clique em **OK**.

Salvando saída

O conteúdo do Visualizador pode ser salvo.

- **Objeto de saída (*.cou)**. Esse formato salva o contêiner de saída inteiro, incluindo gráfico, guias, anotações, entre outros. Esse formato pode ser aberto e visualizado no IBM SPSS Modeler, incluído em projetos e publicado e controlado usando IBM SPSS Collaboration and Deployment Services Repository. Esse formato não é compatível com IBM SPSS Statistics.
- **Arquivos do Visualizador (*.spv)**. O formato que é usado para exibir arquivos na janela Visualizador. Quando você salva nesse formato a partir de um nugget do modelo no IBM SPSS Modeler, somente o conteúdo do Visualizador da guia **Modelo** é salvo.

Para controlar opções para salvar relatórios da web ou salvar resultados em outros formatos (por exemplo, texto, Word, Excel), use **Exportar** no menu **Arquivo**.

Para salvar um documento do Visualizador

1. Nos menus da janela do Visualizador, escolha:
Arquivo > Salvar
2. Insira o nome do documento e, em seguida, clique em **Salvar**.
Opcionalmente, é possível fazer o seguinte:

Bloquear arquivos para evitar a edição no IBM SPSS Smartreader

Se um documento do Visualizador estiver bloqueado, é possível manipular tabelas dinâmicas (trocar linhas e colunas, alterar a camada exibida, etc.), mas não é possível editar nenhuma saída ou salvar quaisquer mudanças no documento do Visualizador no IBM SPSS Smartreader (um produto separado para trabalhar com documentos do Visualizador). Essa configuração não tem efeito sobre os documentos do Visualizador abertos no IBM SPSS Statistics ou IBM SPSS Modeler.

Criptografar arquivos com uma senha

É possível proteger informações confidenciais armazenadas em um documento do Visualizador criptografando o documento com uma senha. Após a criptografia, o documento só pode ser aberto fornecendo a senha. Os usuários do IBM SPSS Smartreader também deverão fornecer a senha para abrir o arquivo.

Para criptografar um documento do Visualizador:

- a. Selecione **Criptografar Arquivo com Senha** na caixa de diálogo Salvar Saída Como.
- b. Clique em **Salvar**.
- c. Na caixa de diálogo Criptografar Arquivo, forneça uma senha e insira-a novamente na caixa de texto Confirmar Senha. As senhas são limitadas a 10 caracteres e fazem distinção entre maiúsculas e minúsculas.

Aviso: Senhas não podem ser recuperadas se forem perdidas. Se a senha for perdida, o arquivo não poderá ser aberto.

Criando senhas fortes

- Use oito ou mais caracteres.
- Inclua números, símbolos e até pontuação em sua senha.
- Evite sequências de números ou caracteres, como "123" e "abc", e repetição, como "111aaa".
- Não crie senhas que usem informações pessoais, como datas de aniversário ou apelidos.
- Altere a senha periodicamente.

Nota: O armazenamento de arquivos criptografados em um IBM SPSS Collaboration and Deployment Services Repository não é suportado.

Modificando arquivos criptografados

- Se você abrir um arquivo criptografado, fazer modificações nele e escolher Arquivo > Salvar, o arquivo modificado será salvo com a mesma senha.
- É possível alterar a senha em um arquivo criptografado abrindo o arquivo, repetindo os passos para sua criptografia e especificando uma senha diferente na caixa de diálogo Criptografar Arquivo.
- É possível salvar uma versão não criptografada de um arquivo criptografado abrindo o arquivo, escolhendo Arquivo > Salvar Como e cancelando a seleção de **Criptografar Arquivo com Senha** na caixa de diálogo Salvar Saída Como.

Nota: Arquivos de dados criptografados e documentos de saída não podem ser abertos em versões do IBM SPSS Statistics anteriores à versão 21. Arquivos de sintaxe criptografados não podem ser abertos em versões anteriores à 22.

Armazenar informações de modelo necessárias com o documento de saída

Esta opção se aplica somente quando há itens do visualizador de modelo no documento de saída que requerem informações auxiliares para ativar algumas das variáveis interativas. Clique em **Mais Informações** para exibir uma lista desses itens do visualizador de modelo e as variáveis interativas que requerem informações auxiliares. O armazenamento dessas informações com o documento de saída podem aumentar substancialmente o tamanho do documento. Se você optar por não armazenar essas informações, ainda será possível abrir esses itens de saída, mas as variáveis interativas especificadas não estarão disponíveis.

Tabelas dinâmicas

Tabelas dinâmicas

Muitos resultados são apresentados em tabelas que podem ser colocadas sobre o eixo interativamente. Ou seja, é possível reorganizar as linhas, colunas e camadas.

Manipulando uma tabela dinâmica

Opções para manipular uma tabela dinâmica incluem:

- Transpor linhas e colunas
- Mover linhas e colunas
- Criar camadas multidimensionais
- Agrupar e desagrupar linhas e colunas
- Mostrar e ocultar linhas, colunas e outras informações
- Girar rótulos de linhas e colunas
- Localizar definições de termos

Ativando uma tabela dinâmica

Antes de poder manipular ou modificar uma tabela dinâmica, você precisa **ativar** a tabela. Para ativar uma tabela:

1. Dê um clique duplo na tabela.
ou
2. Clique com o botão direito na tabela e, no menu pop-up, escolha **Editar Conteúdo**.
3. No submenu, escolha **No Visualizador** ou **Na Janela Separada**.

Colocando uma tabela sobre o eixo

Uma tabela tem três dimensões: linhas, colunas e camadas. Uma dimensão pode conter diversos elementos (ou nenhum). É possível alterar a organização da tabela movendo elementos entre ou dentro de dimensões. Para mover um elemento, basta arrastá-lo e soltá-lo onde desejar.

Alterando a ordem de exibição dos elementos dentro de uma dimensão

Para alterar a ordem de exibição dos elementos dentro de uma dimensão de tabela (linha, coluna ou camada):

1. Se as bandejas de dinamização ainda não estiverem ativadas, no menu Tabela Dinâmica, escolha:
Tabela Dinâmica > Bandejas de Dinamização
2. Arraste e solte os elementos dentro da dimensão na bandeja de dinamização.

Movendo linhas e colunas dentro de um elemento de dimensão

1. Na tabela em si (e não nas bandejas de dinamização), clique no rótulo para a linha ou coluna que deseja mover.
2. Arraste o rótulo para a nova posição.

Transpor linhas e colunas

Se você quiser apenas inverter as linhas e colunas, há uma alternativa simples para usar as bandejas de dinamização:

1. Nos menus, escolha:
Tabela Dinâmica > Transpor Linhas e Colunas

Isso tem o mesmo efeito que arrastar todos os elementos da linha para a dimensão da coluna e arrastar todos os elementos da coluna para a dimensão da linha.

Agrupando linhas ou colunas

1. Selecione os rótulos para as linhas ou colunas que deseja agrupar (clique e arraste ou Shift+clique para selecionar vários rótulos).
2. Nos menus, escolha:
Editar > Grupo

Um rótulo do grupo é inserido automaticamente. Dê um clique duplo no rótulo do grupo para editar o texto do rótulo.

Nota: Para incluir linhas ou colunas em um grupo existente, primeiro deve-se desagrupar os itens que estão atualmente no grupo. Em seguida, é possível criar um novo grupo que inclua os itens adicionais.

Desagrupando linhas ou colunas

O desagrupamento automático exclui o rótulo do grupo.

Girando rótulos de linha e coluna

É possível girar rótulos entre as exibições horizontal e vertical para os rótulos de coluna mais internos e os rótulos de linha mais externos em uma tabela.

1. Nos menus, escolha:
Formatar > Girar Rótulos de Coluna Internos
ou
Formatar > Girar Rótulos de Coluna Externos

Somente os rótulos de coluna mais internos e os rótulos de linha mais externos podem ser girados.

Ordenando linhas

Para ordenar linhas de uma tabela dinâmica:

1. Ative a tabela.
2. Selecione qualquer célula na coluna que deseja usar para ordenar. Para ordenar somente um grupo selecionado de linhas, selecione duas ou mais células contíguas na coluna que deseja usar para ordenar.
3. Nos menus, escolha:
Editar > Ordenar Linhas
4. Selecione **Crescente** ou **Decrescente** no submenu.
 - Se a dimensão da linha contiver grupos, a ordenação afetará somente o grupo contendo a seleção.
 - Não é possível ordenar nos limites do grupo.
 - Não é possível ordenar tabelas com mais de um item na dimensão da linha.

Inserindo linhas e colunas

Para inserir uma linha ou coluna em uma tabela dinâmica:

1. Ative a tabela.
2. Selecione qualquer célula na tabela.
3. Nos menus, escolha:
Inserir Antes
ou
Inserir Depois
Nos submenu, escolha:
Linha
ou
Coluna
 - Um sinal de mais (+) é inserido em cada célula da nova linha ou coluna para evitar que a nova linha ou coluna seja automaticamente oculta por estar vazia.
 - Em uma tabela com dimensões aninhadas ou em camadas, uma coluna ou linha é inserida em cada nível de dimensão correspondente.

Controlando a exibição dos rótulos de valor e variável

Se a variável contiver rótulos de valor ou variável descritivos, é possível controlar a exibição dos nomes e rótulos de variáveis e os valores de dados e rótulos de valor em tabelas dinâmicas.

1. Ative a tabela dinâmica.
2. Nos menus, escolha:
Visualizar > Rótulos de Variáveis
ou
Visualizar > Rótulos de Valor
3. Selecione uma das opções a seguir no submenu:
 - **Nome** ou **Valor**. Somente nomes de variável (ou valores) são exibidos. Rótulos descritivos não são exibidos.
 - **Rótulo**. Somente rótulos descritivos são exibidos. Nomes de variável (ou valores) não são exibidos.
 - **Ambos**. Ambos os nomes (ou valores) e rótulos descritivos são exibidos.

Alterando a linguagem de saída

Para alterar a linguagem de saída em uma tabela dinâmica:

1. Ative a tabela
2. Nos menus, escolha:

Visualizar > Linguagem

3. Selecione uma das linguagens disponíveis.

A mudança da linguagem afeta somente o texto que é gerado pelo aplicativo, como títulos de tabela, rótulos de linha e coluna e texto da nota de rodapé. Nomes de variáveis e rótulos de valor e variável descritivos não são afetados.

Navegando em tabelas grandes

Para usar a janela de navegação para navegar em tabelas grandes:

1. Ative a tabela.
2. Nos menus, escolha:

Visualizar > Navegação

Desfazendo mudanças

É possível desfazer a mudança mais recente ou todas as mudanças em uma tabela dinâmica ativada. Ambas as ações se aplicam somente a mudanças feitas desde a ativação mais recente da tabela.

Para desfazer a mudança mais recente:

1. Nos menus, escolha:

Editar > Desfazer

Para desfazer todas as mudanças:

2. Nos menus, escolha:

Editar > Restaurar

Trabalhando com camadas

É possível exibir uma tabela bidimensional separada para cada categoria ou combinação de categorias. A tabela pode ser considerada como empilhada em camadas, com somente a camada superior visível.

Criando e exibindo camadas

Para criar camadas:

A movimentação dos elementos para a dimensão da camada cria uma tabela multidimensional, mas somente uma "fatia" bidimensional é exibida. A tabela visível é a da camada superior. Por exemplo, se uma variável categórica yes/no estiver na dimensão da camada, a tabela multidimensional terá duas camadas: uma para a categoria "yes" e uma para a categoria "no".

Alterando a camada exibida

1. Escolha uma categoria na lista suspensa de camadas (na própria tabela dinâmica, e não na bandeja de dinamização).

Ir para Categoria de Camada

Ir para Categoria de Camada permite alterar camadas em uma tabela dinâmica. Essa caixa de diálogo é particularmente útil quando há muitas camadas ou a camada selecionada possui muitas categorias.

Mostrando e ocultando itens

Muitos tipos de células podem ser ocultos, incluindo:

- Rótulos de dimensão
- Categorias, incluindo a célula do rótulo e as células de dados em uma linha ou coluna
- Rótulos de categoria (sem ocultar as células de dados)
- Notas de rodapé, títulos e legendas

Ocultando linhas e colunas em uma tabela

Mostrando linhas e colunas ocultas em uma tabela

1. Nos menus, escolha:

Visualizar > Mostrar Todas as Categorias

Isso exibe todas as linhas e colunas ocultas na tabela. (Se **Ocultar Linhas e Colunas Vazias** estiver selecionada em Propriedades da Tarefa para essa tabela, uma linha ou coluna completamente nova permanecerá oculta.)

Ocultando e mostrando rótulos de dimensão

1. Selecione o rótulo de dimensão ou qualquer rótulo de categoria dentro da dimensão.
2. No menu Visualizar ou no menu pop-up, escolha **Ocultar Rótulo de Dimensão** ou **Mostrar Rótulo de Dimensão**.

Ocultando e mostrando títulos de tabelas

Para ocultar um título:

TableLooks

Uma TableLook é um conjunto de propriedades que definem a aparência de uma tabela. É possível selecionar uma TableLook definida anteriormente ou criar sua própria TableLook.

- Antes ou depois de uma TableLook ser aplicada, é possível alterar formatos de célula para células individuais ou grupos de células usando propriedades de célula. Os formatos de célula editados permanecem intactos, mesmo quando você aplica uma nova TableLook.
- Opcionalmente, é possível reconfigurar todas as células para os formatos de célula que são definidos pela TableLook atual. Essa opção reconfigura todas as células que foram editadas. Se **Como Exibido** for selecionada na lista Arquivos TableLook, quaisquer células editadas serão reconfiguradas para as atuais propriedades da tabela.
- Somente propriedades da tabela definidas no diálogo Propriedades da Tabela são salvas em TableLooks. TableLooks não incluem modificações de células individuais.

Para aplicar uma TableLook

1. Ative a tabela dinâmica.
2. Nos menus, escolha:
Formatar > TableLooks...
3. Selecione uma TableLook na lista de arquivos. Para selecionar um arquivo de outro diretório, clique em **Navegar**.
4. Clique em **OK** para aplicar a TableLook na tabela dinâmica selecionada.

Para editar ou criar uma TableLook

1. Na caixa de diálogo TableLooks, selecione uma TableLook da lista de arquivos.
 2. Clique em **Editar Aparência**.
 3. Ajuste as propriedades da tabela para os atributos desejados e clique em **OK**.
 4. Clique em **Salvar Aparência** para salvar a TableLook editada ou clique em **Salvar Como** para salvá-la como uma nova TableLook.
- A edição da TableLook afeta somente a tabela dinâmica selecionada. Uma TableLook editada não é aplicada a nenhuma outra tabela que usa essa TableLook, a menos que você selecione essas tabelas e reaplique a TableLook.
 - Somente propriedades da tabela definidas no diálogo Propriedades da Tabela são salvas em TableLooks. TableLooks não incluem modificações de células individuais.

Propriedades de tabela

Propriedades da Tabela permite configurar propriedades gerais de uma tabela, e configurar estilos de célula para várias partes de uma tabela. É possível:

- Controlar propriedades gerais, como ocultar linhas ou colunas vazias e ajustar propriedades de impressão.
- Controlar o formato e a posição dos marcadores de nota de rodapé.
- Determinar os formatos específicos para células na área de dados, para rótulos de linha e coluna, e para outras áreas da tabela.
- Controlar a largura e a cor das linhas que formam as bordas de cada área da tabela.

Para alterar propriedades da tabela dinâmica

As novas propriedades são aplicadas à tabela dinâmica selecionada.

Propriedades da tabela: geral

Várias propriedades se aplicam à tabela como um todo. É possível:

- Mostrar ou ocultar linhas e colunas vazias. (Uma linha ou coluna vazia não tem nada em nenhuma das células de dados.)
- Controle a localização dos rótulos de linha, que podem estar no canto superior esquerdo ou aninhados.
- Controle as larguras máxima e mínima da coluna (expressas em pontos).

Para alterar as propriedades gerais da tabela:

1. Clique na guia **Geral**.
2. Selecione as opções desejadas.
3. Clique em **OK** ou **Aplicar**.

Configurar linhas para exibição: *Nota:* Esta variável se aplica somente às tabelas anteriores.

Por padrão, tabelas com muitas linhas são exibidas em seções de 100 linhas. Para controlar o número de linhas exibidas em uma tabela:

1. Selecione **Exibir Tabela por Linhas**.
2. Clique em **Configurar Linhas para Exibir**.
ou
3. No menu **Visualizar** de uma tabela dinâmica ativada, escolha **Exibir Tabela por Linhas** e **Configurar Linhas para Exibição**.

Linhas para Exibir. Controla o número máximo de linhas para exibir de uma vez. Os controles de navegação lhe permitem mover para diferentes seções da tabela. O valor mínimo é 10. O padrão é 100.

Tolerância Viúva/Órfã. Controla o número máximo de linhas na dimensão da linha mais interna da tabela para dividir entre as visualizações exibidas da tabela. Por exemplo, se houvesse seis categorias em cada grupo da dimensão da linha mais interna, a especificação de um valor de seis evitaria que qualquer grupo fosse dividido entre as visualizações exibidas. Essa configuração pode fazer com que o número total de linhas em uma visualização exibida exceda o número máximo especificado de linhas para exibição.

Propriedades da tabela: notas

A guia **Notas** do diálogo **Propriedades da Tabela** controla a formatação da nota de rodapé e o texto de comentário da tabela.

Notas de rodapé. As propriedades dos marcadores de notas de rodapé incluem estilo e posição em relação ao texto.

- O estilo dos marcadores de notas de rodapé é numérico (1, 2, 3,...) ou letras (a, b, c, ...).

- Os marcadores de nota de rodapé podem ser anexados ao texto como sobrescritos ou subscritos.

Texto de Comentário. É possível incluir texto de comentário para cada tabela.

- O texto do comentário é exibido em uma dica de ferramenta quando você passa o mouse sobre um tabela no Visualizador.
- Leitores de tela lêem o texto do comentário quando a tabela está em foco.
- A dica de ferramenta no Visualizador exibe somente os primeiros 200 caracteres do comentário, mas os leitores de tela lêem o texto inteiro.
- Quando você exporta uma saída para HTML ou relatório da web, o texto do comentário é usado como um texto alternativo.

Propriedades da tabela: formatos de célula

Para formatação, uma tabela é dividida em áreas: título, camadas, rótulos de canto, rótulos de linha, rótulos de coluna, dados, legenda e notas de rodapé. Para cada área de uma tabela, é possível modificar os formatos de célula associados. Formatos de célula incluem características de texto (como fonte, tamanho, cor e estilo), alinhamentos horizontal e vertical, cores do plano de fundo e margens da célula interna.

Formatos de célula são aplicados a áreas (categorias de informações). Eles não são características de células individuais. Essa distinção é uma consideração importante durante a colocação de uma tabela sobre o eixo.

Por exemplo,

- Se você especificar uma fonte em negrito como um formato de célula dos rótulos da coluna, os rótulos da coluna aparecerão em negrito, independentemente de quais informações estão sendo exibidas atualmente na dimensão da coluna. Se você mover um item da dimensão da coluna para outra dimensão, isso não reterá a característica de negrito dos rótulos da coluna.
- Se você deixar os rótulos da coluna em negrito simplesmente destacando as células em uma tabela dinâmica ativada e clicando no botão Negrito na barra de ferramentas, o conteúdo dessas células permanecerá em negrito, independentemente da dimensão para a qual ele será movido, e os rótulos da coluna não reterão a característica de negrito para outros itens movidos para a dimensão da coluna.

Para alterar formatos de célula:

1. Selecione a guia **Formatos de Célula**.
2. Selecione uma Área na lista suspensa ou clique em uma área da amostra.
3. Selecione características para a área. Suas seleções serão refletidas na amostra.
4. Clique em **OK** ou **Aplicar**.

Alternando cores de linhas

Para aplicar uma cor de plano de fundo e/ou de texto para alternar linhas na área Dados da tabela:

1. Selecione **Dados** na lista suspensa Área.
2. Selecione (marque) **Cor de Linha Alternativa** no grupo Cor do Plano de Fundo.
3. Selecione as cores para usar para o texto e o plano de fundo da linha alternativos.

Cores de linha alternativas afetam somente a área de Dados da tabela. Elas não afetam as áreas do rótulo da coluna ou linha.

Propriedades da tabela: bordas

Para cada localização da borda em uma tabela, é possível selecionar um estilo de linha e uma cor. Se você selecionar **Nenhum** como estilo, não haverá linha na localização selecionada.

Para alterar bordas da tabela:

1. Clique na guia **Bordas**.
2. Selecione uma localização da borda, clicando em seu nome na lista ou clicando em uma linha na área Amostra.
3. Selecione um estilo de linha ou selecione **Nenhum**.
4. Selecione uma cor.
5. Clique em **OK** ou **Aplicar**.

Propriedades da tabela: impressão

É possível controlar as seguintes propriedades para tabelas dinâmicas impressas:

- Imprimir todas as camadas ou apenas a camada superior da tabela e imprimir cada camada em uma página separada.
- Reduzir uma tabela horizontalmente ou verticalmente para ajustar à página para impressão.
- Controlar linhas viúvas/órfãs controlando o número mínimo de linhas e colunas que estarão contidas em qualquer seção impressa de uma tabela, caso a tabela seja muito grande e/ou muito longa para o tamanho de página definido.
Nota: Se uma tabela for muito longa para se ajustar à página atual porque há outra saída acima dela, mas ela se ajustar ao comprimento de página definido, a tabela será impressa automaticamente em uma nova página, independentemente da configuração viúva/órfã.
- Incluir texto de continuação para tabelas que não se ajustam a uma única página. É possível exibir texto de continuação na parte inferior de cada página e na parte superior de cada página. Se nenhuma opção for selecionada, o texto de continuação não será exibido.

Para controlar propriedades de impressão da tabela dinâmica:

1. Clique na guia **Impressão**.
2. Selecione as opções de impressão que deseja.
3. Clique em **OK** ou **Aplicar**.

Propriedades de célula

Propriedades de célula são aplicadas a uma célula selecionada. É possível alterar a fonte, formato do valor, alinhamento, margens e cores. Propriedades de célula substituem propriedades de tabela; portanto, se você alterar as propriedades da tabela, não altere nenhuma propriedade de célula aplicada individualmente.

Para alterar as propriedades da célula:

Fonte e segundo plano

A guia Fonte e Segundo Plano controla o estilo de fonte, a cor e a cor do plano de fundo das células selecionadas na tabela.

Formatar valor

A guia Formatar Valor controla os formatos de valor para as células selecionadas. É possível selecionar formatos para números, datas, horas ou moedas, além de ser possível ajustar o número de dígitos decimais exibido.

Alinhamento e margens

A guia Alinhamento e Margens controla os alinhamentos horizontal e vertical dos valores e as margens superior, inferior, esquerda e direita dos alinhamentos selecionados. O alinhamento horizontal **Misturado** alinha o conteúdo de cada célula de acordo com seu tipo. Por exemplo, datas são alinhadas à direita e valores de texto são alinhados à esquerda.

Notas de rodapé e legendas

É possível incluir notas de rodapé e legendas em uma tabela. Também é possível ocultar notas de rodapé ou legendas, alterar marcadores de nota de rodapé e renumerar notas de rodapé.

Incluindo notas de rodapé e legendas

Para incluir uma legenda em uma tabela:

Uma nota de rodapé pode ser anexada a qualquer item em uma tabela. Para incluir uma nota de rodapé:

1. Clique em um título, célula ou legenda dentro de uma tabela dinâmica ativada.
2. No menu Inserir, escolha **Nota de Rodapé**.
3. Insira o texto da nota de rodapé na área fornecida.

Para ocultar ou mostrar uma legenda

Para ocultar uma legenda:

1. Selecione a legenda.
2. No menu Visualizar, escolha **Ocultar**.

Para mostrar legendas ocultas:

1. No menu Visualizar, escolha **Mostrar Todos**.

Para ocultar ou mostrar uma nota de rodapé em uma tabela

Para ocultar uma nota de rodapé:

1. Clique com o botão direito na célula que contém a referência de nota de rodapé e selecione **Ocultar Notas de Rodapé** no menu pop-up
ou
2. Selecione a nota de rodapé na área da nota de rodapé da tabela e selecione **Ocultar** no menu pop-up.
Nota: Para tabelas anteriores, selecione a área da nota de rodapé da tabela, selecione **Editar Nota de Rodapé** no menu pop-up e, em seguida, cancele a seleção (desmarque) da propriedade Visible para quaisquer notas de rodapé que deseja ocultar.
Se uma célula contiver diversas notas de rodapé, use o último método para ocultar seletivamente as notas de rodapé.

Para ocultar todas as notas de rodapé na tabela:

1. Selecione todas as notas de rodapé na área de nota de rodapé da tabela (use clicar e arrastar ou Shift+clique para selecionar as notas de rodapé) e selecione **Ocultar** no menu Visualizar.

Para mostrar notas de rodapé ocultas:

1. Selecione **Mostrar Todas as Notas de Rodapé** no menu Visualizar.

Marcador de Nota de Rodapé

O Marcador de Nota de Rodapé altera os caracteres que podem ser usados para marcar uma nota de rodapé. Por padrão, marcadores de nota de rodapé padrão são letras ou números sequenciais, dependendo das configurações de propriedades da tabela. Também é possível designar um marcador especial. Marcadores especiais não são afetados quando você renumera notas de rodapé ou se alterna entre números e letras para marcadores padrão. A exibição dos números ou letras para marcadores padrão e as posições subscrita ou sobrescrita dos marcadores de nota de rodapé são controladas pela guia Notas de Rodapé do diálogo Propriedades da Tabela.

Para alterar marcadores de nota de rodapé:

Marcadores especiais são limitados a 2 caracteres. Notas de rodapé com marcadores especiais precedem notas de rodapé com letras ou números sequenciais na área da nota de rodapé da tabela; portanto, a mudança para um marcador especial pode reordenar a lista de notas de rodapé.

Renumerando notas de rodapé

Quando você tiver colocado uma tabela sobre o eixo alternando linhas, colunas e camadas, as notas de rodapé podem ficar fora de ordem. Para renumerar as notas de rodapé:

Editando notas de rodapé em tabelas anteriores

Para tabelas anteriores, é possível usar o diálogo Editar Notas de Rodapé para inserir e modificar configurações de fonte e texto de nota de rodapé, alterar marcadores de nota de rodapé e ocultar ou excluir notas de rodapé seletivamente.

Quando você insere uma nova nota de rodapé em uma tabela anterior, o diálogo Editar Notas de Rodapé abre automaticamente. Para usar o diálogo Editar Notas de Rodapé para editar notas de rodapé existentes (sem criar uma nova nota de rodapé):

Marcador. Por padrão, marcadores de nota de rodapé padrão são letras ou números sequenciais, dependendo das configurações de propriedades da tabela. Para designar um marcador especial, basta inserir o novo valor do marcador na coluna Marcador. Marcadores especiais não são afetados quando você renumera notas de rodapé ou se alterna entre números e letras para marcadores padrão. A exibição dos números ou letras para marcadores padrão e as posições subscrita ou sobrescrita dos marcadores de nota de rodapé são controladas pela guia Notas de Rodapé do diálogo Propriedades da Tabela. Consulte o tópico “Propriedades da tabela: notas” na página 88 para obter mais informações.

Para alterar um marcador de especial de volta para um marcador padrão, clique nele com o botão direito no diálogo Editar Notas de Rodapé, selecione **Marcador de Nota de Rodapé** no menu pop-up e selecione Marcador Padrão na caixa de diálogo Marcador de Nota de Rodapé.

Nota de Rodapé. O conteúdo da nota de rodapé. A exibição reflete as atuais configurações de fonte e segundo plano. As configurações de fonte podem ser alteradas para notas de rodapé individuais usando o subdiálogo Formatar. Consulte o tópico “Configurações de cor e fonte da nota de rodapé” para obter mais informações. Uma única cor de plano de fundo é aplicada a todas as notas de rodapé e pode ser alterada na guia Fonte e Segundo Plano do diálogo Propriedades da Célula. Consulte o tópico “Fonte e segundo plano” na página 90 para obter mais informações.

Visível. Todas as notas de rodapé ficam visíveis por padrão. Cancele a seleção (desmarque) da caixa de seleção para ocultar uma nota de rodapé.

Configurações de cor e fonte da nota de rodapé: Para tabelas anteriores, é possível usar o diálogo Formatar para alterar a família da fonte, seu estilo, tamanho e cor para uma ou mais das notas de rodapé selecionadas:

A família da fonte, o estilo, o tamanho e as cores selecionados são aplicados a todas as notas de rodapé selecionadas.

A cor do plano de fundo, o alinhamento e as margens podem ser configurados no diálogo Propriedades da Célula e aplicados a todas as notas de rodapé. Não é possível alterar essas configurações para notas de rodapé individuais. Consulte o tópico “Fonte e segundo plano” na página 90 para obter mais informações.

Larguras de células de dados

Configurar Largura da Célula de Dados é usada para configurar todas as células de dados para a mesma largura.

Para configurar a largura para todas as células de dados:

1. Nos menus, escolha:
Formatar > Configurar Largura da Célula de Dados...
2. Insira um valor para a largura da célula.

Alterando a largura da coluna

Exibindo bordas ocultas em uma tabela dinâmica

Para tabelas sem muitas bordas visíveis, é possível exibir as bordas ocultas. Isso pode simplificar tarefas como alterar as larguras das colunas.

1. No menu Visualizar, escolha **Linhas de Grade**.

Selecionando linhas, colunas e células em uma tabela dinâmica

É possível selecionar uma linha ou coluna inteira ou um conjunto especificado de células de rótulos ou dados.

Para selecionar várias células:

Selecione > Células de Rótulos e Dados

Imprimindo tabelas dinâmicas

Vários fatores podem afetar a aparência de tabelas dinâmicas impressas, e esses fatores podem ser controlados alterando atributos de tabela dinâmica.

- Para tabelas dinâmicas multidimensionais (tabelas com camadas), é possível imprimir todas as camadas ou somente a superior (visível). Consulte o tópico “Propriedades da tabela: impressão” na página 90 para obter mais informações.
- Para tabelas dinâmicas longas ou grandes, é possível redimensionar automaticamente a tabela para ajustar à página ou controlar a localização de quebras de tabela e quebras de página. Consulte o tópico “Propriedades da tabela: impressão” na página 90 para obter mais informações.

Use Visualização de Impressão no menu Arquivo para ver a aparência das tabelas dinâmicas impressas.

Controlando quebras de tabela para tabelas grandes e longas

Tabelas dinâmicas que são muito grandes ou muito longas para imprimir dentro do tamanho de página definido são automaticamente divididas e impressas em várias seções. É possível:

- Controlar localizações de linha e coluna onde tabelas grades são divididas.
- Especificar linhas e colunas que devem ser mantidas juntas quando tabelas forem divididas.
- Escalar novamente as tabelas grandes para ajustá-las ao tamanho de página definido.

Para especificar quebras de linha e coluna para tabelas dinâmicas impressas:

1. Ative a tabela dinâmica.
2. Clique em qualquer célula na coluna à esquerda de onde deseja inserir a quebra ou clique em qualquer célula na linha antes da linha na qual deseja inserir a quebra.
3. Nos menus, escolha:

Formatar > Pontos de Interrupção > Ponto de Interrupção Vertical

ou

Formatar > Pontos de Interrupção > Ponto de Interrupção Horizontal

Para especificar linhas ou colunas para manter juntas:

1. Selecione os rótulos das linhas ou colunas que deseja manter juntas. Clique e arraste ou pressione Shift+clique para selecionar vários rótulos de linha ou coluna.
2. Nos menus, escolha:

Formatar > Pontos de Interrupção > Manter Junto

Para visualizar pontos de interrupção e manter os grupos juntos:

1. Nos menus, escolha:

Formatar > Pontos de Interrupção > Exibir Pontos de Interrupção

Pontos de interrupção são mostrados como linhas verticais ou horizontais. Os grupos mantidos juntos aparecem como regiões retangulares esmaecidas dentro de uma borda mais escura.

Nota: A exibição de pontos de interrupção e de grupos mantidos juntos não é suportada para tabelas anteriores.

Para limpar pontos de interrupção e grupos mantidos juntos

Para limpar um ponto de interrupção:

1. Clique em qualquer célula na coluna à esquerda de um ponto de interrupção vertical ou clique em qualquer célula na linha acima de um ponto de interrupção horizontal.

2. Nos menus, escolha:

Formatar > Pontos de Interrupção > Limpar Ponto de Interrupção ou Grupo

Para limpar um grupo mantido junto:

3. Selecione os rótulos de coluna ou linha que especificam o grupo.

4. Nos menus, escolha:

Formatar > Pontos de Interrupção > Limpar Ponto de Interrupção ou Grupo

Todos os pontos de interrupção e grupos mantidos juntos são limpos automaticamente quando você reordena ou coloca sobre o eixo qualquer linha ou coluna. Esse comportamento não se aplica a tabelas anteriores.

Criando um gráfico a partir de uma tabela dinâmica

1. Dê um clique duplo na tabela dinâmica para ativá-la.
2. Selecione as linhas, colunas ou células que deseja exibir no gráfico.
3. Clique com o botão direito em qualquer lugar na área selecionada.
4. Escolha **Criar Gráfico** no menu pop-up e selecione um tipo de gráfico.

Tabelas anteriores

É possível escolher renderizar tabelas como tabelas anteriores (chamadas tabelas com recursos completos na liberação 19) que sejam totalmente compatíveis com liberações do IBM SPSS Statistics anteriores à 20. Tabelas anteriores podem ser renderizadas lentamente e são recomendadas somente se a compatibilidade com versões anteriores à 20 for necessária. Para obter informações sobre como criar tabelas anteriores, consulte “Opções de tabela dinâmica” na página 95.

Opções

Opções

Opções controlam várias configurações.

Para alterar configurações de opções

1. Nos menus, escolha:
Editar > Opções...
2. Clique nas guias para as configurações que deseja alterar.
3. Altere as configurações.
4. Clique em **OK** ou **Aplicar**.

Opções gerais

Número Máximo de Encadeamentos

O número de encadeamentos que os procedimentos multiencadeados usam quando calculam resultados. A configuração **Automática** é baseada no número de núcleos de processamento disponíveis. Especifique um valor inferior se deseja tornar disponíveis mais recursos de processamento para outros aplicativos enquanto os procedimentos multiencadeados estão sendo executados. Esta opção está desativada no modo de análise distribuída.

Out

Exibe um zero à esquerda para valores decimais. Exibe zeros à esquerda para valores numéricos que consistem apenas de uma parte decimal. Por exemplo, quando zeros à esquerda são exibidos, o valor .123 é exibido como 0.123. Esta configuração não se aplica a valores numéricos que tenham um formato de porcentagem ou moeda. Exceto por arquivos ASCII (*.dat) corrigidos, zeros à esquerda não são incluídos quando os dados forem salvos em um arquivo externo.

Sistema de medição. O sistema de medição usado (pontos, polegadas ou centímetros) para especificar atributos como margens da célula da tabela dinâmica, larguras da célula e espaço entre tabelas para impressão.

Opções do Visualizador

As opções de exibição de saída do Visualizador afetam somente a nova saída que é produzida após a mudança das configurações. A saída que já é exibida no Visualizador não é afetada pelas mudanças nessas configurações.

Estado de Saída Inicial. Controla quais itens são exibidos ou ocultos automaticamente cada vez que você executa um procedimento e como os itens são alinhados inicialmente. É possível controlar a exibição dos seguintes itens: log, avisos, notas, títulos, tabelas dinâmicas, gráficos, diagramas de árvore e saída de texto. Também é possível ativar ou desativar a exibição de comandos no log. É possível copiar a sintaxe de comando do log e salvar em um arquivo de sintaxe.

Nota: Todos os itens de saída são exibidos alinhados à esquerda no Visualizador. Somente o alinhamento da saída impressa é afetado pelas configurações de justificação. Os itens centralizados e alinhados à direita são identificados por um pequeno símbolo.

Título. Controla o estilo de fonte, o tamanho e a cor dos novos títulos de saída.

Título da Página. Controla o estilo de fonte, o tamanho e a cor dos novos títulos de página e títulos de página que são gerados pela sintaxe de comando TITLE e SUBTITLE ou criados por **Novo Título da Página** no menu **Inserir**.

Saída de Texto Fonte usada para saída de texto. A saída de texto é projetada para uso com uma fonte monoespaçada (densidade fixa). Se você selecionar uma fonte proporcional, o resultado tabular não será adequadamente alinhado.

Configuração da Página Padrão. Controla as opções padrão para orientação e margens para impressão.

Opções de tabela dinâmica

As opções Tabela Dinâmica configuram várias opções para a exibição de tabelas dinâmicas.

TableLook

Selecione uma TableLook na lista de arquivos e clique em **OK** ou **Aplicar**. É possível usar uma das TableLooks fornecidas com o IBM SPSS Statistics ou criar a sua própria no Editor de Tabela Dinâmica (escolha **TableLooks** no menu **Formatar**).

- **Navegar**. Permite selecionar uma TableLook de outro diretório.
- **Configurar Diretório da TableLook**. Permite alterar o diretório da TableLook padrão. Use **Navegar** para navegar para o diretório que deseja usar, selecione uma TableLook nesse diretório e, em seguida, selecione **Configurar Diretório da TableLook**.

Nota: TableLooks criadas em versões anteriores do IBM SPSS Statistics não podem ser usadas na versão 16.0 ou posterior.

Larguras da coluna

Estas opções controlam o ajustamento automático das larguras de colunas em tabelas dinâmicas.

- **Ajustar somente para rótulos**. Ajusta a largura da coluna para a largura do rótulo da coluna. Isso produz tabelas mais compactas, mas valores de dados maiores que o rótulo podem ser truncados.
- **Ajustar para rótulos e dados para todas as tabelas**. Ajusta a largura da coluna para o que for maior: o rótulo da coluna ou o maior valor de dados. Isso produz tabelas maiores, mas assegura que todos os valores sejam exibidos.

Modo de edição padrão

Esta opção controla a ativação de tabelas dinâmicas na janela Visualizador ou em uma janela separada. Por padrão, um clique duplo em uma tabela dinâmica ativa todas as tabelas, exceto as muito grandes, na janela Visualizador. É possível escolher ativar tabelas dinâmicas em uma janela separada ou selecionar uma configuração de tamanho que abra tabelas dinâmicas menores na janela Visualizador e tabelas dinâmicas maiores em uma janela separada.

Copiando tabelas grandes na área de transferência em formato Rich Text

Quando tabelas dinâmicas são coladas em formato Word/RTF, as tabelas muito grandes para a largura do documento são quebradas, diminuídas para se ajustarem à largura do documento ou deixadas inalteradas.

Opções de saída

Opções de saída controlam a configuração padrão para inúmeras opções de saída.

Acessibilidade do Leitor de Tela. Controla como rótulos de linha e coluna da tabela dinâmica são lidos pelos leitores de tela. É possível ler os rótulos de linha e coluna completos para cada célula de dados ou ler somente os rótulos que mudam conforme você se move entre células de dados na tabela.

Capítulo 7. Manipulando valores omissos

Visão geral de valores omissos

Durante a fase Preparação de Dados da mineração de dados, geralmente você deseja substituir valores omissos nos dados. **Valores omissos** são valores no conjunto de dados que são desconhecidos, não coletados ou inseridos incorretamente. Geralmente esses valores são inválidos para seus campos. Por exemplo, o campo *Sexo* deve conter os valores *M* e *F*. Se descobrir os valores *Y* ou *Z* no campo, você pode assumir seguramente que esses valores são inválidos e, portanto, devem ser interpretados como em branco. Da mesma forma, um valor negativo para o campo *Idade* não faz sentido e também deve ser interpretado como em branco. Muitas vezes, esses valores obviamente incorretos são inseridos, ou campos deixados em branco, propositalmente durante um questionário para indicar a ausência de uma resposta. Às vezes, você pode querer examinar esses campos em branco mais de perto para determinar se uma ausência de resposta, como uma recusa em fornecer a idade de alguém, é um fator na predição de um resultado específico.

Algumas técnicas de modelagem manipulam dados omissos melhor que outras. Por exemplo, C5.0 e Apriori lidam bem com valores declarados explicitamente como "omissos" em um nó Tipo. Outras técnicas de modelagem têm problemas para lidar com valores omissos e enfrentam tempos de treinamento mais longos, resultando em modelos menos precisos.

Há vários tipos de valores omissos reconhecidos pelo IBM SPSS Modeler:

- **Valores omissos do sistema ou nulos.** Esses são valores de não sequência de caracteres que foram deixados em branco no arquivo de origem ou banco de dados e não foram definidos especificamente como "omissos" em um nó Tipo ou de origem. Valores omissos do sistema são exibidos como **\$null\$**. Observe que sequências de caracteres vazias não são consideradas nulas no IBM SPSS Modeler, embora elas possam ser tratadas como nulas por certos bancos de dados.
- **Sequências de caracteres vazias e espaço em branco.** Valores de sequência de caracteres vazia e espaços em branco (sequências de caracteres sem caracteres visíveis) são tratados como sendo diferentes de valores nulos. Sequências de caracteres vazias são tratadas como sendo equivalentes a espaço em branco para a maioria dos propósitos. Por exemplo, se você selecionar a opção para tratar um espaço em branco como brancos em um nó Tipo ou de origem, essa configuração se aplicará também às sequências de caracteres vazias.
- **Valores omissos definidos pelo usuário ou em branco.** São valores como desconhecido, 99 ou -1 que são definidos explicitamente em um nó de origem ou nó de Tipo como omissos. Opcionalmente, também é possível escolher tratar nulos e espaços em branco como brancos, o que permite que eles sejam sinalizados para tratamento especial e sejam excluídos da maioria dos cálculos. Por exemplo, é possível usar a função @BLANK para tratar esses valores, com outros tipos de valores omissos, como brancos.

Leitura de dados combinados. Observe que quando você está lendo campos com armazenamento numérico (número inteiro, real, hora, registro de data e hora ou data), quaisquer valores não numéricos são configurados para *nulo* ou *omisso do sistema*. Isso acontece porque, ao contrário de alguns aplicativos, não é permitido combinar tipos de armazenamento dentro de um campo. Para evitar isso, quaisquer campos com dados combinados devem ser lidos como sequências de caracteres alterando o tipo de armazenamento no nó de origem ou aplicativo externo, conforme necessário.

Lendo sequências de caracteres vazias do Oracle. Durante a leitura ou gravação em um banco de dados Oracle, saiba que, ao contrário do IBM SPSS Modeler e da maioria dos outros bancos de dados, o Oracle trata e armazena valores de sequência de caracteres vazios como equivalentes a valores nulos. Isso

significa que os mesmos dados extraídos de um banco de dados Oracle podem se comportar diferentemente quando extraídos de um arquivo ou outro banco de dados, e os dados podem retornar resultados diferentes.

Manipulando valores omissos

Você deve decidir como tratar valores omissos considerando seus conhecimentos de negócios ou domínio. Para facilitar o tempo de treinamento e aumentar a precisão, você pode querer remover espaços em branco do conjunto de dados. Por outro lado, a presença de valores em branco pode levar a novas oportunidades de negócios ou insights adicionais. Ao escolher a melhor técnica, é necessário considerar os seguintes aspectos de seus dados:

- Tamanho do conjunto de dados
- Número de campos contendo espaços em branco
- Quantidade de informações omissas

Em termos gerais, há duas abordagens que podem ser seguidas:

- É possível excluir campos ou registros com valores omissos
- É possível imputar, substituir ou forçar valores omissos usando uma variedade de métodos

Ambas as abordagens podem ser amplamente automatizadas usando o nó Auditoria de Dados. Por exemplo, é possível gerar um nó de Filtro que exclua campos com muitos valores omissos para ser útil na modelagem e gerar um Supernode que impute valores omissos para qualquer ou para todos os campos restantes. É daí que vem o verdadeiro poder da auditoria, permitindo não só avaliar o estado atual de seus dados, mas também tomar uma ação com base na avaliação.

Manipulando registros com valores omissos

Se a maioria dos valores omissos estiver concentrada em um pequeno número de registros, é possível simplesmente excluir esses registros. Por exemplo, um banco geralmente mantém registros completos e detalhados sobre seus clientes de empréstimo. No entanto, se o banco for menos restritivo para aprovar empréstimos para os membros de sua própria equipe, os dados reunidos para empréstimos da equipe provavelmente terão vários campos em branco. Nesse caso, há duas opções para manipular esses valores omissos:

- É possível usar um nó de Seleção para remover registros da equipe.
- Se os dados configurados forem grandes, é possível descartar todos os registros com valores em branco.

Manipulando campos com valores omissos

Se a maioria dos valores omissos estiver concentrada em um pequeno número de campos, é possível abordá-la em nível de campo em vez de em nível de registro. Essa abordagem também permite experimentar a importância relativa de determinados campos antes de você decidir sobre uma abordagem para manipular valores omissos. Se um campo for irrelevante na modelagem, provavelmente não valerá a pena mantê-lo, independentemente de quantos valores omissos ele tem.

Por exemplo, uma empresa de pesquisa de mercado pode coletar dados de um questionário geral com 50 questões. Duas das questões abordam idade e opinião política, informações que muitas pessoas relutam em fornecer. Nesse caso, *Idade* e *Opinião Política* podem ter muitos valores omissos.

Nível de Medição de Campo

Ao determinar qual método usar, também é necessário considerar o nível de medição dos campos com valores omissos.

Campos Numéricos. Para tipos de campo numéricos, como *Contínuo*, é sempre necessário eliminar quaisquer valores não numéricos antes da construção de um modelo, pois muitos modelos não funcionarão se espaços em branco forem incluídos em campos numéricos.

Campos Categóricos. Para campos categóricos, como *Nominal* e *Flag*, a alteração de valores omissos não é necessária, mas aumenta a precisão do modelo. Por exemplo, um modelo que usa o campo *Sexo* ainda funcionará com valores insignificantes, como *Y* e *Z*, mas a remoção de todos os valores além de *M* e *F* aumentará a precisão do modelo.

Rastreamento ou Removendo Campos

Para rastrear campos com muitos valores omissos, você tem várias opções:

- É possível usar um nó Auditoria de Dados para filtrar campos com base na qualidade.
- É possível usar um nó Seleção de Variável para rastrear campos com mais de uma porcentagem especificada de valores omissos e ranquear campos com base na importância relativa a uma resposta especificada.
- Em vez de remover os campos, é possível usar um nó Tipo para configurar o papel do campo para **Nenhum**. Isso manterá os campos no conjunto de dados, mas os excluirá dos processos de modelagem.

Tratamento de registros com valores omissos do sistema

O que são valores omissos do sistema?

Valores omissos do sistema representam valores de dados que não são conhecidos ou não são aplicáveis. Em bancos de dados, esses valores costumam ser chamados de valores *NULL*.

Valores omissos do sistema são diferentes de valores em branco. Valores em branco normalmente são definidos no nó Tipo como valores específicos, ou intervalos de valores, que podem ser considerados omissos definidos pelo usuário. Valores em branco são manipulados diferentemente no contexto de modelagem.

Construindo valores omissos do sistema

Valores omissos do sistema podem estar presentes nos dados que são lidos a partir de uma origem de dados (por exemplo, tabelas de banco de dados podem conter valores *NULL*).

Valores omissos do sistema podem ser construídos usando o valor **undef** em expressões. Por exemplo, a expressão do CLEM a seguir retorna Idade, se for menor que ou igual a 30, ou um valor omissos, se for maior que 30:

```
if Age > 30 then undef else Age endif
```

Valores omissos também podem ser criados quando uma junção externa é executada, quando um número é dividido por zero, quando a raiz quadrada de um número negativo é calculada e em outras situações.

Exibindo valores omissos do sistema

Valores omissos do sistema são exibidos em tabelas e outras saídas como `$null$`.

Testando valores omissos do sistema

Use a função especial **@NULL** para retornar true se o valor de argumento for um valor omissos do sistema, por exemplo:

```
if @NULL(MyFieldName) then 'It is null' else 'It is not null' endif
```

Valores omissos do sistema passados para funções

Valores omissos do sistema que são passados para funções geralmente propagam valores omissos para a saída. Por exemplo, se o valor do campo f1 for um valor omissos do sistema em uma determinada linha, a expressão $\log(f1)$ também será avaliada para um valor omissos do sistema para essa linha. Uma exceção é a função @NULL.

Valores omissos do sistema em expressões que envolvem operadores aritméticos

A aplicação de operadores aritméticos em valores que incluem um valor omissos do sistema resulta em um valor omissos do sistema. Por exemplo, se o valor do campo f1 for um valor omissos do sistema em uma determinada linha, a expressão $f1 + 10$ também será avaliada para um valor omissos do sistema para essa linha.

Valores omissos do sistema em expressões que envolvem operadores lógicos

Quando você trabalha com valores omissos do sistema em expressões que envolvem operadores lógicos, as regras de três lógicas de valor (*verdadeiro*, *falso* e *omisso*) se aplicam e podem ser descritas em tabelas de verdade. As tabelas de verdade para os operadores lógicos comuns de *not*, *and* e *or* são mostradas nas seguintes tabelas.

Tabela 4. Tabela de verdade para NOT

Operando	Operando NOT
verdadeiro	falso
falso	verdadeiro
omisso	omisso

Tabela 5. Tabela de verdade para AND

Operand1	Operand2	Operand1 AND Operand2
verdadeiro	verdadeiro	verdadeiro
verdadeiro	falso	falso
verdadeiro	omisso	omisso
falso	verdadeiro	falso
falso	falso	falso
falso	omisso	falso
omisso	verdadeiro	omisso
omisso	falso	falso
omisso	omisso	omisso

Tabela 6. Tabela de verdade para OR

Operand1	Operand2	Operand1 OR Operand2
verdadeiro	verdadeiro	verdadeiro
verdadeiro	falso	verdadeiro
verdadeiro	omisso	verdadeiro
falso	verdadeiro	verdadeiro
falso	falso	falso
falso	omisso	omisso
omisso	verdadeiro	verdadeiro

Tabela 6. Tabela de verdade para OR (continuação)

Operand1	Operand2	Operand1 OR Operand2
omisso	falso	omisso
omisso	omisso	omisso

Valores omissos do sistema em expressões que envolvem operadores de comparação

Quando você compara um valor omisso do sistema e um valor não omisso do sistema, o resultado é avaliado para um valor omisso do sistema em vez de para um resultado verdadeiro ou falso. Valores omissos do sistema podem ser comparados uns com os outros; dois valores omissos do sistema são considerados iguais.

Valores omissos do sistema em expressões if/then/else/endif

Quando você usa uma expressão condicional, e a expressão condicional retorna um valor omisso do sistema, o valor da cláusula else é retornado da expressão condicional.

Valores omissos do sistema no nó de Seleção

Quando, para um determinado registro, a expressão de seleção é avaliada para um valor omisso, o registro não é emitido do nó de Seleção (essa ação se aplica aos modos Inclusão e Descarte).

Valores omissos do sistema no nó de Mesclagem

Quando você mescla usando uma chave, nenhum registro que tem valores omissos do sistema em um campo-chave é mesclado.

Valores omissos do sistema em agregação

Quando você agrega dados em colunas, valores omissos não são incluídos no cálculo. Por exemplo, em uma coluna com três valores { 1, 2 e undef }, a soma dos valores na coluna é calculada como 3; o valor médio é calculado como 1.5.

Imputando ou preenchendo valores omissos

Em casos em que há somente alguns valores omissos, talvez seja útil inserir valores para substituir espaços em branco. Isso pode ser feito a partir do relatório Auditoria de Dados, que permite especificar opções para campos específicos conforme apropriado e gerar um SuperNode que impute valores usando inúmeros métodos. Esse é o método mais flexível, além de permitir especificar a manipulação de grandes números de campos em um único nó.

Os métodos a seguir estão disponíveis para imputar valores omissos:

Fixo. Substitui um valor fixo (a média do campo, o ponto médio do intervalo ou a constante especificada).

Aleatório. Substitui um valor aleatório com base em uma distribuição uniforme ou normal.

Expressão. Permite especificar uma expressão customizada. Por exemplo, você poderia substituir valores por uma variável global criada pelo nó Configurar Globais.

Algoritmo. Substitui um valor predito por um modelo baseado no algoritmo C&RT. Para cada campo imputado usando esse método, haverá um modelo C&RT separado, com um nó de Preenchimento que substitui brancos e nulos pelo valor predito pelo modelo. Um nó de Filtro é usado para remover os campos de predição gerados pelo modelo.

Alternativamente, para forçar valores para campos específicos, é possível usar um nó Tipo para assegurar que os tipos de campo cubram somente valores legais e configurar a coluna *Verificação* para **Forçar** para os campos cujos valores em branco precisam de substituição.

Funções do CLEM para valores omissos

Há várias funções usadas para manipular valores omissos. As funções a seguir geralmente são usadas nos nós de Seleção e Preenchimento para descartar ou preencher valores omissos:

- `count_nulls(LIST)`
- `@BLANK(FIELD)`
- `@NULL(FIELD)`
- `undef`

As funções @ podem ser usadas com a função @FIELD para identificar a presença de valores em branco ou nulos em um ou mais campos. Os campos podem simplesmente ser sinalizados quando houver valores em branco ou nulos presentes, ou podem ser preenchidos com valores de substituição ou usados em várias outras operações.

É possível contar os nulos em uma lista de campos, como a seguir:

```
count_nulls(['cardtenure' 'card2tenure' 'card3tenure'])
```

Durante o uso de qualquer uma das funções que aceitam uma lista de campos como entrada, as funções especiais @FIELDS_BETWEEN e @FIELDS_MATCHING podem ser usadas, conforme mostrado no exemplo a seguir:

```
count_nulls(@FIELDS_MATCHING('card*'))
```

É possível usar a função undef para preencher campos com o valor omissivo do sistema, exibido como **\$null\$**. Por exemplo, para substituir qualquer valor numérico, você poderia usar uma instrução condicional, como:

```
if not(Age > 17) or not(Age < 66) then undef else Age endif
```

Isso substitui qualquer coisa que não esteja no intervalo com um valor omissivo do sistema, exibido como **\$null\$**. Usando a função not(), é possível capturar todos os outros valores numéricos, incluindo quaisquer negativos. Consulte o tópico “Funções manipulando valores em branco e nulos” na página 150 para obter mais informações.

Nota sobre Descarte de Registros

Durante o uso de um nó de Seleção para descartar registros, observe que a sintaxe usa lógica de três valores e inclui automaticamente valores nulos em instruções de seleção. Para excluir valores nulos (omissos do sistema) em uma expressão de seleção, deve-se especificar isso explicitamente usando and not na expressão. Por exemplo, para selecionar e incluir todos os registros nos quais a droga do tipo de prescrição é Droga C, você usaria a seguinte instrução de seleção:

```
Drug = 'drugC' and not(@NULL(Drug))
```

Versões anteriores de valores nulos excluídos nessas situações.

Capítulo 8. Construindo expressões do CLEM

Sobre o CLEM

Control Language for Expression Manipulation (CLEM) é uma poderosa linguagem para analisar e manipular os dados que fluem ao longo dos fluxos do IBM SPSS Modeler. Os mineradores de dados usam CLEM extensivamente em operações de fluxo para executar tarefas tão simples quanto derivar lucros dos dados de custo e renda ou tão complexas quanto transformar dados de log da web em um conjunto de campos e registros com informações utilizáveis.

CLEM é usado dentro do IBM SPSS Modeler para:

- Comparar e avaliar condições em campos de registro.
- Derivar valores para novos campos.
- Derivar novos valores para campos existentes.
- Motivar a sequência de registros.
- Inserir dados de registros em relatórios.

Expressões do CLEM são indispensáveis para preparação de dados no IBM SPSS Modeler e podem ser usadas em uma grande variedade de nós-de operações de registro e campo (Seleção, Balanceamento, Preenchimento) a plots e saída (Análise, Relatório, Tabela). Por exemplo, é possível usar CLEM em um nó de Derivação para criar um novo campo em uma fórmula, como razão.

Expressões do CLEM também podem ser usadas para operações de procura e substituição globais. Por exemplo, a expressão @NULL(@FIELD) pode ser usada em um nó de Preenchimento para substituir **valores omissos do sistema** com o valor de número inteiro 0. (Para substituir **valores omissos do sistema**, também chamados em branco, use a função @BLANK.)

Expressões do CLEM mais complexas também podem ser criadas. Por exemplo, é possível derivar novos campos com base em um conjunto condicional de regras, como uma nova categoria de valor criada usando as expressões a seguir: If: CardID = @OFFSET(CardID,1), Then: @OFFSET(ValueCategory,1), Else: 'exclude'.

Esse exemplo usa a função @OFFSET para dizer "se o valor do campo *CardID* para um determinado registro for igual para o registro anterior, retorne o valor do campo denominado *ValueCategory* para o registro anterior. Caso contrário, designe a sequência de caracteres "exclude." Em outras palavras, se *CardIDs* para registros adjacentes forem os mesmos, eles deverão ser designados à mesma categoria de valor. (Registros com a sequência de caracteres exclude podem ser selecionados posteriormente usando um nó de Seleção.)

CLEMExemplos

Para ilustrar a sintaxe correta, bem como os tipos de expressões possíveis com o CLEM, algumas expressões de exemplo são mostradas a seguir.

Expressões Simples

Fórmulas podem ser tão simples quanto estas, que derivam um novo campo com base nos valores dos campos *After* e *Before*:

(After - Before) / Before * 100.0

Observe que os nomes de campos não são citados entre aspas quando se referem aos valores do campo.

Da mesma forma, a expressão a seguir simplesmente retorna o log de cada valor para o campo *salary*.

```
log(salary)
```

Expressões Complexas

Expressões também podem ser longas e mais complexas. As expressões a seguir retornam *true* se o valor dos dois campos (*\$KX-Kohonen* e *\$KY-Kohonen*) estiver dentro dos intervalos especificados. Observe que aqui os nomes de campo são colocados entre aspas simples, pois os nomes de campo contêm caracteres especiais.

```
(' $KX-Kohonen' >= -0.2635771036148072 and ' $KX-Kohonen' <= 0.3146203637123107  
and ' $KY-Kohonen' >= -0.18975617885589602 and  
' $KY-Kohonen' <= 0.17674794197082522) -> T
```

Várias funções, como funções de sequência de caracteres, requerem a inserção de vários parâmetros usando a sintaxe correta. No exemplo a seguir, a função *subscrs* é usada para retornar o primeiro caractere de um campo *produce_ID*, indicando se um item é orgânico, modificado geneticamente ou convencional. Os resultados de uma expressão são descritos por `-> `result``.

```
subscrs(1,produce_ID) -> `c`
```

Da mesma forma, a expressão a seguir é:

```
stripchar(`3`,`123`) -> `12`
```

É importante observar que os caracteres estão sempre contidos entre crases simples.

Combinando funções em uma expressão

Frequentemente, expressões do CLEM consistem em uma combinação de funções. A função a seguir combina *subscr* e *lowertoupper* para retornar o primeiro caractere de *produce_ID* e convertê-lo em maiúscula.

```
lowertoupper(subscr(1,produce_ID)) -> `C`
```

Essa mesma expressão pode ser escrita com taquigrafia como:

```
lowertoupper(produce_ID(1)) -> `C`
```

Outra combinação de funções bastante usada é:

```
locchar_back(`n`, (length(web_page)), web_page)
```

Essa expressão localiza o caractere ``n`` dentro dos valores do campo *web_page*, fazendo a leitura reversa a partir do último caractere do valor do campo. Incluindo a função *length* também, a expressão calcula dinamicamente o comprimento do valor atual em vez de usar um número estático, como 7, que será inválido para valores com menos de sete caracteres.

Funções especiais

Há inúmeras funções especiais (precedidas por um símbolo @) disponíveis. As funções mais usadas incluem:

```
@BLANK('referrer ID') -> T
```

Frequentemente, funções especiais são usadas em combinação, que é um método comumente usado de flag de espaços em branco em mais de um campo por vez.

```
@BLANK(@FIELD)-> T
```

Exemplos adicionais são discutidos em toda a documentação do CLEM. Consulte o tópico “Visão geral de referência do CLEM” na página 119 para obter mais informações.

Valores e tipos de dados

Expressões do CLEM são semelhantes às fórmulas construídas a partir de valores, nomes de campo, operadores e funções. A expressão do CLEM válida mais simples é um valor ou um nome de campo. Exemplos de valores válidos são:

```
3
1.79
'banana'
```

Exemplos de nomes de campo são:

```
Product_ID
'$P-NextField'
```

em que *Product* é o nome de um campo de um conjunto de dados de cesta de mercado, '*\$P-NextField*' é o nome de um parâmetro e o valor da expressão é o valor do campo nomeado. Normalmente, nomes de campo começam com uma letra e também podem conter dígitos e sublinhados (_). É possível usar nomes que não sigam essas regras se você colocar o nome entre aspas. Os valores do CLEM podem ser os seguintes:

- Sequências de caracteres--por exemplo, "c1", "Type 2", "a piece of free text"
- Números inteiros--por exemplo, 12, 0, -189
- Números reais--por exemplo, 12.34, 0.0, -0.0045
- Campos de data/hora--por exemplo, 05/12/2002, 12/05/2002, 12/05/02

Também é possível usar os elementos a seguir:

- Códigos de caractere--por exemplo, `a` ou 3
- Listas de itens--por exemplo, [1 2 3], ['Type 1' 'Type 2']

Códigos de caractere e listas geralmente não ocorrem como valores de campo. Geralmente eles são usados como argumentos de funções do CLEM.

Regras de Citação

Embora o software seja flexível durante a determinação de campos, valores, parâmetros e sequências de caracteres usados em uma expressão do CLEM, as regras gerais a seguir fornecem uma lista de "melhores práticas" para usar durante a criação de expressões:

- **Sequências de caracteres**—Sempre usam aspas duplas durante a gravação de sequências de caractere ("Type 2" ou "value"). Aspas simples podem ser usadas, mas há um risco de confusão com os campos citados.
- **Caracteres**—Sempre usam crase simples como esta ` . Por exemplo, observe o caractere d na função `stripchar(`d`, "drugA")`. A única exceção para isso é quando você está usando um número inteiro para se referir a um caractere específico em uma sequência de caracteres. Por exemplo, observe o caractere 5 na função `lowertoupper("drugA"(5)) -> "A"`. *Nota:* Em um teclado padrão do Reino Unido ou dos EUA, a tecla para o caractere de acento crase (acento grave, Unicode 0060) pode estar localizada bem abaixo da tecla Esc.
- **Campos**—Normalmente os campos não têm aspas quando usados em expressões do CLEM (`subscr(2, arrayID)`) -> CHAR). É possível usar aspas simples quando necessário para cercar espaços ou outros caracteres especiais ('Order Number'). Os campos que estiverem entre aspas, mas estiverem indefinidos no conjunto de dados, serão mal interpretados como sequências de caracteres.
- **Parâmetros**—Sempre usam aspas simples ('\$P-threshold').

Expressões e condições

As expressões do CLEM podem retornar um resultado (usado na derivação de novos valores)--por exemplo:

```
Weight * 2.2
Age + 1
sqrt(Signal-Echo)
```

Ou podem ser avaliadas como *verdadeiras* ou *falsas* (usado durante a seleção de uma condição)--por exemplo:

```
Drug = "drugA"
Age < 16
not(PowerFlux) and Power > 2000
```

É possível combinar operadores e funções de forma arbitrária em expressões do CLEM--por exemplo:

```
sqrt(abs(Signal)) * max(T1, T2) + Baseline
```

Os colchetes e a precedência do operador determinam a ordem em que a expressão é avaliada. Nesse exemplo, a ordem de avaliação é:

- `abs(Signal)` é avaliada e `sqrt` é aplicada a seu resultado.
- `max(T1, T2)` é avaliada.
- Os dois resultados são multiplicados: `x` tem precedência maior que `+`.
- Por fim, `Baseline` é incluída no resultado.

A ordem decrescente de precedência (ou seja, operações que são executadas primeiro para operações executadas por último) é a seguinte:

- Argumentos da função
- Chamadas de função
- **xx**
- **x / mod div rem**
- + -
- > < >= <= /== == = /=

Se desejar substituir a precedência, ou se você estiver em dúvida quanto à ordem de avaliação, é possível usar parênteses para deixar isso explícito--por exemplo,

```
sqrt(abs(Signal)) * (max(T1, T2) + Baseline)
```

Parâmetros Stream, Session e SuperNode

Parâmetros podem ser definidos para uso em expressões do CLEM e em scripts. Na verdade, eles são variáveis definidas pelo usuário que são salvas e persistem no fluxo, sessão ou SuperNode atuais e que podem ser acessadas a partir da interface com o usuário, bem como por meio de scripts. Se você salvar um fluxo, por exemplo, quaisquer parâmetros configurados para esse fluxo também serão salvos. (Isso os distingue das variáveis do script local, que podem ser usadas somente no script no qual elas são declaradas.) Geralmente os parâmetros são usados em scripts para controlar o comportamento do script, fornecendo informações sobre campos e valores que não precisam ser codificados permanentemente no script.

O escopo do parâmetro depende de onde ele está configurado:

- Parâmetros de fluxo podem ser configurados em um script de fluxo ou na caixa de diálogo de propriedades do fluxo e estão disponíveis para todos os nós no fluxo. Eles são exibidos na lista Parâmetros no construtor de expressões.
- Parâmetros de sessão podem ser configurados em um script independente ou na caixa de diálogo de parâmetros de sessão. Eles estão disponíveis para todos os fluxos usados na sessão atual (todos os fluxos listados na guia Fluxos na área de janela de gerenciadores).

Parâmetros também podem ser configurados para SuperNodes, em cujo caso ficam visíveis somente para os nós contidos nesse SuperNode.

Usando parâmetros em expressões do CLEM

Parâmetros são representados em expressões do CLEM por `$P-pname`, em que `pname` é o nome do parâmetro. Quando usados em expressões do CLEM, os parâmetros devem ser colocados dentro de aspas simples—por exemplo, `'$P-scale'`.

Os parâmetros disponíveis são facilmente visualizados por meio do construtor de expressões. Para visualizar parâmetros atuais:

1. Em qualquer caixa de diálogo aceitando expressões do CLEM, clique no botão **Construtor de Expressões**.
2. Na lista **Campos**, selecione **Parâmetros**.

É possível selecionar parâmetros da lista para inserção na expressão do CLEM. Consulte o tópico “Selecionando campos, parâmetros e variáveis globais” na página 114 para obter mais informações.

Trabalhando com sequências de caracteres

Há inúmeras operações disponíveis para sequências de caracteres, incluindo:

- Converter uma sequência de caracteres em maiúscula ou minúscula—`uppertolower(CHAR)`.
- Remover caracteres especificados, como ``ID_`` ou ``$`` de uma variável de sequência de caracteres—`stripchar(CHAR, STRING)`.
- Determinar o comprimento (número de caracteres) para uma variável de sequência de caracteres—`length(STRING)`.
- Verificar a ordem alfabética dos valores de sequência de caracteres—`alphabefore(STRING1, STRING2)`.
- Remover espaço em branco inicial ou final dos valores—`trim(STRING)`, `trim_start(STRING)` ou `trimend(STRING)`.
- Extrair o primeiro ou último caractere *n* de uma sequência de caracteres—`startstring(LENGTH, STRING)` ou `endstring(LENGTH, STRING)`. Por exemplo, suponha que você tenha um campo denominado *item* que combine um nome de produto com um código de ID de quatro dígitos (ACME CAMERA-D109). Para criar um novo campo que contenha somente o código de quatro dígitos, especifique a fórmula a seguir em um nó de Derivação:

```
endstring(4, item)
```
- Corresponder a um padrão específico—`STRING matches PATTERN`. Por exemplo, para selecionar pessoas que tenham a palavra "mercado" em algum lugar de seu cargo, você poderia especificar o seguinte em um nó de Seleção:

```
job_title matches "*market*"
```
- Substituir todas as instâncias de uma subsequência de caracteres dentro de uma sequência de caracteres—`replace(SUBSTRING, NEWSUBSTRING, STRING)`. Por exemplo, para substituir todas as instâncias de um caractere não suportado, como uma barra vertical (|), com um ponto e vírgula antes da mineração de texto, use a função `replace` em um nó de Preenchimento. Sob **Preencher Campos**, selecione todos os campos nos quais o caractere pode ocorrer. Para a condição **Substituir**, selecione **Sempre** e especifique a condição a seguir sob **Substituir por**:

```
replace('|',';',@FIELD)
```
- Derivar um campo de flag com base na presença de uma subsequência de caracteres específica. Por exemplo, você poderia usar uma função de sequência de caracteres em um nó de Derivação para gerar um campo de flag separado para cada resposta com uma expressão como:

```
hassubstring(museums,"museum_of_design")
```

Consulte o tópico “Funções de sequência de caracteres” na página 135 para obter mais informações.

Manipulando valores omissos e em branco

A substituição de valores omissos ou em branco é uma tarefa de preparação de dados comum para mineradores de dados. CLEM fornece inúmeras ferramentas para automatizar o tratamento de brancos. O nó de Preenchimento é o local mais comum para se trabalhar com espaços em branco; no entanto, as funções a seguir podem ser usadas em qualquer nó que aceite expressões do CLEM:

- @BLANK(FIELD) pode ser usado para determinar registros cujos valores estão em branco para um determinado campo, como *Idade*.
- @NULL(FIELD) pode ser usado para determinar registros cujos valores são omissos do sistema para o(s) campo(s) especificado(s). No IBM SPSS Modeler, valores omissos do sistema são exibidos como valores \$null\$.

Consulte o tópico “Funções manipulando valores em branco e nulos” na página 150 para obter mais informações.

Trabalhando com números

Várias operações padrão em valores numéricos estão disponíveis no IBM SPSS Modeler, como:

- Calcular o seno do ângulo especificado—`sin(NUM)`
- Calcular o logaritmo natural dos campos numéricos—`log(NUM)`
- Calcular a soma de dois números—`NUM1 + NUM2`

Consulte o tópico “Funções numéricas” na página 130 para obter mais informações.

Trabalhando com datas e horas

Os formatos de data e hora podem variar, dependendo de sua origem de dados e do seu código de idioma. Os formatos de data e hora são específicos para cada fluxo e configurados na caixa de diálogo de propriedades do fluxo. Os exemplos a seguir são funções comumente usadas para se trabalhar com campos de data/hora.

Calculando o Tempo que Passou

É possível calcular facilmente o tempo que passou a partir de uma data de linha de base usando uma família de funções semelhantes à seguinte. Essa função retorna o tempo em meses a partir da data de linha de base até a data representada pela sequência de datas DATE como um número real. Essa é uma figura aproximada com base em um mês de 30 dias.

```
date_in_months(Date)
```

Comparando Valores de Data/Hora

Valores de campos de data/hora podem ser comparados entre registros usando funções semelhantes à seguinte. Essa função retorna um valor de *true* se a sequência de datas DATE1 representar uma data anterior à representada pela sequência de datas DATE2. Caso contrário, esta função retornará um valor 0.

```
date_before(Date1, Date2)
```

Calculando Diferenças

Também é possível calcular a diferença entre dois tempos e duas datas usando funções, como:

```
date_weeks_difference(Date1, Date2)
```

Esta função retorna o tempo em semanas a partir da data representado pela sequência de datas DATE1 até a data representada pela sequência de datas DATE2 como um número real. Isso é baseado em uma semana de 7 dias. Se DATE2 for anterior à DATE1, essa função retornará um número negativo.

Data de Hoje

A data atual pode ser incluída no conjunto de dados usando a função @TODAY. A data de hora é incluída como uma sequência de caracteres no campo especificado ou no novo campo usando o formato de data selecionado na caixa de diálogo de propriedades do fluxo. Consulte o tópico “Funções de data e hora” na página 140 para obter mais informações.

Resumindo diversos campos

A linguagem do CLEM inclui inúmeras funções que retornam estatísticas básicas de diversos campos. Essas funções podem ser úteis principalmente na análise de dados de pesquisas de opinião, em que várias respostas para uma questão podem ser armazenadas em diversos campos. Consulte o tópico “Trabalhando com dados de múltiplas respostas” na página 110 para obter mais informações.

Funções de comparação

É possível comparar valores em vários campos usando as funções `min_n` e `max_n`-por exemplo:

```
max_n(['card1fee' 'card2fee' 'card3fee' 'card4fee'])
```

Também é possível usar inúmeras funções de contagem para obter contagens de valores que atendem a critérios específicos, mesmo quando esses valores são armazenados em vários campos. Por exemplo, para contar o número de cartões que ficaram retidos por mais de cinco anos:

```
count_greater_than(5, ['cardtenure' 'card2tenure' 'card3tenure'])
```

Para contar valores nulos no mesmo conjunto de dados:

```
count_nulls(['cardtenure' 'card2tenure' 'card3tenure'])
```

Observe que esse exemplo conta o número de cartões sendo retidos, e não o número de pessoas retendo-os. Consulte o tópico “Funções de comparação” na página 128 para obter mais informações.

Para contar o número de vezes que um valor especificado ocorre em vários campos, é possível usar a função `count_equal`. O exemplo a seguir conta o número de campos na lista que contém o valor Y.

```
count_equal("Y", [Answer1, Answer2, Answer3])
```

Dados os valores a seguir para os campos na lista, a função retorna os resultados para o valor Y conforme mostrado.

Tabela 7. Valores de função.

Answer1	Answer2	Answer3	Contagem
E	N	E	2
E	N	N	1

Funções numéricas

É possível obter estatísticas de vários campos usando funções `sum_n`, `mean_n` e `sdev_n`-por exemplo:

```
sum_n(['card1ba1' 'card2ba1' 'card3ba1'])  
mean_n(['card1ba1' 'card2ba1' 'card3ba1'])
```

Consulte o tópico “Funções numéricas” na página 130 para obter mais informações.

Gerando listas de campos

Durante o uso de qualquer uma das funções que aceitam uma lista de campos como entrada, as funções especiais @FIELDS_BETWEEN(start, end) e @FIELDS_MATCHING(pattern) podem ser usadas como entrada. Por exemplo, supondo que a ordem dos campos é como a mostrada no exemplo sum_n anterior, o seguinte seria equivalente:

```
sum_n(@FIELDS_BETWEEN(card1ba1, card3ba1))
```

Alternativamente, para contar o número de valores nulos em todos os campos começando com "card":

```
count_nulls(@FIELDS_MATCHING('card*'))
```

Consulte o tópico “Campos Especiais” na página 151 para obter mais informações.

Trabalhando com dados de múltiplas respostas

Inúmeras funções de comparação podem ser usadas para analisar dados de múltiplas respostas, incluindo:

- value_at
- first_index / last_index
- first_non_null / last_non_null
- first_non_null_index / last_non_null_index
- min_index / max_index

Por exemplo, imagine uma questão de múltiplas respostas feita pelo primeiro, segundo e terceiro motivos mais importantes para se decidir sobre uma determinada compra (por exemplo, preço, recomendação pessoal, revisão, fornecedor local, etc). Nesse caso, você pode determinar a importância do preço derivando o índice do campo no qual ele foi incluído primeiro:

```
first_index("price", [Reason1 Reason2 Reason3])
```

Da mesma forma, suponha que você tenha pedido aos clientes para ranquear três carros por ordem de probabilidade de compra e codificado as respostas em três campos separados, como a seguir:

Tabela 8. Exemplo de ranqueamento de carros.

ID do cliente	car1	car2	car3
101	1	3	2
102	3	2	1
103	2	3	1

Nesse caso, você poderia determinar o índice do campo para o carro que eles mais gostaram (ranqueamento #1 ou mais baixo) usando a função min_index:

```
min_index(['car1' 'car2' 'car3'])
```

Consulte o tópico “Funções de comparação” na página 128 para obter mais informações.

Referenciando conjuntos de múltiplas respostas

A função especial @MULTI_RESPONSE_SET pode ser usada para referenciar todos os campos em um conjunto de múltiplas respostas. Por exemplo, se os três campos car no exemplo anterior fossem incluídos em um conjunto de múltiplas respostas denominado car_rankings, o seguinte retornaria o mesmo resultado:

```
max_index(@MULTI_RESPONSE_SET("car_rankings"))
```

O construtor de expressões

É possível digitar expressões do CLEM manualmente ou usar o construtor de expressões, que exibe uma lista completa de funções e operadores do CLEM, bem como campos de dados do fluxo atual, permitindo construir expressões rapidamente sem memorizar os nomes exatos de campos e funções. Além disso, o construtor controla automaticamente a inclusão das aspas adequadas para campos e valores, facilitando a criação das expressões corretas sintaticamente.

Nota: O construtor de expressões não é suportado em configurações de parâmetro ou script.

Nota: Se desejar alterar sua origem de dados, antes de alterá-la, é necessário verificar se o construtor de expressões ainda pode suportar as funções que você selecionou. Como nem todos os bancos de dados suportam todas as funções, talvez você encontre um erro se executar com relação a uma nova origem de dados.

Acessando o construtor de expressões

O construtor de expressões está disponível em todos os nós nos quais expressões do CLEM são usadas, incluindo os nós de Seleção, Balanceamento, Derivação, Preenchimento, Análise, Relatório e Tabela. É possível abri-lo clicando no botão de calculadora à direita do campo de fórmula.

Criando expressões

O construtor de expressões fornece não apenas listas completas de campos, funções e operadores, mas também acesso aos valores de dados, caso seus dados sejam instanciados.

Para criar uma expressão usando o construtor de expressões

1. Digite no campo de expressão usando as listas de campo e função como referência.
ou
2. Selecione os campos obrigatórios e as funções das listas de rolagem.
3. Dê um clique duplo ou clique no botão de seta amarelo para incluir o campo ou função no campo de expressão.
4. Use os botões de operando no centro da caixa de diálogo para inserir as operações na expressão.

Selecionando funções

A lista de funções exibe todas as funções e operadores do CLEM disponíveis. Role para selecionar uma função da 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 maioria dessas categorias são descritas na seção Referência da linguagem de descrição do CLEM. Para obter informações adicionais, consulte “Referência de função” na página 124.

As outras categorias são as seguintes.

- **Funções Gerais** contém uma seleção de algumas das funções usadas mais comumente.
- **Usado Recentemente** contém uma lista de funções do CLEM usadas dentro da sessão atual.
- **Funções @** contém uma lista de todas as funções especiais, que têm nomes precedidos por um sinal “@”.
- **Funções de Banco de Dados.** Se o fluxo incluir uma conexão de banco de dados (por meio de um nó de origem de banco de dados), essa seleção listará as funções disponíveis de dentro desse banco de dados, incluindo funções definidas pelo usuário (UDFs). Para obter informações adicionais, consulte “Funções de banco de dados” na página 112.
- **Agregados do banco de dados.** Se o fluxo inclui uma conexão ao banco de dados (por meio de um nó de origem da base de dados), esta seleção lista as opções de agregação disponíveis a partir desse banco de dados. Essas opções estão disponíveis no nó Construtor de expressões de agregado.

- **Agregados da janela do banco de dados.** Se o fluxo inclui uma conexão ao banco de dados (por meio de um nó de origem da base de dados), esta seleção lista a janela de opções de agregação que podem ser usadas nesse banco de dados. Essas opções estão disponíveis no Construtor de expressões dentro dos nós na paleta **Operações de campo** apenas.

Nota: Como o SPSS Modeler obtém as **Funções Agregadas de Janela** a partir da Visualização do sistema de banco de dados, as opções disponíveis são dependentes do comportamento do banco de dados.

Embora chamadas "agregadas", essas opções não são projetadas para serem usadas no nó Agregar, elas são mais aplicáveis aos nós como Derivar ou Selecionar. Isso ocorre porque a sua produção é escalar em vez de um agregado verdadeiro, ou seja, elas não reduzem a quantidade de dados mostrados na saída da mesma maneira que o nó Agregar não o faz. Por exemplo, você poderia usar este tipo de agregação para fornecer uma média móvel para baixo através de linhas de dados, como "média da linha atual e todas as linhas anteriores".

- **Agregados integrados.** Contém uma lista dos possíveis modos de agregação que podem ser usados.
- **Operadores** lista todos os operadores que podem ser usados ao construir expressões. Os operadores também estão disponíveis nos botões no centro da caixa de diálogo.
- **Todos as funções** contém uma lista completa de funções disponíveis do CLEM .

Depois de ter selecionado um grupo de funções, dê um clique duplo para inserir as funções no campo expressão no ponto indicado pela posição do cursor.

Funções de banco de dados

Funções de banco de dados podem ser listadas em vários locais diferentes; a tabela a seguir mostra os locais onde o SPSS Modeler procura detalhes da função. Esta tabela pode ser usada pelos administradores de banco de dados para assegurar que os usuários tenham privilégios de acesso às áreas necessárias para poderem usar as diferentes funções.

Além disso, a tabela lista as condições que são usadas para filtrar quando uma função está disponível para uso com base no banco de dados e no tipo de função.

Tabela 9. Funções de banco de dados no Construtor de Expressões

Banco de dados	Tipo de função	Onde localizar funções	Condições usadas para filtrar funções
DB2 LUW	UDF	SYSCAT.ROUTINES SYSCAT.ROUTINEPARMS	ROUTINETYPE é F e FUNCTIONTYPE é S
DB2 LUW	UDA	SYSCAT.ROUTINES SYSCAT.ROUTINEPARMS	ROUTINETYPE é F e FUNCTIONTYPE é C
DB2 iSeries	UDF	QSYS2.SYSROUTINES QSYS2.SYSPARMS	ROUTINE_TYPE é F e FUNCTION_TYPE é S
DB2 iSeries	UDA	QSYS2.SYSROUTINES QSYS2.SYSPARMS	ROUTINE_TYPE é F e FUNCTION_TYPE é C
DB2 z/OS	UDF	SYSIBM.SYSROUTINES SYSIBM.SYSPARMS	ROUTINETYPE é F e FUNCTIONTYPE é S
DB2 z/OS	UDA	SYSIBM.SYSROUTINES SYSIBM.SYSPARMS	ROUTINETYPE é F e FUNCTIONTYPE é C
GreenPlum	UDF	PG_PROC PG_NAMESPACE	Todas as condições a seguir são atendidas: <ul style="list-style-type: none"> • PRORETTYPE não é igual a 2278 • PRORETTSET é FALSE • PROISAGG é TRUE ou PROISWIN é TRUE • PROALLARGTYPES é NULL

Tabela 9. Funções de banco de dados no Construtor de Expressões (continuação)

Banco de dados	Tipo de função	Onde localizar funções	Condições usadas para filtrar funções
GreenPlum	UDA	PG_PROC PG_NAMESPACE	Todas as condições a seguir são atendidas: <ul style="list-style-type: none"> • PRORETTYPE não é 16, 1000 ou 2278 • PROISAGG é TRUE • PRORETSET é FALSE • PROISWIN é FALSE • PROALLARGTYPES é NULL
GreenPlum	WUDA	PG_PROC PG_NAMESPACE	Todas as condições a seguir são atendidas: <ul style="list-style-type: none"> • PRORETTYPE não é 16, 1000 ou 2278 • PROISAGG é FALSE • PRORETSET é FALSE • PROISWIN é FALSE • PROALLARGTYPES é NULL
SQL Server	UDF	SYS.ALL_OBJECTS SYS.ALL_PARAMETERS SYS.TYPES	TYPE é FN ou FS
SQL Server	UDA	SYS.ALL_OBJECTS SYS.ALL_PARAMETERS SYS.TYPES	TYPE é AF
Oracle	UDF	ALL_ARGUMENTS ALL_PROCEDURES	Todas as condições a seguir são atendidas: <ul style="list-style-type: none"> • OBJECT_TYPE é FUNCTION • AGGREGATE é NO • PLS_TYPE não é NULL
Oracle	UDA	ALL_ARGUMENTS ALL_PROCEDURES	Todas as condições a seguir são atendidas: <ul style="list-style-type: none"> • ARGUMENT_NAME é NULL • AGGREGATE é YES • PLS_TYPE não é NULL
Sybase	UDF	SYSROCPARM SYSDOMAIN SYSPROCEDURE	As condições a seguir se aplicam: <ul style="list-style-type: none"> • PARM_TYPE é igual a 4 • PROC_DEFN tem um valor semelhante a: 'CREATE FUNCTION%' e PARM_ID é igual a 1
Sybase	UDA	SYSROCPARM SYSDOMAIN SYSPROCEDURE	As condições a seguir se aplicam: <ul style="list-style-type: none"> • PARM_TYPE é igual a 4 • PROC_DEFN tem um valor semelhante a: 'CREATE AGGREGATE FUNCTION%' e PARM_ID é igual a 1
Teradata	UDF	DBC.FUNCTIONS DBC.ALLRIGHTS	Todas as condições a seguir são atendidas: <ul style="list-style-type: none"> • FUNCTIONTYPE é F • COLUMNNAME é RETURN0 • SPPARAMETERTYPE é 0 • ACCESSRIGHT é EF
Teradata	UDA	DBC.FUNCTIONS DBC.ALLRIGHTS	Todas as condições a seguir são atendidas: <ul style="list-style-type: none"> • FUNCTIONTYPE é A • COLUMNNAME é RETURN0 • SPPARAMETERTYPE é 0 • ACCESSRIGHT é EF

Tabela 9. Funções de banco de dados no Construtor de Expressões (continuação)

Banco de dados	Tipo de função	Onde localizar funções	Condições usadas para filtrar funções
Netezza	UDF	####._V_FUNCTION NZA._V_FUNCTION INZA._V_FUNCTION	<p>Para ####._V_FUNCTION, as condições a seguir se aplicam:</p> <ul style="list-style-type: none"> RESULT não contém uma sequência de caracteres com valores como: TABLE% FUNCTION não contém uma sequência de caracteres com valores como: '/'_%' escape '/' VARARGS é FALSE <p>Para NZA._V_FUNCTION e INZA._V_FUNCTION, as condições a seguir se aplicam:</p> <ul style="list-style-type: none"> RESULT não contém uma sequência de caracteres com valores como: TABLE% FUNCTION não contém uma sequência de caracteres com valores como: '/'_%' escape '/' BUILTIN é f VARARGS é FALSE
Netezza	UDA	####._V_AGGREGATE NZA._V_FUNCTION INZA._V_FUNCTION	<p>Ambas as condições a seguir são atendidas:</p> <ul style="list-style-type: none"> AGGTYPE é ANY ou GROUPED VARARGS é FALSE
Netezza	WUDA	####._V_AGGREGATE NZA._V_FUNCTION INZA._V_FUNCTION	<p>Para ####._V_AGGREGATE, as condições a seguir se aplicam:</p> <ul style="list-style-type: none"> AGGTYPE é ANY ou ANALYTIC AGGREGATE não é MAX_LABEL VARARGS é FALSE <p>Para NZA._V_FUNCTION e INZA._V_FUNCTION, as condições a seguir se aplicam:</p> <ul style="list-style-type: none"> AGGTYPE é ANY ou ANALYTIC BUILTIN é f VARARGS é FALSE

Chave para termos usados na tabela

- UDF Função Definida pelo Usuário
- UDA Função Agregada Definida pelo Usuário
- WUDA Função Agregada da Janela Definida pelo Usuário
- ##### O banco de dados ao qual você está atualmente conectado.

Selecionando campos, parâmetros e variáveis globais

A lista de campo exibe todos os campos disponíveis neste ponto no fluxo de dados. Role para selecionar um campo da lista. Dê um clique duplo ou clique no botão de seta amarela para incluir um campo na expressão.

Consulte o tópico “Parâmetros Stream, Session e SuperNode” na página 106 para obter mais informações.

Além dos campos, também é possível escolher entre os seguintes itens:

Conjuntos de múltiplas respostas. Para obter mais informações, consulte o guia *IBM SPSS Modeler Source, Process e Output Nodes*.

Usado recentemente contém uma lista de campos, conjuntos de múltiplas respostas, parâmetros e valores globais usados dentro da sessão atual.

Parâmetros. Consulte o tópico “Parâmetros Stream, Session e SuperNode” na página 106 para obter mais informações.

Valores globais. Para obter mais informações, consulte o guia *IBM SPSS Modeler Source, Process e Output Nodes*.

Visualizando ou selecionando valores

Valores de campo podem ser visualizados a partir de inúmeros locais no sistema, incluindo o construtor de expressões, relatórios de auditoria de dados e durante a edição de valores futuros em um nó de Intervalos de Tempo. Observe que os dados devem ser completamente instanciados em um nó de origem ou Tipo para usar essa variável, de modo que o armazenamento, os tipos e os valores sejam conhecidos.

Para visualizar valores para um campo a partir do construtor de expressões ou de um nó Intervalos de Tempo, selecione o campo obrigatório e clique no botão do selecionador de valor para abrir uma caixa de diálogo listando valores para o campo selecionado. Então é possível selecionar um valor e clicar em **Inserir** para colar o valor na atual expressão ou lista.



Figura 14. Botão do selecionador de valor

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

Verificando expressões do CLEM

Clique em **Verificar** no construtor de expressões (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 do operador em vez de digitar a expressão manualmente. Esse método inclui automaticamente as aspas 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

Localizar e substituir

A caixa de diálogo Localizar/Substituir está disponível em locais onde você edita script ou texto de expressão, incluindo o editor de script, construtor de expressões do CLEM ou quando você define um modelo no nó de Relatório. Durante a edição de texto em qualquer uma dessas áreas, pressione Ctrl+F para acessar a caixa de diálogo, assegurando que o foco do cursor esteja na área de texto. Se estiver trabalhando em um nó de Preenchimento, por exemplo, é possível acessar a caixa de diálogo a partir de qualquer uma das áreas de texto na guia Configurações ou do campo de texto no construtor de expressões.

1. Com o cursor em uma área de texto, pressione Ctrl+F para acessar a caixa de diálogo Localizar/Substituir.
2. Insira o texto que deseja procurar ou escolha na lista suspensa de itens procurados recentemente.
3. Insira o texto de substituição, se houver.
4. Clique em **Localizar Próximo** para iniciar a procura.
5. Clique em **Substituir** para substituir a seleção atual ou em **Substituir Todos** para atualizar todas as instâncias ou as selecionadas.
6. A caixa de diálogo fecha após cada operação. Pressione F3 em qualquer área de texto para repetir a última operação Localizar ou pressione Ctrl+F para acessar a caixa de diálogo novamente.

Opções de procura

Respeitar maiúsculas e minúsculas. Especifica se a operação Localizar faz distinção entre maiúsculas e minúsculas; por exemplo, se *myvar* corresponde a *myVar*. O texto de substituição é sempre inserido exatamente como digitado, independentemente dessa configuração.

Somente palavras inteiras. Especifica se a operação Localizar corresponde ao texto integrado às palavras. Se for selecionada, por exemplo, uma procura por *spider* não corresponderá a *spiderman* ou *spider-man*.

Expressões regulares. Especifica se a sintaxe de expressão regular é usada (consulte a próxima seção). Quando selecionada, a opção **Somente palavras inteiras** é desativada e seu valor é ignorado.

Somente texto selecionado. Controla o escopo da procura durante o uso da opção **Substituir Todos**.

Sintaxe de expressão regular

Expressões regulares permitem procurar caracteres especiais, como guias ou caracteres de nova linha, classes ou intervalos de caracteres, como de *a* a *d*, qualquer dígito ou não dígito e limites, como início ou fim de uma linha. Os tipos de expressões a seguir são suportados.

Tabela 10. Correspondências de caractere.

Caracteres	Correspondências
x	O caractere x
\\	O caractere de barra invertida
\\0n	O caractere com valor octal 0n (0 <= n <= 7)
\\0nn	O caractere com valor octal 0nn (0 <= n <= 7)

Tabela 10. Correspondências de caractere (continuação).

Caracteres	Correspondências
\0mnn	O caractere com valor octal 0mnn (0 <= m <= 3, 0 <= n <= 7)
\xhh	O caractere com valor hexadecimal 0xhh
\uhhhh	O caractere com valor hexadecimal 0xhhhh
\t	O caractere de tabulação ('\u0009')
\n	O caractere de nova linha (feed de linha) ('\u000A')
\r	O caractere de retorno de linha ('\u000D')
\f	O caractere de feed de formulário ('\u000C')
\a	O caractere de alerta (sino) ('\u0007')
\e	O caractere de escape ('\u001B')
\cx	O caractere de controle correspondente a x

Tabela 11. Classes de caracteres correspondentes.

Classes de caractere	Correspondências
[abc]	a, b ou c (classe simples)
[^abc]	Qualquer caractere, exceto a, b ou c (subtração)
[a-zA-Z]	a a z ou A a Z, inclusivo (intervalo)
[a-d[m-p]]	a a d ou m a p (união). Alternativamente, poderia ser especificado como [a-dm-p]
[a-z&&[def]]	a a z e d, e ou f (intersecção)
[a-z&&[^bc]]	a a z, exceto para b e c (subtração). Alternativamente, poderia ser especificado como [ad-z]
[a-z&&[^m-p]]	a a z, e não m a p (subtração). Alternativamente, poderia ser especificado como [a-lq-z]

Tabela 12. Classes de caractere predefinidas.

Classes de caractere predefinidas	Correspondências
.	Qualquer caractere (pode ou não corresponder aos terminadores de linha)
\d	Qualquer dígito: [0-9]
\D	Um não dígito: [^0-9]
\s	Um caractere de espaço em branco: [\t\n\x0B\f\r]
\S	Um caractere de espaço não em branco: [^\s]
\w	Um caractere de palavra: [a-zA-Z_0-9]
\W	Um caractere de não palavra: [^\w]

Tabela 13. Correspondências de limite.

Correspondentes de limite	Correspondências
^	O início de uma linha
\$	O fim de uma linha
\b	Um limite de palavra
\B	Um limite de não palavra
\A	O início da entrada

Tabela 13. Correspondências de limite (continuação).

Correspondentes de limite	Correspondências
\Z	O fim da entrada, mas para o terminador final, se houver
\z	O fim da entrada

Capítulo 9. Referência de linguagem do CLEM

Visão geral de referência do CLEM

Esta seção descreve a Linguagem de Controle para Manipulação de Expressão (CLEM), que é uma poderosa ferramenta usada para analisar e manipular dados usados em fluxos do IBM SPSS Modeler. É possível usar CLEM dentro dos nós para executar tarefas que vão da avaliação de condições ou derivação de valores à inserção de dados em relatórios.

As expressões do CLEM consistem em valores, nomes de campo, operadores e funções. Usando a sintaxe correta, é possível criar uma grande variedade de operações de dados poderosas.

Tipos de dados do CLEM

Os tipos de dados do CLEM podem ser compostos do seguinte:

- Números Inteiros
- Reais
- Caracteres
- Sequência de caracteres
- Listas
- Campos
- Data/hora

Regras para Citação

Embora o IBM SPSS Modeler seja flexível quando você está determinando os campos, os valores, os parâmetros e as sequências de caracteres usados em uma expressão do CLEM, as regras gerais a seguir fornecem uma lista de "boas práticas" para serem usadas na criação de expressões:

- Sequências de caracteres—Sempre usam aspas duplas durante a gravação de sequências de caracteres, como "Tipo 2". Aspas simples podem ser usadas, mas há um risco de confusão com os campos citados.
- Campos—Usam aspas simples somente quando for necessário incluir espaços ou outros caracteres especiais, como 'Número de Ordem'. Os campos que estiverem entre aspas, mas estiverem indefinidos no conjunto de dados, serão mal interpretados como sequências de caracteres.
- Parâmetros—Sempre usam aspas simples durante o uso de parâmetros, como '\$P-threshold'.
- Caracteres—Sempre usam crases simples (`), como `stripchar(`d`, "drugA")`.

Essas regras são cobertas com mais detalhes nos tópicos a seguir.

Números Inteiros

Números inteiros são representados como uma sequência de dígitos decimais. Opcionalmente, é possível colocar um sinal de menos (-) antes do número inteiro para denotar um número negativo—por exemplo, 1234, 999, -77.

A linguagem do CLEM manipula números inteiros de precisão arbitrária. O tamanho máximo do número inteiro depende de sua plataforma. Se os valores forem muito grandes para serem exibidos em um campo de número inteiro, a mudança do tipo de campo para Real geralmente restaura o valor.

Reais

Real refere-se a um número de vírgula flutuante. Reais são representados por um ou mais dígitos, seguidos por um separador decimal, seguido por um ou mais dígitos. Os reais do CLEM são mantidos em precisão dupla.

Opcionalmente, é possível colocar um sinal de menos (-) antes do real para denotar um número negativo—por exemplo, 1.234, 0.999, -77.001. Use a forma <número> e <expoente> para expressar um número real em notação exponencial—por exemplo, 1234.0e5, 1.7e-2. Quando o aplicativo IBM SPSS Modeler lê sequências de caracteres numéricos dos arquivos e as converte automaticamente em números, os números sem dígito inicial antes do ponto decimal ou sem dígito após o ponto são aceitos—por exemplo, 999. ou .11. No entanto, essas formas são ilegais em expressões do CLEM.

Nota: Quando você fizer referência a números reais em expressões do CLEM, um período deverá ser usado como um separador decimal, independentemente de quaisquer configurações para o fluxo ou código de idioma atual. Por exemplo, especifique

```
Na > 0.6
```

em vez de

```
Na > 0,6
```

Isso se aplica mesmo se uma vírgula for selecionada como o símbolo decimal na caixa de diálogo de propriedades do fluxo e é consistente com a diretriz geral de que a sintaxe de código deve ser independente de qualquer convenção ou código de idioma específico.

Caracteres

Caracteres (geralmente mostrados como CHAR) normalmente são usados dentro de uma expressão do CLEM para executar testes em sequências de caracteres. Por exemplo, é possível usar a função `isuppercode` para determinar se o primeiro caractere de uma sequência de caracteres é maiúsculo. A expressão do CLEM a seguir usa um caractere para indicar se o teste deve ser executado no primeiro caractere da sequência de caracteres:

```
isuppercode(subscrs(1, "MyString"))
```

Para expressar o código (em contraste com a localização) de um determinado caractere em uma expressão do CLEM, use crase simples no formato ``<caractere>``—por exemplo, ``A``, ``Z``.

Nota: Não há um tipo de armazenamento CHAR para um campo, portanto, um campo é derivado ou preenchido com uma expressão que resulta em um CHAR, então esse resultado será convertido em uma sequência de caracteres.

Sequência de caracteres

Geralmente é necessário colocar sequências de caracteres entre aspas duplas. Exemplos de sequências de caracteres são "c35product2" e "referrerID". Para indicar caracteres especiais em uma sequência de caracteres, use uma barra invertida—por exemplo, "\\$65443". (Para indicar um caractere de barra invertida, use uma barra invertida dupla, \\.) É possível usar aspas simples em uma sequência de caracteres, mas o resultado é indistinguível a partir de um campo entre aspas ('referrerID'). Consulte o tópico "Funções de sequência de caracteres" na página 135 para obter mais informações.

Listas

Uma lista é uma sequência ordenada de elementos, os quais podem ser do tipo combinado. As listas são colocadas entre colchetes ([]). Exemplos de listas são [1 2 4 16] e ["abc" "def"]. Listas não são usadas como o valor dos campos do IBM SPSS Modeler. Elas são usadas para fornecer argumentos para funções, como `member` e `oneof`.

Nota: Listas podem ser compostas somente por objetos estáticos (por exemplo, uma sequência de caracteres, um número ou um nome de campo), e não chamadas para funções.

Campos

Nomes em expressões do CLEM que não são nomes de funções são considerados nomes de campo. É possível gravá-los simplesmente como `Power`, `val27`, `state_flag`, e assim por diante, mas se o nome começar com um dígito ou incluir caracteres não alfabéticos, como espaços (com a exceção do sublinhado), coloque o nome entre aspas simples--por exemplo, `'Power Increase'`, `'2nd answer'`, `'#101'`, `'$P-NextField'`.

Nota: Os campos que estiverem entre aspas, mas estiverem indefinidos no conjunto de dados, serão mal interpretados como sequências de caracteres.

Datas

Cálculos de datas são baseados em uma data de "linha de base", que é especificada na caixa de diálogo de propriedades do fluxo. A data de linha de base padrão é 1 de janeiro de 1900.

A linguagem do CLEM suporta os formatos de data a seguir.

Tabela 14. Formatos de data da linguagem do CLEM

Formato	Exemplos
DDMMAA	150163
MMDDAA	011563
AAMMDD	630115
AAAAMMDD	19630115
AAAADDD	Ano de quatro dígitos seguido por um número de três dígitos representando o dia do ano - por exemplo, 2000032 representa o 32º dia de 2000 ou 1 de fevereiro de 2000.
DAY	Dia da semana no código de idioma atual — por exemplo, Segunda, Terça, ..., em inglês.
MONTH	Mês no código de idioma atual - por exemplo, Janeiro, Fevereiro,
DD/MM/AA	15/01/63
DD/MM/AAAA	15/01/1963
MM/DD/AA	01/15/63
MM/DD/AAAA	01/15/1963
DD-MM-AA	15-01-63
DD-MM-AAAA	15-01-1963
MM-DD-AA	01-15-63
MM-DD-AAAA	01-15-1963
DD.MM.AA	15.01.63
DD.MM.AAAA	15.01.1963
MM.DD.AA	01.15.63
MM.DD.AAAA	01.15.1963
DD-MON-AA	15-JAN-63, 15-jan-63, 15-Jan-63
DD/MES/AA	15/JAN/63, 15/jan/63, 15/Jan/63
DD.MES.AA	15.JAN.63, 15.jan.63, 15.Jan.63

Tabela 14. Formatos de data da linguagem do CLEM (continuação)

Formato	Exemplos
DD-MES-AAAA	15-JAN-1963, 15-jan-1963, 15-Jan-1963
DD/MES/AAAA	15/JAN/1963, 15/jan/1963, 15/Jan/1963
DD.MES.AAAA	15.JAN.1963, 15.jan.1963, 15.Jan.1963
MES AAAA	Jan 2004
t T AAAA	Data representada como um dígito (1-4) representando o trimestre seguido pela letra Q e um ano de quatro dígitos - por exemplo, 25 de dezembro de 2004 seria representado como 4 Q 2004.
ss SM AAAA	Número de dois dígitos representando a semana do ano seguida pelas letras WK e um ano de quatro dígitos. A semana do ano é calculada supondo que o primeiro dia da semana seja segunda e haja pelo menos um dia na primeira semana.

Hora

A linguagem do CLEM suporta os seguintes formatos de hora.

Tabela 15. Formatos de hora da linguagem do CLEM

Formato	Exemplos
HHMMSS	120112, 010101, 221212
HHMM	1223, 0745, 2207
MMSS	5558, 0100
HH:MM:SS	12:01:12, 01:01:01, 22:12:12
HH:MM	12:23, 07:45, 22:07
MM:SS	55:58, 01:00
(H)H:(M)M:(S)S	12:1:12, 1:1:1, 22:12:12
(H)H:(M)M	12:23, 7:45, 22:7
(M)M:(S)S	55:58, 1:0
HH.MM.SS	12.01.12, 01.01.01, 22.12.12
HH.MM	12.23, 07.45, 22.07
MM.SS	55.58, 01.00
(H)H.(M)M.(S)S	12.1.12, 1.1.1, 22.12.12
(H)H.(M)M	12.23, 7.45, 22.7
(M)M.(S)S	55.58, 1.0

Operadores do CLEM

Os operadores a seguir estão disponíveis.

Tabela 16. Operadores de linguagem CLEM.

Operação	Comentários	Precedência (consulte a próxima seção)
ou	Usado entre duas expressões do CLEM. Retorna um valor de verdadeiro se uma delas for verdadeira ou se ambas forem verdadeiras.	10

Tabela 16. Operadores de linguagem CLEM (continuação).

Operação	Comentários	Precedência (consulte a próxima seção)
e	Usado entre duas expressões do CLEM. Retorna um valor de verdadeiro se ambas forem verdadeiras.	9
=	Usado entre dois itens comparáveis quaisquer. Retorna verdadeiro se ITEM1 for igual a ITEM2.	7
==	Idêntico a =.	7
/=	Usado entre dois itens comparáveis quaisquer. Retorna verdadeiro se ITEM1 <i>não</i> for igual a ITEM2.	7
/==	Idêntico a /=.	7
>	Usado entre dois itens comparáveis quaisquer. Retorna verdadeiro se ITEM1 for exatamente maior que ITEM2.	6
>=	Usado entre dois itens comparáveis quaisquer. Retorna verdadeiro se ITEM1 for maior que ou igual a ITEM2.	6
<	Usado entre dois itens comparáveis quaisquer. Retorna verdadeiro se ITEM1 for exatamente menor que ITEM2	6
<=	Usado entre dois itens comparáveis quaisquer. Retorna verdadeiro se ITEM1 for menor que ou igual a ITEM2.	6
&&=_0	Usado entre dois números inteiros. Equivalente à expressão booleana INT1 && INT2 = 0.	6
&&/=_0	Usado entre dois números inteiros. Equivalente à expressão booleana INT1 && INT2 /= 0.	6
+	Inclui dois números: NUM1 + NUM2.	5
><	Concatena duas sequências de caracteres; por exemplo, STRING1 >< STRING2.	5
-	Subtrai um número de outro: NUM1 - NUM2. Também pode ser usado na frente de um número: - NUM.	5
*	Usado para multiplicar dois números: NUM1 * NUM2.	4
&&	Usado entre dois números inteiros. O resultado é o bit a bit 'and' dos números inteiros INT1 e INT2.	4
&&~	Usado entre dois números inteiros. O resultado é o bit a bit 'and' de INT1 e o complemento bit a bit de INT2.	4
	Usado entre dois números inteiros. O resultado é o bit a bit 'inclusive or' de INT1 e INT2.	4
~	Usado na frente de um número inteiro. Produz o complemento bit a bit do INT.	4
/&	Usado entre dois números inteiros. O resultado é o bit a bit 'exclusive or' de INT1 e INT2.	4

Tabela 16. Operadores de linguagem CLEM (continuação).

Operação	Comentários	Precedência (consulte a próxima seção)
INT1 << N	Usado entre dois números inteiros. Produz o padrão de bits de INT deslocado para a esquerda N posições.	4
INT1 >> N	Usado entre dois números inteiros. Produz o padrão de bits de INT deslocado para a direita N posições.	4
/	Usado para dividir um número por outro: NUM1 / NUM2.	4
**	Usado entre dois números: BASE ** POWER. Retorna BASE elevada à potência POWER.	3
rem	Usado entre dois números inteiros: INT1 rem INT2. Retorna o resto, INT1 - (INT1 div INT2) * INT2.	2
div	Usado entre dois números inteiros: INT1 div INT2. Realiza divisão de número inteiro.	2

Precedência do Operador

A precedência determina a análise sintática de expressões complexas, principalmente expressões sem colchetes com mais de um operador infix. Por exemplo,

$3 + 4 * 5$

analisa como $3 + (4 * 5)$ em vez de $(3 + 4) * 5$, pois a precedência relativa determina que * deve ser analisado antes de +. Cada operador na linguagem do CLEM tem um valor de precedência associado a ele; quanto menor esse valor, mais importante ele será na lista de análise sintática, o que significa que ele será processado mais rápido que os outros operadores e com valores de precedência mais altos.

Referência de função

As funções do CLEM a seguir estão disponíveis para o trabalho com dados no IBM SPSS Modeler. É possível inserir essas funções como código em diversas caixas de diálogo, como os nós de Derivação e Configuração para Flag, ou usar o construtor de expressões para criar expressões válidas do CLEM sem memorizar listas de funções ou nomes de campos.

Tabela 17. Funções do CLEM para usar com dados do IBM SPSS Modeler

Tipo de Função	Descrição
Informações	Usado para ganho de insight nos valores de campos. Por exemplo, a função <code>is_string</code> retorne true para todos os registros cujo tipo é uma sequência de caracteres.
Conversão	Usado para construir novos campos ou converter o tipo de armazenamento. Por exemplo, a função <code>to_timestamp</code> converte o campo selecionado em um registro de data e hora.
Comparação	Usado para comparar valores de campo uns com os outros ou com uma sequência de caracteres especificada. Por exemplo, <code><=</code> é usado para comparar se os valores dos dois campos são menores ou iguais.
Lógico	Usado para executar operações lógicas, como <code>if</code> , <code>then</code> , <code>else</code> .
Numérico	Usado para executar cálculos numéricos, como o logaritmo natural dos valores de campo.

Tabela 17. Funções do CLEM para usar com dados do IBM SPSS Modeler (continuação)

Tipo de Função	Descrição
Trigonométrico	Usado para executar cálculos trigonométricos, como o arco cosseno de um ângulo especificado.
Probabilidade	Retorna probabilidades baseadas em várias distribuições, como a probabilidade de um valor da distribuição <i>t</i> do Estudante ser menor que um valor específico.
Espacial	Usado para executar cálculos espaciais em dados geoespaciais.
Bit a bit	Usado para manipular números inteiros como padrões de bits.
Aleatório	Usado para selecionar itens ou gerar números aleatoriamente.
Sequência de caracteres	Usado para executar várias operações em sequências de caracteres, como <code>stripchar</code> , o que permite remover um caractere especificado.
SoundEx	Usado para localizar sequências de caracteres quando a ortografia exata não é conhecida, com base em suposições fonéticas sobre como determinadas letras são pronunciadas.
data e hora	Usado para executar várias operações em campos de data, hora e registro de data e hora.
Sequência	Usado para ganho de insight sobre a sequência de registro de um conjunto de dados ou execução de operações que são baseadas nessa sequência.
Global	Usado para acessar valores globais criados por um nó Configurar Globais. Por exemplo, <code>@MEAN</code> é usado para se referir à média de todos os valores para um campo em todo o conjunto de dados.
Branco e nulos	Usado para acessar, flag e, frequentemente, preencher os espaços em branco especificados pelo usuário ou valores omissos do sistema. Por exemplo, <code>@BLANK(FIELD)</code> é usado para levantar um flag true para registros nos quais há valores em branco presentes.
Campos especiais	Usado para denotar os campos específicos sendo examinados. Por exemplo, <code>@FIELD</code> é usado na derivação de campos múltiplos.

Convenções em descrições de função

As convenções a seguir são usadas em todo este guia ao referir-se aos itens em uma função.

Tabela 18. Convenções em descrições de função.

Convenção	Descrição
<i>BOOL</i>	Um booleano ou flag como true ou false.
<i>NUM, NUM1, NUM2</i>	Qualquer número.
<i>REAL, REAL1, REAL2</i>	Qualquer número real, como 1.234 ou -77.01.
<i>INT, INT1, INT2</i>	Qualquer número inteiro, como 1 ou -77.
<i>CHAR</i>	Um código de caractere, como `A`.
<i>STRING</i>	Uma sequência de caracteres, como "referrerID".
<i>LIST</i>	Uma lista de itens, como ["abc" "def"].
<i>ITEM</i>	Um campo, como Customer ou extract_concept.
<i>DATE</i>	Um campo de data, como start_date, em que os valores estão em um formato como DD-MON-YYYY.
<i>TIME</i>	Um campo de tempo, como power_flux, no qual os valores estão em um formato como HHMMSS.

As funções neste guia são listadas com a função em uma coluna, o tipo de resultado (número inteiro, sequência de caracteres, etc) em outra e uma descrição (quando disponível) em uma terceira coluna. Por exemplo, a seguir está a descrição da função rem.

Tabela 19. Descrição da função rem.

Função	Resultado	Descrição
INT1 rem INT2	Número	Retorna o restante de INT1 dividido por INT2. Por exemplo, INT1 – (INT1 div INT2) * INT2.

Os detalhes sobre convenções de uso, por exemplo, como listar itens ou especificar caracteres em uma função, são descritos em outro lugar. Consulte o tópico “Tipos de dados do CLEM” na página 119 para obter mais informações.

Funções de informações

Funções de informações são usadas para obter um insight dos valores de um determinado campo. Normalmente elas são usadas para derivar campos de flag. Por exemplo, é possível usar a função @BLANK para criar um campo de flag indicando registros cujos valores estão em branco para o campo selecionado. Da mesma forma, é possível verificar o tipo de armazenamento para um campo usando qualquer uma das funções de tipo de armazenamento, como is_string.

Tabela 20. Funções de informações do CLEM.

Função	Resultado	Descrição
@BLANK(FIELD)	Booleano	Retorna verdadeiro para todos os registros cujos valores estão em branco de acordo com o conjunto de regras de manipulação de brancos em um nó Tipo de envio de dados ou nó de origem (guia Tipos).
@NULL(ITEM)	Booleano	Retorna true para todos os registros cujos valores estão indefinidos. Valores indefinidos são valores nulos do sistema, exibidos no IBM SPSS Modeler como \$null\$.
is_date(ITEM)	Booleano	Retorna true para todos os registros cujo tipo é uma data.
is_datetime(ITEM)	Booleano	Retorna true para todos os registros cujo tipo é uma data, hora ou registro de data e hora.
is_integer(ITEM)	Booleano	Retorna true para todos os registros cujo tipo é um número inteiro.
is_number(ITEM)	Booleano	Retorna true para todos os registros cujo tipo é um número.
is_real(ITEM)	Booleano	Retorna true para todos os registros cujo tipo é um real.
is_string(ITEM)	Booleano	Retorna true para todos os registros cujo tipo é uma sequência de caracteres.
is_time(ITEM)	Booleano	Retorna true para todos os registros cujo tipo é uma hora.
is_timestamp(ITEM)	Booleano	Retorna true para todos os registros cujo tipo é um registro de data e hora.

Funções de conversão

Funções de conversão permitem construir novos campos e converter o tipo de armazenamento dos arquivos existentes. Por exemplo, é possível formar novas sequências de caracteres associando ou separando sequências de caracteres. Para associar duas sequências de caracteres, use o operador ><. Por exemplo, se o campo Site tiver o valor "BRAMLEY", então "xx" >< Site retornará "xxBRAMLEY". O resultado de >< é sempre uma sequência de caracteres, mesmo se os argumentos não forem sequências de caracteres. Assim, se o campo V1 for 3 e o campo V2 for 5, então V1 >< V2 retornará "35" (uma sequência de caracteres, não um número).

Funções de conversão (e quaisquer outras funções que requerem um tipo específico de entrada, como valor de data e hora) dependem dos atuais formatos especificados na caixa de diálogo Opções de Fluxo. Por exemplo, se desejar converter um campo de sequência de caracteres com valores *Jan 2003, Fev 2003*, e assim por diante, selecione o formato de data correspondente **MÊS AAAA** como o formato de data padrão para o fluxo.

Tabela 21. Funções de conversão do CLEM

Função	Resultado	Descrição
ITEM1 >< ITEM2	<i>Sequência de caracteres</i>	Concatena valores para dois campos e retorna a sequência de caracteres resultante como <i>ITEM1ITEM2</i> .
to_integer(ITEM)	<i>Número inteiro</i>	Converte o armazenamento do campo especificado em um número inteiro.
to_real(ITEM)	<i>Real</i>	Converte o armazenamento do campo especificado em um real.
to_number(ITEM)	<i>Número</i>	Converte o armazenamento do campo especificado em um número.
to_string(ITEM)	<i>Sequência de caracteres</i>	Converte o armazenamento do campo especificado em uma sequência de caracteres. Quando um real é convertido em uma sequência de caracteres usando essa função, ele retorna um valor com 6 dígitos após o ponto radix.
to_time(ITEM)	<i>Hora</i>	Converte o armazenamento do campo especificado em hora.
to_date(ITEM)	<i>Date</i>	Converte o armazenamento do campo especificado em uma data.
to_timestamp(ITEM)	<i>Registro de data e hora</i>	Converte o armazenamento do campo especificado em um registro de data e hora.
to_datetime(ITEM)	<i>Data/hora</i>	Converte o armazenamento do campo especificado em um valor de data, hora ou registro de data e hora.
datetime_date(ITEM)	<i>Date</i>	Retorna o valor de data para um <i>número</i> , uma <i>sequência de caracteres</i> ou um <i>registro de data e hora</i> . Observe que essa é a única função que permite converter um número (em segundos) de volta em uma data. Se ITEM for uma sequência de caracteres, cria uma data analisando uma sequência de caracteres no formato de data atual. O formato de data especificado na caixa de diálogo de propriedades de fluxo deve estar correto para que essa função seja bem-sucedida. Se ITEM for um número, ele será interpretado como um número de segundos desde a data (ou época) de base. Frações de um dia são truncadas. Se ITEM for um registro de data e hora, a parte de data do registro de data e hora será retornada. Se ITEM for uma data, ela será retornada inalterada.
stb_centroid_latitude(ITEM)	<i>Número inteiro</i>	Retorna um valor de número inteiro para latitude correspondente ao centroide do argumento geohash.
stb_centroid_longitude(ITEM)	<i>Número inteiro</i>	Retorna um valor de número inteiro para longitude correspondente ao centroide do argumento geohash.

Tabela 21. Funções de conversão do CLEM (continuação)

Função	Resultado	Descrição
to_geohash(ITEM)	Sequência de caracteres	<p>Retorna a sequência de caracteres geohashed correspondente à latitude e longitude usando o número de bits especificado para a densidade.</p> <p>Um geohash é um código usado para identificar um conjunto de coordenadas geográficas com base nos detalhes de latitude e longitude. Os parâmetros de árvore para to_geohash são:</p> <ul style="list-style-type: none"> • <i>latitude</i>: Intervalo (-180, 180) e unidades são graus no sistema de coordenadas WGS84 • <i>longitude</i>: Intervalo (-90, 90) e unidades são graus no sistema de coordenadas WGS84 • <i>bits</i>: O número de bits a serem usados para armazenar o hash. Intervalo [1,75]. Isso afeta tanto o comprimento da sequência retornada (1 caractere é usado para cada 5 bits) e a precisão do hash. Por exemplo, 5 bits (1 caractere) representa aproximadamente 2500 quilômetros ou 45 bits (9 caracteres), representa aproximadamente 2.3 metros.

Funções de comparação

Funções de comparação são usadas para comparar valores de campo uns com os outros ou com uma sequência de caracteres especificada. Por exemplo, é possível verificar a igualdade das sequências de caracteres usando =. Um exemplo de verificação da igualdade da sequência de caracteres é: `Class = "class 1"`.

Para propósitos de comparação numérica, *maior* significa mais próximo do infinito positivo e *menor* significa mais próximo do infinito negativo. Ou seja, todos os números negativos são menores que qualquer número positivo.

Tabela 22. Funções de comparação do CLEM

Função	Resultado	Descrição
count_equal(ITEM1, LIST)	Número inteiro	Retorna o número de valores de uma lista de campos que são iguais a ITEM1 ou nulos se ITEM1 for nulo.
count_greater_than(ITEM1, LIST)	Número inteiro	Retorna o número de valores de uma lista de campos que são maiores que ITEM1 ou nulos se ITEM1 for nulo.
count_less_than(ITEM1, LIST)	Número inteiro	Retorna o número de valores de uma lista de campos que são menores que ITEM1 ou nulos se ITEM1 for nulo.
count_not_equal(ITEM1, LIST)	Número inteiro	Retorna o número de valores de uma lista de campos que não são iguais a ITEM1 ou nulos se ITEM1 for nulo.
count_nulls(LIST)	Número inteiro	Retorna o número de valores nulos de uma lista de campos.
count_non_nulls(LIST)	Número inteiro	Retorna o número de valores não nulos de uma lista de campos.
date_before(DATE1, DATE2)	Boolean	Usado para verificar a ordem dos valores de data. Retorna um valor real se DATE1 for anterior a DATE2.
first_index(ITEM, LIST)	Número inteiro	Retorna um índice do primeiro campo contendo um ITEM de LIST de campos ou 0 se o valor não for localizado. Suportado somente para os tipos sequência de caracteres, número inteiro e real.
first_non_null(LIST)	Qualquer	Retorna o primeiro valor não nulo na lista de campos fornecida. Todos os tipos de armazenamento suportados.

Tabela 22. Funções de comparação do CLEM (continuação)

Função	Resultado	Descrição
first_non_null_index(LIST)	Número inteiro	Retorna o índice do primeiro campo no LIST especificado contendo um valor não nulo ou 0 se todos os valores forem nulos. Todos os tipos de armazenamento são suportados.
ITEM1 = ITEM2	Boolean	Retorna true para registros nos quais ITEM1 é igual a ITEM2.
ITEM1 /= ITEM2	Boolean	Retorna true se duas sequências de caracteres não forem idênticas ou 0 se elas forem idênticas.
ITEM1 < ITEM2	Boolean	Retorna true para registros nos quais ITEM1 é menor que ITEM2.
ITEM1 <= ITEM2	Boolean	Retorna true para registros nos quais ITEM1 é menor que ou igual a ITEM2.
ITEM1 > ITEM2	Boolean	Retorna true para registros nos quais ITEM1 é maior que ITEM2.
ITEM1 >= ITEM2	Boolean	Retorna true para registros nos quais ITEM1 é maior que ou igual a ITEM2.
last_index(ITEM, LIST)	Número inteiro	Retorna um índice do último campo contendo um ITEM de LIST de campos ou 0 se o valor não for localizado. Suportado somente para os tipos sequência de caracteres, número inteiro e real.
last_non_null(LIST)	Qualquer	Retorna o último valor não nulo na lista de campos fornecida. Todos os tipos de armazenamento suportados.
last_non_null_index(LIST)	Número inteiro	Retorna o índice do último campo no LIST especificado contendo um valor não nulo ou 0 se todos os valores forem nulos. Todos os tipos de armazenamento são suportados.
max(ITEM1, ITEM2)	Qualquer	Retorna o maior dos dois itens--ITEM1 ou ITEM2.
max_index(LIST)	Número inteiro	Retorna o índice do campo contendo o valor máximo de uma lista de campos numéricos ou 0 se todos os valores forem nulos. Por exemplo, se o terceiro campo listado contiver o máximo, o valor de índice 3 será retornado. Se vários campos contiverem o valor máximo, o listado primeiro (primeiro da esquerda) será retornado.
max_n(LIST)	Número	Retorna o valor máximo de uma lista de campos numéricos ou nulo se todos os valores de campo forem nulos.
member(ITEM, LIST)	Boolean	Retorna true se ITEM for membro de um LIST especificado. Caso contrário, retorna false. Uma lista de nomes de campo também pode ser especificada.
min(ITEM1, ITEM2)	Qualquer	Retorna o menor dos dois itens--ITEM1 ou ITEM2.
min_index(LIST)	Número inteiro	Retorna o índice do campo contendo o valor mínimo de uma lista de campos numéricos ou 0 se todos os valores forem nulos. Por exemplo, se o terceiro campo listado contiver o mínimo, o valor de índice 3 será retornado. Se vários campos contiverem o valor mínimo, o listado primeiro (primeiro da esquerda) será retornado.
min_n(LIST)	Número	Retorna o valor mínimo de uma lista de campos numéricos ou nulo se todos os valores de campo forem nulos.
time_before(TIME1, TIME2)	Boolean	Usado para verificar a ordem dos valores de hora. Retorna um valor real se TIME1 for anterior a TIME2.
value_at(INT, LIST)		Retorna o valor de cada campo listado no offset INT ou NULL, caso o offset esteja fora do intervalo de valores válidos (ou seja, menor que 1 ou maior que o número de campos listados). Todos os tipos de armazenamento suportados.

Funções lógicas

Expressões do CLEM podem ser usadas para executar operações lógicas.

Tabela 23. Funções lógicas do CLEM

Função	Resultado	Descrição
COND1 and COND2	Boolean	Esta operação é uma conjunção lógica e retorna um valor verdadeiro se <i>COND1</i> e <i>COND2</i> forem verdadeiras. Se <i>COND1</i> for falsa, <i>COND2</i> não será avaliada; isso possibilita ter conjunções nas quais <i>COND1</i> testa primeiro se a operação em <i>COND2</i> é legal. Por exemplo, <code>length(Label) >=6</code> e <code>Label(6) = 'x'</code> .
COND1 or COND2	Boolean	Esta operação é uma disjunção lógica (inclusiva) e retorna um valor de verdadeiro se <i>COND1</i> ou <i>COND2</i> for verdadeira ou se ambas forem verdadeiras. Se <i>COND1</i> for verdadeira, <i>COND2</i> não será avaliada.
not(COND)	Boolean	Esta operação é uma negação lógica e retorna um valor de verdadeiro se <i>COND</i> for falsa. Caso contrário, essa operação retornará um valor de 0.
if COND then EXPR1 else EXPR2 endif	Qualquer	Esta operação é uma avaliação condicional. Se <i>COND</i> for verdadeira, essa operação retornará o resultado de <i>EXPR1</i> . Caso contrário, o resultado da avaliação de <i>EXPR2</i> será retornado.
if COND1 then EXPR1 elseif COND2 then EXPR2 else EXPR_N endif	Qualquer	Esta operação é uma avaliação condicional com várias ramificações. Se <i>COND1</i> for verdadeira, essa operação retornará o resultado de <i>EXPR1</i> . Caso contrário, se <i>COND2</i> for verdadeira, essa operação retornará o resultado da avaliação de <i>EXPR2</i> . Caso contrário, o resultado da avaliação de <i>EXPR_N</i> será retornado.

Funções numéricas

CLEM contém inúmeras funções numéricas comumente usadas.

Tabela 24. Funções numéricas do CLEM.

Função	Resultado	Descrição
-NUM	Número	Usado para negar <i>NUM</i> . Retorna o número correspondente com o sinal oposto.
NUM1 + NUM2	Número	Retorna a soma de <i>NUM1</i> e <i>NUM2</i> .
NUM1 - NUM2	Número	Retorna o valor de <i>NUM2</i> subtraído de <i>NUM1</i> .
NUM1 * NUM2	Número	Retorna o valor de <i>NUM1</i> multiplicado por <i>NUM2</i> .
NUM1 / NUM2	Número	Retorna o valor de <i>NUM1</i> dividido por <i>NUM2</i> .
INT1 div INT2	Número	Usado para executar divisão de número inteiro. Retorna o valor de <i>INT1</i> dividido por <i>INT2</i> .
INT1 rem INT2	Número	Retorna o restante de <i>INT1</i> dividido por <i>INT2</i> . Por exemplo, <code>INT1 - (INT1 div INT2) * INT2</code> .
INT1 mod INT2	Número	Essa função foi descontinuada. Use a função <code>rem</code> no lugar.

Tabela 24. Funções numéricas do CLEM (continuação).

Função	Resultado	Descrição
$BASE ** POWER$	Número	Retorna $BASE$ elevado à $POWER$ potência, que pode ser qualquer número (exceto que $BASE$ não deve ser zero se $POWER$ for zero de algum tipo além do número inteiro 0). Se $POWER$ for um número inteiro, o cálculo será executado multiplicando-se as potências de $BASE$ sucessivamente. Assim, se $BASE$ for um número inteiro, o resultado será um número inteiro. Se $POWER$ for o número inteiro 0, o resultado será sempre um 1 do mesmo tipo do $BASE$. Caso contrário, se $POWER$ não for um número inteiro, o resultado será calculado como $\exp(POWER * \log(BASE))$.
$abs(NUM)$	Número	Retorna o valor absoluto de NUM , que é sempre um número do mesmo tipo.
$exp(NUM)$	Real	Retorna e elevado à NUM potência, em que e é a base dos logaritmos naturais.
$fracof(NUM)$	Real	Retorna a parte fracionária de NUM , definida como $NUM - \text{intof}(NUM)$.
$\text{intof}(NUM)$	Número inteiro	Trunca seu argumento para um número inteiro. Retorna o número inteiro com o mesmo sinal que NUM e com a maior magnitude para que $abs(INT) \leq abs(NUM)$.
$\log(NUM)$	Real	Retorna o logaritmo natural (base e) de NUM , que não deve ser zero de nenhum tipo.
$\log10(NUM)$	Real	Retorna o logaritmo 10 de base de NUM , que não deve ser zero de nenhum tipo. Essa função é definida como $\log(NUM) / \log(10)$.
$negate(NUM)$	Número	Usado para negar NUM . Retorna o número correspondente com o sinal oposto.
$round(NUM)$	Número inteiro	Usado para arredondar NUM para um número inteiro usando $\text{intof}(NUM+0.5)$ se NUM for positivo ou $\text{intof}(NUM-0.5)$ se NUM for negativo.
$sign(NUM)$	Número	Usado para determinar o sinal de NUM . Essa operação retorna -1 , 0 ou 1 se NUM for um número inteiro. Se NUM for real, ela retornará -1.0 , 0.0 ou 1.0 , dependendo se NUM for negativo, zero ou positivo.
$\text{sqrt}(NUM)$	Real	Retorna a raiz quadrada de NUM . NUM deve ser positivo.
$\text{sum}_n(\text{LIST})$	Número	Retorna a soma dos valores de uma lista de campos numéricos ou nulo se todos os valores de campo forem nulos.
$\text{mean}_n(\text{LIST})$	Número	Retorna o valor médio de uma lista de campos numéricos ou nulo se todos os valores de campo forem nulos.
$\text{sdev}_n(\text{LIST})$	Número	Retorna o desvio padrão de uma lista de campos numéricos ou nulo se todos os valores de campo forem nulos.

Funções trigonométricas

Todas as funções nessa seção usam um ângulo como argumento ou retornam um como resultado. Em ambos os casos, as unidades do ângulo (radianos ou graus) são controladas pela configuração da opção de fluxo relevante.

Tabela 25. Funções trigonométricas do CLEM

Função	Resultado	Descrição
$\arccos(NUM)$	Real	Calcula o arco cosseno do ângulo especificado.
$\text{arccosh}(NUM)$	Real	Calcula o arco cosseno hiperbólico do ângulo especificado.

Tabela 25. Funções trigonométricas do CLEM (continuação)

Função	Resultado	Descrição
arcsin(NUM)	Real	Calcula o arco seno do ângulo especificado.
arcsinh(NUM)	Real	Calcula o arco seno hiperbólico do ângulo especificado.
arctan(NUM)	Real	Calcula o arco tangente do ângulo especificado.
arctan2(NUM_Y, NUM_X)	Real	Calcula o arco tangente do NUM_Y / NUM_X e usa os sinais dos dois números para derivar informações do quadrante. O resultado é um real no intervalo $-\pi < \text{ANGLE} \leq \pi$ (radianos) – $-180 < \text{ANGLE} \leq 180$ (graus)
arctanh(NUM)	Real	Calcula o arco tangente hiperbólico do ângulo especificado.
cos(NUM)	Real	Calcula o cosseno do ângulo especificado.
cosh(NUM)	Real	Calcula o cosseno hiperbólico do ângulo especificado.
pi	Real	Essa constante é a melhor aproximação real para pi.
sin(NUM)	Real	Calcula o seno do ângulo especificado.
sinh(NUM)	Real	Calcula o seno hiperbólico do ângulo especificado.
tan(NUM)	Real	Calcula a tangente do ângulo especificado.
tanh(NUM)	Real	Calcula a tangente hiperbólica do ângulo especificado.

Funções de probabilidade

Funções de probabilidade retornam probabilidades baseadas em várias distribuições, como a probabilidade de um valor da distribuição *t* do Estudante ser menor que um valor específico.

Tabela 26. Funções de probabilidade do CLEM

Função	Resultado	Descrição
cdf_chisq(NUM, DF)	Real	Retorna a probabilidade de um valor da distribuição qui-quadrado com os graus de liberdade especificados serem menores que o número especificado.
cdf_f(NUM, DF1, DF2)	Real	Retorna a probabilidade de um valor da distribuição <i>F</i> , com graus de liberdade <i>DF1</i> e <i>DF2</i> , ser menor que o número especificado.
cdf_normal(NUM, MEAN, STDDEV)	Real	Retorna a probabilidade de um valor da distribuição normal com a média especificada e um desvio padrão ser menor que o número especificado.
cdf_t(NUM, DF)	Real	Retorna a probabilidade de um valor da distribuição <i>t</i> do Estudante com os graus de liberdade especificados serem menores que o número especificado.

Funções espaciais

Funções espaciais podem ser usadas com dados geoespaciais. Por exemplo, elas permitem calcular as distâncias entre dois pontos, a área de um polígono, entre outras coisas. Também pode haver situações que requerem uma mesclagem de diversos conjuntos de dados geoespaciais baseados em um predicado espacial (dentro, perto de, entre outros), o que pode ser feito por meio de uma condição de mesclagem.

Essas funções espaciais trabalham com o sistema de coordenadas especificado em **Ferramentas > Propriedades do Fluxo > Opções > Geoespacial**.

Nota: Essas funções espaciais não se aplicam a dados tridimensionais. Se os dados tridimensionais forem importados no fluxo, somente as duas primeiras dimensões serão usadas por essas funções. Os valores do eixo Z serão ignorados.

Tabela 27. Funções espaciais do CLEM

Função	Resultado	Descrição
<code>close_to(SHAPE,SHAPE,NUM)</code>	<i>Boolean</i>	Testa se 2 formas estão dentro de uma certa DISTANCE uma da outra. Se um sistema de coordenadas projetadas for usado, a DISTANCE será em metros. Se nenhum sistema de coordenadas for usado, essa será uma unidade arbitrária.
<code>crosses(SHAPE,SHAPE)</code>	<i>Boolean</i>	Testa se 2 formas cruzam uma à outra. Essa função é adequada para 2 formas de linestring ou 1 linestring e um 1 polígono.
<code>overlap(SHAPE,SHAPE)</code>	<i>Boolean</i>	Testa se há uma interseção entre 2 polígonos e se a interseção é interior para ambas as formas.
<code>within(SHAPE,SHAPE)</code>	<i>Boolean</i>	Testa se a totalidade de SHAPE1 está contida em um POLYGON.
<code>area(SHAPE)</code>	<i>Real</i>	Retorna a área do POLYGON especificado. Se um sistema planejado for usado, a função retornará metros quadrados. Se nenhum sistema de coordenadas for usado, essa será uma unidade arbitrária. A forma deve ser POLYGON ou a MULTIPOLYGON.
<code>num_points(SHAPE,LIST)</code>	<i>Número inteiro</i>	Retorna o número de pontos de um campo de ponto (MULTIPOINT) que estão contidos dentro dos limites de um POLYGON. SHAPE1 deve ser POLYGON ou MULTIPOLYGON.
<code>distance(SHAPE,SHAPE)</code>	<i>Real</i>	Retorna a distância entre SHAPE1 e SHAPE2. Se um sistema de coordenadas projetadas for usado, a função retornará metros. Se nenhum sistema de coordenadas for usado, essa será uma unidade arbitrária. SHAPE1 e SHAPE2 podem ser qualquer tipo de medição geo.

Operações de número inteiro bit a bit

Essas funções permitem que números inteiros sejam manipulados como padrões de bits representando valores de dois complementos, em que a posição de bit N tem um peso 2^{*N} . Bits são numerados de 0 para cima. Essas operações agem como se o bit de sinal de um número inteiro fosse estendido indefinidamente para a esquerda. Assim, em qualquer lugar acima de seu bit mais significativo, um número inteiro positivo tem 0 bit e um número inteiro negativo tem 1 bit.

Tabela 28. Operações de número inteiro bit a bit do CLEM.

Função	Resultado	Descrição
<code>~ INT1</code>	<i>Número inteiro</i>	Produz o complemento bit a bit do número inteiro INT1. Ou seja, há um 1 no resultado para cada posição de bit para a qual INT1 tem 0. É sempre verdadeiro que $\sim INT = -(INT + 1)$.
<code>INT1 INT2</code>	<i>Número inteiro</i>	O resultado dessa operação é o bit a bit "or inclusivo" de INT1 e INT2. Ou seja, há um 1 no resultado para cada posição de bit para a qual há um 1 em INT1 ou INT2 ou ambos.
<code>INT1 /& INT2</code>	<i>Número inteiro</i>	O resultado dessa operação é o bit a bit "or exclusivo" de INT1 e INT2. Ou seja, há um 1 no resultado para cada posição de bit para a qual há um 1 em INT1 ou INT2, mas não ambos.

Tabela 28. Operações de número inteiro bit a bit do CLEM (continuação).

Função	Resultado	Descrição
$INT1 \ \&\& \ INT2$	Número inteiro	Produz o bit a bit "and" dos números inteiros $INT1$ e $INT2$. Ou seja, há um 1 no resultado para cada posição de bit para a qual há um 1 em $INT1$ e $INT2$.
$INT1 \ \&\&\sim \ INT2$	Número inteiro	Produz o bit a bit "and" de $INT1$ e o complemento bit a bit de $INT2$. Ou seja, há um 1 no resultado para cada posição de bit para a qual há um 1 em $INT1$ e um 0 em $INT2$. É o mesmo que $INT1 \ \&\& \ (\sim INT2)$ e é útil para a verificação de bits de $INT1$ configurados em $INT2$.
$INT \ \ll \ N$	Número inteiro	Produz o padrão de bits de $INT1$ deslocado para a esquerda N posições. Um valor negativo para N produz um deslocamento para a direita.
$INT \ \gg \ N$	Número inteiro	Produz o padrão de bits de $INT1$ deslocado para a direita N posições. Um valor negativo para N produz um deslocamento para a esquerda.
$INT1 \ \&\&=_0 \ INT2$	Booleano	Equivalente à expressão booleana $INT1 \ \&\& \ INT2 \ /== \ 0$, porém mais eficiente
$INT1 \ \&\&/=_0 \ INT2$	Booleano	Equivalente à expressão booleana $INT1 \ \&\& \ INT2 \ == \ 0$, porém mais eficiente.
<code>integer_bitcount(INT)</code>	Número inteiro	Conta o número de 1 ou 0 bits na representação de dois complementos de INT . Se INT for não negativo, N será o número de 1 bit. Se INT for negativo, será o número de 0 bit. Devido à extensão do sinal, há um número infinito de 0 bit em um número inteiro não negativo ou 1 bit em um número inteiro negativo. Esse é sempre o caso de $\text{integer_bitcount}(INT) = \text{integer_bitcount}(- (INT+1))$.
<code>integer_leastbit(INT)</code>	Número inteiro	Retorna a posição de bit N do bit menos significativo configurado no número inteiro INT . N é a potência mais alta de 2 pela qual INT é dividido exatamente.
<code>integer_length(INT)</code>	Número inteiro	Retorna o comprimento em bits de INT como um número inteiro de dois complementos. Ou seja, N é o menor número inteiro, de modo que $INT < (1 \ll N)$ se $INT \geq 0$ $INT \geq (-1 \ll N)$ se $INT < 0$. Se INT for não negativo, a representação de INT como um número inteiro não assinado requer um campo de pelo menos N bits. Alternativamente, no mínimo $N+1$ bits é necessário para representar INT como um número inteiro assinado, independentemente do sinal.
<code>testbit(INT, N)</code>	Booleano	Testa o bit na posição N no número inteiro INT e retorna o estado do bit N como um valor booleano, que é verdadeiro para 1 e falso para 0.

Funções aleatórias

As funções a seguir são usadas para selecionar itens ou gerar números aleatoriamente.

Tabela 29. Funções aleatórias do CLEM.

Função	Resultado	Descrição
<code>oneof(LIST)</code>	Qualquer um	Retorna um elemento escolhido aleatoriamente de $LIST$. Os itens da lista devem ser inseridos como $[ITEM1, ITEM2, \dots, ITEM_N]$. Observe que uma lista de nomes de campo também pode ser especificada.

Tabela 29. Funções aleatórias do CLEM (continuação).

Função	Resultado	Descrição
random(NUM)	Número	Retorna um número aleatório distribuído uniformemente do mesmo tipo (<i>INT</i> ou <i>REAL</i>), começando de 1 a <i>NUM</i> . Se você usar número inteiro, somente números inteiros serão retornados. Se você usar um número real (decimal), números reais serão retornados (precisão decimal determinada pelas opções de fluxo). O maior número aleatório retornado pela função poderia ser igual a <i>NUM</i> .
random0(NUM)	Número	Tem as mesmas propriedades que random(NUM), mas começando em 0. O maior número aleatório retornado pela função nunca será igual a <i>NUM</i> .

Funções de sequência de caracteres

No CLEM, é possível executar as operações a seguir com sequências de caracteres:

- Comparar sequências de caracteres
- Criar sequências de caracteres
- Acessar caracteres

No CLEM, uma sequência de caracteres pode ser qualquer sequência de caracteres entre aspas duplas correspondentes ("aspas de sequência de caracteres"). Caracteres (*CHAR*) podem ser qualquer caractere alfanumérico único. Eles são declarados expressões em CLEM usando crases simples em forma de `<caractere>`, como ``z``, ``A`` ou ``2``. Caracteres que estiverem fora dos limites ou forem índices negativos para uma sequência de caracteres resultarão em um comportamento indefinido.

Nota. As comparações entre sequências de caracteres que usam e não usam SQL pushback podem gerar diferentes resultados onde existem espaços à direita.

Tabela 30. Funções de sequência de caracteres do CLEM.

Função	Resultado	Descrição
allbutfirst(N, STRING)	Sequência de caracteres	Retorna uma sequência de caracteres, que é <i>STRING</i> com os primeiros <i>N</i> caracteres removidos.
allbutlast(N, STRING)	Sequência de caracteres	Retorna uma sequência de caracteres, que é <i>STRING</i> com os últimos caracteres removidos.
alphabefore(STRING1, STRING2)	Booleano	Usado para verificar a ordenação alfabética das sequências de caracteres. Retorna true se <i>STRING1</i> preceder <i>STRING2</i> .
endstring(LENGTH, STRING)	Sequência de caracteres	Extrai os últimos <i>N</i> caracteres da sequência de caracteres especificada. Se o comprimento da sequência de caracteres for menor que ou igual ao comprimento especificado, ele permanecerá inalterado.
hasendstring(STRING, SUBSTRING)	Número inteiro	Essa função é igual à <i>isendstring</i> (SUBSTRING, STRING).
hasmidstring(STRING, SUBSTRING)	Número inteiro	Essa função é igual à <i>ismidstring</i> (SUBSTRING, STRING) (subsequência de caracteres integrada).
hasstartstring(STRING, SUBSTRING)	Número inteiro	Essa função é igual à <i>isstartstring</i> (SUBSTRING, STRING).
hassubstring(STRING, N, SUBSTRING)	Número inteiro	Essa função é igual à <i>issubstring</i> (SUBSTRING, N, STRING), em que <i>N</i> é padronizado para 1.

Tabela 30. Funções de sequência de caracteres do CLEM (continuação).

Função	Resultado	Descrição
count_substring(<i>STRING</i> , <i>SUBSTRING</i>)	Número inteiro	Retorna o número de vezes que a subsequência de caracteres especificada ocorre dentro da sequência de caracteres. Por exemplo, count_substring("foooo.txt", "oo") retorna 3.
hassubstring(<i>STRING</i> , <i>SUBSTRING</i>)	Número inteiro	Essa função é igual à isssubstring(<i>SUBSTRING</i> , 1, <i>STRING</i>), em que <i>N</i> é padronizado para 1.
isalphacode(<i>CHAR</i>)	Booleano	Retorna um valor de true se <i>CHAR</i> for um caractere na sequência de caracteres especificada (geralmente um nome de campo) cujo código de caractere é uma letra. Caso contrário, essa função retorna um valor de 0. Por exemplo, isalphacode(produce_num(1)).
isendstring(<i>SUBSTRING</i> , <i>STRING</i>)	Número inteiro	Se a sequência de caracteres <i>STRING</i> terminar com a subsequência de caracteres <i>SUBSTRING</i> , então essa função retornará o número inteiro subscrito de <i>SUBSTRING</i> em <i>STRING</i> . Caso contrário, esta função retornará um valor 0.
islowercode(<i>CHAR</i>)	Booleano	Retorna um valor de true se <i>CHAR</i> for um caractere de letra minúscula para a sequência de caracteres especificada (geralmente um nome de campo). Caso contrário, essa função retorna um valor de 0. Por exemplo, islowercode(``) e islowercode(country_name(2)) são expressões válidas.
ismidstring(<i>SUBSTRING</i> , <i>STRING</i>)	Número inteiro	Se <i>SUBSTRING</i> for uma subsequência de caracteres de <i>STRING</i> , mas não começar no primeiro caractere de <i>STRING</i> ou terminar no último, essa função retornará o subscrito no qual a subsequência de caracteres começa. Caso contrário, esta função retornará um valor 0.
isnumbercode(<i>CHAR</i>)	Booleano	Retorna um valor de true se <i>CHAR</i> para a sequência de caracteres especificada (geralmente um nome de campo) for um caractere cujo código de caractere é um dígito. Caso contrário, essa função retorna um valor de 0. Por exemplo, isnumbercode(product_id(2)).
isstartstring(<i>SUBSTRING</i> , <i>STRING</i>)	Número inteiro	Se a sequência de caracteres <i>STRING</i> começar com a subsequência de caracteres <i>SUBSTRING</i> , essa função retornará o subscrito 1. Caso contrário, essa função retornará um valor de 0.
issubstring(<i>SUBSTRING</i> , <i>N</i> , <i>STRING</i>)	Número inteiro	Procura a sequência de caracteres <i>STRING</i> , começando de seu <i>N</i> º caractere, para uma sequência de caracteres igual à <i>SUBSTRING</i> . Se for localizada, essa função retornará a inscrição de número inteiro onde começa a subsequência de caracteres correspondente. Caso contrário, essa função retornará um valor 0. Se <i>N</i> não for fornecido, essa função será padronizada para 1.

Tabela 30. Funções de sequência de caracteres do CLEM (continuação).

Função	Resultado	Descrição
<code>issubstring(SUBSTRING, STRING)</code>	Número inteiro	Procura a sequência de caracteres <i>STRING</i> , começando de seu <i>N</i> º caractere, para uma sequência de caracteres igual à <i>SUBSTRING</i> . Se for localizada, essa função retornará a inscrição de número inteiro onde começa a subsequência de caracteres correspondente. Caso contrário, esta função retornará um valor 0. Se <i>N</i> não for fornecido, essa função será padronizada para 1.
<code>issubstring_count(SUBSTRING, N, STRING):</code>	Número inteiro	Retorna o índice da <i>N</i> ª ocorrência de <i>SUBSTRING</i> dentro da <i>STRING</i> especificada. Se houver menos de <i>N</i> ocorrências de <i>SUBSTRING</i> , um 0 será retornado.
<code>issubstring_lim(SUBSTRING, N, STARTLIM, ENDLIM, STRING)</code>	Número inteiro	Essa função é igual à <code>issubstring</code> , mas a correspondência é restringida para iniciar até o subscrito <i>STARTLIM</i> e para terminar até o subscrito <i>ENDLIM</i> . As restrições <i>STARTLIM</i> ou <i>ENDLIM</i> podem ser desativadas fornecendo um valor de <code>false</code> para qualquer argumento - por exemplo, <code>issubstring_lim(SUBSTRING, N, false, false, STRING)</code> é igual à <code>issubstring</code> .
<code>isuppercode(CHAR)</code>	Booleano	Retorna um valor de <code>true</code> se <i>CHAR</i> for um caractere de letra maiúscula. Caso contrário, essa função retorna um valor de 0. Por exemplo, <code>isuppercode(` `)</code> e <code>isuppercode(country_name(2))</code> são expressões válidas.
<code>last(CHAR)</code>	Sequência de caracteres	Retorna o último caractere <i>CHAR</i> de <i>STRING</i> (que deve ter pelo menos um caractere de comprimento).
<code>length(STRING)</code>	Número inteiro	Retorna o comprimento da sequência de caracteres <i>STRING</i> --ou seja, o número de caracteres nele.
<code>locchar(CHAR, N, STRING)</code>	Número inteiro	Usado para identificar a localização dos caracteres em campos simbólicos. A função procura a sequência de caracteres <i>STRING</i> para o caractere <i>CHAR</i> , começando a procura no <i>N</i> º caractere de <i>STRING</i> . Essa função retorna um valor indicando a localização (começando em <i>N</i>) onde o caractere está. Se o caractere não for localizado, essa função retornará um valor de 0. Se a função tiver um offset inválido (<i>N</i>) (por exemplo, um offset além do comprimento da sequência de caracteres), essa função retornará <code>\$null\$</code> . Por exemplo, <code>locchar(`n`, 2, web_page)</code> procura o campo chamado <i>web_page</i> para o caractere <code>`n`</code> , começando no segundo caractere no valor de campo. <i>Nota:</i> Certifique-se de usar crases simples para conter o caractere especificado.

Tabela 30. Funções de sequência de caracteres do CLEM (continuação).

Função	Resultado	Descrição
locchar_back(CHAR, N, STRING)	Número inteiro	Semelhante a locchar, exceto que a procura é executada backward, começando no N ^o caractere. Por exemplo, locchar_back(`n`, 9, web_page) procura o campo <i>web_page</i> , começando no nono caractere e movendo backward para o início da sequência de caracteres. Se a função tiver um offset inválido (por exemplo, um offset além do comprimento da sequência de caracteres), essa função retornará \$null\$. O ideal é usar locchar_back com a função length(<field>) para usar dinamicamente o comprimento do valor atual do campo. Por exemplo, locchar_back(`n`, (length(web_page)), web_page).
lowertoupper(CHAR) lowertoupper (STRING)	CHAR ou String	A entrada pode ser uma sequência de caracteres ou um caractere, que é usada nessa função para retornar um novo item do mesmo tipo, com quaisquer caracteres minúsculos convertidos em seus equivalentes maiúsculos. Por exemplo, lowertoupper(`a`), lowertoupper("My string") e lowertoupper(field_name(2)) são expressões válidas.
matches	Booleano	Retorna true se uma sequência de caracteres corresponder a um padrão especificado. O padrão deve ser uma sequência de caracteres literal; ele não pode ser um nome contendo um padrão. Um ponto de interrogação (?) pode ser incluído no padrão para corresponder a exatamente um caractere; um asterisco (*) corresponde a zero ou mais caracteres. Para corresponder a um ponto de interrogação literal ou asterisco (em vez de usá-los como curinga), uma barra invertida (\) pode ser usada como um caractere de escape.
replace(SUBSTRING, NEWSUBSTRING, STRING)	Sequência de caracteres	Dentro da STRING especificada, substitua todas as instâncias de SUBSTRING por NEWSUBSTRING.
replicate(COUNT, STRING)	Sequência de caracteres	Retorna uma sequência de caracteres que consiste na sequência de caracteres original copiada o número de vezes especificado.
stripchar(CHAR, STRING)	Sequência de caracteres	Permite remover caracteres especificados de uma sequência de caracteres ou campo. É possível usar essa função, por exemplo, para remover símbolos adicionais, como notações de moeda, dos dados para atingir um nome ou número simples. Por exemplo, o uso da sintaxe stripchar(`\$`, 'Cost') retorna um novo campo com o sinal de dólar removido de todos os valores. <i>Nota:</i> Certifique-se de usar crases simples para conter o caractere especificado.

Tabela 30. Funções de sequência de caracteres do CLEM (continuação).

Função	Resultado	Descrição
skipchar(CHAR, N, STRING)	Número inteiro	Procura na sequência de caracteres <i>STRING</i> qualquer caractere além de <i>CHAR</i> , começando no <i>N</i> º caractere. Essa função retorna uma subsequência de caracteres de número inteiro, indicando o ponto em que um deles é localizado ou 0 se cada caractere do <i>N</i> º em diante for <i>CHAR</i> . Se a função tiver um offset inválido (por exemplo, um offset além do comprimento da sequência de caracteres), essa função retornará \$null\$. locchar geralmente é usado com funções skipchar para determinar o valor de <i>N</i> (o ponto em que começa a procura da sequência de caracteres). Por exemplo, skipchar(`s`, (locchar(`s`, 1, "MyString")), "MyString").
skipchar_back(CHAR, N, STRING)	Número inteiro	Semelhante a skipchar, exceto que a procura é executada backward , começando no <i>N</i> º caractere.
startstring(LENGTH, STRING)	Sequência de caracteres	Extraí os primeiros <i>N</i> caracteres da sequência de caracteres especificada. Se o comprimento da sequência de caracteres for menor que ou igual ao comprimento especificado, ele permanecerá inalterado.
strmember(CHAR, STRING)	Número inteiro	Equivalente à locchar(CHAR, 1, STRING). Retorna uma subsequência de caracteres de número inteiro, indicando o ponto em que ocorre o primeiro <i>CHAR</i> , ou 0. Se a função tiver um offset inválido (por exemplo, um offset além do comprimento da subsequência de caracteres), essa função retornará \$null\$.
subscrs(N, STRING)	CHAR	Retorna o <i>N</i> º caractere <i>CHAR</i> da sequência de caracteres de entrada <i>STRING</i> . Essa função também pode ser gravada em forma de taquigrafia como STRING(<i>N</i>). Por exemplo, lowertoupper("name"(1)) é uma expressão válida.
substring(N, LEN, STRING)	Sequência de caracteres	Retorna uma sequência de caracteres <i>SUBSTRING</i> , que consiste nos caracteres <i>LEN</i> da sequência de caracteres <i>STRING</i> , começando no caractere no subscrito <i>N</i> .
substring_between(N1, N2, STRING)	Sequência de caracteres	Retorna a subsequência de caracteres de <i>STRING</i> , que começa no subscrito <i>N1</i> e termina no subscrito <i>N2</i> .
trim(STRING)	Sequência de caracteres	Remove os caracteres de espaço em branco inicial e final da sequência de caracteres especificada.
trim_start(STRING)	Sequência de caracteres	Remove os caracteres de espaço em branco iniciais da sequência de caracteres especificada.
trimend(STRING)	Sequência de caracteres	Remove os caracteres de espaço em branco finais da sequência de caracteres especificada.
unicode_char(NUM)	CHAR	Entrada deve ser valores decimais, e não hexadecimais. Retorna o caractere com valor Unicode <i>NUM</i> .
unicode_value(CHAR)	NUM	Retorna o valor Unicode de <i>CHAR</i>

Tabela 30. Funções de sequência de caracteres do CLEM (continuação).

Função	Resultado	Descrição
uppertolower(CHAR) uppertolower (STRING)	CHAR ou String	A entrada pode ser uma sequência de caracteres ou um caractere e é usada nessa função para retornar um novo item do mesmo tipo, com quaisquer caracteres maiúsculos convertidos em seus equivalentes minúsculos. <i>Nota:</i> Lembre-se de especificar sequências de caracteres com aspas duplas e caracteres com crases simples. Nomes de campos simples devem ser especificados sem aspas.

Funções SoundEx

SoundEx é um método usado para localizar sequências quando o som é conhecido, mas a ortografia precisa não. Desenvolvido em 1918, ele procura palavras com sons similares com base em suposições fonéticas sobre como certas letras são pronunciadas. Ele pode ser usado para procurar nomes em um banco de dados, por exemplo, no qual ortografias e pronúncias para nomes similares podem variar. O algoritmo SoundEx básico é documentado em várias origens e, apesar das limitações conhecidas (por exemplo, combinações de letra inicial como ph e f não corresponderão, mesmo que soarem iguais), é suportado de alguma forma pela maioria dos bancos de dados.

Tabela 31. Funções soundex do CLEM

Função	Resultado	Descrição
soundex(STRING)	Número inteiro	Retorna o código SoundEx de quatro caracteres para a <i>STRING</i> especificada.
soundex_difference(STRING1, STRING2)	Número inteiro	Retorna um número inteiro entre 0 e 4 indicando o número de caracteres que são iguais na mesma codificação SoundEx para as duas sequências de caracteres, em que 0 indica nenhuma semelhança e 4 indica grande semelhança ou sequências de caracteres idênticas.

Funções de data e hora

CLEM inclui uma família de funções para o tratamento de campos com armazenamento de data/hora das variáveis de sequência de caracteres representando datas e horas. Os formatos de data e hora usados são específicos para cada fluxo e especificados na caixa de diálogo de propriedades do fluxo. As funções de data e hora analisam sequências de caracteres de data e hora de acordo com o formato atualmente selecionado.

Quando você especifica um ano em uma data que usa somente dois dígitos (ou seja, o século não é especificado), IBM SPSS Modeler usa o século padrão que é especificado na caixa de diálogo de propriedades do fluxo.

Nota: Se a função de dados for enviada de volta para SQL ou IBM SPSS Analytic Server, em uma ramificação que segue uma origem de dados do Analytic Server, quaisquer sequências de formatações de data (to_date) dentro dos dados deverão corresponder ao formato de data especificado no fluxo do SPSS Modeler.

Tabela 32. Funções de data e hora do CLEM.

Função	Resultado	Descrição
@TODAY	<i>Sequência de caracteres</i>	Se você selecionar Dias/Minutos de Rolagem na caixa de diálogo de propriedades do fluxo, essa função retornará a data atual como uma sequência de caracteres no formato de data atual. Se você usar um formato de data de dois dígitos e não selecionar Dias/Minutos de Rolagem , essa função retornará \$null\$ no servidor atual.
to_time(ITEM)	<i>Time</i>	Converte o armazenamento do campo especificado em hora.
to_date(ITEM)	<i>Data</i>	Converte o armazenamento do campo especificado em uma data.
to_timestamp(ITEM)	<i>Registro de data e hora</i>	Converte o armazenamento do campo especificado em um registro de data e hora.
to_datetime(ITEM)	<i>Data/hora</i>	Converte o armazenamento do campo especificado em um valor de data, hora ou registro de data e hora.
datetime_date(ITEM)	<i>Data</i>	Retorna o valor de data para um <i>número</i> , uma <i>sequência de caracteres</i> ou um <i>registro de data e hora</i> . Observe que essa é a única função que permite converter um número (em segundos) de volta em uma data. Se ITEM for uma sequência de caracteres, cria uma data analisando uma sequência de caracteres no formato de data atual. O formato de data especificado na caixa de diálogo de propriedades de fluxo deve estar correto para que essa função seja bem-sucedida. Se ITEM for um número, ele será interpretado como um número de segundos desde a data (ou época) de base. Frações de um dia são truncadas. Se ITEM for registro de data e hora, a parte de data do registro de data e hora será retornada. Se ITEM for uma data, ela será retornada inalterada.
date_before(DATE1, DATE2)	<i>Booleano</i>	Retorna um valor de true se DATE1 representar uma data ou um registro de data e hora anteriores aos representados por DATE2. Caso contrário, esta função retornará um valor 0.
date_days_difference(DATE1, DATE2)	<i>Número inteiro</i>	Retorna o tempo, em dias, a partir da data ou do registro de data e hora representados por DATE1 até os representados por DATE2 como um número inteiro. Se DATE2 for anterior a DATE1, essa função retornará um número negativo.
date_in_days(DATE)	<i>Número inteiro</i>	Retorna o tempo, em dias, a partir da data de linha de base até a data ou o registro de data e hora representados por DATE como um número inteiro. Se DATE for anterior à data da linha de base, essa função retornará um número negativo. Deve-se incluir uma data válida para o cálculo funcionar de maneira adequada. Por exemplo, você não deve especificar 29 de fevereiro de 2001 como a data. Como 2001 não é um ano bissexto, essa data não existe.

Tabela 32. Funções de data e hora do CLEM (continuação).

Função	Resultado	Descrição
date_in_months(<i>DATE</i>)	<i>Real</i>	Retorna o tempo, em meses, a partir da data de linha de base até a data ou o registro de data e hora representados por <i>DATE</i> como um número real. Essa é uma figura aproximada com base em um mês de 30,4375 dias. Se <i>DATE</i> for anterior à data da linha de base, essa função retornará um número negativo. Deve-se incluir uma data válida para o cálculo funcionar de maneira adequada. Por exemplo, você não deve especificar 29 de fevereiro de 2001 como a data. Como 2001 não é um ano bissexto, essa data não existe.
date_in_weeks(<i>DATE</i>)	<i>Real</i>	Retorna o tempo, em semanas, a partir da data de linha de base até a data ou o registro de data e hora representados por <i>DATE</i> como um número real. Isso é baseado em uma semana de 7 dias. Se <i>DATE</i> for anterior à data da linha de base, essa função retornará um número negativo. Deve-se incluir uma data válida para o cálculo funcionar de maneira adequada. Por exemplo, você não deve especificar 29 de fevereiro de 2001 como a data. Como 2001 não é um ano bissexto, essa data não existe.
date_in_years(<i>DATE</i>)	<i>Real</i>	Retorna o tempo, em anos, a partir da data de linha de base até a data ou o registro de data e hora representados por <i>DATE</i> como um número real. Essa é uma figura aproximada com base em um ano de 365,25 dias. Se <i>DATE</i> for anterior à data da linha de base, essa função retornará um número negativo. Deve-se incluir uma data válida para o cálculo funcionar de maneira adequada. Por exemplo, você não deve especificar 29 de fevereiro de 2001 como a data. Como 2001 não é um ano bissexto, essa data não existe.
date_months_difference (<i>DATE1</i> , <i>DATE2</i>)	<i>Real</i>	Retorna o tempo, em meses, a partir da data ou do registro de data e hora representados por <i>DATE1</i> até os representados por <i>DATE2</i> como um número real. Essa é uma figura aproximada com base em um mês de 30,4375 dias. Se <i>DATE2</i> for anterior a <i>DATE1</i> , essa função retornará um número negativo.
datetime_date(<i>YEAR</i> , <i>MONTH</i> , <i>DAY</i>)	<i>Data</i>	Cria um valor de data para <i>YEAR</i> , <i>MONTH</i> e <i>DAY</i> fornecidos. Os argumentos devem ser números inteiros.
datetime_day(<i>DATE</i>)	<i>Número inteiro</i>	Retorna o dia do mês de um <i>DATE</i> ou registro de data e hora fornecidos. O resultado é um número inteiro no intervalo de 1 a 31.
datetime_day_name(<i>DAY</i>)	<i>Sequência de caracteres</i>	Retorna o nome completo do <i>DAY</i> fornecido. O argumento deve ser um número inteiro no intervalo de 1 (domingo) a 7 (sábado).
datetime_hour(<i>TIME</i>)	<i>Número inteiro</i>	Retorna a hora de um <i>TIME</i> ou registro de data e hora. O resultado é um número inteiro no intervalo de 0 a 23.
datetime_in_seconds(<i>TIME</i>)	<i>Real</i>	Retorna a parte de segundos armazenada em <i>TIME</i> .
datetime_in_seconds(<i>DATE</i>), datetime_in_seconds(<i>DATETIME</i>)	<i>Real</i>	Retorna o número acumulado, convertido em segundos, da diferença entre <i>DATE</i> ou <i>DATETIME</i> atuais e a data de linha de base (1900-01-01).
datetime_minute(<i>TIME</i>)	<i>Número inteiro</i>	Retorna o minuto de um <i>TIME</i> ou registro de data e hora. O resultado é um número inteiro no intervalo de 0 a 59.
datetime_month(<i>DATE</i>)	<i>Número inteiro</i>	Retorna o mês de um <i>DATE</i> ou registro de data e hora. O resultado é um número inteiro no intervalo de 1 a 12.

Tabela 32. Funções de data e hora do CLEM (continuação).

Função	Resultado	Descrição
datetime_month_name (MONTH)	Sequência de caracteres	Retorna o nome completo do <i>MONTH</i> fornecido. O argumento deve ser um número inteiro no intervalo de 1 a 12.
datetime_now	Registro de data e hora	Retorna a hora atual como um registro de data e hora.
datetime_second(TIME)	Número inteiro	Retorna o segundo de um <i>TIME</i> ou registro de data e hora. O resultado é um número inteiro no intervalo de 0 a 59.
datetime_day_short_name (DAY)	Sequência de caracteres	Retorna o nome abreviado do <i>DAY</i> fornecido. O argumento deve ser um número inteiro no intervalo de 1 (domingo) a 7 (sábado).
datetime_month_short_name (MONTH)	Sequência de caracteres	Retorna o nome abreviado do <i>MONTH</i> fornecido. O argumento deve ser um número inteiro no intervalo de 1 a 12.
datetime_time(HOUR, MINUTE, SECOND)	Time	Retorna o valor do <i>HOUR</i> , <i>MINUTE</i> e <i>SECOND</i> especificados. Os argumentos devem ser números inteiros.
datetime_time(ITEM)	Time	Retorna o valor de hora do <i>ITEM</i> fornecido.
datetime_timestamp(YEAR, MONTH, DAY, HOUR, MINUTE, SECOND)	Registro de data e hora	Retorna o valor de registro de data e hora para o <i>YEAR</i> , <i>MONTH</i> , <i>DAY</i> , <i>HOUR</i> , <i>MINUTE</i> e <i>SECOND</i> fornecidos.
datetime_timestamp(DATE, TIME)	Registro de data e hora	Retorna o valor de registro de data e hora para o <i>DATE</i> e <i>TIME</i> fornecidos.
datetime_timestamp (NUMBER)	Registro de data e hora	Retorna o valor de registro de data e hora do número de segundos fornecido.
datetime_weekday(DATE)	Número inteiro	Retorna o dia da semana de <i>DATE</i> ou registro de data e hora fornecidos.
datetime_year(DATE)	Número inteiro	Retorna o ano de um <i>DATE</i> ou registro de data e hora fornecidos. O resultado é um número inteiro como 2002.
date_weeks_difference (DATE1, DATE2)	Real	Retorna o tempo em semanas da data ou registro de data e hora representados por <i>DATE1</i> até os representados por <i>DATE2</i> como um número real. Isso é baseado em uma semana de 7 dias. Se <i>DATE2</i> for anterior a <i>DATE1</i> , essa função retornará um número negativo.
date_years_difference (DATE1, DATE2)	Real	Retorna o tempo, em anos, a partir da data ou do registro de data e hora representados por <i>DATE1</i> até os representados por <i>DATE2</i> como um número real. Essa é uma figura aproximada com base em um ano de 365,25 dias. Se <i>DATE2</i> for anterior a <i>DATE1</i> , essa função retornará um número negativo.
date_from_ywd(YEAR, WEEK, DAY)	Número inteiro	Converte o ano, a semana em ano e o dia em semana, para uma data usando o padrão ISO 8601.
date_iso_day(DATE)	Número inteiro	Retorna o dia da semana a partir da data usando o padrão ISO 8601.
date_iso_week(DATE)	Número inteiro	Retorna a semana do ano a partir da data usando o padrão ISO 8601.
date_iso_year(DATE)	Número inteiro	Retorna o ano da data usando o padrão ISO 8601.
time_before(TIME1, TIME2)	Booleano	Retorna um valor de true se <i>TIME1</i> representar uma hora ou registro de data e hora anteriores aos representados por <i>TIME2</i> . Caso contrário, esta função retornará um valor 0.

Tabela 32. Funções de data e hora do CLEM (continuação).

Função	Resultado	Descrição
time_hours_difference (TIME1, TIME2)	Real	Retorna a diferença de tempo em horas entre as horas ou registros de data e hora representados por <i>TIME1</i> e <i>TIME2</i> como um número real. Se você selecionar Dias/Minutos de Rolagem na caixa de diálogo de propriedades do fluxo, um valor superior de <i>TIME1</i> será usado para fazer referência ao dia anterior. Se você não selecionar a opção de rolagem, um valor superior de <i>TIME1</i> fará com que o valor retornado seja negativo.
time_in_hours (TIME)	Real	Retorna o tempo em horas representado por <i>TIME</i> como um número real. Por exemplo, sob o formato de hora HHMM, a expressão <code>time_in_hours('0130')</code> é avaliada para 1,5. <i>TIME</i> pode representar uma hora ou registro de data e hora.
time_in_mins (TIME)	Real	Retorna o tempo em minutos representado por <i>TIME</i> como um número real. <i>TIME</i> pode representar uma hora ou registro de data e hora.
time_in_secs (TIME)	Número inteiro	Retorna o tempo em segundos representado por <i>TIME</i> como um número inteiro. <i>TIME</i> pode representar uma hora ou registro de data e hora.
time_mins_difference (TIME1, TIME2)	Real	Retorna a diferença de tempo em minutos entre as horas ou registros de data e hora representados por <i>TIME1</i> e <i>TIME2</i> como um número real. Se você selecionar Dias/Minutos de Rolagem na caixa de diálogo de propriedades do fluxo, um valor superior de <i>TIME1</i> será usado para fazer referência ao dia anterior (ou à hora anterior, caso apenas minutos e segundos estejam especificados no formato atual). Se você não selecionar a opção de rolagem, um valor superior de <i>TIME1</i> fará com que o valor retornado seja negativo.
time_secs_difference (TIME1, TIME2)	Número inteiro	Retorna a diferença de tempo em segundos entre as horas ou registros de data e hora representados por <i>TIME1</i> e <i>TIME2</i> como um número inteiro. Se você selecionar Dias/Minutos de Rolagem na caixa de diálogo de propriedades do fluxo, um valor superior de <i>TIME1</i> será usado para fazer referência ao dia anterior (ou à hora anterior, caso apenas minutos e segundos estejam especificados no formato atual). Se você não selecionar a opção de rolagem, um valor superior de <i>TIME1</i> fará com que o valor retornado seja negativo.

Convertendo valores de data e hora

Observe que as funções de conversão (e quaisquer outras funções que requerem um tipo específico de entrada, como um valor de data e hora) dependem dos atuais formatos especificados na caixa de diálogo Opções de Fluxo. Por exemplo, se você tivesse um campo denominado *DATE* que é armazenado como uma sequência de caracteres com valores *Jan 2003*, *Fev 2003*, e assim por diante, você poderia convertê-lo em um armazenamento de datas da seguinte forma:

```
to_date (DATE)
```

Para essa conversão funcionar, selecione o formato de data correspondente **MÊS AAAA** como formato de data padrão para o fluxo.

Para ver um exemplo que converta valores da sequência de caracteres usando um nó de Preenchimento, consulte o fluxo *broadband_create_models.str*, instalado na pasta *\Demos* sob a subpasta *streams*.

Datas armazenadas como números. Observe que *DATE* no exemplo anterior é o nome de um campo, enquanto *to_date* é uma função do CLEM. Se você tiver datas armazenadas como números, é possível convertê-las usando a função *datetime_date*, em que o número é interpretado como um número de segundos desde a data (ou época) de base.

```
datetime_date(DATE)
```

Ao converter uma data em um número de segundos (e voltar), é possível executar cálculos como data atual mais ou menos um número fixo de dias, por exemplo:

```
datetime_date((date_in_days(DATE)-7)*60*60*24)
```

Funções de sequência

Para algumas operações, a sequência de eventos é importante. O aplicativo permite trabalhar com as sequências de registro a seguir:

- Sequências e séries temporais
- Funções de sequência
- Indexação de registro
- Valores de média, soma e comparação
- Mudança--diferenciação de monitoramento
- @SINCE
- Valores de offset
- Recursos de sequência adicionais

Para muitos aplicativos, cada registro passando por um fluxo pode ser considerado um caso individual, independente de todos os outros. Nessas situações, a ordem dos registros geralmente é irrelevante.

Para algumas classes de problemas, no entanto, a sequência de registro é muito importante. Normalmente são situações de séries temporais nas quais a sequência de registros representa uma sequência ordenada de eventos ou ocorrências. Cada registro representa uma captura instantânea em um determinado instante no tempo; muitas das informações mais ricas, porém, podem estar contidas não nos valores instantâneos, mas na maneira como esses valores estão mudando e se comportando ao longo do tempo.

É claro, o parâmetro relevante pode ser qualquer coisa além de tempo. Por exemplo, os registros poderiam representar análises executadas a distâncias ao longo de uma linha, mas os mesmos princípios se aplicariam.

Funções especiais e de sequência são imediatamente reconhecíveis pelas seguintes características:

- Todas elas são prefixadas por @.
- Seus nomes estão em letras maiúsculas.

Funções de sequência podem se referir ao registro atualmente sendo processado por um nó, aos registros que já passaram por um nó e, até, em alguns casos, aos registros que ainda têm que passar por um nó. Funções de sequência podem ser combinadas livremente com outros componentes de expressões do CLEM, embora algumas tenham restrições quanto ao que pode ser usado como seus argumentos.

Exemplos

Talvez você ache útil saber quanto tempo passou desde que ocorreu um certo evento ou que uma condição era verdadeira. Use a função @SINCE para fazer isto--por exemplo:

```
@SINCE(Income > Outgoings)
```

Essa função retorna o offset do último registro no qual essa condição era verdadeira--ou seja, o número de registros antes desse no qual a condição era verdadeira. Se a condição nunca foi verdadeira, @SINCE retornará @INDEX + 1.

Às vezes você pode querer se referir a um valor do registro atual na expressão usada por @SINCE. Isso pode ser feito usando a função @THIS, que especifica que um nome de campo sempre se aplica ao registro atual. Para localizar o offset do último registro que tinha um valor de campo Concentration duas vezes maior que o do registro atual, você poderia usar:

```
@SINCE(Concentration > 2 * @THIS(Concentration))
```

Em alguns casos, é verdadeira a condição dada a @SINCE do registro atual por definição—por exemplo:

```
@SINCE(ID == @THIS(ID))
```

Por esse motivo, @SINCE não avalia sua condição para o registro atual. Use uma função semelhante, @SINCE0, se desejar avaliar a condição para o registro atual, bem como anteriores; se a condição for verdadeira no registro atual, @SINCE0 retornará 0.

Tabela 33. Funções de sequência do CLEM

Função	Resultado	Descrição
MEAN(FIELD)	Real	Retorna a média dos valores para <i>FIELD</i> ou <i>FIELDS</i> especificados.
@MEAN(FIELD, EXPR)	Real	Retorna a média de valores para <i>FIELD</i> sobre os últimos registros <i>EXPR</i> recebidos pelo nó atual, incluindo o registro atual. <i>FIELD</i> deve ser o nome de um campo numérico. <i>EXPR</i> pode ser qualquer expressão avaliada para um número inteiro maior que 0. Se <i>EXPR</i> for omitido ou exceder o número de registros recebidos até o momento, a média sobre todos os registros recebidos até o momento será retornada.
@MEAN(FIELD, EXPR, INT)	Real	Retorna a média de valores para <i>FIELD</i> sobre os últimos registros <i>EXPR</i> recebidos pelo nó atual, incluindo o registro atual. <i>FIELD</i> deve ser o nome de um campo numérico. <i>EXPR</i> pode ser qualquer expressão avaliada para um número inteiro maior que 0. Se <i>EXPR</i> for omitido ou exceder o número de registros recebidos até o momento, a média sobre todos os registros recebidos até o momento será retornada. <i>INT</i> especifica o número máximo de valores a reconsiderar. Isso é muito mais eficiente que usar apenas dois argumentos.
@DIFF1(FIELD)	Real	Retorna o primeiro diferenciado de <i>FIELD</i> . O formulário de argumento único simplesmente retorna a diferença entre o valor atual e o valor anterior do campo. Retorna \$null\$ se registros relevantes anteriores não existirem.
@DIFF1(FIELD1, FIELD2)	Real	O formulário de dois argumentos fornece o primeiro diferenciado de <i>FIELD1</i> com relação ao <i>FIELD2</i> . Retorna \$null\$ se registros relevantes anteriores não existirem. É calculado como @DIFF1(FIELD1)/@DIFF1(FIELD2).
@DIFF2(FIELD)	Real	Retorna o segundo diferenciado de <i>FIELD</i> . O formulário de argumento único simplesmente retorna a diferença entre o valor atual e o valor anterior do campo. Retorna \$null\$ se registros relevantes anteriores não existirem. @DIFF2 é calculado como @DIFF(@DIFF(FIELD)).
@DIFF2(FIELD1, FIELD2)	Real	O formulário de dois argumentos fornece o segundo diferenciado de <i>FIELD1</i> com relação ao <i>FIELD2</i> . Retorna \$null\$ se registros relevantes anteriores não existirem. Este é um cálculo complexo -- @DIFF1(FIELD1)/@DIFF1(FIELD2) - @OFFSET(@DIFF1(FIELD1),1)/@OFFSET(@DIFF1(FIELD2)) / @DIFF1(FIELD2).

Tabela 33. Funções de sequência do CLEM (continuação)

Função	Resultado	Descrição
@INDEX	Número inteiro	Retorna o índice do registro atual. Os índices são alocados a registros conforme chegam ao nó atual. Ao primeiro registro é dado o índice 1, e o índice é incrementado em 1 para cada registro subsequente.
@LAST_NON_BLANK(FIELD)	Qualquer	Retorna o último valor para <i>FIELD</i> que não estava em branco, conforme definido na origem de envio de dados ou nó Tipo. Se não houver valores não em branco para <i>FIELD</i> nos registros lidos até o momento, \$null\$ será retornado. Observe que os valores não em branco, também chamados valores omissos de usuário, podem ser definidos separadamente para cada campo.
@MAX(FIELD)	Número	Retorna o valor máximo para o <i>FIELD</i> especificado.
@MAX(FIELD, EXPR)	Número	Retorna o valor máximo para <i>FIELD</i> sobre os últimos registros <i>EXPR</i> recebidos até o momento, incluindo o registro atual. <i>FIELD</i> deve ser o nome de um campo numérico. <i>EXPR</i> pode ser qualquer expressão avaliada para um número inteiro maior que 0.
@MAX(FIELD, EXPR, INT)	Número	Retorna o valor máximo para <i>FIELD</i> sobre os últimos registros <i>EXPR</i> recebidos até o momento, incluindo o registro atual. <i>FIELD</i> deve ser o nome de um campo numérico. <i>EXPR</i> pode ser qualquer expressão avaliada para um número inteiro maior que 0. Se <i>EXPR</i> for omitido ou exceder o número de registros recebidos até o momento, o valor máximo sobre todos os registros recebidos até o momento será retornado. <i>INT</i> especifica o número máximo de valores a reconsiderar. Isso é muito mais eficiente que usar apenas dois argumentos.
@MIN(FIELD)	Número	Retorna o valor mínimo para o <i>FIELD</i> especificado.
@MIN(FIELD, EXPR)	Número	Retorna o valor mínimo para <i>FIELD</i> sobre os últimos registros <i>EXPR</i> recebidos até o momento, incluindo o registro atual. <i>FIELD</i> deve ser o nome de um campo numérico. <i>EXPR</i> pode ser qualquer expressão avaliada para um número inteiro maior que 0.
@MIN(FIELD, EXPR, INT)	Número	Retorna o valor mínimo para <i>FIELD</i> sobre os últimos registros <i>EXPR</i> recebidos até o momento, incluindo o registro atual. <i>FIELD</i> deve ser o nome de um campo numérico. <i>EXPR</i> pode ser qualquer expressão avaliada para um número inteiro maior que 0. Se <i>EXPR</i> for omitido ou exceder o número de registros recebidos até o momento, o valor mínimo sobre todos os registros recebidos até o momento será retornado. <i>INT</i> especifica o número máximo de valores a reconsiderar. Isso é muito mais eficiente que usar apenas dois argumentos.

Tabela 33. Funções de sequência do CLEM (continuação)

Função	Resultado	Descrição
@OFFSET(FIELD, EXPR)	Qualquer	<p>Retorna o valor de <i>FIELD</i> no offset de registro do registro atual pelo valor de <i>EXPR</i>. Um offset positivo se refere a um registro que já passou (um "lookback"), enquanto um negativo especifica um "lookahead" para um registro que ainda vai chegar. Por exemplo, @OFFSET(Status, 1) retorna o valor do campo Status no registro anterior, enquanto @OFFSET(Status, -4) representa um "lookahead" de quatro registros na sequência (ou seja, para registros que ainda não passaram por esse nó) para obter o valor. <i>Observe que um offset negativo (lookahead) deve ser especificado como uma constante.</i> Somente para offsets positivos, <i>EXPR</i> também pode ser uma expressão do CLEM arbitrária avaliada para o registro atual para fornecer o offset. Nesse caso, a versão de três argumentos dessa função deve melhorar o desempenho (consulte a próxima função). Se a expressão retornar qualquer coisa que não seja um número inteiro não negativo, isso causará um erro—ou seja, não é legal ter offsets lookahead calculados.</p> <p><i>Nota:</i> Uma função @OFFSET de autorreferência não pode usar lookahead literal. Por exemplo, em um nó de Preenchimento, não é possível substituir o valor de field1 usando uma expressão como @OFFSET(field1,-2).</p> <p><i>Nota:</i> No nó de Preenchimento, durante o preenchimento de um campo, há dois valores efetivamente diferentes desse campo, a saber, o valor pré-preenchido e o valor pós-preenchido. Quando @OFFSET se refere a si mesmo, ele se refere ao valor pré-preenchido. Esse valor pós-preenchido existe somente para linhas antigas, portanto, o @OFFSET de autorreferência só pode se referir a linhas antigas. Como o @OFFSET de autorreferência não pode se referir ao futuro, ele executa as seguintes verificações de offset:</p> <ul style="list-style-type: none"> • Se o offset for literal, e no futuro, um erro será relatado antes do início da execução. • Se o offset for uma expressão e for avaliado para o futuro no tempo de execução, @OFFSET retornará \$null\$. <p><i>Nota:</i> O uso de "lookahead" e "lookback" dentro de um nó não é suportado.</p>
@OFFSET(FIELD, EXPR, INT)	Qualquer	<p>Executa a mesma operação que a função @OFFSET com a adição de um terceiro argumento, <i>INT</i>, o que especifica o número máximo de valor para lookback. Nos casos em que o offset é calculado a partir de uma expressão, o terceiro argumento deve melhorar o desempenho.</p> <p>Por exemplo, em uma expressão como @OFFSET(Foo, Month, 12), o sistema sabe manter somente os últimos doze valores de Foo; caso contrário, ele terá que armazenar cada valor se for o caso. Nos casos em que o valor do offset for uma constante—incluindo offsets "lookahead" negativos, que devem ser constantes—o terceiro argumento será sem pontos e a versão de dois argumentos dessa função deverá ser usada. Consulte também a nota sobre funções de autorreferência na versão de dois argumentos descrita anteriormente.</p> <p><i>Nota:</i> O uso de "lookahead" e "lookback" dentro de um nó não é suportado.</p>

Tabela 33. Funções de sequência do CLEM (continuação)

Função	Resultado	Descrição
@SDEV(FIELD)	Real	Retorna o desvio padrão dos valores para <i>FIELD</i> ou <i>FIELDS</i> especificados.
@SDEV(FIELD, EXPR)	Real	Retorna o desvio padrão de valores para <i>FIELD</i> sobre os últimos registros <i>EXPR</i> recebidos pelo nó atual, incluindo o registro atual. <i>FIELD</i> deve ser o nome de um campo numérico. <i>EXPR</i> pode ser qualquer expressão avaliada para um número inteiro maior que 0. Se <i>EXPR</i> for omitido ou exceder o número de registros recebidos até o momento, o desvio padrão sobre todos os registros recebidos até o momento será retornado.
@SDEV(FIELD, EXPR, INT)	Real	Retorna o desvio padrão de valores para <i>FIELD</i> sobre os últimos registros <i>EXPR</i> recebidos pelo nó atual, incluindo o registro atual. <i>FIELD</i> deve ser o nome de um campo numérico. <i>EXPR</i> pode ser qualquer expressão avaliada para um número inteiro maior que 0. Se <i>EXPR</i> for omitido ou exceder o número de registros recebidos até o momento, o desvio padrão sobre todos os registros recebidos até o momento será retornado. <i>INT</i> especifica o número máximo de valores a reconsiderar. Isso é muito mais eficiente que usar apenas dois argumentos.
@SINCE(EXPR)	Qualquer	Retorna o número de registros que passaram desde que <i>EXPR</i> , uma expressão do CLEM arbitrária, era verdadeira.
@SINCE(EXPR, INT)	Qualquer	A inclusão do segundo argumento, <i>INT</i> , especifica o número máximo de registros para lookback. Se <i>EXPR</i> nunca foi verdadeira, <i>INT</i> será @INDEX+1.
@SINCE0(EXPR)	Qualquer	Considera o registro atual, enquanto @SINCE não; @SINCE0 retornará 0 se <i>EXPR</i> for verdadeira para o registro atual.
@SINCE0(EXPR, INT)	Qualquer	A inclusão do segundo argumento, <i>INT</i> , especifica o número máximo de registros para o lookback.
@SUM(FIELD)	Número	Retorna a soma de valores para o <i>FIELD</i> ou <i>FIELDS</i> especificados.
@SUM(FIELD, EXPR)	Número	Retorna a soma de valores para <i>FIELD</i> sobre os últimos registros <i>EXPR</i> recebidos pelo nó atual, incluindo o registro atual. <i>FIELD</i> deve ser o nome de um campo numérico. <i>EXPR</i> pode ser qualquer expressão avaliada para um número inteiro maior que 0. Se <i>EXPR</i> for omitido ou exceder o número de registros recebidos até o momento, a soma sobre todos os registros recebidos até o momento será retornada.
@SUM(FIELD, EXPR, INT)	Número	Retorna a soma de valores para <i>FIELD</i> sobre os últimos registros <i>EXPR</i> recebidos pelo nó atual, incluindo o registro atual. <i>FIELD</i> deve ser o nome de um campo numérico. <i>EXPR</i> pode ser qualquer expressão avaliada para um número inteiro maior que 0. Se <i>EXPR</i> for omitido ou exceder o número de registros recebidos até o momento, a soma sobre todos os registros recebidos até o momento será retornada. <i>INT</i> especifica o número máximo de valores a reconsiderar. Isso é muito mais eficiente que usar apenas dois argumentos.
@THIS(FIELD)	Qualquer	Retorna o valor do campo denominado <i>FIELD</i> no registro atual. Usado somente em expressões @SINCE.

Funções globais

As funções @MEAN, @SUM, @MIN, @MAX e @SDEV trabalham, no máximo, em todos os registros lidos até, e incluindo o, atual. Em alguns casos, no entanto, é útil poder trabalhar para descobrir como os valores no registro atual se comparam com os valores vistos no conjunto de dados inteiro. Usando um nó Configurar Globais para gerar valores no conjunto de dados inteiro, é possível acessar esses valores em uma expressão do CLEM usando as funções globais.

Por exemplo,

```
@GLOBAL_MAX(Age)
```

retorna o valor mais alto de Age no conjunto de dados, enquanto a expressão

```
(Value - @GLOBAL_MEAN(Value)) / @GLOBAL_SDEV(Value)
```

expressa a diferença entre o Value desse registro e a média global como um número de desvios padrão. É possível usar valores globais somente após eles terem sido calculados por um nó Configurar Globais. Todos os valores globais atuais podem ser cancelados clicando no botão **Limpar Valores Globais** na guia Globais na caixa de diálogo de propriedades do fluxo.

Tabela 34. Funções globais do CLEM.

Função	Resultado	Descrição
@GLOBAL_MAX(FIELD)	Número	Retorna o valor máximo para <i>FIELD</i> sobre o conjunto de dados inteiro conforme gerado anteriormente por um nó Configurar Globais. <i>FIELD</i> deve ser o nome de um campo numérico, de data / hora / data/hora ou de sequência de caracteres. Se o valor global correspondente não tiver sido configurado, ocorrerá um erro.
@GLOBAL_MIN(FIELD)	Número	Retorna o valor mínimo para <i>FIELD</i> sobre o conjunto de dados inteiro conforme gerado anteriormente por um nó Configurar Globais. <i>FIELD</i> deve ser o nome de um campo numérico, de data / hora / data/hora ou de sequência de caracteres. Se o valor global correspondente não tiver sido configurado, ocorrerá um erro.
@GLOBAL_SDEV(FIELD)	Número	Retorna o desvio padrão de valores para <i>FIELD</i> sobre o conjunto de dados inteiro conforme gerado anteriormente por um nó Configurar Globais. <i>FIELD</i> deve ser o nome de um campo numérico. Se o valor global correspondente não tiver sido configurado, ocorrerá um erro.
@GLOBAL_MEAN(FIELD)	Número	Retorna a média de valores para <i>FIELD</i> sobre o conjunto de dados inteiro conforme gerado anteriormente por um nó Configurar Globais. <i>FIELD</i> deve ser o nome de um campo numérico. Se o valor global correspondente não tiver sido configurado, ocorrerá um erro.
@GLOBAL_SUM(FIELD)	Número	Retorna a soma de valores para <i>FIELD</i> sobre o conjunto de dados inteiro conforme gerado anteriormente por um nó Configurar Globais. <i>FIELD</i> deve ser o nome de um campo numérico. Se o valor global correspondente não tiver sido configurado, ocorrerá um erro.

Funções manipulando valores em branco e nulos

Usando CLEM, é possível especificar que certos valores em um campo devem ser considerados valores "em branco" ou omissos. As funções a seguir trabalham com os em branco.

Tabela 35. Funções de valor nulo e em branco do CLEM.

Função	Resultado	Descrição
@BLANK(FIELD)	Booleano	Retorna verdadeiro para todos os registros cujos valores estão em branco de acordo com o conjunto de regras de manipulação de brancos em um nó Tipo de envio de dados ou nó de origem (guia Tipos).
@LAST_NON_BLANK(FIELD)	Qualquer um	Retorna o último valor para <i>FIELD</i> que não estava em branco, conforme definido na origem de envio de dados ou nó Tipo. Se não houver valores não em branco para <i>FIELD</i> nos registros lidos até o momento, \$null\$ será retornado. Observe que os valores não em branco, também chamados valores omissos de usuário, podem ser definidos separadamente para cada campo.
@NULL(FIELD)	Booleano	Retorna verdadeiro se o valor de <i>FIELD</i> for \$null\$ omissos do sistema. Retorna falso para todos os outros valores, incluindo os em branco definidos pelo usuário. Se desejar verificar ambos, use @BLANK(FIELD) e @NULL(FIELD).
undef	Qualquer um	Usado geralmente no CLEM para inserir um valor \$null\$ — por exemplo, para preencher valores em branco com nulos no nó de Preenchimento.

Campos em branco devem ser "preenchidos" com o nó de Preenchimento. Nos nós de Preenchimento e Derivação (somente modo múltiplo), a função especial @FIELD do CLEM se refere ao(s) atual(is) campo(s) sendo examinado(s).

Campos Especiais

Funções especiais são usadas para denotar os campos específicos sendo examinados ou para gerar uma lista de campos como entrada. Por exemplo, ao derivar vários campos de uma vez, é necessário usar @FIELD para denotar "executar esta ação de derivação nos campos selecionados". O uso da expressão log(@FIELD) deriva um novo campo de log para cada campo selecionado.

Tabela 36. Campos especiais do CLEM.

Função	Resultado	Descrição
@FIELD	Qualquer um	Executa uma ação em todos os campos especificados no contexto de expressão.
@TARGET	Qualquer um	Quando uma expressão do CLEM é usada em uma função de análise definida pelo usuário, @TARGET representa o campo de destino ou o "valor correto" para o par de destino/previsto sendo analisado. Geralmente essa função é usada em um nó Análise.
@PREDICTED	Qualquer um	Quando uma expressão do CLEM é usada em uma função de análise definida pelo usuário, @PREDICTED representa o valor previsto para o par de destino/previsto sendo analisado. Geralmente essa função é usada em um nó Análise.
@PARTITION_FIELD	Qualquer um	Substitui o nome do atual campo de partição.

Tabela 36. Campos especiais do CLEM (continuação).

Função	Resultado	Descrição
@TRAINING_PARTITION	Qualquer um	Retorna o valor da atual partição de treinamento. Por exemplo, para selecionar registros de treinamento usando um nó de Seleção, use a expressão do CLEM: @PARTITION_FIELD = @TRAINING_PARTITION. Isso assegura que o nó de Seleção trabalhe sempre independentemente dos valores usados para representar cada partição nos dados.
@TESTING_PARTITION	Qualquer um	Retorna o o valor da atual partição de teste.
@VALIDATION_PARTITION	Qualquer um	Retorna o valor da atual partição de validação.
@FIELDS_BETWEEN(start, end)	Qualquer um	Retorna a lista de nomes de campo entre os campos de início e encerramento especificados (inclusivos) com base na ordem natural (ou seja, inserção) dos campos nos dados.
@FIELDS_MATCHING(pattern)	Qualquer um	Retorna uma lista de nomes de campos correspondentes a um padrão especificado. Um ponto de interrogação (?) pode ser incluído no padrão para corresponder a exatamente um caractere; um asterisco (*) corresponde a zero ou mais caracteres. Para corresponder a um ponto de interrogação literal ou asterisco (em vez de usá-los como curinga), uma barra invertida (\) pode ser usada como um caractere de escape.
@MULTI_RESPONSE_SET	Qualquer um	Retorna a lista de campos no conjunto de múltiplas respostas.

Capítulo 10. Usando IBM SPSS Modeler com um repositório

Sobre o IBM SPSS Collaboration and Deployment Services Repository

O IBM SPSS Modeler pode ser usado em conjunção com um repositório do IBM SPSS Collaboration and Deployment Services, permitindo que você gerencie o ciclo de vida de modelos de mineração de dados e objetos preditivos relacionados, e permitindo que esses objetos sejam usados por aplicativos corporativos, ferramentas e soluções. Os objetos do IBM SPSS Modeler que podem ser compartilhados nesta maneira incluem fluxos, nós, saídas de fluxos, cenários, projetos e modelos. Os objetos são armazenados no repositório central, de onde eles podem ser compartilhados com outros aplicativos e controlados usando versão estendida, metadados e recursos de procura.

Antes que possa usar o IBM SPSS Modeler com o repositório, é necessário instalar um adaptador no host do repositório. Sem esse adaptador, você poderá ver a seguinte mensagem ao tentar acessar os objetos do repositório a partir de determinados nós ou modelos do IBM SPSS Modeler:

O repositório poderá precisar ser atualizado para suportar os novos tipos de nó, modelo e saída.

Para obter instruções sobre como instalar o adaptador, consulte o guia *Instalação de Implementação do IBM SPSS Modeler*, disponível na DVD do IBM SPSS Modeler Deployment. Os detalhes de como acessar os objetos de repositório do IBM SPSS Modeler a partir do IBM SPSS Collaboration and Deployment Services Deployment Manager são fornecidos no *Guia de Implementação do IBM SPSS Modeler*.

As seções a seguir fornecem informações sobre como acessar o repositório de dentro do IBM SPSS Modeler.

Versão Extensiva e Suporte de Procura

O repositório fornece a versão do objeto e os recursos de procura abrangentes. Por exemplo, suponha que você crie um fluxo e armazene-o no repositório no qual ele pode ser compartilhado com pesquisadores de outras divisões. Se você atualizar o fluxo posteriormente no IBM SPSS Modeler, poderá incluir a versão atualizada no repositório sem sobrescrever a versão anterior. Todas as versões permanecem acessíveis e podem ser procuradas pelo nome, rótulo, campos usados ou outros atributos. Você pode, por exemplo, procurar por todas as versões do modelo, que usam receita líquida como uma entrada ou todos os modelos criados por um autor específico. (Para fazer isso com um sistema de arquivos tradicional, você teria que salvar cada versão com um nome de arquivo diferente, e os relacionamentos entre as versões seriam desconhecidos para o software).

Conexão Única

A variável de conexão única permite que os usuários se conectem ao repositório sem precisar inserir detalhes do nome de usuário e senha todas as vezes. Os detalhes de login da rede local existente do usuário fornecem a autenticação necessária para IBM SPSS Collaboration and Deployment Services. Essa variável depende do seguinte:

- O IBM SPSS Collaboration and Deployment Services deve ser configurado para usar um provedor de conexão única.
- O usuário deve estar com login efetuado em um host que seja compatível com o provedor.

Consulte o tópico “Conectando ao repositório” na página 155 para obter mais informações.

Armazenando e implementando objetos do repositório

Fluxos criados no IBM SPSS Modeler podem ser **armazenados** no repositório exatamente como eles são, como arquivos com a extensão *.str*. Dessa maneira, um único fluxo pode ser acessado por diversos usuários na empresa. Consulte o tópico “Armazenando objetos no repositório” na página 156 para obter mais informações.

Também é possível implementar um fluxo no repositório. Um fluxo implementado é armazenado como um arquivo com metadados adicionais. Um fluxo implementado pode tirar total proveito das variáveis de nível corporativo do IBM SPSS Collaboration and Deployment Services, como escoragem automatizada e atualização de modelo. Por exemplo, um modelo pode ser automaticamente atualizado em intervalos planejados regularmente conforme novos dados se tornam disponíveis. Alternativamente, um conjunto de fluxos pode ser implementado para análise do Champion Challenger, na qual os fluxos são comparados para determinar quais contêm o modelo preditivo mais efetivo.

Nota: Os nós de modelagem Regras de Associação, STP e TCM não suportam os passos Avaliação de Modelo ou Champion Challenger no IBM SPSS Collaboration and Deployment Services.

É possível implementar um fluxo de uma de duas formas: como um fluxo (com a extensão *.str*) ou como um cenário (com a extensão *.scn*). A implementação como um fluxo permite que o fluxo seja usado pelo aplicativo thin client IBM SPSS Modeler Advantage. Consulte o tópico “Abrindo um fluxo no IBM SPSS Modeler Advantage” na página 173 para obter mais informações. A implementação como um cenário permite que o fluxo seja usado pelo Predictive Applications versão 5, o predecessor do IBM SPSS Modeler Advantage.

Para obter informações adicionais, consulte “Opções de Implementação do Fluxo” na página 166.

Requisitos para fluxos implementados como cenários

- Para assegurar acesso consistente aos dados corporativos, fluxos implementados como cenários devem ser acessados por meio do componente Visualização Corporativa do IBM SPSS Collaboration and Deployment Services. Isso significa que no IBM SPSS Modeler deve haver pelo menos um nó de origem da Visualização Corporativa dentro de cada ramificação de escoragem ou modelagem designada no fluxo.
- Para usar o nó da Visualização Corporativa, o IBM SPSS Collaboration and Deployment Services deve ser instalado, configurado e acessível a partir de seu site, com uma Visualização Corporativa, Visualizações de Aplicativo e Definições de Provedor de Dados (DPDs) já definidas. Para obter mais informações, entre em contato com o administrador local ou consulte o website corporativo em .
- Uma DPD é definida com relação a uma determinada origem de dados ODBC. Para usar uma DPD a partir do IBM SPSS Modeler, deve-se ter uma origem de dados ODBC definida no host do servidor IBM SPSS Modeler que tenha o mesmo nome e que se conecte ao mesmo armazenamento de dados que o referenciado na DPD.
- Além disso, o IBM SPSS Collaboration and Deployment Services Enterprise View Driver deve ser instalado em cada computador usado para modificar ou executar o fluxo. Para Windows, simplesmente instale o driver no computador onde o SPSS Modeler ou SPSS Modeler Server está instalado e nenhuma configuração adicional do driver será necessária. No UNIX, uma referência ao script *pev.sh* deve ser incluída no script de inicialização. Entre em contato com o administrador local para obter detalhes sobre como instalar o IBM SPSS Collaboration and Deployment Services Enterprise View Driver.

Outras opções de implementação

Enquanto o IBM SPSS Collaboration and Deployment Services oferece as variáveis mais extensivas para gerenciar o conteúdo corporativo, inúmeros outros mecanismos para implementar ou exportar fluxos também estão disponíveis, incluindo:

- Exportar o fluxo e o modelo para uso posterior com o IBM SPSS Modeler Solution Publisher Runtime.

- Exportar um ou mais modelos em PMML, um formato baseado em XML para codificar informações de modelo. Consulte o tópico “Importando e exportando modelos como PMML” na página 174 para obter mais informações.

Conectando ao repositório

1. Para se conectar a um repositório, no menu principal do IBM SPSS Modeler, clique em:
Ferramentas > Repositório > Opções...
2. No campo **RepositoryURL**, insira ou selecione o caminho do diretório para, ou URL da, instalação do repositório que deseja acessar. Só é possível se conectar a um repositório por vez.
As configurações são específicas de cada site ou instalação. Para detalhes de login específicos, entre em contato com o administrador do sistema local.

Configurar Credenciais. Deixe essa caixa desmarcada para ativar a variável **conexão única**, que tenta efetuar seu login usando detalhes de nome de usuário e senha do computador local. Se uma conexão única não for possível ou se você marcar essa caixa para desativar a conexão única (por exemplo, para efetuar login em uma conta do administrador), uma tela adicional será exibida para você inserir suas credenciais.

Inserindo credenciais para o repositório

Dependendo de suas configurações, os campos a seguir podem ser necessários na caixa de diálogo Repositório: Credenciais:

ID do usuário e senha. Especifique um nome de usuário válido e uma senha para efetuar login. Se for necessário, entre em contato com o administrador local para obter mais informações.

Provedor. Escolha um provedor de segurança para autenticação. O repositório pode ser configurado para usar diferentes provedores de segurança; se for necessário, entre em contato com seu administrador local para obter mais informações.

Lembrar repositório e ID do usuário. Salva as atuais configurações como o padrão para que não seja preciso reinseri-las cada vez que você desejar se conectar.

Navegação pelas credenciais do repositório

Quando você se conecta a um repositório a partir de um nó de origem do Analytic Server, Cognos, ODBC ou TM1, é possível selecionar credenciais registradas anteriormente para se conectar a um repositório. Essas credenciais são listadas na caixa de diálogo Selecionar Credencial do Repositório. Para marcar essa caixa de diálogo, clique em Navegar próximo do campo **Credencial**.

Na caixa de diálogo Selecionar Credencial do Repositório, destaque as credenciais na lista fornecida e clique em OK. Se a lista for muito grande, use o campo **Filtro** para inserir um nome, ou parte dele, para localizar as credenciais necessárias.

Navegando no conteúdo do repositório

O repositório permite navegar no conteúdo armazenado de uma maneira semelhante ao Windows Explorer; também é possível navegar em *versões* de cada objeto armazenado.

1. Para abrir a janela do IBM SPSS Collaboration and Deployment Services Repository, nos menus do SPSS Modeler, clique em:

Ferramentas > Repositório > Explorar...

1. Especifique configurações de conexão para o repositório, caso seja necessário. Consulte o tópico “Conectando ao repositório” para obter mais informações. Para porta, senha e outros detalhes de conexão específicos, entre em contato com o administrador do sistema local.

A janela do Explorer exibe inicialmente uma visualização em árvore da hierarquia de pastas. Clique no nome da pasta para exibir seu conteúdo.

Objetos que correspondem aos atuais critérios de procura ou seleção são listados na área de janela direita, com informações detalhadas sobre a versão selecionada exibida na área de janela inferior direita. Os atributos exibidos se aplicam à versão mais recente.

Armazenando objetos no repositório

É possível armazenar fluxos, nós, modelos, paletas de modelos, projetos e objetos de saída no repositório, onde eles podem ser acessados por outros usuários e aplicativos.

Também é possível publicar saída de fluxo no repositório em um formato que permita que outros usuários a visualizem pela Internet usando IBM SPSS Collaboration and Deployment Services Deployment Portal.

Configurando propriedades do objeto

Quando você armazena um objeto, a caixa de diálogo Repositório: Armazenamento é exibida, permitindo configurar os valores de inúmeras propriedades para o objeto. É possível:

- Escolher o nome e a pasta do repositório sob os quais o objeto deve ser armazenado
- Incluir informações sobre o objeto, como rótulo de versão e outras propriedades pesquisáveis
- Designar um ou mais tópicos de classificação ao objeto
- Configurar opções de segurança para o objeto

As seções a seguir descrevem as propriedades que podem ser configuradas.

Escolhendo a localização para armazenar objetos

Na caixa de diálogo Repositório: Armazenamento, insira o seguinte.

Salvar em. Mostra a pasta atual--a localização onde o objeto será armazenado. Dê um clique duplo no nome da pasta na lista para configurar essa pasta como a atual. Use o botão Pasta para Cima para navegar para a pasta pai. Use o botão Nova Pasta para criar uma pasta no nível atual.

Nome do arquivo. O nome sob o qual o objeto será armazenado.

Armazenamento. Armazena o objeto na localização atual.

Incluindo informações sobre objetos armazenados

Todos os campos na guia Informações da caixa de diálogo Repositório: Armazenamento são opcionais.

Autor. O nome de usuário do usuário criando o objeto no repositório. Por padrão, isso mostra o nome de usuário usado para conexão do repositório, mas é possível alterar esse nome aqui.

Rótulo de Versão. Selecione um rótulo da lista para indicar a versão do objeto ou clique em **Incluir** para criar um novo rótulo. Evite usar o caractere "[" no rótulo. Assegure-se de que nenhuma caixa esteja marcada se não desejar designar um rótulo a essa versão do objeto. Consulte o tópico "Visualizando e editando propriedades do objeto" na página 164 para obter mais informações.

Descrição. Uma descrição do objeto. Os usuários podem procurar objetos por descrição (consulte a nota).

Palavras-chave. Uma ou mais palavras-chave relacionadas ao objeto e que podem ser usadas para propósitos de procura (consulte a nota).

Expiração. Uma data após a qual o objeto não estará mais visível para usuários gerais, embora ainda possa ser visto por seu proprietário e pelo administrador do repositório. Para configurar uma data de expiração, selecione a opção **Data** e insira a data ou escolha uma usando o botão de calendário.

Armazenamento. Armazena o objeto na localização atual.

Nota: Informações nos campos **Descrição** e **Palavras-chave** são tratadas como distintas de tudo que é inserido no SPSS Modeler na guia Anotações do objeto. Uma procura no repositório por descrição ou palavra-chave não retorna informações da guia Anotações. Consulte o tópico “Procurando objetos no repositório” na página 161 para obter mais informações.

Designando tópicos a um objeto armazenado

Tópicos são um sistema de classificação hierárquico para o conteúdo armazenado no repositório. É possível escolher entre os tópicos disponíveis durante o armazenamento de objetos, e os usuários também podem procurar objetos por tópico. A lista de tópicos disponíveis é configurada pelos usuários do repositório com os privilégios apropriados (para obter mais informações, consulte *Deployment Manager User's Guide*).

Para designar um tópico ao objeto, na guia Tópicos da caixa de diálogo Repositório: Armazenamento:

1. Clique no botão **Incluir**.
2. Clique em um nome de tópico da lista de tópicos disponíveis.
3. Clique em **OK**.

Para remover uma designação de tópico:

4. Selecione o tópico na lista de tópicos designados.
5. Clique em **Excluir**.

Configurando opções de segurança para objetos armazenados

É possível configurar ou alterar inúmeras opções de segurança para um objeto armazenado na guia Segurança da caixa de diálogo Repositório: Armazenamento. Para um ou mais **principais** (ou seja, usuários ou grupos de usuários), é possível:

- Designar direitos de acesso ao objeto
- Modificar direitos de acesso ao objeto
- Remover direitos de acesso ao objeto

Principal. O nome de usuário do repositório do usuário ou grupo de tem direitos de acesso nesse objeto.

Permissões. Os direitos de acesso que este usuário ou grupo tem para o objeto.

Incluir. Permite incluir um ou mais usuários ou grupos na lista daqueles que têm direitos de acesso neste objeto. Consulte o tópico “Incluindo um usuário na lista de permissões” para obter mais informações.

Modificar. Permite modificar os direitos de acesso do usuário ou grupo selecionado para este objeto. O acesso de leitura é concedido por padrão. Essa opção permite conceder direitos de acesso adicionais, ou seja, Permissões de Proprietário, Gravar, Excluir e Modificar.

Excluir. Exclui o usuário ou grupo selecionado da lista de permissões para este objeto.

Incluindo um usuário na lista de permissões: Os campos a seguir estão disponíveis quando você seleciona **Incluir** na guia Segurança da caixa de diálogo Repositório: Armazenamento:

Selecionar provedor. Escolha um provedor de segurança para autenticação. O repositório pode ser configurado para usar diferentes provedores de segurança; se for necessário, entre em contato com seu administrador local para obter mais informações.

Localizar. Insira o nome de usuário do repositório do usuário ou grupo que deseja incluir e clique em **Procurar** para exibir o nome na lista de usuários. Para incluir mais de um nome de usuário por vez, deixe esse campo em branco e clique em **Procurar** para exibir uma lista de todos os nomes de usuário do repositório.

Lista de usuários. Selecione um ou mais nomes de usuário da lista e clique em OK para incluí-los na lista de permissões.

Modificando direitos de acesso para um objeto: Os campos a seguir estão disponíveis quando você seleciona **Modificar** na guia Segurança da caixa de diálogo Repositório: Armazenamento:

Proprietário. Selecione esta opção para dar a esse usuário ou grupo direitos de acesso de proprietário ao objeto. O proprietário tem controle total sobre o objeto, incluindo os direitos de acesso Excluir e Modificar.

Leitura. Por padrão, um usuário ou grupo que não é proprietário do objeto tem somente direitos de acesso de Leitura ao objeto. Marque as caixas de seleção apropriadas para incluir direitos de acesso de Gravação, Exclusão e Modificação de Permissões para esse usuário ou grupo.

Armazenando fluxos

É possível armazenar um fluxo como um arquivo *.str* no repositório, de onde ele pode ser acessado por outros usuários.

Nota: Para obter informações sobre como implementar um fluxo, para aproveitar as vantagens das variáveis de repositório adicionais, consulte “Implementando fluxos” na página 166.

Para armazenar o fluxo atual:

1. No menu principal, clique em:
Arquivo > Armazenar > Armazenar como Fluxo...
2. Especifique configurações de conexão para o repositório, caso seja necessário. Consulte o tópico “Conectando ao repositório” na página 155 para obter mais informações. Para porta, senha e outros detalhes de conexão específicos, entre em contato com o administrador do sistema local.
3. Na caixa de diálogo Repositório: Armazenamento, escolha a pasta na qual deseja armazenar o objeto, especifique quaisquer outras informações que deseja registrar e clique no botão **Armazenamento**. Consulte o tópico “Configurando propriedades do objeto” na página 156 para obter mais informações.

Armazenando projetos

É possível armazenar um projeto do IBM SPSS Modeler completo como um arquivo *.cpj* no repositório para que ele possa ser acessado por outros usuários.

Como um arquivo de projeto é um contêiner para outros objetos do IBM SPSS Modeler, você precisa dizer ao IBM SPSS Modeler para armazenar os objetos do projeto no repositório. Isso é feito usando uma configuração na caixa de diálogo Propriedades do Projeto. Consulte o tópico “Configurando propriedades do projeto” na página 180 para obter mais informações.

Após configurar um projeto para armazenar objetos no repositório, sempre que você incluir um novo objeto no projeto, o IBM SPSS Modeler automaticamente solicitará que você armazene o objeto.

Após a conclusão da sessão do IBM SPSS Modeler, deve-se armazenar uma nova versão do arquivo de projeto para que ele se lembre de suas adições. O arquivo de projeto contém (e recupera) automaticamente as versões mais recentes de seus objetos. Se você não incluiu nenhum objeto em um projeto durante uma sessão do IBM SPSS Modeler, não será necessário restaurar o arquivo de projeto. No entanto, deve-se armazenar novas versões para os objetos do projeto (fluxos, saída, e assim por diante) que você alterou.

Para armazenar um projeto

1. Selecione o projeto na guia CRISP-DM ou Classes na área de janela de gerenciadores no IBM SPSS Modeler e, no menu principal, clique em:
Arquivo > Projeto > Armazenar Projeto...
2. Especifique configurações de conexão para o repositório, caso seja necessário. Consulte o tópico “Conectando ao repositório” na página 155 para obter mais informações. Para porta, senha e outros detalhes de conexão específicos, entre em contato com o administrador do sistema local.
3. Na caixa de diálogo Repositório: Armazenamento, escolha a pasta na qual deseja armazenar o objeto, especifique quaisquer outras informações que deseja registrar e clique no botão **Armazenamento**. Consulte o tópico “Configurando propriedades do objeto” na página 156 para obter mais informações.

Armazenando nós

É possível armazenar uma definição de nó individual a partir do fluxo atual como um arquivo *.nod* no repositório, de onde ele pode ser acessado por outros usuários.

Para armazenar um nó:

1. Clique com o botão direito na tela de fluxo e clique em **Armazenar Nó**.
2. Especifique configurações de conexão para o repositório, caso seja necessário. Consulte o tópico “Conectando ao repositório” na página 155 para obter mais informações. Para porta, senha e outros detalhes de conexão específicos, entre em contato com o administrador do sistema local.
3. Na caixa de diálogo Repositório: Armazenamento, escolha a pasta na qual deseja armazenar o objeto, especifique quaisquer outras informações que deseja registrar e clique no botão **Armazenamento**. Consulte o tópico “Configurando propriedades do objeto” na página 156 para obter mais informações.

Armazenando objetos de saída

É possível armazenar um objeto de saída a partir do fluxo atual como um arquivo *.cou* no repositório, onde ele pode ser acessado por outros usuários.

Para armazenar um objeto de saída:

1. Clique no objeto na guia Saídas da área de janela de gerenciadores no SPSS Modeler e, no menu principal, clique em:
Arquivo > Saídas > Armazenar Saída...
2. Alternativamente, clique com o botão direito em um objeto na guia Saídas e clique em **Armazenar**.
3. Especifique configurações de conexão para o repositório, caso seja necessário. Consulte o tópico “Conectando ao repositório” na página 155 para obter mais informações. Para porta, senha e outros detalhes de conexão específicos, entre em contato com o administrador do sistema local.
4. Na caixa de diálogo Repositório: Armazenamento, escolha a pasta na qual deseja armazenar o objeto, especifique quaisquer outras informações que deseja registrar e clique no botão **Armazenamento**. Consulte o tópico “Configurando propriedades do objeto” na página 156 para obter mais informações.

Armazenando modelos e paletas de modelos

É possível armazenar um modelo individual como um arquivo *.gm* no repositório, de onde ele pode ser acessado por outros usuários. Também é possível armazenar o conteúdo completo da paleta Modelos como um arquivo *.gen* no repositório.

Armazenando um modelo:

1. Clique no objeto na paleta Modelos no SPSS Modeler e, no menu principal, clique em:
Arquivo > Modelos > Armazenar Modelo...
2. Alternativamente, clique com o botão direito em um objeto na paleta Modelos e clique em **Armazenar Modelo**.

3. Continue em "Concluindo o procedimento de armazenamento".

Armazenando uma paleta Modelos:

1. Clique com o botão direito no segundo plano da paleta Modelos.
2. No menu pop-up, clique em **Armazenar Paleta**.
3. Continue em "Concluindo o procedimento de armazenamento".

Concluindo o procedimento de armazenamento:

1. Especifique configurações de conexão para o repositório, caso seja necessário. Consulte o tópico "Conectando ao repositório" na página 155 para obter mais informações. Para porta, senha e outros detalhes de conexão específicos, entre em contato com o administrador do sistema local.
2. Na caixa de diálogo Repositório: Armazenamento, escolha a pasta na qual deseja armazenar o objeto, especifique quaisquer outras informações que deseja registrar e clique no botão **Armazenamento**. Consulte o tópico "Configurando propriedades do objeto" na página 156 para obter mais informações.

Recuperando objetos do repositório

É possível recuperar fluxos, modelos, paletas de modelos, nós, projetos e objetos de saída que foram armazenados no repositório.

Nota: Além de usar as opções de menu conforme descrito aqui, também é possível recuperar fluxos, objetos de saída, modelos e paletas de modelo clicando com o botão direito na guia apropriada da área de janela de gerenciadores na parte superior direita da janela do SPSS Modeler.

1. Para recuperar um fluxo, no menu principal do IBM SPSS Modeler, clique em:
Arquivo > Recuperar Fluxo...
2. Para recuperar um modelo, paleta de modelos, projeto ou objeto de saída, no menu principal do IBM SPSS Modeler, clique em:
Arquivo > Modelos > Recuperar Modelo...
ou
Arquivo > Modelos > Recuperar Paleta de Modelos...
ou
Arquivo > Projetos > Recuperar Projeto...
ou
Arquivo > Saídas > Recuperar Saída...
3. Alternativamente, clique com o botão direito na área de janela do projeto ou gerenciadores e clique em **Recuperar** no menu pop-up.
4. Para recuperar um nó, no menu principal do IBM SPSS Modeler, clique em:
Inserir > Nó (ou SuperNode) do Repositório...
 - a. Especifique configurações de conexão para o repositório, caso seja necessário. Consulte o tópico "Conectando ao repositório" na página 155 para obter mais informações. Para porta, senha e outros detalhes de conexão específicos, entre em contato com o administrador do sistema local.
5. Na caixa de diálogo Repositório: Recuperar, navegue para o objeto, selecione-o e clique no botão **Recuperar**. Consulte o tópico para obter mais informações.

Escolhendo um objeto para recuperar

Os campos a seguir estão disponíveis na caixa de diálogo Repositório: Recuperar/Procurar:

Procurar em. Mostra a hierarquia de pastas para a pasta atual. Para navegar para uma pasta diferente, selecione uma na lista para navegar diretamente ou navegue usando a lista de objetos abaixo desse campo.

Botão Pasta para Cima. Navega um nível acima da pasta atual na hierarquia.

Botão Nova Pasta. Cria uma nova pasta no nível atual na hierarquia.

Nome do arquivo. O nome do arquivo de repositório do objeto selecionado. Para recuperar esse objeto, clique em **Recuperar**.

Arquivos de tipo. O tipo de objeto que você escolheu para recuperar. Somente objetos desse tipo, com pastas, são mostrados na lista de objetos. Para exibir objetos de um tipo diferente para recuperação, selecione o tipo de objeto da lista.

Abrir como bloqueado. Por padrão, quando um objeto é recuperado, ele é bloqueado no repositório para que outras pessoas não possam atualizá-lo. Se não desejar que o objeto seja bloqueado na recuperação, desmarque essa caixa.

Descrição, Palavras-chave. Se detalhes adicionais sobre o objeto foram definidos quando o objeto foi armazenado, esses detalhes serão exibidos aqui. Consulte o tópico “Incluindo informações sobre objetos armazenados” na página 156 para obter mais informações.

Versão. Para recuperar uma versão do objeto além da mais recente, clique neste botão. Informações detalhadas para todas as versões são exibidas, permitindo que você escolha a versão desejada.

Selecionando versão de um objeto

Para selecionar uma versão específica de um objeto do repositório, na caixa de diálogo Repositório: Selecionar Versão:

1. (Opcional) Ordene a lista por versão, rótulo, tamanho, data de criação ou usuário criador dando um clique duplo no cabeçalho da coluna apropriada.
2. Selecione a versão do objeto com o qual deseja trabalhar.
3. Clique em Continuar.

Procurando objetos no repositório

É possível procurar objetos por nome, pasta, tipo, rótulo, data ou outros critérios.

Procurando por Nome

Para procurar objetos por nome:

1. No menu principal do IBM SPSS Modeler, clique em:
Ferramentas > Repositório > Explorar...
 - a. Especifique configurações de conexão para o repositório, caso seja necessário. Consulte o tópico “Conectando ao repositório” na página 155 para obter mais informações. Para porta, senha e outros detalhes de conexão específicos, entre em contato com o administrador do sistema local.
2. Clique na guia **Procura**.
3. No campo **Procurar Objetos Nomeados**, especifique o nome do objeto que deseja localizar.

Durante a procura de objetos por nome, um asterisco (*) pode ser usado como um caractere curinga para corresponder a qualquer sequência de caracteres e um ponto de interrogação (?) corresponde a qualquer caractere simples. Por exemplo, **cluster** corresponde a todos os objetos que incluem o *cluster* de sequência de caracteres em qualquer lugar no nome. A sequência de caracteres de procura *m0?_** corresponde a *M01_cluster.str* e a *M02_cluster.str*, mas não *M01a_cluster.str*. As procuras não fazem distinção entre maiúsculas e minúsculas (*cluster* corresponde a *Cluster* e a *CLUSTER*).

Nota: Se o número de objetos for maior, as procuras podem levar alguns minutos.

Procurando por outros critérios

É possível executar uma procura com base em título, rótulos, datas, autor, palavras-chave, conteúdo indexado ou descrição. Somente em objetos que correspondem a *todos* os critérios de procura especificados serão localizados. Por exemplo, você poderia localizar todos os fluxos contendo um ou mais modelos de armazenamento em cluster que também tenham um rótulo específico aplicado e que foram modificados após uma data específica.

Tipos de Objeto. É possível restringir a procura a modelos, fluxos, saídas, nós, SuperNodes, projetos, paletas de modelo, cenários ou outros tipos de objetos.

- **Modelos.** É possível procurar modelos por categoria (classificação, aproximação, armazenamento em cluster, etc.) ou por um algoritmo de modelagem específico, como Kohonen.
Também é possível procurar por campos usados-por exemplo, todos os modelos que usam um campo denominado *receita* como um campo de entrada ou saída (resposta).
- **Fluxos.** Para fluxos, é possível restringir a procura por campos usados ou tipo de modelo (categoria ou algoritmo) contidos no fluxo.

Tópicos. É possível procurar modelos associados a tópicos específicos de uma lista configurada por usuários do repositório com os privilégios apropriados (para obter mais informações, consulte *Guia do Usuário do Deployment Manager*). Para obter a lista, marque essa caixa, clique no botão Incluir Tópicos que é exibido, selecione um ou mais tópicos na lista e clique em OK.

Rótulo. Restringe a procura a rótulos de versão do objeto específicos.

Datas. É possível especificar uma data de criação ou modificação e procurar objetos antes, depois ou entre o intervalo de dados especificado.

Autor. Restringe a procura a objetos criados por um usuário específico.

Palavras-chave. Procura por palavras-chave específicas. No IBM SPSS Modeler, palavras-chave são especificadas na guia Anotação para um fluxo, modelo ou objeto de saída.

Descrição. Procura por termos específicos no campo de descrição. No IBM SPSS Modeler, a descrição é especificada na guia Anotação para um modelo, fluxo ou objeto de saída. Frases de procura múltiplas podem ser separadas por pontos e vírgulas-por exemplo, *receita; tipo de agricultura; valor da indenização*. (Observe que dentro da frase de procura, os espaços são importantes. Por exemplo, *tipo de agricultura* com um espaço e *tipo de agricultura* com dois espaços não são a mesma coisa.)

Modificando objetos do repositório

É possível modificar objetos existentes no repositório diretamente do SPSS Modeler. É possível:

- Criar, renomear ou excluir pastas
- Bloquear ou desbloquear objetos
- Excluir objetos

Criando, renomeando e excluindo pastas

1. Para executar operações em pastas no repositório, no menu principal do SPSS Modeler, clique em: **Ferramentas > Repositório > Explorar...**
 - a. Especifique configurações de conexão para o repositório, caso seja necessário. Consulte o tópico “Conectando ao repositório” na página 155 para obter mais informações. Para porta, senha e outros detalhes de conexão específicos, entre em contato com o administrador do sistema local.
2. Assegure que a guia **Pastas** esteja ativa.
3. Para criar uma nova pasta, clique com o botão direito na pasta pai e clique em **Nova Pasta**.

4. Para renomear uma pasta, clique nela com o botão direito e clique em **Renomear Pasta**.
5. Para excluir uma pasta, clique nela com o botão direito e clique em **Excluir Pasta**.

Bloqueando e desbloqueando objetos do repositório

É possível bloquear um objeto para evitar que outros usuários atualizem qualquer uma de suas versões existentes ou criem novas versões. Um objeto bloqueado é indicado por um símbolo de cadeado sobre o ícone do objeto.



Figura 15. Objeto bloqueado

Para bloquear um objeto

1. Na janela do explorador de repositório, clique com o botão direito no objeto obrigatório.
2. Clique em **Bloquear**.

Para desbloquear um objeto

1. Na janela do explorador de repositório, clique com o botão direito no objeto obrigatório.
2. Clique em **Desbloquear**.

Excluindo objetos do repositório

Antes de excluir um objeto do repositório, deve-se decidir se você deseja excluir todas as versões do objeto ou somente uma versão específica.

Para Excluir Todas as Versões de um Objeto

1. Na janela do explorador de repositório, clique com o botão direito no objeto obrigatório.
2. Clique em **Excluir Objetos**.

Para Excluir a Versão Mais Recente de um Objeto

1. Na janela do explorador de repositório, clique com o botão direito no objeto obrigatório.
2. Clique em **Excluir**.

Para Excluir uma Versão Anterior de um Objeto

1. Na janela do explorador de repositório, clique com o botão direito no objeto obrigatório.
2. Clique em **Excluir Versões**.
3. Selecione a(s) versão(ões) para excluir e clique em **OK**.

Gerenciando propriedades dos objetos do repositório

É possível controlar várias propriedades de objeto a partir do SPSS Modeler. É possível:

- Visualizar as propriedades de uma pasta
- Visualizar e editar as propriedades de um objeto
- Criar, aplicar e excluir rótulos de versão para um objeto

Visualizando propriedades da pasta

Para visualizar propriedades para qualquer pasta na janela de repositório, clique com o botão direito na pasta necessária. Clique em **Propriedades da Pasta**.

guia Geral

Esta guia exibe o nome da pasta e as datas de criação e modificação.

Guia Permissões

Nesta guia, você especifica permissões de leitura e gravação para a pasta. Todos os usuários e grupos com acesso à pasta pai são listados. Permissões seguem uma hierarquia. Por exemplo, se você não tiver permissão de leitura, não será possível ter permissão de gravação. Se você não tiver permissão de gravação, não será possível ter permissão de exclusão.

Usuários e Grupos. Lista os usuários e grupos do repositório que têm pelo menos acesso de Leitura a essa pasta. Marque as caixas de seleção Gravar e Excluir para incluir esses direitos de acesso para essa pasta em um determinado usuário ou grupo. Clique no ícone **Incluir Usuários/Grupos** no lado direito da guia Permissões para designar acesso a usuários e grupos adicionais. A lista de usuários e grupos disponíveis é controlada pelo administrador.

Permissões em Cascata. Escolha uma opção para controlar como as mudanças feitas na pasta atual são aplicadas em suas pastas filhas, se houver.

- **Todas as Permissões em Cascata.** Coloca em cascata todas as configurações de permissão, da pasta atual a todas as pastas filhas e descendentes. Essa é uma maneira rápida de configurar permissões para várias pastas de uma vez. Configure permissões conforme necessário para a pasta pai e, em seguida, forme a cascata.
- **Cascata Somente de Mudanças.** Coloca em cascata somente as mudanças feitas desde que as últimas mudanças foram aplicadas. Por exemplo, se um novo grupo tiver sido incluído e você desejar dar a ele acesso a todas as pastas sob a ramificação Vendas, é possível dar ao grupo acesso à pasta raiz Vendas e formar uma cascata das mudanças para todas as subpastas. Todas as outras permissões para subpastas existentes permanecem como antes.
- **Não Formar Cascata.** Quaisquer mudanças feitas se aplicam apenas à pasta atual e não formam uma cascata para pastas filhas.

Visualizando e editando propriedades do objeto

Na caixa de diálogo Propriedades do Objeto, é possível visualizar e editar propriedades. Embora algumas propriedades não possam ser alteradas, sempre é possível atualizar um objeto incluindo uma nova versão.

1. Na janela do repositório, clique com o botão direito no objeto necessário.
2. Clique em **Propriedades do Objeto**.

Guia Geral

Nome. O nome do objeto como ele é visualizado no repositório.

Criado em. Data em que o objeto (e não a versão) foi criado.

Última modificação. Data em que a versão mais recente foi modificada.

Autor. O nome de login do usuário.

Descrição. Por padrão, contém a descrição especificada na guia Anotação do objeto no SPSS Modeler.

Tópicos vinculados. O repositório permite que modelos e objetos relacionados sejam organizados por tópicos, se necessário. A lista de tópicos disponíveis é configurada pelos usuários do repositório com os privilégios apropriados (para obter mais informações, consulte *Deployment Manager User's Guide*).

Palavras-chave. Você especifica palavras-chave na guia Anotação para um fluxo, modelo ou objeto de saída. Palavras-chave múltiplas devem ser separadas por espaços, com até no máximo 255 caracteres. (Se as palavras-chave contiverem espaços, use aspas para separá-las.)

Guia Versões

Objetos armazenados no repositório podem ter versões múltiplas. A guia Versões exibe informações sobre cada versão.

As propriedades a seguir podem ser especificadas ou modificadas para versões específicas de um objeto armazenado:

Versão. Identificador exclusivo para a versão gerada com base no horário em que a versão foi armazenada.

Rótulo. O rótulo atual da versão, se houver. Ao contrário do identificador de versão, os rótulos podem ser movidos de uma versão do objeto para outra.

O tamanho do arquivo, a data de criação e o autor também são exibidos para cada versão.

Editar Rótulos. Clique no ícone **Editar Rótulos** na parte superior direita da guia Versões para definir, aplicar, remover rótulos para objetos armazenados. Consulte o tópico “Gerenciando rótulos de versão do objeto” para obter mais informações.

Guia Permissões

Na guia Permissões, é possível configurar permissões de leitura e gravação para o objeto. Todos os usuários e grupos com acesso ao objeto atual são listados. Permissões seguem uma hierarquia. Por exemplo, se você não tiver permissão de leitura, não será possível ter permissão de gravação. Se você não tiver permissão de gravação, não será possível ter permissão de exclusão.

Usuários e Grupos. Lista os usuários e grupos do repositório que têm pelo menos acesso de Leitura a esse objeto. Marque as caixas de seleção Gravar e Excluir para incluir esses direitos de acesso para esse objeto em um determinado usuário ou grupo. Clique no ícone **Incluir Usuários/Grupos** no lado direito da guia Permissões para designar acesso a usuários e grupos adicionais. A lista de usuários e grupos disponíveis é controlada pelo administrador.

Gerenciando rótulos de versão do objeto

A caixa de diálogo Editar Rótulos de Versão permite:

- Aplicar rótulos no objeto selecionado
- Remover rótulos do objeto selecionado
- Definir um novo rótulo e aplicá-lo ao objeto

Para aplicar rótulos no objeto

1. Selecione um ou mais rótulos na lista **Rótulos Disponíveis**.
2. Clique com o botão Seta à Direita para mover os rótulos selecionados para a lista **Rótulos Aplicados**.
3. Clique em **OK**.

Para remover rótulos do objeto

1. Selecione um ou mais rótulos na lista **Rótulos Aplicados**.
2. Clique no botão Seta à Esquerda para mover os rótulos selecionados para a lista **Rótulos Disponíveis**.
3. Clique em **OK**.

Para definir um novo rótulo e aplicá-lo ao objeto

1. Digite o nome do rótulo no campo **Novo Rótulo**.
2. Clique com o botão Seta à Direita para mover o novo rótulo para a lista **Rótulos Aplicados**.
3. Clique em **OK**.

Implementando fluxos

Para permitir que um fluxo seja usado com o aplicativo thin client IBM SPSS Modeler Advantage, ele deve ser implementado como um fluxo (arquivo *.str*) no repositório.

Nota: Não é possível implementar um fluxo que tenha mais de um nó de origem na ramificação de escoragem.

Independentemente de um fluxo estar implementado como um fluxo (arquivo *.str*) ou como um cenário (arquivo *.scn*), o objeto poderá tirar total proveito das variáveis de nível corporativo do IBM SPSS Collaboration and Deployment Services. Consulte o tópico “Armazenando e implementando objetos do repositório” na página 154 para obter mais informações.

Para implementar o fluxo atual (método do menu Arquivo)

1. No menu principal, clique em:
Arquivo > Armazenamento > Implementar
2. Escolha o tipo de implementação e conclua o restante da caixa de diálogo conforme necessário.
3. Clique em **Implementar como Fluxo** para implementar o fluxo para usar com IBM SPSS Modeler Advantage ou IBM SPSS Collaboration and Deployment Services. Clique em **Implementar como Cenário** para implementar o fluxo para usar com IBM SPSS Collaboration and Deployment Services ou Predictive Applications versão 5.
4. Clique em **Armazenamento**. Para obter mais informações, clique em **Ajuda**.
5. Continue em "Concluindo o processo de implementação".

Para implementar o fluxo atual (método do menu Ferramentas)

1. No menu principal, clique em:
Ferramentas > Propriedades do Fluxo > Implementação
2. Escolha o tipo de implementação, conclua o restante da guia Implementação conforme necessário e clique em **Armazenamento**. Consulte o tópico “Opções de Implementação do Fluxo” para obter mais informações.

Concluindo o processo de implementação

1. Especifique configurações de conexão para o repositório, caso seja necessário. Consulte o tópico “Conectando ao repositório” na página 155 para obter mais informações. Para porta, senha e outros detalhes de conexão específicos, entre em contato com o administrador do sistema local.
2. Na caixa de diálogo Repositório: Armazenamento, escolha a pasta na qual deseja armazenar o objeto, especifique quaisquer outras informações que deseja registrar e clique no botão **Armazenamento**. Consulte o tópico “Configurando propriedades do objeto” na página 156 para obter mais informações.

Opções de Implementação do Fluxo

A guia Implementação na caixa de diálogo Opções de Fluxo permite especificar opções para implementar o fluxo. É possível implementar como um fluxo ou como um cenário.

Quando você implementa como um fluxo, é possível abrir e modificar o fluxo no aplicativo thin client IBM SPSS Modeler Advantage. O fluxo é armazenado no repositório como um arquivo com a extensão *.str*.

A implementação como um cenário armazena o fluxo no repositório como um arquivo com a extensão *.scn*. Isso também permite que um fluxo seja usado pelo Predictive Applications versão 5.

Independentemente de você implementar como um fluxo ou como cenário, é possível aproveitar as vantagens da funcionalidade adicional disponível com IBM SPSS Collaboration and Deployment Services, como acesso multiusuário, escoragem automatizada, atualização de modelo e análise de Champion Challenger.

Nota: Os nós de modelagem Regras de Associação, STP e TCM não suportam os passos Avaliação de Modelo ou Champion Challenger no IBM SPSS Collaboration and Deployment Services.

Na guia Implementação, também é possível visualizar a descrição de fluxo que o IBM SPSS Modeler cria para o fluxo. Consulte o tópico “Descrições de fluxo” na página 52 para obter mais informações.

Nota: Para assegurar o acesso consistente aos dados corporativos, um fluxo que é implementado como um cenário deve acessar seus dados de origem por meio de IBM SPSS Collaboration and Deployment Services Enterprise View, portanto, nesse caso, o fluxo deve incluir pelo menos um nó de origem Visualização Corporativa dentro de cada ramificação de escoragem ou modelagem designada conforme aplicável.

Tipo de implementação. Escolha como deseja implementar o fluxo. Todos os fluxos requerem um nó de escoragem designado para que possam ser implementados; as opções e os requisitos adicionais dependem do tipo de implementação.

Nota: Os nós de modelagem Regras de Associação, STP e TCM não suportam os passos Avaliação de Modelo ou Champion Challenger no IBM SPSS Collaboration and Deployment Services.

- **<nenhum>**. O fluxo não será implementado no repositório. Todas as opções ficam desativadas, exceto a visualização da descrição do fluxo.
- **Somente Escoragem.** O fluxo é implementado no repositório quando você clica no botão **Armazenamento**. Dados podem ser escorados usando o nó designado no campo **Nó de Escoragem**.
- **Atualização de Modelo.** O mesmo que Somente Escoragem mas, além disso, pode ser atualizado no repositório usando os objetos que você designar nos campos **Nó de Modelagem** e **Nugget de Modelo**.
Nota: A atualização de modelo automática não é suportada por padrão no IBM SPSS Collaboration and Deployment Services, portanto, deve-se escolher esse tipo de implementação se você desejar usar essa variável durante a execução de um fluxo a partir do repositório. Consulte o tópico “Atualizar modelo” na página 169 para obter mais informações.

Nó de Escoragem. Selecione um gráfico, uma saída ou um nó de exportação para identificar a ramificação de fluxo a ser usada para escoragem dos dados. Embora o fluxo possa realmente conter qualquer número válido de ramificações, modelos e nós terminais, uma, e somente uma, ramificação de escoragem deve ser designada para propósitos de implementação. Esse é o requisito mais básico para implementar qualquer fluxo.

Parâmetros de Escoragem. Permite especificar parâmetros que podem ser modificados quando a ramificação de escoragem é executada. Consulte o tópico “Parâmetros de escoragem e modelagem” na página 168 para obter mais informações.

Nó de Modelagem. Para atualização de modelo, especifica o nó de modelagem usado para gerar novamente ou atualizar o modelo no repositório. Deve ser um nó de modelagem do mesmo tipo que o especificado para **Nugget do Modelo**.

Parâmetros de Construção de Modelo. Permite especificar parâmetros que podem ser modificados quando o nó de modelagem é executado. Consulte o tópico “Parâmetros de escoragem e modelagem” na página 168 para obter mais informações.

Nugget do Modelo. Para atualização de modelo, especifica o nugget do modelo que será atualizado ou gerado novamente cada vez que o fluxo for atualizado no repositório (geralmente como parte de uma tarefa planejada). O modelo deve estar localizado na ramificação de escoragem. Embora possa existir vários modelos na ramificação de escoragem, somente um pode ser designado. Observe que quando o fluxo é criado inicialmente, ele pode ser efetivamente um modelo de item temporário que é atualizado ou gerado novamente conforme novos dados são disponibilizados.

Implementar como Fluxo. Clique nesta opção se desejar usar o fluxo com IBM SPSS Modeler Advantage ou IBM SPSS Collaboration and Deployment Services (consulte a nota a seguir).

Implementar como Cenário. Clique nesta opção se desejar usar o fluxo com IBM SPSS Collaboration and Deployment Services ou Predictive Applications versão 5 (consulte a nota a seguir).

Verificação. Clique neste botão para verificar se este é um fluxo válido para implementação. Todos os fluxos devem ter um nó de escoragem designado para que possam ser implementados. Se você estiver implementando como um cenário, o fluxo também deverá conter um nó de origem Visualização Corporativa válido. Mensagens de erro serão exibidas se essas condições não forem satisfeitas.

Armazenamento. Implementa o fluxo se ele for válido. Se não, uma mensagem de erro será exibida. Clique no botão **Correção**, corrija o erro e tente novamente.

Descrição do Fluxo de Visualização. Permite visualizar o conteúdo da descrição do fluxo que o IBM SPSS Modeler cria para o fluxo. Consulte o tópico “Descrições de fluxo” na página 52 para obter mais informações.

Nota: (Implementar como fluxo ou cenário) Vários nós de Visualização Corporativa podem ser usados dentro da ramificação de modelagem. Nesse caso, o uso de uma única conexão de dados para todos os nós de Visualização Corporativa dentro da ramificação é preferível na maioria dos casos, e é necessário para análise do Champion Challenger.

Nota: Os nós de modelagem Regras de Associação, STP e TCM não suportam os passos Avaliação de Modelo ou Champion Challenger no IBM SPSS Collaboration and Deployment Services.

- Se o suporte ao Champion Challenger não for necessário, diferentes conexões de Visualização Corporativa poderão ser usadas dentro da mesma ramificação, contanto que as conexões variem somente por definição de provedor de dados (DPD).
- Essas limitações se aplicam somente dentro de uma determinada ramificação. Entre ramificações de construção de modelo e escoragem, diferentes conexões de Visualização Corporativa podem ser usadas sem essas restrições.

Parâmetros de escoragem e modelagem

Durante a implementação de um fluxo no IBM SPSS Collaboration and Deployment Services, é possível escolher quais parâmetros podem ser visualizados ou editados cada vez que o modelo é atualizado ou escorado. Por exemplo, você pode especificar os valores máximo e mínimo ou algum outro valor que possa estar sujeito à mudança cada vez que uma tarefa for executada.

1. Para deixar um parâmetro visível para que ele possa ser visualizado ou editado após o fluxo ser implementado, selecione-o na lista na caixa de diálogo Parâmetros de Escoragem.

A lista de parâmetros disponíveis é definida na guia Parâmetros na caixa de diálogo de propriedades do fluxo. Consulte o tópico “Configurando parâmetros de sessão e fluxo” na página 47 para obter mais informações.

A ramificação de escoragem

Se você estiver implementando um fluxo, uma ramificação do fluxo deverá ser designada como **ramificação de escoragem** (ou seja, aquela que contém o nó de escoragem). Quando você designa uma ramificação como ramificação de escoragem, ela é destacada na tela de fluxo, como na ligação do modelo

com o nugget na ramificação de escoragem. Essa representação visual é particularmente útil em fluxos complexos com diversas ramificações nas quais a ramificação de escoragem pode não ser imediatamente óbvia.

Nota: Somente uma ramificação de fluxo pode ser designada como ramificação de escoragem.

Se o fluxo já tinha uma ramificação de escoragem definida, a ramificação recém-designada a substituirá como ramificação de escoragem. É possível configurar a cor da indicação de ramificação de escoragem por meio de uma opção Cor Customizada. Consulte o tópico “Configurando opções de exibição” na página 188 para obter mais informações.

É possível mostrar ou ocultar a indicação de ramificação de escoragem por meio do botão da barra de ferramentas Mostrar/Ocultar Marcação de Fluxo.



Figura 16. Botão da barra de ferramentas Mostrar/Ocultar Marcação de Fluxo

Identificando a ramificação de escoragem para implementação

É possível designar a ramificação de escoragem a partir do menu pop-up de um nó terminal ou do menu Ferramentas. Se você usar o menu pop-up, o nó de escoragem será configurado automaticamente na guia Implementação das propriedades do fluxo.

Para designar uma ramificação como a ramificação de escoragem (menu pop-up)

1. Conecte o nugget do modelo a um nó terminal (recebimento de dados do nó de saída ou de processamento do nugget).
2. Clique com o botão direito no nó terminal.
3. No menu, clique em **Usar como Ramificação de Escoragem**.

Para designar uma ramificação como ramificação de escoragem (menu Ferramentas)

1. Conecte o nugget do modelo a um nó terminal (recebimento de dados do nó de saída ou de processamento do nugget).
2. No menu principal, clique em:
Ferramentas > Propriedades do Fluxo > Implementação
3. Na lista **Tipo de Implementação**, clique em **Somente Escoragem** ou **Atualização de Modelo** conforme necessário. Consulte o tópico “Opções de Implementação do Fluxo” na página 166 para obter mais informações.
4. Clique no campo **Nó de Escoragem** e selecione um nó terminal da lista.
5. Clique **OK**.

Atualizar modelo

Atualização de modelo é o processo de reconstrução de um modelo existente em um fluxo usando dados mais novos. O fluxo em si não muda no repositório. Por exemplo, o tipo de algoritmo e as configurações específicas do fluxo permanecem iguais, mas o modelo é retreinado nos novos dados e atualizados, caso a nova versão do modelo funcione melhor que a antiga.

Somente um nugget do modelo em um fluxo pode ser configurado--isso é conhecido como **modelo de atualização**. Se você clicar na opção **Atualização de Modelo** na guia Implementação das propriedades do fluxo (consulte “Opções de Implementação do Fluxo” na página 166), o nugget do modelo designado

nesse momento se tornará o modelo de atualização. Também é possível designar um modelo como modelo de atualização no menu pop-up de um nugget do modelo. O nugget já deve estar na ramificação de escoragem para que isso seja possível.

Se você desativar o status do "modelo de atualização" de um nugget, isso será equivalente à configuração do tipo de implementação do fluxo para Somente Escoragem, e a guia Implementação da caixa de diálogo de propriedades do fluxo será atualizada de acordo. É possível ativar e desativar esse status por meio da opção **Usar como Modelo de Atualização** no menu pop-up do nugget na ramificação de escoragem atual.

A remoção da ligação do modelo de um nugget na ramificação de escoragem também remove o status "modelo de atualização" do nugget. É possível desfazer a remoção da ligação do modelo por meio do menu Editar ou da barra de ferramentas; isso também restabelece o status do "modelo de atualização" do nugget.

Como Atualizar Modelo é selecionada

Assim como a ramificação de escoragem, a ligação com um modelo de atualização também é destacada no fluxo. O nugget do modelo escolhido como o modelo de atualização e, portanto, a ligação que é destacada, depende de quantos nuggets estão no fluxo.

Modelo Único no Fluxo

Se houver um único nugget do modelo vinculado na ramificação de escoragem quando ele for identificado como tal, esse nugget se tornará o modelo de atualização para o fluxo.

Vários Modelos no Fluxo

Se houver mais de um nugget vinculado no fluxo, o modelo de atualização será escolhido como tal.

Se um nugget do modelo tiver sido definido na guia Implementação da caixa de diálogo de propriedades do fluxo e também no fluxo, ele se tornará o modelo de atualização.

Se nenhum nugget tiver sido definido na guia Implementação, ou se tiver sido definido, mas não na ramificação de escoragem, o nugget mais próximo do nó terminal se tornará o modelo de atualização.

Se você cancelar subsequente a seleção de todas as ligações de modelo como ligações de atualização, somente a ramificação de escoragem será destacada, não as ligações. O tipo de implementação é configurado para Somente Escoragem.

Nota: É possível escolher configurar uma das ligações para o status Substituir, mas não o outro. Nesse caso, o nugget do modelo escolhido como o modelo de atualização é aquele que tem uma ligação de atualização e que está mais perto do nó terminal quando a ramificação de escoragem é designada.

Nenhum Modelo no Fluxo

Se não houver modelos no fluxo, ou somente modelos sem ligações de modelo, o tipo de implementação será configurado para Somente Escoragem.

Verificando erros em uma ramificação de escoragem

Quando você designa a ramificação de escoragem, ela é verificada para ver se existem erros, como ausência de um nó da Visualização Corporativa no fluxo durante a implementação como um cenário.

Se um erro for localizado, a ramificação de escoragem será destacada na cor do erro de ramificação de escoragem e uma mensagem de erro será exibida. É possível configurar a cor do erro por meio de uma opção Cor Customizada. Consulte o tópico "Configurando opções de exibição" na página 188 para obter mais informações.

Se um erro for localizado, continue da seguinte forma:

1. Corrija o erro de acordo com o conteúdo da mensagem de erro.
2. Execute uma das seguintes ações:
 - Clique com o botão direito no nó terminal e clique em **Verificar Cenário** no menu pop-up.
 - No menu principal, clique em:
Ferramentas > Propriedades do Fluxo > Implementação
e clique em **Verificar**.
3. Se for necessário, repita esse processo até que nenhum erro seja localizado.

Capítulo 11. Exportando para aplicativos externos

Sobre Exportando para aplicativos externos

IBM SPSS Modeler fornece inúmeros mecanismos para você exportar o processo de mineração de dados inteiro para aplicativos externos, de modo que o trabalho realizado para preparar dados e construir modelos possa ser usado em sua vantagem fora do IBM SPSS Modeler também.

A seção anterior mostrou como é possível implementar fluxos em um repositório do IBM SPSS Collaboration and Deployment Services para aproveitar a vantagem de seu acesso multiusuário, planejamento de tarefas e outras variáveis. Da forma semelhante, os fluxos do IBM SPSS Modeler também podem ser usados com:

- IBM SPSS Modeler Advantage
- Aplicativos que podem importar e exportar arquivos em formato PMML

Para obter mais informações sobre como usar fluxos com IBM SPSS Modeler Advantage, consulte “Abrindo um fluxo no IBM SPSS Modeler Advantage”.

Para obter informações sobre como exportar e importar modelos como arquivos PMML, possibilitando o compartilhamento de modelos com quaisquer outros aplicativos que suportam esse formato, consulte “Importando e exportando modelos como PMML” na página 174.

Abrindo um fluxo no IBM SPSS Modeler Advantage

Os fluxos do IBM SPSS Modeler podem ser usados com o aplicativo thin client IBM SPSS Modeler Advantage. Embora seja possível criar aplicativos customizados completamente dentro do IBM SPSS Modeler Advantage, também é possível usar um fluxo já criado no IBM SPSS Modeler como a base de um fluxo de trabalho de aplicativo.

Para abrir um fluxo no IBM SPSS Modeler Advantage:

1. Implemente o fluxo no repositório do IBM SPSS Collaboration and Deployment Services, certificando-se de clicar na opção **Implementar como Fluxo**. Consulte o tópico “Implementando fluxos” na página 166 para obter mais informações.
2. Clique no botão da barra de ferramentas Abrir no IBM SPSS Modeler Advantage, ou no menu principal, clique em:

Arquivo > Abrir no IBM SPSS Modeler Advantage

1. Especifique configurações de conexão para o repositório, caso seja necessário. Consulte o tópico “Conectando ao repositório” na página 155 para obter mais informações. Para porta, senha e outros detalhes de conexão específicos, entre em contato com o administrador do sistema local.

Nota: O servidor de repositório também deve ter o software IBM SPSS Modeler Advantage instalado.

1. Na caixa de diálogo Repositório: Armazenamento, escolha a pasta na qual deseja armazenar o objeto, especifique quaisquer outras informações que deseja registrar e clique no botão **Armazenamento**. Consulte o tópico “Configurando propriedades do objeto” na página 156 para obter mais informações.

Isso ativa o IBM SPSS Modeler Advantage com o fluxo já aberto. O fluxo é fechado no IBM SPSS Modeler.

Importando e exportando modelos como PMML

PMML, ou Predictive Model Markup Language, é um formato XML para descrever modelos estatísticos e mineração de dados, incluindo entradas para os modelos, transformações usadas para preparar dados para mineração de dados e os parâmetros que definem os modelos em si. IBM SPSS Modeler pode importar e exportar PMML, possibilitando o compartilhamento de modelos com outros aplicativos que suportem esse formato, como IBM SPSS Statistics.

Para obter mais informações sobre PMML, consulte o website Grupo de Mineração de Dados (<http://www.dmg.org>).

Para Exportar um Modelo

A exportação de PMML é suportada para a maioria dos tipos de modelo gerados no IBM SPSS Modeler. Consulte o tópico “Tipos de modelo que suportam PMML” na página 175 para obter mais informações.

1. Clique com o botão direito em um nugget do modelo na paleta de modelos. (Alternativamente, dê um clique duplo em um nugget do modelo na tela e selecione o menu Arquivo.)
2. No menu, clique em **Exportar PMML**.
3. Na caixa de diálogo Exportar (ou Salvar), especifique um diretório de resposta e um nome exclusivo para o modelo.

Nota: É possível alterar opções para exportação de PMML na caixa de diálogo Opções de Usuário. No menu principal, clique em:

Ferramentas > Opções > Opções de Usuário

e clique na guia PMML.

Consulte o tópico “Configurando opções de exportação de PMML” na página 189 para obter mais informações.

Para Importar um Modelo Salvo como PMML

Modelos exportados como PMML do IBM SPSS Modeler ou outro aplicativo podem ser importados na paleta de modelos. Consulte o tópico “Tipos de modelo que suportam PMML” na página 175 para obter mais informações.

1. Na paleta de modelos, clique com o botão direito do mouse na paleta e selecione **Importar PMML** no menu.
2. Selecione o arquivo para importar e especifique opções para rótulos de variáveis conforme necessário.
3. Clique em **Abrir**.

Usar rótulos de variáveis se presentes no modelo. O PMML pode especificar nomes de variáveis e rótulos de variáveis (como ID do Referenciador para *RefID*) para variáveis no dicionário de dados. Selecione essa opção para usar rótulos de variáveis se eles estiverem presentes no PMML exportado originalmente.

Se você tiver selecionado a opção de rótulo de variáveis, mas não houver rótulos de variáveis no PMML, os nomes de variáveis serão usados como normalmente.

Tipos de modelo que suportam PMML

Exportação de PMML

Modelos do IBM SPSS Modeler. Os modelos a seguir criados no IBM SPSS Modeler podem ser exportados como PMML 4.0:

- Árvore C e R
- QUEST
- CHAID
- Regressão Linear
- Rede neural
- C5.0
- Regressão Logística
- Genlin
- SVM
- A priori
- Carma
- K-Médias
- Kohonen
- TwoStep
- GLMM (suporte somente para modelos GLMM Somente de Efeito Fixo)
- Lista de Decisão
- Cox
- Sequência (escoragem para modelos PMML de Sequência não é suportada)
- Modelo Statistics

Modelos nativos de banco de dados. Para modelos gerados usando algoritmos nativos do banco de dados, a exportação de PMML está disponível somente para modelos IBM InfoSphere Warehouse. Modelos criados usando Serviços de Análise da Microsoft ou Oracle Data Miner não podem ser exportados. Além disso, observe que os modelos da IBM exportados como PMML não podem ser importados de volta no IBM SPSS Modeler.

Importação de PMML

IBM SPSS Modeler pode importar e escorar modelos PMML gerados pelas versões atuais de todos os produtos IBM SPSS Statistics, incluindo modelos exportados do IBM SPSS Modeler, bem como PMML de transformação ou modelo gerado pelo IBM SPSS Statistics 17.0 ou posterior. Essencialmente, isso significa qualquer PMML que o mecanismo de escoragem pode escorar com as exceções a seguir:

- A Priori, CARMA, Detecção de Anomalias, Sequência e Modelos de Regra de Associação não podem ser importados.
- Modelos PMML não podem ser navegados após importação no IBM SPSS Modeler, mesmo que possam ser usados na escoragem. (Observe que isso inclui modelos que foram exportados do IBM SPSS Modeler com os quais começar. Para evitar essa limitação, exporte o modelo como um arquivo de modelo gerado [**.gm*] em vez de PMML.)
- Modelos do IBM InfoSphere Warehouse exportados como PMML não podem ser importados.
- A validação limitada ocorre na importação, mas a validação integral é executada na tentativa de escorar o modelo. Assim, é possível que a importação seja bem-sucedida, mas a escoragem falhará ou produzirá resultados incorretos.

Nota: Para PMML de terceiros importados no IBM SPSS Modeler, IBM SPSS Modeler tentará escorar PMML válido que pode ser reconhecido e escorado. Mas não é garantido que todo PMML escorará ou que escorará da mesma maneira que o aplicativo que o gerou.

Capítulo 12. Projetos e relatórios

Introdução aos Projetos

Um **projeto** é um grupo de arquivos relacionados a uma tarefa de mineração de dados. Os projetos incluem fluxos de dados, gráficos, modelos gerados, relatórios e mais alguma coisa que você tenha criado no IBM SPSS Modeler. À primeira vista, pode parecer que os projetos do IBM SPSS Modeler sejam simplesmente uma maneira de organizar a saída, mas são realmente capazes de muito mais. Usando os projetos, é possível:

- Anotar cada objeto no arquivo de projeto.
- Usar a metodologia CRISP-DM para orientar seus esforços na mineração de dados. Os projetos também contêm um Sistema de Ajuda CRISP-DM que fornece detalhes e exemplos reais sobre mineração de dados com CRISP-DM.
- Incluir objetos não IBM SPSS Modeler no projeto, como a exibição de slides do PowerPoint usados para apresentar os objetivos da mineração de dados ou White Papers sobre os algoritmos que você planeja usar.
- Produzir relatórios de atualização abrangentes e simples com base em suas anotações. Esses relatórios podem ser gerados em HTML para publicação fácil na intranet de sua organização.

Nota: se a área de janela do projeto não estiver visível na janela do IBM SPSS Modeler, clique em **Projeto** no menu Visualizar.

Os objetos incluídos em um projeto podem ser visualizados de duas maneiras: **Visualização de classes** e **Visualização de CRISP-DM**. Qualquer coisa que for incluída em um projeto será incluída em ambas as visualizações, e é possível alternar entre as visualizações para criar a organização que funcione melhor.

Visualização CRISP-DM

Suportando o processo padrão de vários segmentos de mercados para mineração de dados (CRISP-DM), os projetos do IBM SPSS Modeler fornecem uma maneira não proprietária e comprovada pela indústria de organizar as partes de seus esforços de mineração de dados. CRISP-DM usa seis fases para descrever o processo do início (reunindo necessidades de negócios) ao fim (implementando seus resultados). Mesmo que algumas fases normalmente não envolvam trabalho no IBM SPSS Modeler, a área de janela do projeto inclui todas as seis fases para que você tenha uma localização central para armazenar e controlar todos os materiais associados ao projeto. Por exemplo, a fase Entendimento de Negócios geralmente envolve reunir necessidades e se encontrar com colegas para determinar objetivos em vez de trabalhar com dados no IBM SPSS Modeler. A área de janela do projeto permite armazenar suas observações feitas nessas reuniões na pasta *Entendimento de Negócios* para futura referência e inclusão em relatórios.

A visualização CRISP-DM na área de janela do projeto também é equipada com seu próprio sistema de Ajuda para lhe guiar pelo ciclo de vida de mineração de dados. No IBM SPSS Modeler, essa ajuda pode ser acessada clicando na **Ajuda do CRISP-DM** no menu Ajuda.

Nota: Se a área de janela do projeto não estiver visível na janela, clique em **Projeto** no menu Visualizar.

Configurando a fase do projeto padrão

Objetos incluídos em um projeto são incluídos em uma fase padrão do CRISP-DM. Isso significa que é necessário organizar objetos manualmente de acordo com a fase de mineração de dados em que eles foram usados. É sábio configurar a pasta padrão para a fase na qual você está trabalhando atualmente.

Para selecionar qual fase usar como padrão:

1. Na visualização CRISP-DM, clique com o botão direito na pasta para a fase para configurar como padrão.
2. No menu, clique em **Configurar como Padrão**.

A pasta padrão é exibida em negrito.

Visualização Classes

A visualização Classes na área de janela do projeto organiza seu trabalho no IBM SPSS Modeler categoricamente por tipos de objetos criados. Objetos salvos podem ser incluídos em qualquer uma das categorias a seguir:

- Fluxos
- Nós
- Modelos
- Tabelas, gráficos e relatórios
- Outros (arquivos não IBM SPSS Modeler, como exibição de slides ou White Papers relevantes para seu trabalho de mineração de dados)

A inclusão de objetos na visualização Classes também os inclui na pasta da fase padrão na visualização CRISP-DM.

Nota: Se a área de janela do projeto não estiver visível na janela, clique em **Projeto** no menu Visualizar.

Construindo um projeto

Essencialmente, um projeto é um arquivo contendo referências a todos os arquivos associados ao projeto. Isso significa que os itens do projeto são salvos individualmente e como uma referência no arquivo de projeto (.cpj). Devido a essa estrutura referencial, observe o seguinte:

- Primeiro os itens do projeto devem ser salvos individualmente antes de serem incluídos em um projeto. Se um item não estiver salvo, será solicitado que você o salve antes de incluí-lo no projeto atual.
- Objetos atualizados individualmente, como fluxos, também são atualizados no arquivo de projeto.
- Mover ou excluir objetos manualmente (como fluxos, nós e objetos de saída) do sistema de arquivos renderizará ligações no arquivo de projeto inválido.

Criando um novo projeto

Projetos novos são fáceis de criar na janela IBM SPSS Modeler. É possível começar construindo um, se não houver nenhum aberto, ou é possível fechar um projeto existente e começar do zero.

No menu principal, clique em:

Arquivo > Projeto > Novo Projeto...

Incluindo em um projeto

Após você ter criado ou aberto um projeto, é possível incluir objetos, como fluxos de dados, nós e relatórios, usando diversos métodos.

Incluindo objetos a partir de gerenciadores

Usando os gerenciadores no canto superior direito da janela do IBM SPSS Modeler, é possível incluir fluxos ou saída.

1. Selecione um objeto, como uma tabela ou um fluxo, de uma das guias do gerenciador.
2. Clique com o botão direito e clique em **Incluir no Projeto**.

Se o objeto tiver sido salvo anteriormente, ele será automaticamente incluído na pasta de objetos apropriada (na visualização Classes) ou na pasta de fase padrão (na visualização CRISP-DM).

3. Alternativamente, é possível arrastar e soltar objetos dos gerenciadores para a área de janela do projeto.

Nota: Talvez seja solicitado que você salve o objeto primeiro. Durante o salvamento, certifique-se de que **Incluir Arquivo no Projeto** esteja selecionada na caixa de diálogo Salvar. Isso incluirá o objeto automaticamente no projeto após você salvá-lo.

Incluindo nós da tela

É possível incluir nós individuais da tela de fluxo usando a caixa de diálogo Salvar.

1. Selecione um nó na tela.
2. Clique com o botão direito e clique em **Salvar Nó**. Alternativamente, no menu principal, clique em: **Editar > Nó > Salvar Nó...**
3. Na caixa de diálogo Salvar, selecione **Incluir Arquivo no Projeto**.
4. Crie um nome para o nó e clique em **Salvar**.

Isso salva o arquivo e o inclui no projeto. Nós são incluídos na pasta *Nós* na visualização Classes e na pasta de fase padrão na visualização CRISP-DM.

Incluindo arquivos externos

É possível incluir uma grande variedade de objetos não IBM SPSS Modeler em um projeto. Isso é útil quando você está gerenciando o processo de mineração de dados inteiro dentro do IBM SPSS Modeler. Por exemplo, é possível armazenar ligações para dados, notas, apresentações e gráficos em um projeto. Na visualização CRISP-DM, arquivos externos podem ser incluídos na pasta de sua escolha. Na visualização Classes, arquivos externos podem ser salvos somente na pasta *Outro*.

Para incluir arquivos externos em um projeto:

1. Arraste arquivos da área de trabalho para o projeto.
ou
2. Clique com o botão direito na pasta de resposta na visualização CRISP-DM ou Classes.
3. No menu, clique em **Incluir na Pasta**.
4. Selecione um arquivo na caixa de diálogo e clique em **Abrir**.

Isso incluirá uma referência ao objeto selecionado dentro dos projetos do IBM SPSS Modeler.

Transferindo projetos para o IBM SPSS Collaboration and Deployment Services Repository

É possível transferir um projeto inteiro, incluindo todos os arquivos do componente, para o IBM SPSS Collaboration and Deployment Services Repository em um passo. Nenhum objeto que já está no local de resposta será movido. Essa variável trabalha de maneira inversa: é possível transferir projetos inteiros do IBM SPSS Collaboration and Deployment Services Repository para seu sistema de arquivos local.

Transferindo um Projeto

Certifique-se de que o projeto que deseja transferir esteja aberto na área de janela do projeto.

Para transferir um projeto:

1. Clique com o botão direito na pasta do projeto raiz e clique em **Transferir Projeto**.
2. Se for solicitado, efetue login no IBM SPSS Collaboration and Deployment Services Repository.

3. Especifique a nova localização para o projeto e clique em **OK**.

Configurando propriedades do projeto

É possível customizar a documentação e o conteúdo do projeto usando a caixa de diálogo de propriedades do projeto. Para acessar as propriedades do projeto:

1. Clique com o botão direito em um objeto ou pasta na área de janela do projeto e clique em **Propriedades do Projeto**.
2. Clique na guia **Projeto** para especificar informações básicas sobre o projeto.

Criado. Mostra a data de criação do projeto (não editável).

Sumarização. É possível inserir uma sumarização de seu projeto de mineração de dados que será exibida no relatório do projeto.

Conteúdo. Lista o tipo e o número de componentes referenciados pelo arquivo do projeto (não editável).

Salvar objeto não salvo como. Especifica se objetos não salvos devem ser salvos no sistema de arquivos local ou armazenados no repositório. Consulte o tópico "Sobre o IBM SPSS Collaboration and Deployment Services Repository" na página 153 para obter mais informações.

Atualizar referências do objeto durante o carregamento do projeto. Selecione esta opção para atualizar as referências do projeto para seus componentes. *Nota:* Os arquivos incluídos em um projeto não são salvos no arquivo de projeto em si. Em vez disso, uma referência aos arquivos é armazenada no projeto. Isso significa que mover ou excluir um arquivo removerá o objeto do projeto.

Anotando um projeto

A área de janela do projeto fornece inúmeras maneiras de anotar seus esforços de mineração de dados. Anotações em nível de projeto normalmente são usadas para controlar objetivos e decisões de uma "perspectiva completa", enquanto anotações de pasta ou nó fornecem detalhes adicionais. A guia Anotações fornece espaço suficiente para você documentar detalhes em nível de projeto, como a exclusão de dados com dados omissos irrecuperáveis ou hipóteses promissoras formadas durante a exploração de dados.

Para anotar um projeto:

1. Selecione a pasta do projeto na visualização CRISP-DM ou Classes.
2. Clique com o botão direito na pasta e clique em **Propriedades do Projeto**.
3. Clique na guia **Anotações**.
4. Insira palavras-chave e texto para descrever o projeto.

Propriedades e anotações de pastas

Pastas do projeto individuais (nas visualizações CRISP-DM e Classes) podem ser anotadas. Na visualização CRISP-DM, essa pode ser uma maneira extremamente efetiva de documentar os objetivos de sua organização para cada fase da mineração de dados. Por exemplo, usando a ferramenta de anotação para a pasta *Entendimento dos Negócios*, é possível incluir documentação como "O objetivo de negócios para este estudo é reduzir a perda de clientes entre clientes de alto valor." Esse texto poderia ser incluído automaticamente no relatório do projeto selecionando a opção **Incluir no Relatório**.

Para anotar uma pasta:

1. Selecione uma pasta na área de janela do projeto.
2. Clique com o botão direito na pasta e clique em **Propriedades da Pasta**.

Na visualização CRISP-DM, pastas são anotadas com uma sumarização do propósito de cada fase, bem como uma orientação sobre como concluir tarefas de mineração de dados relevantes. É possível remover ou editar qualquer uma dessas anotações.

Nome. Essa área exibe o nome do campo selecionado.

Texto de dica de ferramenta. Crie Dicas de ferramenta customizadas que serão exibidas quando você passar o mouse sobre uma pasta do projeto. Isso é útil na visualização CRISP-DM, por exemplo, para fornecer uma visão geral rápida dos objetivos de cada fase ou para marcar o status de uma fase, como "Em Andamento ou "Concluído."

Campo de anotação. Use este campo para anotações mais longas que possam ser intercaladas no relatório do projeto. A visualização CRISP-DM inclui uma descrição de cada fase da mineração de dados na anotação, mas você pode ficar à vontade para customizá-la para seu próprio projeto.

Incluir no relatório. Para incluir a anotação em relatórios, selecione **Incluir no Relatório**.

Propriedades do Objeto

É possível visualizar propriedades do objeto e escolher se você deseja incluir objetos individuais no relatório do projeto. Para acessar propriedades do objeto:

1. Clique com o botão direito em um objeto na área de janela do projeto.
2. No menu, clique em **Propriedades do Objeto**.

Nome. Essa área lista o nome do objeto salvo.

Caminho. Essa área lista a localização do objeto salvo.

Incluir no relatório. Selecione esta opção para incluir os detalhes do objeto em um relatório gerado.

Encerrando um projeto

Quando você sai do IBM SPSS Modeler ou abre um novo projeto, o arquivo de projeto existente (*.cpj*) é encerrado.

Alguns arquivos associados ao projeto (como fluxos, nós ou gráficos) podem continuar abertos. Se desejar deixar esses arquivos abertos, responda Não para a mensagem ... **Deseja salvar e fechar esses arquivos?**

Se você modificar e salvar quaisquer arquivos associados após o encerramento de um projeto, essas versões atualizadas serão incluídas no projeto na próxima vez que você abri-lo. Para evitar esse comportamento, remova o arquivo do projeto ou salve-o sob um nome de arquivo diferente.

Gerando um relatório

Uma das variáveis mais úteis dos projetos é a capacidade de gerar relatórios com base nos itens e nas anotações do projeto. Isso é um componente crítico da mineração de dados efetiva, conforme discutido em toda metodologia CRISP-DM. É possível gerar um relatório diretamente em um dos vários tipos de arquivos ou para uma janela de saída na tela para visualização imediata. Aqui, é possível imprimir, salvar ou visualizar o relatório em um navegador da web. É possível distribuir relatórios salvos para outras pessoas em sua organização.

Geralmente relatórios são gerados a partir de arquivos de projeto várias vezes durante o processo de mineração de dados para distribuição para aqueles envolvidos no projeto. O relatório escolhe informações sobre os objetos referenciados a partir do arquivo de projeto, bem como quaisquer anotações criadas. É possível criar relatórios com base em sua visualização Classes ou visualização CRISP-DM.

Para gerar um relatório:

1. Selecione a pasta do projeto na visualização CRISP-DM ou Classes.
2. Clique com o botão direito na pasta e clique em **Relatório do Projeto**.
3. Especifique as opções de relatório e clique em **Gerar Relatório**.

As opções na caixa de diálogo de relatório fornecem várias maneiras de você gerar o tipo de relatório necessário:

Nome de saída. Especifique o nome da janela de saída se escolher enviar a saída do relatório para a tela. É possível especificar um nome customizado ou deixar o IBM SPSS Modeler nomear a janela automaticamente para você.

Saída para tela. Selecione esta opção para gerar e exibir o relatório em uma janela de saída. Observe que você tem a opção de exportar o relatório para vários tipos de arquivo a partir da janela de saída.

Saída para arquivo. Selecione esta opção para gerar e salvar o relatório como um arquivo do tipo especificado na lista Tipo de Arquivo.

Nome do arquivo. Especifique um nome do arquivo para o relatório gerado. Arquivos são salvos por padrão no diretório `\bin` do IBM SPSS Modeler. Use o botão de reticências (...) para especificar uma localização diferente.

Tipo de arquivo. Os tipos de arquivo disponíveis são:

- **Documento HTML.** O relatório é salvo como um único arquivo HTML. Se seu relatório contiver gráficos, eles serão salvos como arquivos PNG e referenciados pelo arquivo HTML. Durante a publicação do seu relatório na Internet, certifique-se de fazer upload do arquivo HTML e de quaisquer imagens referenciadas por ele.
- **Documento de texto.** O relatório é salvo como um único arquivo de texto. Se seu relatório contiver gráficos, somente as referências de caminho e nome do arquivo serão incluídas no relatório.
- **Documento do Microsoft Word.** O relatório é salvo como um único documento com quaisquer gráficos integrados diretamente ao documento.
- **Documento do Microsoft Excel.** O relatório é salvo como uma única planilha com quaisquer gráficos integrados diretamente à planilha.
- **Documento do Microsoft PowerPoint.** Cada fase é mostrada em um novo slide. Quaisquer gráficos serão integrados diretamente aos slides do PowerPoint.
- **Objeto de saída.** Quando aberto no IBM SPSS Modeler, esse arquivo (.cou) é igual à opção **Saída para tela** no grupo **Formato do Relatório**.

Nota: Para exportar para um arquivo do Microsoft Office, deve-se ter o aplicativo correspondente instalado.

Título. Especifique um título para o relatório.

Estrutura de relatório. Selecione **CRISP-DM** ou **Classes**. A visualização CRISP-DM fornece um relatório de status com uma sinopse "geral", bem como detalhes sobre cada fase da mineração de dados. A visualização Classes é uma visualização baseada em objeto mais apropriada para o rastreamento interno de dados e fluxos.

Autor. O nome de usuário padrão é exibido, mas é possível alterá-lo.

Inclusões de relatório. Selecione um método para incluir objetos no relatório. Selecione **todas as pastas e objetos** para incluir todos os itens incluídos no arquivo de projeto. Também é possível incluir itens

dependendo se **Incluir no Relatório** estiver selecionada nas propriedades do objeto. Alternativamente, para verificar itens não relatados, é possível escolher incluir somente itens marcados para exclusão (em que **Incluir no Relatório** não é marcada).

Selecionar. Esta opção permite fornecer atualizações de projeto selecionando somente **itens recentes** no relatório. Alternativamente, é possível controlar problemas mais antigos e talvez não resolvidos configurando parâmetros para **itens antigos**. Selecione **todos os itens** para indeferir tempo como um parâmetro para o relatório.

Ordenar por. É possível selecionar uma combinação das seguintes características do objeto para ordená-las da seguinte forma:

- **Tipo.** Agrupe objetos por tipo.
- **Nome.** Organize objetos em ordem alfabética.
- **Data de inclusão.** Ordene objetos usando a data em que eles foram incluídos no projeto.

Salvando e exportando relatórios gerados

Um relatório gerado para a tela é exibido em uma nova janela de saída. Quaisquer gráficos incluídos no relatório são exibidos como imagens sequenciais.

Terminologia do relatório

O número total de nós em cada fluxo é listado dentro do relatório. Os números são mostrados sob os seguintes títulos, que usam a terminologia do IBM SPSS Modeler, e não a terminologia do CRISP-DM:

- **Leitores de dados.** Nós de origem.
- **Gravadores de dados.** Nós de exportação.
- **Construtores de modelo.** Nós de construção ou modelagem.
- **Aplicadores de modelo.** Modelos gerados, também conhecidos como nuggets.
- **Construtores de saída.** Nós de gráfico ou saída.
- **Outro.** Quaisquer outros nós relacionados ao projeto. Por exemplo, aqueles que estão disponíveis na guia Operações de Campo ou Operações de Registro na Paleta de Nós.

Para salvar um relatório:

1. No menu Arquivo, clique em **Salvar**.
2. Especifique um nome de arquivo.
O relatório é salvo como um objeto de saída.

Para exportar um relatório:

3. No menu Arquivo, clique em **Exportar** e o tipo de arquivo para o qual deseja exportar.
4. Especifique um nome de arquivo.

O relatório é salvo no formato que você escolheu.

É possível exportar para os seguintes tipos de arquivo:

- HTML
- Texto
- Microsoft Word
- Microsoft Excel
- Microsoft PowerPoint

Nota: Para exportar para um arquivo do Microsoft Office, deve-se ter o aplicativo correspondente instalado.

Use os botões na parte superior da janela para:

- Imprimir o relatório.
- Visualizar o relatório como HTML em um navegador da web externo.

Capítulo 13. Customizando o IBM SPSS Modeler

Customizando opções do IBM SPSS Modeler

Há inúmeras operações que podem ser executadas para customizar o IBM SPSS Modeler para suas necessidades. Primeiro, essa customização consiste na configuração de opções do usuário específico, como alocação de memória, diretórios padrão e uso de som e cor. Também é possível customizar a Paleta de Nós localizada na parte inferior da janela IBM SPSS Modeler.

Configurando opções do IBM SPSS Modeler

Há várias maneiras de customizar e configurar opções para o IBM SPSS Modeler:

- Configure opções do sistema, como uso de memória e código de idioma, clicando em **Opções do Sistema** no menu **Ferramentas > Opções**.
- Configure opções do usuário, como fontes e cores de exibição, clicando em **Opções do Usuário** no menu **Ferramentas > Opções**.
- Especifique a localização dos aplicativos que trabalham com IBM SPSS Modeler clicando em **Aplicativos Auxiliares** no menu **Ferramentas > Opções**.
- Especifique os diretórios padrão usados no IBM SPSS Modeler clicando em **Configurar Diretório** ou **Configurar Diretório do Servidor** no menu Arquivo.

Também é possível configurar opções que se apliquem a alguns ou todos os seus fluxos. Consulte o tópico “Configurando opções para fluxos” na página 39 para obter mais informações.

Opções de Sistema

É possível especificar o idioma ou código de idioma preferencial para o IBM SPSS Modeler clicando em **Opções do Sistema** no menu **Ferramentas > Opções**. Aqui também é possível configurar o uso de memória máximo para SPSS Modeler. Observe que as mudanças feitas nessa caixa de diálogo não entrarão em vigor até você reiniciar o SPSS Modeler.

Máximo de memória. Selecione para impor um limite em megabytes no uso de memória do IBM SPSS Modeler. Em algumas plataformas, o SPSS Modeler limita o tamanho do seu processo para reduzir a taxa em computadores com recursos limitados ou cargas pesadas. Se você estiver lidando com grandes quantidades de dados, isso pode causar um erro de "falta de memória". É possível facilitar a carga de memória especificando um novo limite.

Por exemplo, a tentativa de exibir uma árvore de decisão muito grande pode causar um erro de memória; nesse caso, é recomendado aumentar a memória para o valor máximo de 4096 MB. Em casos como esses, nos quais é provável que você esteja processando grandes quantidades de dados, após você aumentar a concessão de memória, encerre o SPSS Modeler e inicie-o a partir de uma linha de comandos para assegurar que a quantidade máxima de memória seja usada durante o processamento dos seus dados.

Para iniciar a partir de uma linha de comandos (supondo que o SPSS Modeler esteja instalado na localização padrão), em uma janela Prompt de Comandos, insira o seguinte:

```
C:\Program Files\IBM\SPSS\Modeler\17.1\bin\modelerclient.exe" -J-Xss4096M
```

Usar código de idioma do sistema. Esta opção fica selecionada por padrão e configurada para inglês (Estados Unidos). Cancele sua seleção para especificar outro idioma a partir da lista de idiomas e códigos de idioma disponíveis.

Gerenciando memória

Além da configuração **Máximo de Memória** especificada na caixa de diálogo Opções do Sistema, há várias maneiras de se otimizar o uso de memória:

- Ajuste a opção **Máximo de Membros para Campos Nominais** na caixa de diálogo de propriedades do fluxo. Essa opção especifica um número máximo de membros para campos nominais após o qual o nível de medição do campo se torna *Sem Tipo*. Consulte o tópico “Configurando opções gerais para fluxos” na página 40 para obter mais informações.
- Force o IBM SPSS Modeler a liberar memória clicando no canto inferior direito da janela, onde a memória que o IBM SPSS Modeler está usando e a quantidade alocada são exibidas (*xxMB / xxMB*). Um clique nessa região a deixa com uma tonalidade mais escura, após o qual as figuras de alocação de memória caem. Após a região retornar para sua cor regular, o IBM SPSS Modeler liberou toda memória possível.

Configurando diretórios padrão

É possível especificar o diretório padrão usado para navegadores de arquivo e saída selecionando **Configurar Diretório** ou **Configurar Diretório do Servidor** no menu Arquivo.

- **Configurar Diretório.** É possível usar essa opção para configurar o diretório ativo. O diretório ativo padrão é baseado no caminho da instalação para sua versão do IBM SPSS Modeler ou a partir do caminho da linha de comandos usado para ativar o IBM SPSS Modeler. No modo local, o diretório ativo é o caminho usado para todas as operações do lado do cliente e arquivos de saída (se eles forem referenciados com caminhos relativos).
- **Configurar Diretório do Servidor.** A opção Configurar Diretório do Servidor no menu Arquivo é ativada sempre que há uma conexão do servidor remoto. Use essa opção para especificar o diretório padrão para todos os arquivos do servidor e arquivos de dados especificados para entrada ou saída. O diretório do servidor padrão é `$CLEO/data`, em que `$CLEO` é o diretório no qual a versão de Servidor do IBM SPSS Modeler está instalada. Usando a linha de comandos, também é possível substituir esse padrão usando o flag `-server_directory` com o argumento de linha de comandos `modelerclient`.

Configurando opções do usuário

É possível configurar opções gerais para IBM SPSS Modeler selecionando **Opções do Usuário** no menu **Ferramentas > Opções**. Essas opções se aplicam a todos os fluxos usados no IBM SPSS Modeler.

Os tipos de opções a seguir podem ser configurados clicando na guia correspondente:

- Opções de notificação, como sobrescrição de modelo e mensagens de erro.
- Opções de exibição, como gráfico e cores do plano de fundo.
- Opções de exibição de cores de sintaxe.
- Opções de exportação de PMML usadas durante a exportação de modelos para Predictive Model Markup Language (PMML).
- Informações do usuário ou autor, como nome, iniciais e endereço de email. Essas informações podem ser exibidas na guia Anotações para nós e para outros objetos que você criar.

Para configurar opções específicas do fluxo, como separadores decimais, formatos de data e hora, otimização, layout de fluxo e scripts de fluxo, use a caixa de diálogo Propriedades do Fluxo disponível nos menus Arquivo e Ferramentas.

Configurando opções de notificação

Usando a guia Notificações da caixa de diálogo Opções de Usuário, é possível configurar diversas opções referentes à ocorrência e ao tipo de avisos e janelas de confirmação no IBM SPSS Modeler. Também é possível especificar o comportamento das guias Saídas e Modelos na área de janela de gerenciadores quando novos modelos e saída são gerados.

Mostrar diálogo de feedback de execução de fluxo Selecione para exibir uma caixa de diálogo que inclua um indicador de progresso quando o fluxo estiver em execução por três segundos. A caixa de diálogo também inclui detalhes dos objetos de saída criados pelo fluxo.

- **Fechar diálogo na conclusão** Por padrão, a caixa de diálogo fecha quando o fluxo termina a execução. Limpe essa caixa de seleção se desejar que a caixa de diálogo permaneça visível quando o fluxo terminar.

Avisar quando um nó sobrescrever um arquivo Selecione para avisar com uma mensagem de erro quando operações de nó sobrescreverem um arquivo existente.

Avisar quando um nó sobrescrever uma tabela de base de dados Selecione para avisar com uma mensagem de erro quando operações de nó sobrescreverem uma tabela de base de dados existente.

Notificações de som

Use a lista para especificar se você receberá notificações de som quando ocorrer um evento ou erro. Há inúmeros sons disponíveis. Use o botão Play (altofalante) para reproduzir um som selecionado. Use o botão de reticências (...) para navegar e selecionar um som.

Nota: Os arquivos .wav usados para criar sons no IBM SPSS Modeler são armazenados no diretório /media/sounds de sua instalação.

- **Silenciar todos os sons** Selecione para desativar a notificação de som para todos os eventos.

Notificações visuais

As opções nesse grupo são usadas para especificar o comportamento das guias Saídas e Modelos na área de janela de gerenciadores na parte superior direita da exibição quando novos itens são gerados. Selecione **Novo Modelo** ou **Nova Saída** na lista para especificar o comportamento da guia correspondente.

A opção a seguir está disponível para **Novo Modelo**:

Substituir modelo anterior Se selecionada (padrão), sobrescreve um modelo existente desse fluxo na guia Modelos e na tela de fluxo. Se essa caixa estiver desmarcada, o modelo será incluído nos modelos existentes na guia e na tela. Observe que essa configuração é sobrescrita pela configuração de substituição de modelo em uma ligação de modelo.

A opção a seguir está disponível para **Nova Saída**:

Avisar quando saídas excederem [n] Selecione se deseja exibir um aviso quando o número de itens na guia Saídas exceder uma quantidade pré-especificada. A quantidade padrão é 20; no entanto, é possível alterá-la se necessário.

As opções a seguir estão disponíveis em todos os casos:

Guia Selecionar Escolha se você alternará para a guia Saídas ou Modelos quando o objeto correspondente for gerado enquanto o fluxo é executado.

- Selecione **Sempre** para alternar para a guia correspondente na área de janela de gerenciadores.
- Selecione **Se Gerado por Fluxo Atual** para alternar para a guia correspondente somente para objetos gerados pelo fluxo atualmente visível na tela.
- Selecione **Nunca** para impedir que o software alterne para a guia correspondente para notificá-lo de saídas ou modelos gerados.

Guia Atualizar Selecione se deseja atualizar a guia Saídas ou Modelos na área de janela de gerenciadores quando novas saídas ou modelos tiverem sido gerados.

- Selecione **Se Não Selecionado** para atualizar a guia correspondente (se ela ainda não estiver selecionada) sempre que novos objetos forem gerados na área de janela de gerenciadores.
- Selecione **Nunca** para impedir que o software atualize a guia correspondente para notificá-lo de objetos gerados.

Rolar paleta para tornar visível (somente **Novo Modelo**). Selecione se deseja rolar automaticamente a guia Modelos na área de janela de gerenciadores para deixar o modelo mais recente visível.

- Selecione **Sempre** para ativar a rolagem.
- Selecione **Se Gerado por Fluxo Atual** para rolar somente para objetos gerados pelo fluxo atualmente visível na tela.
- Selecione **Nunca** para impedir que o software role automaticamente a guia Modelos.

Abrir Janela (somente **Nova Saída**). Selecione se deseja abrir automaticamente uma janela de saída na geração.

- Selecione **Sempre** para sempre abrir uma nova janela.
- Selecione **Se Gerado por Fluxo Atual** para abrir uma nova janela para saída gerada pelo fluxo atualmente visível na tela.
- Selecione **Nunca** para impedir que o software abra as novas janelas automaticamente para a saída gerada.

Para reverter para as configurações padrão do sistema para essa guia, clique em **Valores Padrão**.

Configurando opções de exibição

Usando a guia Exibição da caixa de diálogo Opções de Usuário, é possível configurar opções para a exibição de fontes e cores no IBM SPSS Modeler.

Mostrar diálogo bem-vindo na inicialização Selecione para que a caixa de diálogo bem-vindo seja exibida na inicialização. A caixa de diálogo bem-vindo tem opções para ativar o tutorial de exemplos do aplicativo, abrir um fluxo de demonstração ou um fluxo ou projeto existente ou criar um novo fluxo.

Mostrar marcações de fluxo e SuperNode Se selecionada, faz com que a marcação (se houver) nos fluxos e SuperNodes seja exibida por padrão. A marcação inclui comentários de fluxo, ligações de modelo e destaque de ramificações de escoragem.

Fontes e Cores Padrão (efetivo na reinicialização) Opções nesta caixa de controle são usadas para especificar o design da tela do IBM SPSS Modeler, o esquema de cores e o tamanho das fontes exibidas. As opções que você seleciona aqui não entram em vigor até você fechar e reiniciar o IBM SPSS Modeler.

- **Aparência** Selecione um esquema de cores e um design de tela padrão. É possível escolher entre:
 - SPSS Padrão** (o padrão), um design comum em produtos IBM SPSS.
 - SPSS Clássico**, um design familiar para os usuários de versões anteriores do IBM SPSS Modeler.
 - Windows**, um design do Windows que pode ser útil para um contraste aumentado na tela de fluxo e nas paletas.
- **Tamanho de fonte padrão para nós** Especifique um tamanho de fonte para ser usado nas paletas de nó e para os nós que são exibidos na tela de fluxo.
- **Especificar fonte de largura fixa** Para selecionar uma fonte de largura fixa, e o **Tamanho** da fonte associado para uso em controles de expressão do CLEM e script, marque esta caixa de seleção. A fonte padrão é Monoespaçamento Simple; clique em **Alterar...** para exibir uma lista de outras fontes que podem ser selecionadas.

Nota: É possível configurar o tamanho dos ícones de nó para um fluxo na área de janela Layout da guia Opções da caixa de diálogo de propriedades do fluxo. No menu principal, escolha **Ferramentas > Propriedades do Fluxo > Opções > Layout**.

Cores Customizadas Esta tabela lista as cores atualmente selecionadas que são usadas para vários itens de exibição. Para cada um dos itens que são listados na tabela, é possível alterar a cor atual dando um clique duplo na linha correspondente na coluna **Cor** e selecionando uma cor da lista. Para especificar uma cor customizada, role para o fim da lista e clique na entrada **Cor...**

Ordem de Cor de Categoria do Gráfico Esta tabela lista as cores atualmente selecionadas que são usadas para exibição em gráficos recém-criados. A ordem das cores reflete a ordem em que elas são usadas no gráfico. Por exemplo, se um campo nominal usado como uma sobreposição de cor contiver quatro valores exclusivos, somente as quatro primeiras cores listadas aqui serão usadas. Para cada um dos itens que são listados na tabela, é possível alterar a cor atual dando um clique duplo na linha correspondente na coluna **Cor** e selecionando uma cor da lista. Para especificar uma cor customizada, role para o fim da lista e clique na entrada **Cor...** As mudanças feitas aqui não afetam gráficos criados anteriormente.

Para reverter para as configurações padrão do sistema para essa guia, clique em **Valores Padrão**.

Configurando opções de exibição de sintaxe

Usando a guia Sintaxe da caixa de diálogo Opções do Usuário, é possível configurar opções para os atributos de fonte e exibir cores nos scripts que você criar no IBM SPSS Modeler.

Destaque da Sintaxe. Esta tabela lista as cores atualmente selecionadas usadas para vários itens de sintaxe, incluindo fonte e janela na qual ela é exibida. Para cada um dos itens listados na tabela, é possível alterar a cor clicando na lista suspensa correspondente na linha e selecionando uma cor da lista. Além disso, para itens de fonte, é possível escolher incluir ênfase para negrito e itálico.

Visualizar. Esta tabela mostra uma exibição de sintaxe de exemplo que usa os atributos de cores e fontes selecionados na tabela **Destaque da Sintaxe**. Esta visualização é atualizada assim que você altera qualquer seleção.

Clique em **Valores Padrão** para reverter para as configurações padrão do sistema para esta guia.

Configurando opções de exportação de PMML

Na guia PMML, é possível controlar como o IBM SPSS Modeler exporta modelos para Predictive Model Markup Language (PMML). Consulte o tópico “Importando e exportando modelos como PMML” na página 174 para obter mais informações.

Exportar PMML. Aqui é possível configurar variações de PMML que funcionam melhor com seu aplicativo de resposta.

- Selecione **Com Extensões** para permitir extensões de PMML para casos especiais nos quais não há um PMML padrão equivalente. Observe que na maioria dos casos isso produzirá o mesmo resultado que o PMML padrão.
- Selecione **Como PMML Padrão...** para exportar PMML que seja o mais próximo possível do PMML padrão.

Opções de PMML Padrão. Quando a opção **Como PMML Padrão...** está selecionada, é possível escolher uma das duas maneiras válidas de exportar modelos de regressão logística e linear:

- Como **modelos PMML <GeneralRegression>**
- Como **modelos PMML <Regression>**

Para obter mais informações sobre PMML, consulte o website Grupo de Mineração de Dados em <http://www.dmg.org>.

Configurando informações do usuário

Informações do Usuário/Autor. As informações inseridas aqui podem ser exibidas na guia Anotações dos nós e outros objetos que você criar.

Customizando a Paleta de Nós

Fluxos são construídos usando nós. A Paleta de Nós na parte inferior da janela do IBM SPSS Modeler contém todos os nós que podem ser usados na construção do fluxo. Consulte o tópico “Paleta de Nós” na página 14 para obter mais informações.

É possível reorganizar a Paleta de Nós de duas maneiras:

- Customizar o Gerenciador de Paleta. Consulte o tópico “Customizando o Gerenciador de Paleta” para obter mais informações.
- Alterar como as guias da paleta contendo subpaletas são exibidas na Paleta de Nós. Consulte o tópico “Criando uma subpaleta” na página 192 para obter mais informações.

Customizando o Gerenciador de Paleta

O Gerenciador de Paleta pode ser customizado para acomodar seu uso do IBM SPSS Modeler. Por exemplo, se você analisar frequentemente dados de séries temporais a partir de um banco de dados, talvez você queira se certificar de que o nó de Origem da Base de Dados, nó de Intervalos de Tempo, nó de Séries Temporais e nó de Gráfico de Plotagem de Tempo estejam disponíveis juntos em uma única guia da paleta. O Gerenciador de Paleta permite fazer esses ajustes facilmente criando suas guias da paleta customizadas na Paleta de Nós.

O Gerenciador de Paleta permite executar várias tarefas:

- Controlar quais guias da paleta são mostradas na Paleta de Nós abaixo da tela de fluxo.
- Alterar a ordem em que as guias da paleta são mostradas na Paleta de Nós.
- Criar e editar suas próprias guias da paleta e quaisquer subpaletas associadas.
- Editar as seleções de nó padrão em sua guia Favoritos.

Para acessar o Gerenciador de Paleta, no menu Ferramentas, clique em **Gerenciar Paletas**.

Nome da Paleta. Cada guia da paleta disponível, independentemente de ser ou não mostrada na Paleta de Nós, é listada. Isso inclui quaisquer guias da paleta que você criou. Consulte o tópico “Criando uma guia de paleta” para obter mais informações.

Número de nós. O número de nós exibido em cada guia da paleta. Um número mais alto aqui significa que você pode achar mais conveniente criar subpaletas para dividir os nós na guia. Consulte o tópico “Criando uma subpaleta” na página 192 para obter mais informações.

Mostrado. Selecione este campo para exibir a guia da paleta na Paleta de Nós. Consulte o tópico “Exibindo guias de paleta na Paleta de Nós” na página 191 para obter mais informações.

Subpaletas. Para selecionar subpaletas para exibir em uma guia da paleta, destaque o **Nome da Paleta** necessário e clique nesse botão para exibir a caixa de diálogo Subpaletas. Consulte o tópico “Criando uma subpaleta” na página 192 para obter mais informações.

Restaurar Padrões. Para remover completamente todas as mudanças e adições feitas nas paletas e subpaletas e retornar para as configurações de paleta padrão, clique nesse botão.

Criando uma guia de paleta

Para criar uma guia de paleta customizada:

1. No menu Ferramentas, abra o Gerenciador de Paleta.
2. À direita da coluna *Mostrado*, clique no botão Incluir Paleta; a caixa de diálogo Criar/Editar Paleta é exibida.
3. Digite um **Nome de Paleta** exclusivo.
4. Na área **Nós Disponíveis**, selecione o nó a ser incluído na guia da paleta.

5. Clique no botão Seta à Direita Incluir Nó para mover o nó destacado para the área **Nós Selecionados**. Repita até ter incluído todos os nós desejados.
Após você ter incluído todos os nós necessários, é possível alterar a ordem em que eles são exibidos na guia de paleta:
6. Use os botões de seta simples para mover um nó uma linha para cima ou para baixo.
7. Use os botões de seta de linha para mover um nó para a parte inferior ou parte superior da lista.
8. Para remover um nó de uma paleta, destaque o nó e clique no botão Excluir à direita da área **Nós Selecionados**.

Exibindo guias de paleta na Paleta de Nós

Pode haver opções disponíveis dentro do IBM SPSS Modeler que você nunca usa; nesse caso, é possível usar o Gerenciador de Paleta para ocultar as guias contendo esses nós.

Para selecionar quais guias devem ser mostradas na Paleta de Nós:

1. No menu Ferramentas, abra o Gerenciador de Paleta.
2. Usando as caixas de seleção na coluna *Mostrado*, selecione se deseja incluir ou ocultar cada guia da paleta.

Para remover permanentemente uma guia de paletas da Paleta do Nó, destaque o nó e clique no botão Excluir à direita da coluna *Mostrado*. Após a exclusão, uma guia da paleta não pode ser recuperada.

Nota: Não é possível excluir a paleta padrão fornecida com o IBM SPSS Modeler, exceto a guia Favoritos.

Alterando a ordem de exibição na Paleta de Nós

Após ter selecionado quais guias da paleta deseja exibir, é possível alterar a ordem em que elas são exibidas na Paleta de Nós:

1. Use os botões de seta simples para mover uma guia da paleta uma linha para cima ou para baixo. Movê-las para cima as move para a esquerda da Paleta de Nós e vice-versa.
2. Use os botões de seta de linha para mover uma guia da paleta para cima ou para baixo na lista. Aquelas na parte superior da lista serão mostradas à esquerda da Paleta de Nós.

Exibindo subpaletas em uma guia de paleta

Da mesma maneira que é possível controlar quais guias de paleta são exibidas na Paleta de Nós, é possível controlar quais subpaletas estão disponíveis a partir da guia de paleta pai.

Para selecionar várias subpaletas para exibição em uma guia da paleta:

1. No menu Ferramentas, abra o Gerenciador de Paleta.
2. Selecione a paleta requerida.
3. Clique no botão Subpaletas; a caixa de diálogo Subpaletas é exibida.
4. Usando as caixas de seleção na coluna *Mostrado*, selecione se deseja incluir cada subpaleta na guia de paleta. A subpaleta **Todos** é sempre mostrada e não pode ser excluída.
5. Para remover permanentemente uma subpaleta da guia de paletas, destaque a subpaleta e clique no botão Excluir à direita da coluna *Mostrado*.

Nota: Não é possível excluir as subpaletas padrão fornecidas com a guia de paleta Modelagem.

Alterando a ordem de exibição na guia de paleta

Após você ter selecionado quais subpaletas deseja exibir, é possível alterar a ordem em que elas são exibidas na guia de paleta pai:

1. Use os botões de seta simples para mover uma subpaleta uma linha para cima ou para baixo.

2. Use os botões de seta de linha para mover uma subpaleta para a parte inferior ou parte superior da lista.

As subpaletas que você cria são exibidas na Paleta do Nó quando você seleciona sua guia de paleta pai. Consulte o tópico “Alterando uma visualização da guia paleta” para obter mais informações.

Criando uma subpaleta

Como é possível incluir qualquer nó existente nas guias de paleta customizadas que você cria, é provável que você selecionará mais nós do que o que pode ser facilmente exibido na tela sem rolagem. Para evitar essa rolagem, é possível criar subpaletas nas quais você coloca os nós escolhidos na guia da paleta. Por exemplo, se você criasse uma guia da paleta contendo os nós que devem ser usados com mais frequência para criar seus fluxos, seria possível criar quatro subpaletas que dividiriam as seleções por nó de origem, operações de campo, modelagem e saída.

Nota: Só é possível selecionar nós de subpaleta entre aqueles incluídos na guia da paleta pai.

Para criar uma subpaleta:

1. No menu Ferramentas, abra o Gerenciador de Paleta.
2. Selecione a paleta na qual deseja incluir subpaletas.
3. Clique no botão Subpaletas; a caixa de diálogo Subpaletas é exibida.
4. À direita da coluna *Mostrado*, clique no botão Incluir Subpaleta; a caixa de diálogo Criar/Editar Subpaleta é exibida.
5. Digite um **Nome de Subpaleta** exclusivo.
6. Na área **Nós Disponíveis**, selecione o nó a ser incluído na subpaleta.
7. Clique no botão Seta à Direita Incluir Nó para mover um nó selecionado para a área **Nós Selecionados**.
8. Quando tiver incluído os nós necessários, clique em **OK** para retornar para a caixa de diálogo Subpaletas.

As subpaletas que você cria são exibidas na Paleta do Nó quando você seleciona sua guia de paleta pai. Consulte o tópico “Alterando uma visualização da guia paleta” para obter mais informações.

Alterando uma visualização da guia paleta

Devido ao grande número de nós disponíveis no IBM SPSS Modeler, eles podem não estar todos visíveis em telas menores, sem rolar para a esquerda ou para a direita na Paleta de Nós; isso é perceptível principalmente na guia Paleta de Modelagem. Para reduzir a necessidade de rolagem, é possível escolher a exibição somente dos nós contidos em uma subpaleta (onde disponível). Consulte o tópico “Criando uma subpaleta” para obter mais informações.

Para alterar os nós mostrados em uma guia paleta, selecione a guia paleta e, em seguida, no menu à esquerda, selecione a exibição de todos os nós ou apenas daqueles que estão em uma subpaleta específica.

Gerenciamento do nó CEMI

Agora o CEMI foi descontinuado e substituído pelo CLEF, que oferece um conjunto de variáveis fácil de usar e muito mais flexível. Para obter mais informações, consulte *IBM SPSS Modeler 17.1 CLEF Developer's Guide* fornecido com esta liberação.

Capítulo 14. Considerações de desempenho para fluxos e nós

É possível projetar seus fluxos para maximizar o desempenho organizando os nós na configuração mais eficiente, ativando caches de nó quando apropriado e prestando atenção em outras considerações conforme detalhado nesta seção.

Com exceção das considerações discutidas aqui, melhorias de desempenho adicionais e mais substanciais normalmente podem ser obtidas com o uso efetivo de seu banco de dados, principalmente por meio de otimização de SQL.

Ordem dos nós

Mesmo quando você não está usando otimização de SQL, a ordem dos nós em um fluxo pode afetar o desempenho. O objetivo geral é minimizar o processamento do recebimento de dados; portanto, quando você tiver nós que reduzem a quantidade de dados, coloque-os próximo do início do fluxo. IBM SPSS Modeler Server pode aplicar algumas regras de reordenação automaticamente durante a compilação para apresentar certos nós quando isso for comprovadamente seguro. (Essa variável fica ativada por padrão. Verifique com seu administrador do sistema para se certificar de que ela esteja ativada em sua instalação.)

Ao usar otimização de SQL, você deseja maximizar sua disponibilidade e eficiência. Como a otimização é parada quando o fluxo contém uma operação que não pode ser executada no banco de dados, é melhor agrupar operações otimizadas por SQL no início do fluxo. Essa estratégia mantém mais do processamento no banco de dados e, então, menos dados são executados no IBM SPSS Modeler.

As operações a seguir podem ser executadas na maioria dos bancos de dados. Tente agrupá-las no *início* do fluxo:

- Mesclar por chave (junção)
- Selecionar
- Agregado
- Classificar
- Amostra
- Anexar
- Operações distintas no modo *incluir*, no qual todos os campos são selecionados
- Operações de preenchimento
- Operações básicas de derivação usando manipulação de sequência ou aritmética padrão (dependendo de quais operações são suportadas pelos bancos de dados)
- Configurar para flag

As operações a seguir não podem ser executadas na maioria dos bancos de dados. Elas devem ser colocadas no fluxo *após* as operações na lista precedente:

- Operações em quaisquer dados que não são do banco de dados, como arquivos simples
- Mesclar por ordem
- Equilíbrio
- Operações distintas no modo *descartar* ou onde somente um subconjunto de campos é selecionado como distinto
- Qualquer operação que requer acesso a dados de registros além daqueles sendo processados
- Derivações de campo de contagem e estado
- Operações do nó de histórico

- Operações envolvendo funções @" (séries temporais)
 - Modos de verificação de tipo *Avisar* e *Interromper*
 - Construção, aplicação e análise de modelo
- Nota:* Árvores de decisão, conjuntos de regras, regressão linear e modelos gerados por fator podem gerar SQL e, portanto, ser enviadas de volta para o banco de dados.
- Saída de dados para qualquer lugar além do mesmo banco de dados que está processando os dados

Caches de nó

Para otimizar a execução do fluxo, é possível configurar um **cache** em um nó não terminal. Quando você configura um cache em um nó, o cache é preenchido com os dados que são passados pelo nó na próxima vez que você executa o fluxo de dados. Daí em diante, os dados são lidos do cache (que é armazenado em disco em um diretório temporário) e não da origem de dados.

O armazenamento em cache é mais útil após uma operação demorada, como uma ordenação, mesclagem ou agregação. Por exemplo, suponha que você tenha um nó de origem configurado para ler dados de vendas de um banco de dados e um nó agregado que resuma as vendas por localização. É possível configurar um cache no nó agregado em vez de no nó de origem, pois você deseja que o cache armazene os nós agregados, e não o conjunto de dados inteiro.

Nota: O armazenamento em cache em nós de origem, que simplesmente armazena uma cópia dos dados originais como eles são lidos no IBM SPSS Modeler, não melhora o desempenho na maioria das circunstâncias.

Nós com o armazenamento em cache ativado são exibidos com um ícone de documento pequeno no canto superior direito. Quando os dados são armazenados em cache no nó, o ícone de documento fica verde.

Para ativar um cache

1. Na tela de fluxo, clique com o botão direito no nó e clique em **Cache** no menu.
2. No submenu de armazenamento em cache, clique em **Ativar**.
3. É possível desativar o cache clicando com o botão direito no nó e clicando em **Desativar** no submenu de armazenamento em cache.

Nós de armazenamento em cache em um banco de dados

Para fluxos executados em um banco de dados, dados podem ser armazenados em cache no midstream para uma tabela temporária no banco de dados em vez de para o sistema de arquivo. Quando combinado com otimização de SQL, isso pode resultar em ganhos significativos no desempenho. Por exemplo, a saída de um fluxo que mescla várias tabelas para criar uma visualização de mineração de dados pode ser armazenada em cache e reutilizada conforme necessário. Ao gerar SQL automaticamente para todos os nós de recebimento de dados, o desempenho pode ser melhorado.

Durante o uso do armazenamento em cache do banco de dados com sequências de caracteres com mais de 255 caracteres, assegure que haja um envio de dados de nó Tipo a partir do nó de armazenamento em cache e de que os valores de campo sejam lidos, ou configure o comprimento da sequência de caracteres por meio do parâmetro `default_sql_string_length` no arquivo `options.cfg`. Isso assegura que a coluna correspondente na tabela temporária seja configurada para a largura correta para acomodar as sequências de caracteres.

Para aproveitar a vantagem do armazenamento em cache do banco de dados, a otimização de SQL e o armazenamento em cache do banco de dados devem estar ativados. Observe que as configurações de otimização do Servidor substituem as do Cliente. Consulte o tópico “Configurando opções de otimização para fluxos” na página 42 para obter mais informações.

Com o armazenamento em cache do banco de dados ativado, basta clicar com o botão direito em qualquer nó não terminal para armazenar dados em cache nesse ponto e, automaticamente, o cache será criado diretamente no banco de dados na próxima vez que o fluxo for executado. Se o armazenamento em cache do banco de dados ou a otimização de SQL não estiverem ativados, o cache será gravado no sistema de arquivos.

Nota: Os bancos de dados a seguir suportam tabelas temporárias para o propósito de armazenamento em cache: DB2, Netezza, Oracle, SQL Server e Teradata. Outros bancos de dados usarão uma tabela normal para armazenamento em cache do banco de dados. O código SQL pode ser customizado para bancos de dados específicos - entre em contato com os Serviços para obter assistência.

Desempenho: Nós de Processamento

Ordem. O nó de Ordem deve ler o conjunto inteiro de dados de entrada para que ele possa ser ordenado. Os dados são armazenados na memória até um certo limite, e o excesso é estendido para o disco. O algoritmo de ordenação é um algoritmo de combinação: dados são lidos na memória até um certo limite e ordenados usando um algoritmo de ordenação rápida híbrido. Se todos os dados se ajustarem à memória, a ordenação foi concluída. Caso contrário, um algoritmo de ordenação por mesclagem será aplicado. Os dados ordenados são gravados em arquivo e o chunk de dados seguinte é lido na memória, ordenado e gravado em disco. Isso é repetido até que todos os dados tenham sido lidos; em seguida, os chunks ordenados são mesclados. A mesclagem pode requerer aprovações repetidas para os dados armazenados em disco. No uso de pico, o nó de Ordem terá duas cópias completas dos dados configurados no disco: ordenados e não ordenados.

O tempo de execução geral do algoritmo está na ordem de $N \cdot \log(N)$, em que N é o número de registros. A ordenação na memória é mais rápida que a mesclagem do disco, portanto, o tempo de execução pode ser reduzido pela alocação de mais memória para a ordenação. O algoritmo aloca para si mesmo uma fração de RAM física controlada pela opção de configuração *Multiplificador de Uso de Memória* do IBM SPSS Modeler Server. Para aumentar a memória usada para ordenação, forneça mais RAM física ou aumente esse valor. Observe que quando a proporção de memória usada excede o conjunto de trabalhos do processo, de modo que parte da memória seja paginada para o disco, o desempenho é degradado, pois o padrão de acesso à memória do algoritmo de ordenação na memória é aleatório e pode causar paginação excessiva. O algoritmo de ordenação é usado por vários nós além do nó de Ordem, mas as mesmas regras de desempenho se aplicam.

Categorização. O nó de Categorização lê o conjunto de dados de entrada inteiro para calcular os limites de categoria antes de alocar registros para categorias. O conjunto de dados é armazenado em cache enquanto os limites são calculados; depois ele é varrido novamente para alocação. Quando o método de categorização é *largura fixa* ou *média+desvio padrão*, o conjunto de dados é armazenado em cache diretamente no disco. Esses métodos têm um tempo de execução linear e requerem espaço em disco suficiente para armazenar o conjunto de dados inteiro. Quando o método de categorização é *ranqueamentos* ou *ladrilhos*, o conjunto de dados é ordenado usando o algoritmo de ordenação descrito anteriormente e o conjunto de dados ordenado é usado como o cache. A ordenação dá a esses métodos um tempo de execução de $M \cdot N \cdot \log(N)$, em que M é o número de campos categorizados e N é o número de registros; isso requer espaço em disco igual a duas vezes o tamanho do conjunto de dados.

A geração de um nó Derivar baseada em categorias geradas melhorará o desempenho em passagens subsequentes. Operações de derivação são mais rápidas que categorização.

Mesclagem por Chave (Junção). O nó de Mesclagem, quando o método de mesclagem é *chaves* (equivalente a uma junção de banco de dados), ordena cada um de seus conjuntos de dados de entrada por campos-chave. Essa parte do procedimento tem um tempo de execução de $M \cdot N \cdot \log(N)$, em que M é o número de entradas e N é o número de registros na maior entrada; isso requer espaço em disco suficiente para armazenar todos os seus conjuntos de dados de entrada, mais uma segunda cópia do maior conjunto de dados. O tempo de execução da mesclagem em si é proporcional ao tamanho do conjunto de dados de saída, o que depende da frequência de chaves correspondentes. No pior dos casos, em que a

saída é um produto Cartesiano das entradas, o tempo de execução pode se aproximar de NM. Isso é raro-a maioria das junções tem muitas chaves menos correspondentes. Se um conjunto de dados for relativamente maior que o(s) outro(s), ou se os dados recebidos já estiverem ordenados por uma campo-chave, é possível melhorar o desempenho desse nó usando a guia Otimização.

Agregado. Quando a opção *Chaves são contíguas* não está configurada, esse nó lê (mas não armazena) seu conjunto de dados de entrada inteiro antes de produzir qualquer saída agregada. Nas situações mais extremas, em que o tamanho dos dados agregados atinge um limite (determinado pela opção de configuração *Multiplicador de Uso de Memória* do IBM SPSS Modeler Server), o restante do conjunto de dados é ordenado e processado como se a opção *Chaves são contíguas* estivesse configurada. Quando essa opção é configurada, nenhum dado é armazenado porque os registros de saída agregados são produzidos conforme os dados de entrada são lidos.

Distinto. O nó Distinto armazena todos os campos-chave exclusivos no conjunto de dados de entrada; em casos em que todos os campos são campos-chave e todos os registros são exclusivos, ele armazena o conjunto de dados inteiro. Por padrão, o nó Distinto ordena os dados nos campos-chave e depois seleciona (ou descarta) o primeiro registro distinto de cada grupo. Para conjuntos de dados menores com um baixo número de chaves distintas, ou aqueles que foram pré-ordenados, é possível escolher opções para melhorar a velocidade e a eficiência do processamento.

Tipo. Em algumas ocorrências, o nó Tipo armazena em cache os dados de entrada durante a leitura de valores; o cache é usado para processamento de recebimento de dados. O cache requer espaço em disco suficiente para armazenar o conjunto de dados inteiro, mas acelera o armazenamento.

Avaliação. O nó de Avaliação deve ordenar os dados de entrada para calcular ladrilhos. A ordem é repetida para cada modelo avaliado, pois os escores e a ordem de registro subsequente são diferentes em cada caso. O tempo de execução é $M*N*\log(N)$, em que M é o número de modelos e N é o número de registros.

Desempenho: Nós de Modelagem

Rede Neural e Kohonen. Algoritmos de treinamento de rede neural (incluindo o algoritmo Kohonen) ignoram muitos dados de treinamento. Os dados são armazenados na memória até um certo limite, e o excesso é estendido para o disco. O acesso aos dados de treinamento a partir do disco é caro porque o método de acesso é aleatório, o que pode levar à atividade excessiva do disco. É possível desativar o uso do armazenamento em disco para esses algoritmos, forçando todos os dados a serem armazenados na memória, selecionando a opção **Otimizar para Velocidade** na guia Modelo da caixa de diálogo do nó. Observe que se a quantidade de memória necessária para armazenar os dados for maior que o conjunto de trabalhos do processo do servidor, parte dela será paginada para o disco e o desempenho sofrerá de acordo.

Quando **Otimizar para Memória** está ativada, a porcentagem de RAM física é alocada para o algoritmo de acordo com o valor da opção de configuração do IBM SPSS Modeler Server, *Porcentagem Limite de Memória de Modelagem*. Para usar mais memória para treinamento de redes neurais, forneça mais RAM ou aumente o valor dessa opção, mas observe que a configuração de um valor muito alto causará paginação.

O tempo de execução dos algoritmos de rede neural depende do nível de precisão necessário. É possível controlar o tempo de execução configurando uma condição de parada na caixa de diálogo do nó.

K-médias. O algoritmo de clusterização k-médias tem as mesmas opções para controlar o uso de memória que os algoritmos de rede neural. No entanto, o desempenho nos dados armazenados em disco é melhor, pois o acesso aos dados é sequencial.

Desempenho: expressões do CLEM

As funções em sequência do CLEM (“funções @”) que reconsideram o fluxo de dados devem armazenar dados suficientes para satisfazer o lookback mais longo. Para operações cujo grau de lookback não tem limites, todos os valores do campo devem ser armazenados. Uma operação sem limites é aquela na qual o valor de compensação não é um número inteiro literal; por exemplo, @OFFSET(Sales, Month). O valor de compensação é o nome do campo *Month*, cujo valor é desconhecido até ser executado. O servidor deve salvar todos os valores do campo *Vendas* para assegurar resultados precisos. Onde um limite máximo for conhecido, é necessário fornecê-lo como um argumento adicional; por exemplo, @OFFSET(Sales, Month, 12). Essa operação instrui o servidor a armazenar no máximo os 12 valores mais recentes de *Sales*. As funções em sequência, limitadas ou não, quase sempre inibem a geração de SQL.

Capítulo 15. Acessibilidade no IBM SPSS Modeler

Visão geral de acessibilidade no IBM SPSS Modeler

IBM SPSS Modeler fornece suporte de acessibilidade para todos os usuários, bem como suporte específico para usuários com deficiência visual e outras deficiências funcionais. Esta seção descreve as variáveis e os métodos para se trabalhar usando aprimoramentos de acessibilidade, como leitores de tela e atalhos de teclado.

Tipos de suporte à acessibilidade

Independentemente de você ter uma deficiência visual ou depender do teclado para manipulação, há uma grande variedade de métodos alternativos para o uso desse kit de ferramentas de mineração de dados. Por exemplo, é possível construir fluxos, especificar opções e ler saídas, tudo sem usar o mouse. Os atalhos de teclado disponíveis estão listados nos tópicos a seguir. Além disso, o IBM SPSS Modeler fornece suporte extensivo para leitores de tela, como JAWS para Windows. Também é possível otimizar o esquema de cores para fornecer contraste adicional. Esses tipos de suporte serão discutidos nos tópicos a seguir.

Acessibilidade para deficiência visual

Há inúmeras propriedades que podem ser especificadas no IBM SPSS Modeler que aprimorarão sua capacidade de usar o software.

Exibir opções

É possível selecionar cores para a exibição de gráficos. Também é possível escolher usar suas configurações específicas do Windows para o próprio software. Isso pode ajudar a aumentar o contraste visual.

1. Para configurar opções de exibição no menu Ferramentas, clique em **Opções do Usuário**.
2. Clique na guia **Exibir**. As opções nessa guia incluem o esquema de cores do software, cores de gráficos e tamanhos de fontes para nós.

Uso de Sons para Notificação

Ativando ou desativando sons, é possível controlar a maneira como você é alertado para operações específicas no software. Por exemplo, é possível ativar sons para eventos, como criação e exclusão de nó ou geração de nova saída ou modelos.

1. Para configurar opções de notificação, no menu Ferramentas, clique em **Opções do Usuário**.
2. Clique na guia **Notificações**.

Controlar Ativação Automática de Novas Janelas

A guia Notificações na caixa de diálogo Opções do Usuário também é usada para controlar se uma saída recém-gerada, como tabelas e gráficos, é ativada em uma janela separada. Talvez seja mais fácil para você desativar essa opção e abrir uma janela de saída somente quando necessário.

1. Para configurar essas opções, no menu Ferramentas, clique em **Opções do Usuário**.
2. Clique na guia **Notificações**.
3. Na caixa de diálogo, selecione **Nova Saída** na lista no grupo **Notificações Visuais**.
4. Em **Abrir Janela**, selecione **Nunca**.

Tamanho do Nó

Nós podem ser exibidos usando um tamanho padrão ou pequeno. Talvez você queira ajustar esses tamanhos para de acordo com suas necessidades.

1. Para configurar opções de tamanho de nó, no menu Arquivo, clique em **Propriedades do Fluxo**.
2. Clique na guia **Layout**.
3. Na lista **Tamanho do Ícone**, selecione **Padrão**.

Acessibilidade para usuários cegos

O suporte para usuários cegos é predominantemente dependente do uso de um leitor de tela, como JAWS for Windows. Para otimizar o uso de um leitor de tela com IBM SPSS Modeler, é possível especificar inúmeras configurações.

Exibir opções

Leitores de tela tendem a ter um desempenho melhor quando o contraste visual é maior na tela. Se você já tiver uma configuração do Windows de alto contraste, é possível escolher usar essas configurações do Windows para o software em si.

1. Para configurar opções de exibição no menu Ferramentas, clique em **Opções do Usuário**.
2. Clique na guia **Exibir**.

Uso de Sons para Notificação

Ativando ou desativando sons, é possível controlar a maneira como você é alertado para operações específicas no software. Por exemplo, é possível ativar sons para eventos, como criação e exclusão de nó ou geração de nova saída ou modelos.

1. Para configurar opções de notificação, no menu Ferramentas, clique em **Opções do Usuário**.
2. Clique na guia **Notificações**.

Controlar Ativação Automática de Novas Janelas

A guia Notificações na caixa de diálogo Opções do Usuário também é usada para controlar se uma saída recém-gerada é ativada em uma janela separada. Talvez seja mais fácil para você desativar essa opção e abrir uma janela de saída conforme necessário.

1. Para configurar essas opções, no menu Ferramentas, clique em **Opções do Usuário**.
2. Clique na guia **Notificações**.
3. Na caixa de diálogo, selecione **Nova Saída** na lista no grupo **Notificações Visuais**.
4. Em **Abrir Janela**, selecione **Nunca**.

Teclado de acessibilidade

A funcionalidade do produto está acessível a partir do teclado. No nível mais básico, é possível pressionar Alt mais a tecla apropriada para ativar menus de janela (como Alt+F para acessar o menu Arquivo) ou pressionar a tecla tab para rolar pelos controles da caixa de diálogo. No entanto, há alguns problemas especiais relacionados a cada uma das janelas de menu do produto e sugestões úteis para navegação nas caixas de diálogo.

Esta seção cobrirá os destaques do teclado de acessibilidade, da abertura de um fluxo ao uso de caixas de diálogo de nó e ao trabalho com saída. Além disso, listas de atalhos de teclado são fornecidas para uma navegação muito mais eficiente.

Atalhos para navegação na janela principal

Você faz a maioria do seu trabalho de mineração de dados na janela principal do IBM SPSS Modeler. A área principal é chamada **tela de fluxo** e é usada para construir e executar fluxos de dados. A parte inferior da janela contém as **paletas de nó**, que contêm todos os nós disponíveis. As paletas são organizadas nas guias correspondentes ao tipo de operação de mineração de dados para cada grupo de nós. Por exemplo, nós usados para trazer dados para o IBM SPSS Modeler são agrupados na guia Origens, e nós usados para derivar, filtrar ou digitar campos são agrupados na guia Ops do Campo (abreviação de Operações do Campo).

O lado direito da janela contém várias ferramentas para gerenciar fluxos, saída e projetos. A metade superior direita contém os **gerenciadores** e tem três guias que são usadas para gerenciar fluxos, saída e modelos gerados. É possível acessar esses objetos selecionando a guia e um objeto na lista. A metade inferior direita contém a **área de janela do projeto**, que permite organizar seu trabalho em projetos. Há duas guias nessa área refletindo duas visualizações diferentes de um projeto. A **visualização Classes** ordena objetos do projeto por tipo, enquanto a **visualização CRISP-DM** ordena objetos por fase de mineração de dados relevante, como Preparação de Dados ou Modelagem. Esses vários aspectos da janela do IBM SPSS Modeler são discutidos em todo o Sistema de Ajuda e no Guia do Usuário.

A seguir está uma tabela dos atalhos usados para você se mover dentro da janela principal do IBM SPSS Modeler e construir fluxos. Atalhos para caixas de diálogo e saída são listados nos tópicos que seguem. Observe que essas teclas de atalho estão disponíveis somente na janela principal.

Tabela 37. Atalhos da Janela Principal

Tecla de Atalho	Função
Ctrl+F5	Move o foco para as paletas de nó.
Ctrl+F6	Move o foco para a tela de fluxo.
Ctrl+F7	Move o foco para a área de janela de gerenciadores.
Ctrl+F8	Move o foco para a área de janela do projeto.

Tabela 38. Atalhos de Fluxo e Nó

Tecla de Atalho	Função
Ctrl+N	Cria uma nova tela de fluxo em branco.
Ctrl+O	Exibe a caixa de diálogo Abrir, onde é possível selecionar e abrir um fluxo existente.
Ctrl+teclas de número	Move o foco para a guia correspondente em uma janela ou área de janela. Por exemplo, dentro de uma janela ou área de janela com guias, Ctrl+1 move para a primeira guia, começando da esquerda, Ctrl+2 para a segunda, etc.
Ctrl+Seta para Baixo	Usado na paleta do nó para mover o foco de uma guia da paleta para o primeiro nó sob essa guia.
Ctrl+Seta para Cima	Usado na paleta do nó para mover o foco de um nó para sua guia da paleta.
Enter	Quando um nó é selecionado na paleta do nó (incluindo modelos refinados na paleta de modelos gerada), esse pressionamento de tecla inclui o nó na tela de fluxo. O pressionamento de Enter quando um nó já está selecionado na tela abre a caixa de diálogo para esse nó.
Ctrl+Enter	Quando um nó é selecionado na paleta, inclui esse nó na tela de fluxo sem selecioná-lo e move o foco para o primeiro nó na paleta.
Alt+Enter	Quando um nó é selecionado na paleta, inclui esse nó na tela de fluxo e o seleciona, enquanto move o foco para o primeiro nó na paleta.

Tabela 38. Atalhos de Fluxo e Nó (continuação)

Tecla de Atalho	Função
Shift+Barra de Espaço	Quando um nó ou comentário está em foco na paleta, alterna entre a seleção e o cancelamento da seleção do nó ou comentário. Se algum outro nó ou comentário também estiver selecionado, isso faz com que a seleção seja cancelada.
Ctrl+Shift+Barra de Espaço	Quando um nó ou comentário está em foco no fluxo, ou um nó ou comentário está em foco na paleta, alterna entre a seleção e o cancelamento da seleção do nó ou comentário. Isso não afeta nenhum outro nó ou comentário selecionado.
Seta para Esquerda/Direita	Se a tela de fluxo estiver em foco, move o fluxo inteiro horizontalmente na tela. Se uma guia da paleta estiver em foco, faz ciclos entre as guias. Se um nó da paleta estiver em foco, se move entre os nós na paleta.
Seta para Cima/Baixo	Se a tela de fluxo estiver em foco, move o fluxo inteiro verticalmente na tela. Se um nó da paleta estiver em foco, se move entre os nós na paleta. Se uma subpaleta estiver em foco, se move entre outras subpaletas para essa guia da paleta.
Alt+Seta para Esquerda/Direita	Move nós e comentários selecionados na tela de fluxo horizontalmente na direção da tecla de seta.
Alt+Seta para Cima/Baixo	Move nós e comentários selecionados na tela de fluxo verticalmente na direção da tecla de seta.
Ctrl+A	Seleciona todos os nós em um fluxo.
Ctrl+Q	Quando um nó está em foco, seleciona-o e o recebimento de dados de todos os nós e cancela a seleção do envio de dados de todos os nós.
Ctrl+W	Quando um nó selecionado está em foco, cancela sua seleção e o recebimento de dados de todos os nós selecionados.
Ctrl+Alt+D	Duplica um nó selecionado.
Ctrl+Alt+L	Quando um nugget do modelo é selecionado no fluxo, abra uma caixa de diálogo Inserir para permitir que você carregue um modelo salvo de um arquivo .nod no fluxo.
Ctrl+Alt+R	Exibe a guia Anotações para um nó selecionado, permitindo renomear o nó.
Ctrl+Alt+U	Cria um nó de origem Entrada do Usuário.
Ctrl+Alt+C	Ativa ou desativa o cache para um nó.
Ctrl+Alt+F	Limpa o cache para um nó.
Tab	Na tela de fluxo, faz ciclos por todos os nós de origem e comentários no fluxo atual. Em uma paleta do nó, se move entre os nós na paleta. Em uma subpaleta selecionada, se move para o primeiro nó na subpaleta.
Shift+Tab	Executa a mesma operação que Tab, mas na ordem inversa.
Ctrl+Tab	Com o foco na área de janela dos gerenciadores ou área de janela do projeto, move o foco para a tela de fluxo. Com o foco em uma paleta do nó, move o foco entre um nó e sua guia da paleta.
Qualquer tecla alfabética	Com o foco em um nó no fluxo atual, direciona o foco e faz um ciclo para o próximo nó cujo nome começa com a tecla pressionada.
F1	Abra o Sistema de Ajuda em um tópico relevante para o foco.
F2	Inicia o processo de conexão para um nó selecionado na tela. Use a tecla tab para mover para o nó necessário na tela e pressione Shift+Barra de Espaço para concluir a conexão.

Tabela 38. Atalhos de Fluxo e Nó (continuação)

Tecla de Atalho	Função
F3	Exclui todas as conexões para o nó selecionado na tela.
F6	Move o foco entre a área de janela dos gerenciadores, área de janela do projeto e paletas de nó.
F10	Abre o menu Arquivo.
Shift+F10	Abre o menu pop-up para o nó ou fluxo.
Delete	Exclui um nó selecionado da tela.
Esc	Fecha um menu pop-up ou caixa de diálogo.
Ctrl+Alt+X	Expande um SuperNode.
Ctrl+Alt+Z	Aumenta o zoom em um SuperNode.
Ctrl+Alt+Shift+Z	Diminui o zoom de um SuperNode.
Ctrl+E	Com o foco na tela de fluxo, executa o fluxo atual.

Inúmeras teclas de atalho padrão também são usadas no IBM SPSS Modeler, como Ctrl+C para copiar. Consulte o tópico “Usando teclas de atalho” na página 20 para obter mais informações.

Atalhos para caixas de diálogo e tabelas

Várias teclas de atalho e de leitor de tela são úteis quando você está trabalhando com caixas de diálogo, tabelas e tabelas em caixas de diálogo. Uma lista completa de atalhos de teclado e leitor de tela especiais é mostrada a seguir.

Tabela 39. Atalhos da Caixa de Diálogo e do Construtor de Expressões.

Tecla de Atalho	Função
Alt+4	Usado para indeferir todas as janelas de saída ou caixas de diálogo abertas. A saída pode ser recuperada da guia Saídas na área de janela de gerenciadores.
Ctrl+End	Com o foco em qualquer controle no construtor de expressões, move o ponto de inserção para o final da expressão.
Ctrl+1	No construtor de expressões, move o foco para o controle de edição de expressão.
Ctrl+2	No construtor de expressões, move o foco para a lista de funções.
Ctrl+3	No construtor de expressões, move o foco para a lista de campos.

Atalhos de Tabela

Atalhos de tabela são usados para tabelas de saída, bem como controles de tabela em caixas de diálogo para nós, como Tipo, Filtro e Mesclagem. Normalmente, você usará a tecla tab para se mover entre células da tabela e a Ctrl+Tab para sair do controle de tabela. *Nota:* Ocasionalmente, um leitor de tela pode não começar a ler imediatamente o conteúdo de uma célula. O pressionamento das teclas de seta uma ou duas vezes reconfigurará o software e iniciará o discurso.

Tabela 40. Atalhos de Tabela.

Tecla de Atalho	Função
Ctrl+W	Para tabelas, lê a descrição simples da linha selecionada. Por exemplo, "Os valores da linha 2 selecionada são sexo, flag, m/f, etc."
Ctrl+Alt+W	Para tabelas, lê a descrição detalhada da linha selecionada. Por exemplo, "Os valores da linha 2 selecionada são campo = sexo, tipo = flag, sexo = m/f, etc."

Tabela 40. Atalhos de Tabela (continuação).

Tecla de Atalho	Função
Ctrl+D	Para tabelas, lê a Descrição simples da área selecionada. Por exemplo, "Seleção é uma linha por seis colunas."
Ctrl+Alt+D	Para tabelas, fornece a Descrição detalhada da área selecionada. Por exemplo, "Seleção é uma linha por seis colunas. As colunas selecionadas são Campo, Tipo, Omisso. A linha selecionada é 1."
Ctrl+T	Para tabelas, fornece uma descrição simples das colunas selecionadas. Por exemplo, "Campos, Tipo, Omisso."
Ctrl+Alt+T	Para tabelas, fornece uma descrição detalhada das colunas selecionadas. Por exemplo, "Colunas selecionadas são Campos, Tipo, Omisso."
Ctrl+R	Para tabelas, fornece o número de Registros na tabela.
Ctrl+Alt+R	Para tabelas, fornece o número de Registros na tabela, bem como nomes de coluna.
Ctrl+I	Para tabelas, lê as Informações, ou conteúdo, da célula para a célula que está em foco.
Ctrl+Alt+I	Para tabelas, lê a descrição detalhada das Informações da célula (nome da coluna e conteúdo da célula) para a célula que está em foco.
Ctrl+G	Para tabelas, fornece informações de seleção Gerais simples.
Ctrl+Alt+G	Para tabelas, fornece informações de seleção Gerais detalhadas.
Ctrl+Q	Para tabelas, fornece uma alternância Rápida das células da tabela. Ctrl+Q lê as descrições detalhadas, como "Sexo=Feminino", conforme você se move pela tabela usando as teclas de seta. A seleção de Ctrl+Q novamente alterna para descrições simples (conteúdo da célula).
F8	Para tabelas, quando o foco está na tabela, configura o foco para o cabeçalho da coluna.
Barra de espaço	Para tabelas, quando o foco está no cabeçalho da coluna, permite a ordenação da coluna.

Atalhos para comentários

Ao trabalhar com comentários na tela, é possível usar os atalhos a seguir.

Tabela 41. Atalhos de comentário

Tecla de Atalho	Função
Alt+C	Alterna a variável mostrar/ocultar comentário.
Alt+M	Insere um novo comentário se os comentários estiverem sendo exibidos atualmente; mostra os comentários se eles estiverem atualmente ocultos.
Tab	Na tela de fluxo, faz ciclos por todos os nós de origem e comentários no fluxo atual.
Inserir	Quando um comentário está em foco, indica o início da edição.
Alt+Enter ou Ctrl+Tab	Termina a edição e salva as mudanças na edição.
Esc	Cancela a edição. As mudanças feitas durante a edição são perdidas.
Alt+Shift+Seta para Cima	Reduz a altura da área de texto em uma célula de grade (ou um pixel) se "ajustar à grade" estiver ativado (ativado / desativado).
Alt+Shift+Seta para Baixo	Aumenta a altura da área de texto em uma célula de grade (ou um pixel) se "ajustar à grade" estiver ativado (ativado / desativado).

Tabela 41. Atalhos de comentário (continuação)

Tecla de Atalho	Função
Alt+Shift+Seta para Esquerda	Reduz a largura da área de texto em uma célula de grade (ou um pixel) se "ajustar à grade" estiver ativado (ativado / desativado).
Alt+Shift+Seta para Direita	Aumenta a largura da área de texto em uma célula de grade (ou um pixel) se "ajustar à grade" estiver ativado (ativado / desativado).

Atalhos para o Visualizador de Cluster e Visualizador de Modelo

Há teclas de atalho disponíveis para navegação nas janelas Visualizador de Cluster e Visualizador de Modelo.

Tabela 42. Atalhos Gerais - Visualizador de Cluster e Visualizador de Modelo

Tecla de Atalho	Função
Tab	Movê o foco para o próximo controle de tela.
Shift+Tab	Movê o foco para o controle de tela anterior.
Seta para Baixo	Se uma lista suspensa estiver em foco, abre a lista ou move para o próximo item na lista. Se um menu estiver em foco, move para o próximo item no menu. Se um gráfico miniatura estiver em foco, move para o próximo no conjunto (ou para o primeiro se o último estiver em foco).
Seta para cima	Se uma lista suspensa estiver aberta, move para o item anterior na lista. Se um menu estiver em foco, move para o item anterior no menu. Se um gráfico miniatura estiver em foco, move para o anterior no conjunto (ou para o último se a primeira miniatura estiver em foco).
Enter	Fecha uma lista suspensa aberta ou faz uma seleção em um menu aberto.
F6	Alterna o foco entre as áreas de janela esquerda e direita da janela.
Setas para Esquerda e Direita	Se uma guia estiver em foco, move para a guia anterior ou seguinte. Se um menu estiver em foco, move para o menu anterior ou seguinte.
Alt+letra	Seleciona o botão ou menu com essa letra sublinhada em seu nome.
Esc	Fecha uma lista suspensa ou menu aberto.

Somente Visualizador de Cluster

O Visualizador de Cluster tem uma visualização Clusters que contém uma grade "cluster por variáveis".

Para escolher a visualização Clusters em vez de a visualização Sumarização do Modelo:

1. Pressione Tab repetidamente até o botão **Visualização** ser selecionado.
2. Pressione Seta para Baixo duas vezes para selecionar **Clusters**.
Aqui, é possível selecionar uma célula individual dentro da grade:
3. Pressione Tab repetidamente até chegar ao último ícone na barra de ferramentas de visualização.



Figura 17. Ícone Mostrar Árvore de Visualização

4. Pressione Tab mais uma vez, depois a barra de espaço, depois uma tecla de seta.

Os atalhos de teclado a seguir agora estão disponíveis:

Tabela 43. Atalhos do Visualizador de Cluster

Tecla de Atalho	Função
Tecla de seta	Move o foco entre células individuais na grade. A exibição da distribuição de células na área de janela direita muda conforme o foco é movido.
Ctrl+, (vírgula)	Seleciona ou cancela a seleção da coluna inteira na grade em que a célula está em foco. Para incluir uma coluna na seleção, use as teclas de seta para navegar para uma célula nessa coluna e pressione Ctrl+, novamente.
Tab	Move o foco para fora da grade e para o próximo controle de tela.
Shift+Tab	Move o foco para fora da grade e de volta para o controle de tela anterior.
F2	Entra no modo de edição (somente células de descrição e rótulo).
Enter	Salva mudanças de edição e sai do modo de edição (somente células de descrição e rótulo).
Esc	Sai do modo de edição sem salvar as mudanças (somente células de descrição e rótulo).

Exemplo de teclas de atalho: Construindo Fluxos

Para deixar um processo de construção de fluxo mais claro para os usuários dependentes de teclado ou em um leitor de tela, veja a seguir um exemplo de como construir um fluxo sem usar o mouse. Neste exemplo, você construirá um fluxo contendo um nó de Arquivo Variável, um nó de Derivação e um nó de Histograma usando os passos a seguir:

1. **Inicie o IBM SPSS Modeler.** Quando o IBM SPSS Modeler for iniciado pela primeira vez, o foco estará na guia Favoritos da paleta do nó.
2. **Ctrl+Seta para Baixo.** Move o foco da guia para o corpo da guia.
3. **Seta para a Direita.** Move o foco para o nó de Arquivo Variável.
4. **Barra de Espaço.** Seleciona o nó de Arquivo Variável.
5. **Ctrl+Enter.** Inclui o nó de Arquivo Variável na tela de fluxo. Essa combinação de teclas também mantém a seleção no nó de Arquivo Variável para que o próximo nó incluído seja conectado a ele.
6. **Tab.** Move o foco de volta para a paleta do nó.
7. **Seta para a Direita 4 vezes.** Move para o nó de Derivação.
8. **Barra de Espaço.** Seleciona o nó de Derivação.
9. **Alt+Enter.** Inclui o nó de Derivação na tela e move a seleção para o nó de Derivação. Agora esse nó está pronto para ser conectado ao próximo nó incluído.
10. **Tab.** Move o foco de volta para a paleta do nó.
11. **Seta para a Direita 5 vezes.** Move o foco para o nó de Histograma na paleta.
12. **Barra de Espaço.** Seleciona o nó de Histograma.
13. **Enter.** Inclui o nó no fluxo e move o foco para a tela de fluxo.

Continue com o próximo exemplo ou salve o fluxo se quiser tentar o próximo exemplo depois.

Exemplo de teclas de atalho: Editando Nós

Neste exemplo, você usará o fluxo para construir o exemplo anterior. O fluxo consiste em um nó de Arquivo Variável, um nó de Derivação e um nó de Histograma. As instruções começam com o foco no terceiro nó no fluxo, o nó de Histograma.

1. **Ctrl+Seta para Esquerda 2 vezes.** Move o foco de volta para o nó de Arquivo Variável.
2. **Enter.** Abre a caixa de diálogo Arquivo Variável. Tabule para o campo Arquivo e digite um nome e um caminho de arquivo de texto para selecionar tal arquivo. Pressione Ctrl+Tab para navegar para a parte inferior da caixa de diálogo, tabule para o botão OK e pressione Enter para fechar a caixa de diálogo.
3. **Ctrl+Seta para a Direita.** Direciona o foco para o segundo nó, um nó de Derivação.
4. **Enter.** Abre a caixa de diálogo do nó de Derivação. Tabule para selecionar campos e especificar condições de derivação. Pressione Ctrl+Tab para navegar para o botão OK e pressione Enter para fechar a caixa de diálogo.
5. **Ctrl+Seta para a Direita.** Direciona o foco para o terceiro nó, um nó de Histograma.
6. **Enter.** Abre a caixa de diálogo do nó de Histograma. Tabule para selecionar campos e especificar opções de gráfico. Para listas suspensas, pressione Seta para Baixo para abrir a lista e para destacar um item da lista; em seguida, pressione Enter para selecionar o item da lista. Tabule para o botão OK e pressione Enter para fechar a caixa de diálogo.

Nesse ponto, é possível incluir nós adicionais ou executar o fluxo atual. Lembre-se das dicas a seguir ao construir fluxos:

- Ao conectar nós manualmente, use F2 para criar o ponto de início de conexão, tabule para se mover para o terminal e, em seguida, use Shift+Barra de Espaço para finalizar a conexão.
- Use F3 para destruir todas as conexões para um nó selecionado na tela.
- Após ter criado um fluxo, use Ctrl+E para executar o fluxo atual.

Uma lista completa de teclas de atalho está disponível. Consulte o tópico “Atalhos para navegação na janela principal” na página 201 para obter mais informações.

Usando um leitor de tela

Inúmeros leitores de tela estão disponíveis no mercado. IBM SPSS Modeler é configurado para suportar JAWS for Windows usando Java Access Bridge, que é instalado com IBM SPSS Modeler. Se você tiver o JAWS instalado, basta ativar o JAWS antes de ativar o IBM SPSS Modeler para usar esse produto.

Nota: É recomendado ter pelo menos 6 GB de espaço para executar JAWS com SPSS Modeler.

Devido à natureza da representação gráfica exclusiva do IBM SPSS Modeler do processo de mineração de dados, diagramas e gráficos são melhores usados visualmente. No entanto, é possível que você entenda e tome decisões baseadas na saída e nos modelos visualizados textualmente usando um leitor de tela.

Nota: Com computadores clientes de 64 bits, algumas variáveis de tecnologia assistiva não funcionam. Isso porque o Java Access Bridge não foi projetado para uma operação de 64 bits.

Usando o arquivo de dicionário do IBM SPSS Modeler

Um arquivo de dicionário do IBM SPSS Modeler (*Awt.JDF*) está disponível para inclusão com JAWS. Para usar esse arquivo:

1. Navegue para o subdiretório */accessibility* da instalação do IBM SPSS Modeler e copie o arquivo de dicionário (*Awt.JDF*).
2. Copie-o no diretório com seus scripts JAWS.

Você já pode ter um arquivo denominado *Awt.JDF* em sua máquina se tiver outros aplicativos JAVA em execução. Nesse caso, talvez você não possa usar esse arquivo de dicionário sem editar manualmente o arquivo de dicionário.

Usando um leitor de tela com saída HTML

Durante a visualização da saída exibida como HTML dentro do IBM SPSS Modeler usando um leitor de tela, você pode encontrar algumas dificuldades. Inúmeros tipos de saída são afetados, incluindo:

- Saída visualizada na guia Avançado para os nós Regressão, Regressão Logística e Fator/PCA
- Saída do nó do relatório

Em cada uma dessas janelas ou caixas de diálogo, há uma ferramenta na barra de ferramentas que pode ser usada para ativar a saída em seu navegador padrão, o que fornece suporte ao leitor de tela padrão. É possível usar o leitor de tela para transmitir as informações de saída.

Acessibilidade na janela de árvore interativa

A exibição padrão de um modelo de árvore de decisão na janela Árvore Interativa pode causar problemas para leitores de tela. Para acessar uma versão acessível, nos menus Árvore Interativa, clique em:

Visualizar > Janela Acessível

Isso exibe uma visualização semelhante ao mapa de árvore padrão, mas que o JAWS possa ler corretamente. É possível mover para cima, para baixo, para direita ou para esquerda usando as teclas de seta padrão. Conforme navega na janela acessível, o foco na janela Árvore Interativa é movido de acordo. Use a Barra de Espaço para alterar a seleção ou Ctrl+Barra de Espaço para estender a seleção atual.

Dicas de uso

Há várias dicas para tornar o ambiente do IBM SPSS Modeler mais acessível. A seguir estão sugestões gerais para o trabalho no IBM SPSS Modeler.

- **Sair das caixas de texto estendidas.** Use Ctrl+Tab para sair das caixas de texto estendidas. *Nota:* Ctrl+Tab também é usado para sair dos controles de tabela.
- **Usar a tecla tab em vez de teclas de seta.** Ao selecionar opções para uma caixa de diálogo, use a tecla tab para se mover entre os botões de opções. As teclas de seta não funcionarão nesse contexto.
- **Listas suspensas.** Em uma lista suspensa para caixas de diálogo, é possível usar a tecla Escape ou a barra de espaço para selecionar um item e fechar a lista. Também é possível usar a tecla Escape para fechar as listas suspensas que não são fechadas quando você tabula para outro controle.
- **Status da execução.** Quando você está executando um fluxo em um banco de dados grande, o JAWS pode ficar para trás na leitura do status do fluxo para você. Pressione a tecla Ctrl periodicamente para atualizar o relato de status.
- **Usar paletas de nó.** Quando você entra em uma guia das paletas de nó pela primeira vez, às vezes o JAWS lê "caixa de grupo" em vez de o nome do nó. Nesse caso, é possível usar Ctrl+Seta para a Direita e depois Ctrl+Seta para a Esquerda para reconfigurar o leitor de tela e ouvir o nome do nó.
- **Ler menus.** Ocasionalmente, quando você abre um menu pela primeira vez, o JAWS pode não ler o primeiro item de menu. Se você suspeitar que isso aconteceu, use a Tecla para Baixo e depois a Tecla para Cima para ouvir o primeiro item no menu.
- **Menus em cascata.** JAWS não lê o primeiro nível de um menu em cascata. Se você ouvir uma quebra na fala ao se mover pelo menu, pressione a tecla Seta para a Direita para ouvir os itens de menu filhos.

Além disso, se você tiver o IBM SPSS Modeler Text Analytics instalado, as sugestões a seguir podem deixar a interface do ambiente de trabalho interativa mais acessível.

- **Entrar nas caixas de diálogo.** Talvez seja necessário pressionar a tecla tab para direcionar o foco para o primeiro controle ao entrar na caixa de diálogo.

- **Sair das caixas de texto estendidas.** Use Ctrl+Tab para sair das caixas de texto estendidas e se mover para o próximo controle. *Nota:* Ctrl+Tab também é usado para sair dos controles de tabela.
- **Digitar a primeira letra para localizar o elemento na lista de árvores.** Quando você procura um elemento na área de janela de categorias, área de janela de resultados extraídos ou árvore de biblioteca, é possível digitar a primeira letra do elemento quando o foco estiver na área de janela. Isso irá selecionar a próxima ocorrência de um elemento começando com a letra inserida.
- **Listas suspensas.** Em uma lista suspensa para caixas de diálogo, é possível usar a barra de espaço para selecionar um item e fechar a lista.

Sugestões adicionais de uso são discutidas detalhadamente nos tópicos a seguir.

Interferência de outro software

Durante o teste do IBM SPSS Modeler com leitores de tela, como JAWS, nossa equipe de desenvolvimento descobriu que o uso de um Systems Management Server (SMS) dentro de sua organização pode interferir na capacidade do JAWS de ler aplicativos baseados em Java, como IBM SPSS Modeler. A desativação do SMS corrigirá essa situação. Visite o website da Microsoft para obter mais informações sobre o SMS.

JAWS e Java

Versões diferentes do JAWS fornecem níveis variados de suporte para aplicativos de software baseados em Java. Embora o IBM SPSS Modeler trabalhe com todas as versões recentes de JAWS, algumas versões podem ter pequenos problemas quando usadas com sistemas baseados em Java. Visite o website do JAWS for Windows em <http://www.FreedomScientific.com>.

Usando gráficos no IBM SPSS Modeler

Exibições visuais de informações, como histogramas, gráficos de avaliação, multigráficos e gráficos de dispersão, são difíceis de interpretar com um leitor de tela. Observe, no entanto, que gráficos da web e distribuições podem ser visualizados usando a sumarização textual disponível na janela de saída.

Capítulo 16. Suporte a Unicode

Suporte Unicode no IBM SPSS Modeler

IBM SPSS Modeler é totalmente ativado por Unicode para IBM SPSS Modeler e IBM SPSS Modeler Server. Isso possibilita a troca de dados com outros aplicativos que suportam Unicode, incluindo bancos de dados multi-idiomas, sem qualquer perda de informações que possa ser causada pela conversão em ou de um esquema de codificação específico de um código de idioma.

- IBM SPSS Modeler armazena dados Unicode internamente e pode ler e gravar dados multi-idiomas armazenados como Unicode em bancos de dados sem perda.
- IBM SPSS Modeler pode ler e gravar arquivos de texto codificados em UTF-8. A importação e a exportação de arquivo de texto serão padronizadas para a codificação do código de idioma, mas suportarão UTF-8 como alternativa. Essa configuração pode ser especificada nos nós de importação e exportação de arquivo, ou a codificação padrão pode ser alterada na caixa de diálogo de propriedades do fluxo. Consulte o tópico “Configurando opções gerais para fluxos” na página 40 para obter mais informações.
- Statistics, SAS e arquivos de dados de texto armazenados na codificação de código de idioma serão convertidos em UTF-8 na importação e novamente na exportação. Durante a gravação em um arquivo, se houver caracteres Unicode que não existem no conjunto de caracteres de código de idioma, eles serão substituídos e um aviso será exibido. Isso deve ocorrer somente quando os dados tiverem sido importados de uma origem de dados que suporte Unicode (um arquivo de texto de banco de dados ou UTF-8) e contenha caracteres de um código de idioma diferente ou de vários códigos de idioma ou conjuntos de caracteres.
- Imagens do IBM SPSS Modeler Solution Publisher são codificadas em UTF-8 e são totalmente móveis entre plataformas e códigos de idioma.

Sobre Unicode

O objetivo do padrão Unicode é fornecer uma maneira consistente de codificar um texto multilíngue para que ele possa ser facilmente compartilhado entre bordas, códigos de idioma e aplicativos. O padrão Unicode, agora na versão 4.0.1, define um conjunto de caracteres que é um superconjunto de todos os conjuntos de caracteres de uso comum no mundo atualmente e designa a cada caractere um nome exclusivo e um ponto de código. Os caracteres e seus pontos de código são idênticos aqueles do Conjunto de Caracteres Universais (UCS) definido pelo ISO-10646. Para obter mais informações, consulte Página Inicial do Unicode.

Avisos

Estas informações foram desenvolvidas para produtos e serviços oferecidos no mundo todo.

A 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. Qualquer referência a produtos, programas ou serviços IBM não significa que apenas produtos, programas ou serviços IBM possam ser utilizados. Qualquer produto, programa ou serviço funcionalmente equivalente, que não infrinja nenhum direito de propriedade intelectual da IBM pode ser usado em substituição a este produto, programa ou serviço. Entretanto a avaliação e verificação da operação de qualquer produto, programa ou serviço não IBM são de responsabilidade do Cliente.

A IBM pode ter patentes ou solicitações de patentes pendentes relativas a assuntos tratados nesta publicação. O fornecimento desta publicação não lhe garante direito algum sobre tais patentes. Pedidos de licença devem ser enviados, por escrito, para:

Gerência de Relações Comerciais e Industriais da IBM Brasil
Av. Pasteur, 138-146
Botafogo
Rio de Janeiro, RJ
CEP 22290-240

Para pedidos de licença relacionados a informações de DBCS (Conjunto de Caracteres de Byte Duplo), 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
Legal and Intellectual Property Law
IBM Japan Ltd.
1623-14, Shimotsuruma, Yamato-shi
Kanagawa 242-8502 Japan

O parágrafo a seguir não se aplica ao Reino Unido nem a nenhum país em que tais disposições não estejam de acordo com a legislação local: 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 A ELAS NÃO SE LIMITANDO, AS GARANTIAS IMPLÍCITAS DE NÃO INFRAÇÃO, COMERCIALIZAÇÃO OU ADEQUAÇÃO A UM DETERMINADO PROPÓSITO. Alguns países não permitem a exclusão de garantias expressas ou implícitas em certas transações; portanto, essa disposição pode não se aplicar ao Cliente.

Essas informações podem conter imprecisões técnicas ou erros tipográficos. São feitas mudanças periódicas nas informações aqui contidas; tais mudanças serão incorporadas em futuras edições desta publicação. A IBM pode, a qualquer momento, aperfeiçoar e/ou alterar os produtos e/ou programas descritos nesta publicação, sem aviso prévio.

As referências nestas informações à documentação ou a websites não IBM são fornecidas apenas por conveniência e não representam de forma alguma um endosso a esses documentos ou websites. Os materiais contidos nesses websites não fazem parte dos materiais desse produto IBM e o uso desses websites é de inteira responsabilidade do Cliente.

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

Os licenciados deste programa que desejarem obter informações sobre este assunto com o propósito de permitir: (i) a troca de informações entre programas criados independentemente e outros programas (incluindo este) e (ii) o uso mútuo das informações trocadas, deverão entrar em contato com:

Gerência de Relações Comerciais e Industriais da IBM Brasil
Av. Pasteur, 138-146
Botafogo
Rio de Janeiro, RJ
CEP 22290-240

Tais informações podem estar disponíveis, sujeitas aos termos e condições apropriados, 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 Internacional de Licença do Programa IBM ou de qualquer outro contrato equivalente.

Todos os dados de desempenho aqui contidos foram determinados em um ambiente controlado. Portanto, os resultados obtidos em outros ambientes operacionais podem variar significativamente. Algumas medidas podem ter sido tomadas em sistemas em nível de desenvolvimento e não há garantia de que estas medidas serão iguais em sistemas geralmente disponíveis. Além disso, algumas medidas podem ter sido estimadas por extrapolação. Os resultados reais podem variar. Os usuários deste documento devem verificar os dados aplicáveis para seu ambiente específico.

As informações relativas a produtos não IBM foram obtidas junto aos fornecedores dos respectivos produtos, a partir 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. As dúvidas sobre os recursos de produtos não IBM devem ser encaminhadas diretamente aos seus fornecedores.

Todas as declarações relacionadas aos objetivos e intenções futuras da IBM estão sujeitas a mudanças ou cancelamento sem aviso prévio e representam apenas metas e objetivos.

Estas informações contêm exemplos de dados e relatórios usados nas operações diárias de negócios. Para ilustrá-los da forma mais completa possível, os exemplos podem incluir nomes de indivíduos, empresas, marcas e produtos. Todos estes nomes são fictícios e qualquer semelhança com nomes e endereços usados por uma empresa real é mera coincidência.

Se estas informações estiverem sendo exibidas em formato eletrônico, as fotografias e ilustrações coloridas podem não aparecer.

Marcas comerciais

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

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

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

Java e todas as marcas comerciais e logotipos baseados em Java são marcas comerciais ou marcas registradas da Oracle e/ou de seus afiliados.

Outros nomes de produtos e serviços podem ser marcas comerciais da IBM ou de outras empresa.

Índice Remissivo

A

abrindo
 estados 61
 fluxos 61
 modelos 61
 nós 61
 projetos 178
 saída 61
acessibilidade 199
 dicas no IBM SPSS Modeler 208
 exemplo 206, 207
 variáveis no IBM SPSS Modeler 199
Acessibilidade 209
agrupando 88
 controlando a largura da coluna para texto agrupado 88
agrupando linhas e colunas 84
ajuste de escala
 tabelas dinâmicas 88, 90
ajuste de escala de fluxos para visualização 20
alinhamento 68, 95
 saída 68, 95
análise baseada em árvore
 aplicativos típicos 23
análise do Champion Challenger 154, 166
anotações
 convertendo em comentários 58
 pasta 180
 projeto 180
anotando
 fluxos 54, 59
 nós 54, 59
aplicando zoom 17
aplicativos 23
aplicativos de mineração de dados 24
aplicativos típicos 23
aprendizado por máquina 23
area
 funções espaciais 132
armazenando objetos no Repositório do IBM SPSS Collaboration and Deployment Services 156
arquivo de dicionário 207
arquivos 70
 incluindo um arquivo de texto no Visualizador 70
arquivos BMP 73, 79
 exportando gráficos 73, 79
arquivos de dados
 vários arquivos de dados abertos 95
arquivos de dados de texto
 codificação 211
arquivos de Estatísticas
 codificação 211
arquivos de fluxo de backup
 restaurando 59
arquivos de log
 exibindo SQL gerado 44

arquivos de saída
 salvamento 60
arquivos EPS 73, 80
 exportando gráficos 73, 80
arquivos JPEG 73, 79
 exportando gráficos 73, 79
arquivos PNG 73, 79
 exportando gráficos 73, 79
arquivos PostScript (contidos) 73, 80
 exportando gráficos 73, 80
arquivos PowerPoint 178
arquivos SAS
 codificação 211
arquivos TIFF 80
 exportando gráficos 73, 80
árvores de decisão
 Acessibilidade 208
atalhos
 teclado 20, 200, 201, 203, 204, 205
 uso geral 64
atalhos de teclado 200, 201, 203, 204, 205
ativar nós 35
atributo 23
atualização de modelo 166
atualizando modelos 169, 170
atualizar
 nós de origem 40
automação 103
avisos 46
 configurando opções 186

B

banco de dados
 funções 111, 112
barra de ferramentas 17
bloqueando nós 39
bloqueando objetos do Repositório do IBM SPSS Collaboration and Deployment Services 163
bordas 89, 93
 exibindo bordas ocultas 93
botão do meio do mouse
 simulando 20, 34

C

cabeçalhos 81
cache
 ativação 186
 configurando um cache 36
 limpando 37, 40
 salvamento 37
caixa de diálogo bem-vindo 188
caixa de diálogo inicialização 188
caixas de diálogo 95
 exibindo nomes de variáveis 95
 exibindo rótulos de variáveis 95
 ordem de exibição de variáveis 95
camadas 80, 86, 88, 90

camadas (*continuação*)
 criação 86
 em tabelas dinâmicas 86
 exibindo 86
 imprimindo 80, 88, 90
campos 23, 119, 121
 em expressões do CLEM 114
 visualizando valores 115
campos de hora
 convertendo 144
campos de modelo 63
campos essenciais 62, 63
campos obrigatórios 63
caractere de barra invertida em expressões do CLEM 120
caracteres 119, 120
caracteres especiais
 removendo das sequências de caracteres 107
carregando
 estados 61
 nós 61
casas decimais
 formatos de exibição 42
caso 23
células em tabelas dinâmicas 86, 89, 92
 formatos 89
 larguras 92
 mostrando 86
 ocultando 86
 selecionando 93
cenários 166
 definidos 154
 opções de implementação 166
centralizando saída 68, 95
classes 17, 177, 178
CLEM 111
 construindo expressões 111
 exemplos 103
 expressões 105, 119
 funções 111, 112
 idioma 119
 introdução 22, 103
 tipos de dados 119, 120, 121
 verificando expressões 115
cliente
 diretório padrão 186
close_to
 funções espaciais 132
close_to function 132
codificação 40, 211
 codificação de texto 40
 codificação padrão de fluxo 40
 codificação UTF-8 40, 211
código de idioma
 opções 185
Cognos active report 75
colando saída em outros aplicativos 71
colar 17
colunas 93

- colunas (*continuação*)
 - alterando a largura em tabelas dinâmicas 93
 - selecionando em tabelas dinâmicas 93
- comentários
 - atalhos de teclado 204, 205
 - em nós e fluxos 54
 - listando todos em um fluxo 58
- concatenando sequências de caracteres 126
- condições 105
- conexão única; 10
- conexão única, IBM SPSS Collaboration and Deployment Services Repository 153, 155
- conexões
 - cluster de servidores 11
 - para o IBM SPSS Modeler Server 10, 11
 - para o repositório do IBM SPSS Collaboration and Deployment Services 155
- configuração de página 81
- cabeçalhos e rodapés 81
- tamanho do gráfico 81
- conformidade 508 199
- conjuntos 40
- conjuntos de múltiplas categorias em expressões do CLEM 110
- conjuntos de múltiplas dicotomias em expressões do CLEM 110
- conjuntos de múltiplas respostas em expressões do CLEM 110, 114
- conjuntos de regras avaliando 40
- Construtor de Expressões 203
 - acesso 111
 - usando 111
 - visão geral 111
- controlando o número de linhas para exibir 88
- convenções 125
- coordenadas geoespaciais
 - formato de exibição 46
 - selecionando sistemas 46
- Coordinator of Processes 11
- COP 11
- copiando e colando saída em outros aplicativos 71
- copiar 17
- copiar especial 71
- cor do plano de fundo 90
- cores
 - configuração 188
- cores de linha alternativas tabelas dinâmicas 89
- cores de script
 - configuração 189
- cores em tabelas dinâmicas 89
 - bordas 89
- cortar 17
- credenciais
 - para o Repositório do IBM SPSS Collaboration and Deployment Services 155
- criação de paleta customizada 190

- criação de paleta customizada (*continuação*)
 - criação de subpaleta 192
- CRISP-DM 17, 177
 - visualização de projetos 177
- crosses
 - funções espaciais 132
- customização da guia paleta 192

D

- dados
 - pré-visualização 38
- dados ruidosos 24
- datas
 - convertendo 144
 - manipulando 144
- desativar nós 35, 36
- desbloqueando objetos do Repositório do IBM SPSS Collaboration and Deployment Services 163
- descoberta do conhecimento 23
- descrições de fluxo 52, 53
- desempenho
 - dos nós de modelagem 196
 - dos nós de processamento 195
 - expressões do CLEM 197
- desfazer 17
- dias de sobreposição 41
- dicas
 - para acessibilidade 208
 - uso geral 64
- Dicas de ferramentas anotando nós 59
- diretório
 - padrão 186
- diretório temporário 12
- distance
 - funções espaciais 132
- distribuição f
 - funções de probabilidade 132
- distribuição normal
 - funções de probabilidade 132
- distribuição qui-quadrado
 - funções de probabilidade 132
- distribuição t
 - funções de probabilidade 132
- dividindo tabelas 93
 - controlando quebras de tabela 93
- documentação 3
- DTD 174

E

- Editor de Dados
 - opções de estatísticas descritivas 96
 - vários arquivos de dados abertos 95
- efetuando login no IBM SPSS Modeler Server 10
- em branco 98, 108
- erro na visualização de renderização memória insuficiente 185
- erro no estouro da capacidade 185
- escoragem
 - ramificação 54, 168, 169, 170

- espaço em branco
 - removendo das sequências de caracteres 107, 135
- espaços
 - removendo das sequências de caracteres 107, 135
- estados
 - carregando 61
 - salvamento 60
- estrutura de tópicos 68, 69
 - alterando níveis 69
 - expandindo 69
 - no Visualizador 68
 - reduzindo 69
- excluindo saída 68
- executando fluxos 53
- exemplos
 - Guia de Aplicativos 3
 - visão geral 5
- exemplos de aplicativos 3
- exportando
 - descrições de fluxo 53
 - PMML 174, 175
- exportando gráficos 73, 79, 80
- exportando saída 73, 77, 78
 - formato do Word 73, 75
 - formato Excel 73, 76
 - formato HTML 73
 - formato PDF 73, 77
 - formato PowerPoint 73
 - HTML 74
 - relatório da web 75
- expressões 119
- expressões do CLEM
 - desempenho 197

F

- fator 208
- ferramenta de mapeamento de dados 62
- fluxo 13
- fluxos 9
 - ajuste de escala para visualização 20
 - anotando 54, 59
 - armazenando no Repositório do IBM SPSS Collaboration and Deployment Services 158
 - arquivos de backup 59
 - carregando 61
 - conectando nós 34
 - construindo 33
 - desativando nós 35
 - efetuando bypass em nós 35
 - execução 53
 - incluindo comentários 54
 - incluindo em projetos 178
 - incluindo nós 34, 36
 - opções 39, 40, 41, 42, 44, 45, 46
 - opções de implementação 166
 - renomeando 52, 59
 - salvamento 59
 - sistema de coordenadas geoespaciais 46
 - visualizando tempos de execução 47
- fluxos de dados
 - construindo 33
- fontes 69, 90, 188, 189

fontes (*continuação*)
na área de janela da estrutura de tópicos 69

formato de exibição de moeda 42

formato do Word
exportando saída 73, 75
tabelas grandes 73

formato Excel
exportando saída 73, 76

formato PowerPoint
exportando saída 73

formatos de coordenadas geoespaciais 46

formatos de data 41, 121, 122

formatos de exibição
casas decimais 42
científico 42
coordenadas geoespaciais 46
moeda 42
números 42
símbolo de agrupamento 42

formatos de exibição de número 42

formatos de hora 41, 121, 122

função @BLANK 102, 126, 150

função @DIFF 145

função @FIELD 102, 151

função @FIELDS_BETWEEN 102, 109, 151

função @FIELDS_MATCHING 102, 109, 151

função @INDEX 145

função @LAST_NON_BLANK 145, 150

função @MAX 145

função @MEAN 145

função @MIN 145

função @MULTI_RESPONSE_SET 110, 151

função @NULL 102, 126, 150

função @OFFSET 145
considerações de desempenho 197

função @PARTITION_FIELD 151

função @PREDICTED 151

função @SDEV 145

função @SINCE 145

função @SUM 145

função @TARGET 151

função @TESTING_PARTITION 151

função @THIS 145

função @TODAY 140

função @TRAINING_PARTITION 151

função @VALIDATION_PARTITION 151

função abs 130

função aleatória 134

função allbutfirst 135

função allbutlast 135

função alphabefore 135

função arccos 131

função arccosh 131

função arcsin 131

função arcsinh 131

função arctan 131

função arctan2 131

função arctanh 131

função area 132

função cdf_chisq 132

função cdf_f 132

função cdf_normal 132

função cdf_t 132

função cos 131

função cosh 131

função count_equal 109, 128

função count_greater_than 109, 128

função count_less_than 109, 128

função count_non_nulls 128

função count_not_equal 109, 128

função count_nulls 102, 109, 128

função count_substring 135

função crosses 132

função date_before 128

função datetime_date 126

função DIFF 145

função distance 132

função div 130

função endstring 135

função exponencial 130

função first_index 110, 128

função first_non_null 110, 128

função first_non_null_index 110, 128

função fracof 130

função hasendstring 135

função hasmidstring 135

função hasstartstring 135

função hassubstring 135

função INDEX 145

função integer_bitcount 133

função integer_leastbit 133

função integer_length 133

função intof 130

função is_date 126

função is_datetime 126

função is_integer 126

função is_number 126

função is_real 126

função is_string 126

função is_time 126

função is_timestamp 126

função isalphacode 135

função isendstring 135

função islowercode 135

função ismidstring 135

função isnumbercode 135

função isstartstring 135

função issubstring 135

função issubstring_count 135

função issubstring_lim 135

função isuppercode 135

função last_index 110, 128

função LAST_NON_BLANK 145

função last_non_null 110, 128

função last_non_null_index 110, 128

função length 135

função locchar 135

função locchar_back 135

função log 130

função log10 130

função lowertoupper 135

função matches 135

função max 128

função MAX 145

função max_index 110, 128

função max_n 109, 128

função MEAN 145

função mean_n 109, 130

função member 128

função min 128

função MIN 145

função min_index 110, 128

função min_n 109, 128

função mod 130

função negate 130

função num_points 132

função OFFSET 145

função oneof 134

função overlap 132

função pi 131

função power (exponencial) 130

função random0 134

função rem 130

função replace 135

função replicate 135

função round 130

função SDEV 145

função sdev_n 109, 130

função sign 130

função sin 131

função SINCE 145

função sinh 131

função skipchar 135

função skipchar_back 135

função soundex 140

função soundex_difference 140

função sqrt 130

função startstring 135

função stripchar 135

função strmembre 135

função subscrs 135

função substring 135

função substring_between 135

função SUM 145

função sum_n 109, 130

função tan 131

função tanh 131

função testbit 133

função THIS 145

função time_before 128

função to_date 126, 140

função to_dateline 140

função to_datetime 126

função to_integer 126

função to_number 126

função to_real 126

função to_string 126

função to_time 126, 140

função to_timestamp 126, 140

função trim 135

função trim_start 135

função trimend 135

função undef 150

função unicode_char 135

função unicode_value 135

função uppertolower 135

função value_at 110, 128

função within 132

funções 121, 122, 125, 126, 145
@BLANK 101
@FIELD 111, 151
@GLOBAL_MAX 150
@GLOBAL_MEAN 150
@GLOBAL_MIN 150
@GLOBAL_SDEV 150
@GLOBAL_SUM 150

funções (*continuação*)
 @PARTITION 151
 @PREDICTED 111, 151
 @TARGET 111, 151
 banco de dados 111, 112
 em expressões CLEM 111
 exemplos 103
 funções definidas pelo usuário (UDFs) 111
 manipulando valores omissos 102
 funções bit a bit 133
 funções data/hora
 datetime_date 140
 datetime_day 140
 datetime_day_name 140
 datetime_day_short_name 140
 datetime_hour 140
 datetime_in_seconds 140
 datetime_minute 140
 datetime_month 140
 datetime_month_name 140
 datetime_month_short_name 140
 datetime_now datetime_second 140
 datetime_time 140
 datetime_timestamp 140
 datetime_weekday 140
 datetime_year 140
 funções de banco de dados
 em expressões do CLEM 112
 funções definidas pelo usuário (UDFs) 112
 Funções de CLEM (Idioma Comum para Manipulação de Expressão)
 conversão 126
 lista de disponíveis 124
 funções de comparação 128
 funções de conversão 126
 funções de data 121, 122
 date_before 128, 140
 date_days_difference 140
 date_in_days 140
 date_in_months 140
 date_in_weeks 140
 date_in_years 140
 date_months_difference 140
 date_weeks_difference 140
 date_years_difference 140
 função @TODAY 140
 funções de data e hora 121, 122
 funções de distribuição 132
 funções de hora 121, 122
 time_before 128, 140
 time_hours_difference 140
 time_in_hours 140
 time_in_mins 140
 time_in_secs 140
 time_mins_difference 140
 time_secs_difference 140
 funções de informações 126
 funções de probabilidade 132
 funções de sequência 145
 funções de sequência de caracteres 135
 funções definidas pelo usuário (UDFs) 111, 112
 funções do CLEM
 aleatório 134
 bit a bit 133

funções do CLEM (*continuação*)
 comparação 128
 datetime 140
 em branco e nulos 150
 espaciais 132
 funções especiais 151
 globais 150
 informações 126
 lógicos 130
 numeric 130
 probabilidade 132
 sequência 135, 145
 trigonométricas 131
 valores ausentes 102
 funções espaciais 132
 funções especiais 151
 funções globais 150
 funções if, then, else 130
 funções lógicas 130
 funções numéricas 130
 funções trigonométricas 131

G

geração de SQL
 criação de log 44
 visualizando 44
 gerenciadores 15
 girar rótulos 84
 gráfico de tabela 94
 gráficos 67, 73, 94
 criando a partir de tabelas dinâmicas 94
 exportando 73
 incluindo em projetos 178
 ocultando 67
 salvando saída 60
 graus
 unidades de medida 42

H

HTML 73, 74
 exportando saída 73, 74

I

IBM InfoSphere Warehouse (ISW)
 exportação de PMML 175
 IBM SPSS Collaboration and Deployment Services 154
 IBM SPSS Modeler 1, 13
 dicas e atalhos 64
 documentação 3
 executando a partir da linha de comandos 9
 introdução 9
 opções 185
 variáveis de acessibilidade 199
 visão geral 9, 185
 IBM SPSS Modeler Advantage 154, 166
 IBM SPSS Modeler Server 1
 ID do usuário 10
 nome de domínio (Windows) 10
 nome do host 10, 11
 número da porta 10, 11

IBM SPSS Modeler Server (*continuação*)
 senha 10
 ícones
 configurando opções 20, 45
 ID do usuário
 IBM SPSS Modeler Server 10
 idioma
 alterando a linguagem de saída 85
 opções 185
 implementação 154
 implementando cenários 166
 importando
 PMML 174, 175
 imprimindo 21, 80, 81, 88, 90, 93
 ajuste de escala de tabelas 88, 90
 cabeçalhos e rodapés 81
 camadas 80, 88, 90
 controlando quebras de tabela 93
 espaço entre itens de saída 81
 fluxos 20, 36
 gráficos 80
 números da página 81
 saída de texto 80
 tabelas dinâmicas 80
 tamanho do gráfico 81
 visualização de impressão 80
 incluindo
 em um projeto 178
 incluindo conexões do IBM SPSS Modeler Server 11
 incluindo rótulos do grupo 84
 inserindo rótulos do grupo 84
 introdução 119
 IBM SPSS Modeler 9, 185

J

janela Árvore Interativa
 Acessibilidade 208
 janela principal 13
 Java 209
 JAWS 199, 207, 209
 justificação 68, 95
 saída 68, 95

L

largura da coluna 88, 93, 95
 controlando a largura máxima 88
 controlando a largura para texto agrupado 88
 controlando largura padrão 95
 tabelas dinâmicas 93
 legendas 91
 leitores de tela 201, 203, 207, 208
 exemplo 206, 207
 linha de comandos
 iniciando o IBM SPSS Modeler 9
 linhas 93
 selecionando em tabelas dinâmicas 93
 linhas de grade 93
 tabelas dinâmicas 93
 listando todos os comentários para um fluxo 58
 listas 119, 120

localizar e substituir
documentos do Visualizador 70

M

manipulando em branco
funções do CLEM 150
mapeando campos 62
mapeando dados 63
memória 95
erro no estouro da capacidade 185
gerenciando 185, 186
memória insuficiente 185
mensagens
exibindo SQL gerado 44
mensagens de erro 46
metarquivos 73
exportando gráficos 73
métodos de seleção 93
selecionando linhas e colunas em
tabelas dinâmicas 93
mineração de dados 23
estratégia 25
exemplos de aplicativos 32
minimizando 19
modelagem
ramificação 54
modelo de processo CRISP-DM 25, 26
modelos 54, 62
armazenando no Repositório do IBM
SPSS Collaboration and Deployment
Services 159
atualizando 169, 170
exportando 189
incluindo em projetos 178
substituindo 186
modelos PMML
regressão linear 189
regressão logística 189
mostrando 67, 87
legendas 91
linhas ou colunas 87
notas de rodapé 91
resultados 67
rótulos de dimensão 87
títulos 87
mouse
usando no IBM SPSS Modeler 20, 34
movendo linhas e colunas 84

N

navegando
atalhos de teclado 200
nó Agregado
desempenho 195
nó da Visualização Corporativa 166
nó de auditoria de dados
usar na exploração 23
nó de Auditoria de Dados
usar na mineração de dados 24
nó de Avaliação
desempenho 195
nó de Categorização
desempenho 195

nó de Mesclagem
desempenho 195
nó de Ordem
desempenho 195
nó de preenchimento
valores omissos 101
nó de rede neural
conjuntos grandes 40
nó de Rede Neural
desempenho 196
nó de regra de construção
carregando 61
nó Distinto
desempenho 195
nó do arquivo de cache
carregando 61
nó k-médias
conjuntos grandes 40
desempenho 196
nó Kohonen
conjuntos grandes 40
desempenho 196
nó Qualidade
valores omissos 98
nó Seleção de Variável
valores omissos 98
nó Tipo
desempenho 195
valores omissos 101
nome de domínio (Windows)
IBM SPSS Modeler Server 10
nome do host
IBM SPSS Modeler Server 10, 11
nomeando nós e fluxos 59
nomes de fluxo 59
nomes de nó 59
nomes de variáveis 95
em caixas de diálogo 95
nós 9
anotando 54, 59
armazenando no Repositório do IBM
SPSS Collaboration and Deployment
Services 159
ativação 35
bloqueando 39
carregando 61
conectando em um fluxo 34
configurando opções 36
criação de paleta customizada 190
criação de subpaleta
customizada 192
customização da guia paleta 192
desativando 35, 36
desativando em um fluxo 35
desempenho 195, 196
duplicando 36
edição 36
efetuando bypass em um fluxo 35
excluindo 34
excluindo conexões 36
exibindo na paleta 191
incluindo 34, 36
incluindo comentários em 54
incluindo em projetos 178
introdução 33
ordem dos 193
procurando por 51

nós (*continuação*)
removendo da paleta 191
salvamento 59
tempos de execução 47
visualização de dados 38
visualizando dados 38
nós de modelagem 33
customização da guia paleta de
modelagem 192
desempenho 196
nós de origem 33
atualizando 40
mapeamento de dados 62
nós de processamento 33
desempenho 195
nós de saída 33
nós terminais 33
notação científica 95
formato de exibição 42
suprimindo saída 95
notas de rodapé 88, 91, 92
marcadores 88
renumerando 92
notificações
configurando opções 186
novas variáveis 7
nuggets 54
definidos 15
nuggets do modelo 54
nulos 108
num_points
funções espaciais 132
numeração de página 81
número da porta
IBM SPSS Modeler Server 10, 11
números 108, 119, 120
números inteiros 119

O

objetos
propriedades 181
objetos de saída
armazenando no Repositório do IBM
SPSS Collaboration and Deployment
Services 159
ocultando 67, 68, 87
legendas 91
linhas e colunas 87
notas de rodapé 91
resultados do procedimento 68
rótulos de dimensão 87
títulos 87
opções 95, 96, 185
aparência da tabela dinâmica 95
estatísticas descritivas no Editor de
Dados 96
exibição 188
geral 95
para IBM SPSS Modeler 185
PMML 189
propriedades de fluxo 39, 40, 41, 42,
44, 45, 46, 47
rótulos de saída 96
sintaxe 189
usuário 186
Visualizador 95

- opções de implementação
 - cenários 166
- opções de usuário 186
- operador and 130
- operador de igual 128
- operador maior que 128
- operador menor que 128
- operador não igual 128
- operador not 130
- operador or 130
- operadores
 - associando sequências de caracteres 126
 - em expressões CLEM 111
- ordem de exibição 84
- ordenando
 - linhas da tabela dinâmica 85
- overlap
 - funções espaciais 132

P

- padrão
 - fase do projeto 177
- palavras-chave
 - anotando nós 59
- paleta de modelos 159
- paleta de modelos gerados 15
- paletas 13
 - customizando 190
- parâmetros
 - construção de modelo 168
 - em expressões do CLEM 114
 - escoragem 168
 - fluxo 49
 - prompts de tempo de execução 49
 - sessão 49
 - tipo 49
 - usando em cenários 168
- parâmetros de fluxo 49
- parâmetros de sessão 49
- parar execução 17
- pastas, Repositório do IBM SPSS Collaboration and Deployment Services 162, 163
- PDF
 - exportando saída 73, 77
- período 40
- PMML
 - exportando modelos 174, 175
 - importando modelos 174, 175
 - opções de exportação 189
- PowerPoint 77
 - exportando saída como PowerPoint 77
- precedência 122
- precedência do operador 122
- Predictive Applications 166
- procurando
 - por nós em um fluxo 51
- procurando conexões no COP 11
- procurando objetos no Repositório do IBM SPSS Collaboration and Deployment Services 161
- procurar e substituir
 - documentos do Visualizador 70
- programação visual 13

- projetos 17, 177
 - anotando 180
 - armazenando no Repositório do IBM SPSS Collaboration and Deployment Services 158
 - configurando propriedades 180
 - configurando uma pasta padrão 177
 - construindo 178
 - criando nova 178
 - fechando 181
 - gerando relatórios 181
 - incluindo objetos 178
 - no Repositório do IBM SPSS Collaboration and Deployment Services 179
 - propriedades da pasta 180
 - propriedades do objeto 181
 - Visualização Classes 178
 - Visualização CRISP-DM 177
- prompts, tempo de execução 49
- propriedades 88
 - fases do relatório 181
 - para fluxos de dados 39
 - pasta do projeto 180
 - tabelas 88
 - tabelas dinâmicas 88
- propriedades de célula 90
- propriedades de fluxo
 - Analytic Server 45
- propriedades do objeto, Repositório do IBM SPSS Collaboration and Deployment Services 164

Q

- quebras de tabela 93

R

- radianos
 - unidades de medida 42
- ramificações, modelagem e escoragem 54, 168, 169, 170
- reais 119, 120
- recuperando objetos do Repositório do IBM SPSS Collaboration and Deployment Services 160
- Recurso de produção 95
- Recurso de Produção
 - usando sintaxe de comando do arquivo de diário 95
- redimensionando 19
- registros 23
 - valores omissos 98
 - valores omissos do sistema 99
- regressão 208
- regressão linear
 - exportar como PMML 189
- regressão logística 208
 - exportar como PMML 189
- relatório da web 75
 - exportando saída 75
- relatórios
 - configurando propriedades 181
 - gerando 181
 - incluindo em projetos 178

- relatórios (*continuação*)
 - salvando saída 60
- removendo o mapeamento de campos 62
- removendo rótulos do grupo 84
- renomeando
 - fluxos 52
 - nós 59
- reordenando linhas e colunas 84
- Repositório do IBM SPSS Collaboration and Deployment Services 153, 154
 - armazenando objetos 156
 - bloqueando e desbloqueando objetos 163
 - conectando ao 155
 - conexão única 155
 - credenciais 155
 - excluindo objetos e versões 163
 - navegando 155
 - pastas 162, 163
 - procurando no 161
 - propriedades do objeto 164
 - recuperando objetos 160
 - transferindo projetos para 179
- rodapés 81
- rolagem
 - configurando opções 45
- rótulos 84
 - excluindo 84
 - exibindo 40
 - inserindo rótulos do grupo 84
 - value 174
 - variável 174
- rótulos, objeto do Repositório do IBM SPSS Collaboration and Deployment Services Repository 165
- rótulos de valor 85, 96
 - em tabelas dinâmicas 96
 - na área de janela da estrutura de tópicos 96
- rótulos de variáveis 85, 95, 96
 - em caixas de diálogo 95
 - em tabelas dinâmicas 96
 - na área de janela da estrutura de tópicos 96
- rótulos de versão, objeto do Repositório do IBM SPSS Collaboration and Deployment Services 165
- rótulos do grupo 84

S

- saída 15, 67, 68, 71, 73, 82, 95
 - alinhamento 68, 95
 - alterando a linguagem de saída 85
 - centralização 68, 95
 - colando em outros aplicativos 71
 - copiando 68
 - criptografando 82
 - excluindo 68
 - exportando 73
 - interativa 72
 - mostrando 67
 - movendo 68
 - ocultando 67
 - salvamento 82
 - Visualizador 67

saída HTML
 leitor de tela 208
 saída interativa 72
 salvamento
 estados 60
 fluxos 59
 nós 59
 objetos de saída 60
 objetos múltiplos 60
 salvando gráficos 73, 79, 80
 arquivos BMP 73, 79
 arquivos EMF 73
 arquivos EPS 73, 80
 arquivos JPEG 73, 79
 arquivos PICT 73
 arquivos PNG 79
 arquivos PostScript 80
 arquivos TIFF 80
 metarquivos 73
 salvando saída 73, 77, 78
 formato de texto 73, 78
 formato do Word 73, 75
 formato Excel 73, 76
 formato HTML 73
 formato PDF 73, 77
 formato PowerPoint 73, 77
 HTML 73, 74
 relatório da web 75
 Scripts 22, 103
 seleção da paleta do nó 191
 seleção do sistema de coordenadas
 geoespaciais 46
 senha
 IBM SPSS Modeler Server 10
 seqüências de caracteres 119, 120
 correspondentes 107
 manipulando em expressões do
 CLEM 107
 substituindo 107
 servidor
 diretório padrão 186
 efetuando login 10
 incluindo conexões 11
 procurando servidores no COP 11
 símbolo de agrupamento
 formatos de exibição de número 40
 símbolo decimal
 formatos de exibição de número 40
 sistema
 opções 185
 sistema de medição 95
 subpaleta
 criação 192
 exibindo na guia de paleta 191
 removendo da guia de paleta 191
 substituindo modelos 186
 sugestões
 uso geral 64
 suporte Unicode 211

T

tabelas 93, 203
 alinhamento 90
 controlando quebras de tabela 93
 cor do plano de fundo 90
 fontes 90

tabelas (*continuação*)
 incluindo em projetos 178
 margens 90
 propriedades de célula 90
 salvando saída 60
 tabelas anteriores 94
 tabelas dinâmicas 67, 71, 73, 80, 83, 84,
 86, 87, 88, 89, 90, 92, 93, 94, 95
 agrupando linhas e colunas 84
 ajustamento de largura da coluna
 padrão 95
 ajuste de escala para ajustar à
 página 88, 90
 alinhamento 90
 alterando a aparência 87
 alterando a ordem de exibição 84
 aparência padrão para novas
 tabelas 95
 bordas 89
 camadas 86
 colando como tabelas 71
 colando em outros aplicativos 71
 colocando sobre o eixo 83, 84
 controlando o número de linhas para
 exibir 88
 controlando quebras de tabela 93
 cor do plano de fundo 90
 cores de linha alternativas 89
 criando gráficos a partir de
 tabelas 94
 desagrupando linhas ou colunas 84
 desfazendo mudanças 86
 edição 83
 excluindo rótulos do grupo 84
 exibindo bordas ocultas 93
 exportando como HTML 73
 fontes 90
 formatos de célula 89
 girar rótulos 84
 idioma 85
 imprimindo camadas 80
 imprimindo tabelas grandes 93
 inserindo linhas e colunas 85
 inserindo rótulos do grupo 84
 larguras de células 92
 legendas 91
 linhas de grade 93
 manipulando 83
 margens 90
 mostrando e ocultando células 86
 movendo linhas e colunas 84
 notas de rodapé 91, 92
 ocultando 67
 ordenando linhas 85
 propriedades 88
 propriedades da nota de rodapé 88
 propriedades de célula 90
 propriedades gerais 88
 renderizar tabelas mais rápido 95
 rótulos de valor 85
 rótulos de variáveis 85
 selecionando linhas e colunas 93
 tabelas anteriores 94
 tabelas dinâmicas rápidas 95
 texto de continuação 90
 transpondo linhas e colunas 84
 usando ícones 84

tabelas dinâmicas rápidas 95
 tabelas grandes
 colando no Microsoft Word 71
 TableLooks 87
 aplicando 87
 criação 87
 tamanhos 69
 na estrutura de tópicos 69
 teclas de atalho 20
 tela 13
 tela de fluxo
 configurações 45
 tempos de execução, visualizando 47
 text 70, 73
 texto 78
 exportando saída como texto 73, 78
 incluindo no Visualizador 70
 incluindo um arquivo de texto no
 Visualizador 70
 texto de continuação 90
 para tabelas dinâmicas 90
 texto do rótulo vertical 84
 tipo de implementação 166
 tipos de dados 105
 em parâmetros 49
 títulos 70
 incluindo no Visualizador 70
 transpondo linhas e colunas 84

V

valores 105
 incluindo em expressões do
 CLEM 115
 visualizando a partir de uma
 auditoria de dados 115
 valores de data/hora 108
 valores globais
 em expressões do CLEM 114
 valores omissos 98, 108
 em registros 98
 expressões do CLEM 102
 sistemas 99
 valores omissos do sistema
 em registros 99
 variables
 ordem de exibição nas caixas de
 diálogo 95
 várias sessões do IBM SPSS Modeler 12
 variáveis 23, 95
 vários arquivos de dados abertos 95
 verificando expressões do CLEM 115
 vírgula 40
 visualização
 dados do nó 38
 Visualização Corporativa do IBM SPSS
 Collaboration and Deployment
 Services 154, 166
 Visualizador 67, 68, 69, 81, 82, 95, 96
 alterando a fonte da estrutura de
 tópicos 69
 alterando os níveis da estrutura de
 tópicos 69
 alterando tamanhos da estrutura de
 tópicos 69
 área de janela da estrutura de
 tópicos 67

Visualizador *(continuação)*

- área de janela de resultados 67
- espaço entre itens de saída 81
- estrutura de tópicos 68
- excluindo saída 68
- exibindo nomes de variáveis 96
- exibindo rótulos de valor 96
- exibindo rótulos de variáveis 96
- exibindo valores de dados 96
- expandindo a estrutura de tópicos 69
- localizar e substituir informações 70
- movendo saída 68
- ocultando resultados 67
- opções de exibição 95
- procurar e substituir informações 70
- reduzindo a estrutura de tópicos 69
- salvando documento 82

W

within

- funções espaciais 132



Impresso no Brasil