

IBM SPSS Analytic Server  
Versão 1

*Guia do Usuário*

**IBM**



---

# Índice

## Capítulo 1. Visão Geral . . . . . 1

Arquitetura . . . . . 2

## Capítulo 2. Modeler Client Integration . . 3

Origem do Analytic Server . . . . . 3

    conexões do Analytic Server . . . . . 3

    Selecionar Origem de Dados do Analytic Server . 3

Exportação do Analytic Server . . . . . 4

    Caminho de Destino do Analytic Server . . . . 4

## Capítulo 3. Console do Administrador . . 5

Para Efetuar Logon no Console do Administrador . . 5

Navegando no Console do Administrador . . . . . 5

Consumidores . . . . . 5

Funções . . . . . 7

Permissões . . . . . 8

Projetos . . . . . 9

Origens de Dados . . . . . 11

## Índice Remissivo . . . . . 19



---

## Capítulo 1. Visão Geral

IBM® SPSS Analytic Server é uma solução para analítica de dados grandes que combina tecnologia IBM SPSS com sistemas de dados grandes e lhe permite trabalhar com interfaces com o usuário IBM SPSS familiares para resolver problemas que não foram solucionados anteriormente.

### Porque a Analítica de Big Data É Importante

Os volumes de dados coletados pelas organizações estão crescendo exponencialmente; por exemplo, negócios financeiros e varejistas têm todas as transações dos clientes por um ano (ou dois anos, ou dez anos), provedores de telco têm Call Data Records (CDR) e leituras de sensor de dispositivo e as empresas de Internet têm os resultados de crawls da web.

A analítica de dados grandes é necessária onde existe:

- Um grande volume de dados (terabytes, petabytes, exabytes), principalmente quando há uma mistura de dados estruturados & dados não estruturados
- Dados que são alterados/acumulados rapidamente

A analítica de dados grandes também ajuda quando:

- Um grande número (milhares) de modelos está sendo construído
- Modelos são construídos/atualizados frequentemente

### Desafios

As mesmas organizações que coletam grandes volumes de dados costumam ter dificuldades ao fazerem uso deles por vários motivos:

- A arquitetura dos produtos de analítica tradicionais não é adequada para cálculo distribuído, e
- Algoritmos estatísticos existentes não são projetados para trabalharem com dados grandes (esses algoritmos esperam que os dados venham deles, mas os dados grandes são muito caros para se mover); dessa forma
- Executar uma analítica moderna em dados grandes requer novas qualificações e profundo conhecimento de sistemas de dados grandes. Poucos analistas têm essas qualificações.
- As soluções em memória trabalham para solucionar problemas de médio porte, mas não se adaptam muito bem a dados realmente grandes.

### Solução

Analytic Server fornece:

- Uma arquitetura centrada em dados que alavanca sistemas de dados grandes, como Hadoop Map/Reduce, como dados em HDFS.
- Uma interface definida para incorporar novos algoritmos estatísticos projetados para acessarem dados.
- Interfaces com o usuário do IBM SPSS familiares que ocultam os detalhes dos ambientes de dados grandes para que os analistas possam se concentrar na análise dos dados.
- Uma solução escalável para problemas de qualquer porte.

# Arquitetura

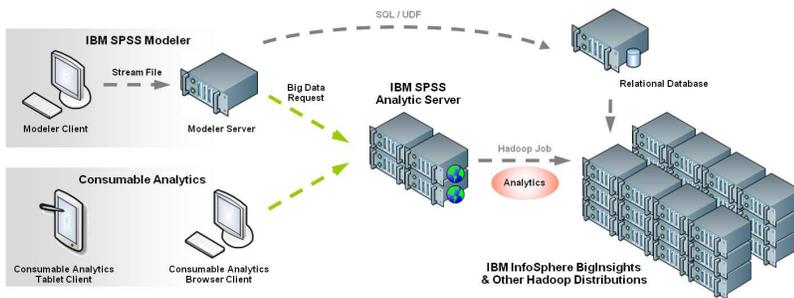


Figura 1. Arquitetura

O Analytic Server fica entre um aplicativo cliente e uma nuvem Hadoop. Supondo que os dados residam na nuvem, a estrutura de tópicos geral para se trabalhar com o Analytic Server é:

1. Definir a análise que você deseja executar no aplicativo cliente. Para a liberação atual, os aplicativos clientes são IBM SPSS Modeler e IBM SPSS Analytic Catalyst.
2. Quando você executa a análise, o aplicativo cliente envia uma solicitação de execução do Analytic Server.
3. O Analytic Server coordena a tarefa para ser executada na nuvem Hadoop e relata os resultados para o aplicativo cliente.
4. É possível usar os resultados para definir outras análises e as repetições de ciclo.

---

## Capítulo 2. Modeler Client Integration

---

### Origem do Analytic Server

A Origem do Analytic Server permite o uso de dados HDFS em sua análise. Normalmente um fluxo com uma Origem do Analytic Server é enviado para Analytic Server para ser executado em um ambiente de "dados grandes". Quando um fluxo com uma Origem do Analytic Server é executado localmente, as primeiras  $N$  linhas são transferidas por download do HDFS e a tarefa é executada localmente.

Em uma conexão do Analytic Server, selecione uma origem de dados contendo os dados que deseja usar. Uma origem de dados contém os arquivos e os metadados associados a essa origem.

**Conexão.** Clique em **Editar** para definir ou editar os detalhes de uma conexão com o servidor onde os dados residem. Consulte o tópico "conexões do Analytic Server" para obter mais informações.

**Origem de dados.** Após ter estabelecido uma conexão, clique em **Selecionar** para exibir uma lista de origens de dados disponíveis. Consulte o tópico "Selecionar Origem de Dados do Analytic Server" para obter mais informações.

**Número de linhas.** Quando você executa localmente, isso determina o número de linhas (registros) para obter do HDFS. Especifique um valor entre 100 e 100.000.

**Gerar ASL para este nó.** Quando selecionada, faz o fluxo ser executado no Analytic Server. Quando não selecionada, o fluxo é executado localmente e os registros são obtidos do HDFS.

### conexões do Analytic Server

O diálogo de conexões do Analytic Server permite selecionar o servidor do qual importar ou exportar dados. Entre em contato com o administrador do sistema se não estiver certo de quais configurações usar.

**URL do Servidor.** Insira a URL do servidor do qual importar ou exportar dados.

**Porta do servidor.** Insira o número da porta a ser usada para se conectar ao host.

**Raiz de contexto.** Insira o nome da raiz do aplicativo, que é o diretório de nível superior de um aplicativo quando ele é implementado em um servidor da web. *NOTA:* esse recurso está atualmente desativado.

**Consumidor.** Insira o consumidor sob o qual efetuar logon no servidor. Consumidores fornecem uma divisão de alto nível dos projetos; um usuário pode pertencer a vários consumidores.

**Nome de usuário.** Insira o nome de usuário com o qual efetuar logon no servidor.

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

**Salvar como Padrão.** Clique nesse botão para armazenar essas configurações como seu padrão para evitar reinseri-las toda vez que você abrir o nó.

### Selecionar Origem de Dados do Analytic Server

O diálogo Selecionar Origem de Dados permite selecionar uma origem de dados para análise.

**Origem de dados.** Uma lista das origens de dados disponíveis do servidor escolhido. Selecione a origem que deseja usar e clique em **OK**.

---

## Exportação do Analytic Server

A Exportação do Analytic Server permite gravar dados de sua análise em HDFS. Normalmente um fluxo com uma Exportação do Analytic Server é enviado para Analytic Server para ser executado em um ambiente de "dados grandes". Ele também pode ser executado localmente para fazer o upload de conjuntos de dados relativamente pequenos (com no máximo 100.000 registros) para HDFS.

Em uma conexão do Analytic Server, selecione uma origem de dados contendo os dados que deseja usar. Uma origem de dados contém os arquivos e os metadados associados a essa origem.

**Conexão.** Clique em **Editar** para definir ou editar os detalhes de uma conexão com o servidor onde os dados residem. Consulte "conexões do Analytic Server" na página 3 para obter mais informações.

**Nome da origem de dados.** Insira o nome da origem de dados que será associada aos dados exportados. A origem de dados recém-criada pode ser acessada por meio do uso de um "Origem do Analytic Server" na página 3.

**Caminho de destino.** Após ter estabelecido uma conexão, clique em **Navegar** para escolher o projeto e a pasta em que os dados serão gravados. Consulte o tópico "Caminho de Destino do Analytic Server" para obter mais informações.

**Descrição.** Insira uma descrição para a origem de dados. Essa descrição persiste como um atributo de origem de dados no servidor e pode ser visualizada no Console do Administrador. Consulte o tópico Capítulo 3, "Console do Administrador", na página 5 para obter mais informações.

**Gerar um nó de Importação para estes dados.** Selecione para gerar um nó de origem para os dados, conforme exportado para a origem de dados especificada. Quando você envia esse fluxo por meio do Console do Executor, ele fornece um link para um arquivo *.nod* que pode ser transferido por download mediante a conclusão bem-sucedida da tarefa. Quando você executa localmente, esse nó é incluído na tela de fluxo.

*Nota sobre modelos de dados:* lembre-se de que uma origem de dados é composta por um ou mais arquivos de dados que contêm os valores dos dados e um arquivo de modelo de dados que define os metadados para a origem de dados. Embora os arquivos de dados não sejam criados até o nó de Exportação do Analytic Server ser executado, no Modelador ou por meio do Console do Executor, o arquivo de modelo de dados é criado assim que você clica em **OK**. Além disso, o nó de Exportação do Analytic Server atualizará o modelo de dados cada vez que você clicar em **OK**, portanto, tome cuidado ao inspecionar um fluxo para não atualizar um arquivo de modelo de dados acidentalmente.

## Caminho de Destino do Analytic Server

O diálogo Selecionar Caminho de Destino permite selecionar um local para a saída.

**Caminho de destino.** A estrutura de projeto e pasta do servidor escolhido. Selecione o caminho que deseja usar e clique em **OK**.

---

## Capítulo 3. Console do Administrador

O Console do Administrador permite configurar o Analytic Server; principalmente para criar origens de dados para usuários, mas também para configurar consumidores e, dentro de um determinado consumidor:

- Designar usuários e grupos ao consumidor
- Criar funções e designá-las a usuários e grupos
- Designar permissões para usuários, grupos e funções
- Criar projetos para uso na análise

---

### Para Efetuar Logon no Console do Administrador

1. Insira a URL do Analytic Server na barra de endereço do seu navegador. Geralmente ela tem a forma *serverAddress:port*, com 8080 como a porta.
2. Insira o consumidor sob o qual você efetuará logon no servidor.
3. Insira o nome de usuário com o qual efetuar logon no servidor.
4. Insira a senha associada ao nome de usuário especificado.

Após o login, o menu sanfona padrão é o Origem de Dados.

---

### Navegando no Console do Administrador

O Console do Administrador tem quatro componentes:

- O cabeçalho exibe o nome do produto e o nome do consumidor / usuário do usuário atualmente conectado.
- A coluna esquerda exibe os menus sanfona disponíveis ou agrupamentos funcionais. O menu sanfona selecionado determina o que é mostrado na área de conteúdo.
- A área de conteúdo. Exibe o conteúdo associado ao menu sanfona atualmente selecionado. Os detalhes do conteúdo de cada menu sanfona estão nas seções abaixo.
- O rodapé contém o link de logout e exibe a versão instalada do Analytic Server.

---

### Consumidores

Consumidores fornecem uma divisão de alto nível de funções, permissões, projetos e origens de dados. Observe que esse menu sanfona só fica visível para os usuários com a permissão *Gerenciar Consumidores*. Consulte o tópico “Permissões” na página 8 para obter mais informações.

A coluna esquerda exibe os consumidores existentes sob o título do menu sanfona. Selecione um consumidor para ter seus detalhes exibidos na área de conteúdo. A digitação na área de procura mostra apenas os consumidores com a sequência de procura em seu nome.

A área de conteúdo é dividida em controles comuns, nas guias *Configurações*, *Configurações* e *Usuário*.

### Controles Comuns

A área de controle comum exibe o nome do consumidor selecionado no menu sanfona. Por padrão, o consumidor especificado no login é selecionado e mostrado.

- Um clique em *Configurar Consumidor Ativo* efetua logout do usuário atual e exibe o diálogo de login, com o consumidor atualmente selecionado pré-preenchido no campo do consumidor e o usuário atual

no campo de usuário. Isso é útil quando você acaba de criar um novo consumidor e deseja editar funções, permissões, projetos e origens de dados dentro dele.

- Um clique em *Novo...* cria um novo consumidor com o nome que você especifica no diálogo Incluir Novo Consumidor. Nomes fazem distinção entre maiúsculas e minúsculas, ignoram espaços em branco iniciais e finais e são protegidos contra injeção de SQL.
- Um clique em *Renomear...* permite renomear o consumidor atualmente selecionado com o nome especificado no diálogo Renomear Consumidor. *NOTA:* Essa funcionalidade está desativada atualmente.
- Um clique em *Excluir* remove o consumidor atualmente selecionado e todas as funções, origens de dados, subpastas e projetos associados.

## Guia Configurações

É possível fornecer uma descrição de texto do consumidor na guia Configurações. A descrição do consumidor é um campo de texto editável que permite fornecer um texto explicativo sobre o consumidor. Um clique em *Salvar* salva o atual estado das configurações.

## guia Segurança

O WebSphere Security Provider possibilita o uso da implementação de segurança do WebSphere para fornecer autenticação do usuário e procura por atributo de usuários e grupos.

Quando esse provedor de segurança for selecionado, os usuários conectados a esse consumidor serão autenticados com relação ao registro do usuário configurado no WebSphere para o aplicativo Analytic Server.

As configurações de Registro de Usuário LDAP são expostas no arquivo Liberty server.xml; por exemplo:

```
<ldapRegistry
  baseDN="CN=Users,DC=dg1-domain,DC=com"
  ldapType="Microsoft Active Directory"
  host="9.30.83.101"
  id="DG1-CONTROL-1"
  bindDN="CN=aeadmin,CN=Users,DC=dg1-domain,DC=com"
  bindPassword="{xor}KzosKw=="
  port="389">
  <activeFilters
    id="DG1-CONTROL-1 Filters"
    groupFilter="(&(objectClass=group)(member=CN=%v,CN=Users,DC=dg1-domain,DC=com))"
    groupMemberIdMap="group:member"
    groupIdMap="group:sAMAccountName" />
  </ldapRegistry>
```

Figura 2. Exemplo: Conteúdo do Arquivo Liberty Server.xml

É possível localizar detalhes sobre a segurança do WebSphere em: [ftp://ftp.software.ibm.com/software/webserver/appserv/library/v85/was85base\\_security.pdf](ftp://ftp.software.ibm.com/software/webserver/appserv/library/v85/was85base_security.pdf).

## Guia Usuários

Na guia Usuários, é possível selecionar quais usuários serão associados ao consumidor.

A lista de usuários Disponíveis exibe os usuários que não estão atualmente associados a esse consumidor. Isso inclui usuários que foram autenticados com relação ao sistema externo e usuários que foram incluídos via diálogo Registrar Novos Usuários.

- A digitação na área de janela de procura filtra os usuários com a sequência de procura em seu nome.
- A seleção do ícone usuários acima da lista mostra os usuários disponíveis. O cancelamento da seleção do ícone oculta os usuários. Esse ícone fica selecionado por padrão.

Os usuários podem ser movidos para os usuários e grupos do Consumidor usando o botão mover. Esses usuários poderão efetuar login com o consumidor ativo. Os usuários que ainda não tiverem sido autenticados com relação ao sistema externo poderão ser incluídos pelo administrador clicando no botão *Incluir Usuários Não Registrados*.

O propósito do diálogo Registrar Novos Usuários é que Administrador possa configurar seus usuários no SecurityProvider ativo e gerenciá-lo no Console de Administrador, sem requerer que cada usuário efetue login primeiro.

1. Digite no campo de texto de filtro e clique em **Procurar** para localizar usuários não registrados que têm a sequência de procura no nome de usuário. Se o campo de texto de filtro Nome de Usuário estiver vazio, todos os usuários não registrados serão exibidos.
2. Selecione um ou mais usuários e clique em **OK** para incluí-los na lista Usuários do Consumidor.

---

## Funções

Funções fornecem uma maneira de agrupar permissões e designar mais facilmente a um usuário um certo nível de acesso ao sistema de acordo com sua "função", independentemente de ser um administrador, usuário avançado ou usuário regular. As funções definidas para um consumidor são independentes das funções definidas para outro. Observe que esse menu sanfona só fica visível para os usuários com a permissão *Gerenciar Funções*. Consulte o tópico "Permissões" na página 8 para obter mais informações.

A coluna esquerda exibe as funções existentes sob o título do menu sanfona. Selecione uma função para que seus detalhes sejam exibidos na área de conteúdo. A digitação na área de procura mostra apenas as funções com a sequência de procura em seu nome.

A área de conteúdo é dividida em controles comuns, nas guias *Configurações* e *Usuário e Grupos*.

## Controles Comuns

Exibe o nome da função selecionada no menu sanfona.

- Um clique em **Novo...** cria uma nova função com o nome que você especifica no diálogo Incluir Nova Função. Nomes fazem distinção entre maiúsculas e minúsculas, ignoram espaços em branco iniciais e finais e são protegidos contra injeção de SQL.
- Um clique em **Renomear...** renomeia a atual função com o nome especificado no diálogo Renomear Função. *Nota:* Essa funcionalidade está atualmente desativada.
- Um clique em **Excluir** remove a função.

## Guia Configurações

É possível fornecer uma descrição de texto da função na guia Configurações. A descrição da função é um campo de texto editável que permite fornecer um texto explicativo sobre a função. Um clique em **Salvar** salva o atual estado das configurações.

## Guia Usuários e Grupos

Na guia Usuários e Grupos, é possível selecionar quais usuários e grupos serão associados à função.

A lista de usuários e grupos disponíveis exibe usuários e grupos no consumidor ativo que não estão associados atualmente a essa função.

- A digitação na área de janela de procura filtra os usuários e os grupos com a sequência de procura em seu nome.
- A seleção do ícone usuários acima da lista mostra os usuários disponíveis. O cancelamento da seleção do ícone oculta os usuários. Esse ícone fica selecionado por padrão.

- A seleção do ícone grupos acima da lista mostra os grupos disponíveis. O cancelamento da seleção do ícone oculta os grupos. Esse ícone fica selecionado por padrão.

Os usuários e os grupos podem ser movidos para usuários e grupos da Função usando o botão mover. Todos os usuários e membros desses grupos têm permissões designadas a essa função no menu sanfona Permissões.

## Permissões

Permissões permitem conceder ou restringir o acesso ao Analytic Server para um usuário, grupo ou função. As permissões designadas a um consumidor são independentes das permissões designadas a outro. Observe que esse menu sanfona só fica visível para usuários com permissão *Gerenciar Permissões*.

Permissão	Descrição
Enviar tarefas	Usuário pode enviar tarefas por meio do Console do Executor.
Gerenciar projetos próprios	Usuário pode criar novos projetos, além de modificar e excluir projetos dos quais é participante.
Gerenciar qualquer projeto	Usuário pode modificar e excluir qualquer projeto, mesmo aqueles nos quais não está especificamente listado como participante ou visualizador.
Gerenciar origens de dados próprias	Usuário pode criar novas origens de dados, além de modificar e excluir origens de dados das quais é proprietário.
Gerenciar qualquer origem de dados	Usuário pode modificar e excluir qualquer origem de dados, mesmo aquelas nas quais não está especificamente listado como proprietário.
Gerenciar funções	Usuário pode criar, modificar, excluir e designar funções.
Gerenciar Permissões	Usuário pode designar permissões.
Gerenciar consumidores	Usuário pode criar, modificar e excluir consumidores.

O conjunto de permissões é fixo e não pode ser criado, modificado ou excluído, apenas designado a usuários, grupos e funções; dessa forma, não há nada na coluna esquerda.

A lista Usuários, Grupos e Funções exibe usuários, grupos e funções no consumidor ativo.

- A digitação na área de janela de procura filtra os usuários, grupos e funções com a sequência de procura em seu nome.
- A seleção do ícone usuários acima da lista mostra os usuários disponíveis. O cancelamento da seleção do ícone oculta os usuários. Esse ícone fica selecionado por padrão.
- A seleção do ícone grupos acima da lista mostra os grupos disponíveis. O cancelamento da seleção do ícone oculta os grupos. Esse ícone fica selecionado por padrão.
- A seleção do ícone funções acima da lista mostra as funções disponíveis. O cancelamento da seleção do ícone oculta as funções. Esse ícone fica selecionado por padrão.

A lista Permissões exibe as permissões disponíveis.

- A seleção de um ou mais itens na lista mostra as permissões atualmente escolhidas para os usuários, grupos e funções selecionados. Quando houver permissões existentes conflitantes entre os usuários, grupos e funções selecionados, as permissões com conflitos aparecerão em itálico.
- A marcação ou desmarcação das caixas seleciona e cancela a seleção de permissões para os usuários, grupos e funções selecionados.
- Um clique em **Salvar** salva o atual estado das permissões.

---

## Projetos

Projetos são a estrutura organizacional de nível superior para conter arquivos e pastas dentro de um consumidor. Projetos podem ser compartilhados com usuários individuais e grupos.

A coluna esquerda exibe os projetos existentes sob o menu sanfona. Selecione um projeto para que seus detalhes sejam exibidos na área de conteúdo. A digitação na área de procura mostra apenas os projetos com a sequência de procura em seu nome.

A área de conteúdo é dividida em controles comuns, nas guias *Configurações*, *Participantes* e *Visualizadores*.

### Controles Comuns

Exibe o nome do projeto selecionado no menu sanfona.

- Um clique em **Novo...** cria um novo projeto com o nome especificado no diálogo Incluir Novo Projeto. Nomes fazem distinção entre maiúsculas e minúsculas, ignoram espaços em branco iniciais e finais e são protegidos contra injeção de SQL.
- Um clique em **Renomear...** renomeia o projeto atual com o nome especificado no diálogo Renomear Projeto (não mostrado). *Nota:* Essa funcionalidade está atualmente desativada.
- Um clique em **Excluir** remove o projeto e todas as subpastas e arquivos.

### Guia Configurações

**Descrição de projeto** é um campo de texto editável que permite fornecer um texto explicativo sobre o projeto.

É **público** é uma caixa de opção que indica se alguém pode ver o projeto (marcado) ou se os usuários e grupos têm que ser incluídos explicitamente na lista de participantes (desmarcado).

Um clique em **Salvar** salva o estado atual das configurações. *Nota:* Quando um projeto é declarado público, as listas de participantes e visualizadores são limpas.

**Origens de dados do projeto** é uma área não editável que lista todas as origens de dados associadas ao projeto.

A área de janela direita mostra a estrutura de projeto/pasta para o projeto atualmente selecionado. É possível navegar na estrutura de pasta, mas ela não é editável, exceto por meio de botões. *Nota:* o conteúdo de um projeto só pode ser modificado por meio do Console do Administrador quando o projeto tiver sido bloqueado na guia *Versões*, e essas mudanças só são finalizadas após o botão **Confirmar** ser clicado na guia *Versões*.

- Um clique em **Fazer Upload de Arquivo** faz o upload de um arquivo para o projeto/pasta atualmente selecionado.
- Um clique em **Fazer Download de Arquivo** faz o download de um arquivo selecionado em um sistema de arquivos local.
- Um clique em **Incluir Pasta** cria uma nova pasta sob a pasta selecionada com o nome que você especifica no diálogo Nome da Nova Pasta.
- Um clique em **Excluir Seleção** remove o arquivo/pasta selecionado.

### Guia Participantes

Participantes são membros integrais de um projeto e podem modificar o projeto, bem como as pastas e os arquivos dentro dele.

A lista de usuários e grupos disponíveis exibe usuários e grupos no consumidor ativo que não estão associados atualmente a esse projeto.

- A digitação na área de janela de procura filtra os usuários e os grupos com a sequência de procura em seu nome.
- A seleção do ícone usuários acima da lista mostra os usuários disponíveis. O cancelamento da seleção do ícone oculta os usuários. Esse ícone fica selecionado por padrão.
- A seleção do ícone grupos acima da lista mostra os grupos disponíveis. O cancelamento da seleção do ícone oculta os grupos. Esse ícone fica selecionado por padrão.

Os usuários e os grupos podem ser movidos para Usuários e Grupos do Projeto usando o botão mover. Esses usuários e membros desses grupos terão acesso de leitura (nó de Origem do Analytic Server) e de gravação (nó de Exportação do Analytic Server) a esse projeto ao se conectarem ao Analytic Server por meio de IBM SPSS Modeler.

*Nota:* Os usuários com a permissão "Gerenciar qualquer Projeto" têm acesso de leitura e gravação a cada projeto, independentemente de estarem ou não especificamente listados como participante na guia Participantes.

## Guia Visualizadores

Visualizadores não podem modificar o projeto, mas têm acesso de leitura para visualizar seu conteúdo.

A lista de usuários e grupos disponíveis exibe usuários e grupos no consumidor ativo que não estão associados atualmente a esse projeto.

- A digitação na área de janela de procura filtra os usuários e os grupos com a sequência de procura em seu nome.
- A seleção do ícone usuários acima da lista mostra os usuários disponíveis. O cancelamento da seleção do ícone oculta os usuários. Esse ícone fica selecionado por padrão.
- A seleção do ícone grupos acima da lista mostra os grupos disponíveis. O cancelamento da seleção do ícone oculta os grupos. Esse ícone fica selecionado por padrão.

Os usuários e os grupos podem ser movidos para Usuários e Grupos do Projeto usando o botão mover. Esses usuários e membros desses grupos terão acesso de leitura (nó de Origem do Analytic Server) mas NÃO de gravação (nó de Exportação do Analytic Server) a esse projeto ao se conectarem por meio do SPSS Modeler.

## Guia Versões

Projetos têm sua versão baseada nas mudanças no conteúdo do arquivo e pasta. As mudanças nos atributos de um projeto, como descrição, se é público e com quem é compartilhado, não requerem uma nova versão. A inclusão, modificação ou exclusão de arquivos ou pastas não requer uma nova versão.

A tabela exibe as versões de projeto existentes, suas datas de criação e confirmação, o(s) usuário(s) responsável(is) e a versão pai. A versão pai é aquela na qual a versão selecionada é baseada.

- Um clique em **Bloquear** permite fazer mudanças no conteúdo da versão do projeto selecionado.
- Um clique em **Confirmar** salva todas as mudanças feitas em um projeto e torna essa versão o estado atual visível do projeto.
- Um clique em **Descartar** descarta todas as mudanças feitas em um projeto bloqueado e retorna o estado visível do projeto para a versão confirmada mais recentemente.
- Um clique em **Excluir** remove a versão selecionada.

**Limpar automaticamente quando o número de versões for excedido** excluirá automaticamente a versão do projeto mais antiga confirmada quando o número de versões exceder o número especificado.

---

## Origens de Dados

Uma origem de dados é uma coleção de registros, mais um modelo de dados, que definem um conjunto de dados para análise. A origem dos registros pode ser um arquivo (texto delimitado, texto de largura fixa, Excel) em HDFS ou banco de dados. O modelo de dados define todos os metadados (nomes de campo, armazenamento, nível de medição, entre outros) necessários para a análise dos dados. Os proprietários da origem de dados podem conceder ou restringir o acesso às origens de dados.

A coluna esquerda exibe as origens de dados existentes sob o título do menu sanfona. Selecione uma origem de dados para ter seus detalhes exibidos na área de conteúdo. A digitação na área de procura mostra apenas as origens de dados com a sequência de procura em seu nome.

A área de conteúdo é dividida em controles comuns, nas guias *Arquivos*, *Modelo de Dados*, *Configurações*, *Projetos* e *Proprietários*.

### Controles Comuns

Exibe o nome da origem de dados selecionada no menu sanfona.

- Um clique em **Novo...** cria uma nova origem de dados com o nome e o tipo de conteúdo especificados no diálogo Incluir Nova Origem de Dados.
  - Os nomes de origem de dados devem fazer distinção entre maiúsculas e minúsculas, ignorar espaços em branco iniciais e finais e serem protegidos contra injeção de SQL.
  - Os tipos de conteúdo disponíveis são *Arquivo* e *Banco de Dados*.
- Um clique em **Renomear...** renomeia a atual origem de dados com o nome especificado no diálogo Renomear Origem de Dados.
- Um clique em **Excluir** remove a origem de dados. Isso deixa intactos todos os arquivos associados à origem de dados.

### Guia Arquivos

A guia de arquivos permite associar arquivos a uma origem de dados que possa conter registros ou um modelo de dados. *Nota:* normalmente é preciso estabelecer as regras para a leitura de uma origem de dados de arquivo na guia Especificações de Entrada.

**Arquivo de modelo de dados** é um texto não editável que lista o arquivo de modelo de dados atualmente selecionado.

**Arquivos associados atuais** é uma lista de todos os arquivos associados à origem de dados. Todos os arquivos contendo registros dentro de uma única origem de dados devem ter um modelo de dados comum.

- Você associa arquivos à origem de dados dando um clique duplo neles na lista de projeto/pasta. Se um arquivo estiver fora da estrutura do projeto de arquivo do Analytic Server, é possível incluí-lo na lista de arquivos digitando seu caminho absoluto na caixa de texto abaixo da lista e clicando em **Incluir Caminho**.
- Um clique duplo em um arquivo na lista o remove dessa lista.
- Um clique em **Configurar como Modelo** identifica o arquivo selecionado como contendo um modelo de dados para a origem de dados, o remove da lista e o inclui em **Arquivo de Modelo de Dados**.

A área de janela Projetos mostra a estrutura do projeto/pasta. Não é diretamente editável.

- Um clique em **Fazer Upload de Arquivo** faz o upload de um arquivo para o projeto/pasta atualmente selecionado.
- Um clique em **Fazer Download de Arquivo** faz o download de um arquivo selecionado em um sistema de arquivos local.

- Um clique em **Incluir Pasta** cria uma nova pasta sob a pasta selecionada com o nome que você especifica no diálogo Nome da Nova Pasta.
- Um clique em **Excluir Seleção** remove o arquivo/pasta selecionado.

## Guia Banco de Dados

Especifique os parâmetros de conexão para o banco de dados com o conteúdo do registro.

*Nota:* Uma origem de dados não pode misturar arquivos e um banco de dados como origens de registros. Durante a criação de uma origem de dados com um banco de dados fornecendo conteúdo de registro, a guia Arquivos só pode ser usada para selecionar um modelo de dados existente.

- **Banco de dados.** Selecione o tipo de banco de dados ao qual deseja se conectar. Escolha: DB2, Oracle, SQL Server, TeraData ou Netezza.
- **Endereço do servidor.** Insira a URL do servidor que hospeda o banco de dados.
- **Porta do servidor.** O número da porta na qual o banco de dados recebe.
- **Nome do banco de dados.** O nome do banco de dados ao qual você deseja se conectar.
- **Nome de usuário.** Se o banco de dados for protegido por senha, insira seu nome de usuário.
- **Senha.** Se o banco de dados for protegido por senha, insira sua senha.
- **Nome da tabela.** Insira o nome de uma tabela do banco de dados que deseja usar.

## Guia Modelo de Dados

Se você tiver especificado um arquivo de modelo na guia Arquivos, a grade será preenchida aqui.

- Um clique em **Gerar Novo Modelo de Dados** varre os primeiros N registros da origem de dados e retorna um modelo de dados “mais provável”, junto com uma visualização dos primeiros 10 registros para que seja possível verificar os resultados. É possível modificar os nomes de campo clicando em uma célula e digitando nele. É possível modificar tipos de armazenamento, níveis de medição e funções de campo selecionando a partir de listas suspensas. Essas células não são editáveis.
- Um clique em **Salvar** salva o modelo de dados e “congela” as células.

## Guia Especificações de Entrada

A guia Especificações de Entrada permite estabelecer regras para a leitura de uma origem de dados do arquivo.

Selecione um arquivo na lista Arquivos e insira a especificação de entrada em formato JSON. Para arquivos separados por vírgula sem cabeçalho, ela pode ser simplesmente:

```
{
  "parser" : {
    "name": "csv",
    "config" : {
      "hasHeader": false
    }
  }
}
```

A configuração do objeto JSON depende do tipo de arquivo. Para arquivos delimitados:

**Nome** Nome do analisador. Sempre "delimited\_parser" para arquivos delimitados. Requerido.

**config** Configurações do analisador.

### conjunto de caracteres

A codificação de caracteres do arquivo. O nome de um conjunto de caracteres Java, como "UTF-8", "ISO-8859-2", "GB18030", etc. Opcional. O padrão é "UTF-8".

**Codificação**

Um sinônimo descontinuado para conjunto de caracteres.

**hasHeader**

Se configurado como verdadeiro, a primeira linha de dados conterá nomes de campo. Opcional. O padrão é falso.

**invalidChars**

Determina como caracteres inválidos (sequências de bytes que não correspondem aos caracteres na codificação) devem ser manipulados. Um valor nulo indica que eles devem ser tratados como erros; uma sequência vazia indica que eles devem ser descartados silenciosamente; e uma sequência não vazia (geralmente um único caractere) indica que eles serão substituídos pelo conteúdo da sequência. Opcional. O padrão é nulo.

**initialCharsToSkip**

O número de caracteres no início do arquivo a ser ignorado. Um número inteiro não negativo. Opcional. O padrão é zero.

**skipFirstRecord**

Se o primeiro registro no arquivo deve ou não ser ignorado. Normalmente isso é usado para ignorar uma linha de cabeçalho contendo nomes de campo. Verdadeiro ou falso (valores JSON verdadeiros ou falsos sem aspas). Opcional. O padrão é falso.

**commentChars**

Uma sequência de caracteres. Caractere(s) marcando comentários de fim de linha, de modo que o caractere e tudo que vier após ele no registro sejam ignorados. Cada caractere é obtido como um marcador de comentário independente. Por exemplo, "/\*" significa que uma barra ou um asterisco inicia um comentário. Não é possível definir marcadores de comentários multicalcaracteres, como "//". A sequência vazia sinaliza que nenhum caractere de comentário está definido. Se definidos, os caracteres de comentário são verificados antes de as aspas serem processadas ou os caracteres no início do arquivo serem ignorados (`initialCharsToSkip>0`). Opcional. O padrão é a sequência vazia.

**delimitByControlChars**

JSON verdadeiro ou falso. Se verdadeiro, todos os caracteres de controle ASCII, exceto LF e CR, serão tratados como delimitadores de campo. Opcional. O padrão é falso.

**fieldDelimiters**

Uma sequência de um ou mais caracteres marcando limites de campo. Cada caractere é obtido como um delimitador independente. Por exemplo, ",\t" significa que uma vírgula ou uma tabulação marca os limites do campo. Se `delimitByControlChars` for verdadeiro, os caracteres na sequência `fieldDelimiters` serão tratados como delimitadores, além de caracteres de controle. Opcional. O padrão é "," se `delimitByControlChars` for falso e a sequência estiver vazia.

**mergeWhitespace**

JSON verdadeiro ou falso. Se for verdadeiro, ocorrências adjacentes múltiplas de espaço e/ou tabulação serão consideradas um único delimitador de campo. Não efeito será surtido se nem um espaço e nem uma tabulação for um delimitador de campo. Opcional. O padrão é verdadeiro.

**singleQuote**

"keep", "drop" ou "pair". Especifica a manipulação de aspas simples (apóstrofes). Se for "keep", as aspas simples não terão um significado especial e serão tratadas como qualquer outro caractere. Se for "drop", elas serão simplesmente excluídas, a menos que estejam entre aspas. Se for "pair", elas serão tratadas como caracteres aspas simples, e os caracteres entre pares de aspas simples perderão qualquer significado especial (serão considerados entre aspas). Se as aspas simples podem ou não ocorrer dentro de sequências com aspas simples é determinado pela configuração `quoteByDoubling`. Opcional. O padrão é "keep".

**doubleQuote**

"keep", "drop" ou "pair". Especifica a manipulação de aspas duplas. Se for "keep", as aspas duplas não terão um significado especial e serão tratadas como qualquer outro caractere. Se for "drop", elas serão simplesmente excluídas, a menos que estejam entre aspas. Se for "pair", elas serão tratadas como caracteres aspas duplas, e os caracteres entre pares de aspas simples perderão qualquer significado especial (serão considerados entre aspas). Se as aspas duplas podem ou não ocorrer dentro de sequências com aspas duplas é determinado pela configuração quoteByDoubling. Opcional. O padrão é "pair".

**quoteByDoubling**

JSON verdadeiro ou falso. Indica se aspas duplas podem ser representadas em sequências de aspas duplas quando doubleQuote está configurado como "pair" e se aspas simples podem ser representadas em sequências de aspas simples quando singleQuote está configurado como "pair". Se for verdadeiro, as aspas duplas serão escapadas dentro das sequências de aspas duplas por duplicação e as aspas simples serão escapadas dentro das sequências de aspas simples por duplicação. Se for falso, não haverá como usar aspas duplas dentro de uma sequência de aspas duplas ou aspas simples dentro de uma sequência de aspas simples. Opcional. O padrão é verdadeiro.

**fieldOrder**

Uma matriz de sequências. Os elementos da matriz são obtidos para serem os nomes dos campos na ordem em que aparecem no registro. Se vazia, a ordem dos campos no registro será assumida como sendo a ordem dos campos no modelo de dados. Opcional. O padrão é a matriz vazia.

```
{
  "parser" : {
    "name": "delimited_parser",
    "config" : {
      "charset" : "UTF-8",
      "invalidChars" : null,
      "initialCharsToSkip" : 0,
      "skipFirstRecord" : false,
      "commentChars" : "",
      "fieldDelimiters" : ",\t",
      "mergeWhitespace" : true,
      "singleQuote" : "keep",
      "doubleQuote" : "pair",
      "quoteByDoubling" : true,
      "fieldOrder" : [ "age", "education", "height" ],
    }
  }
}
```

*Nota:* As opções para manipulação de aspas duplas e simples e para escape de aspas se aplicam independentemente do armazenamento de campo. Por exemplo, o registro

"1234", "56789"

conterá dois valores de número inteiro válido, contanto que a opção doubleQuote seja "drop" ou "pair". Porém, se a opção for "keep", os valores não poderão ser decodificados como um número inteiro válido, já que contêm aspas. Também não há como ler uma sequência com uma nova linha integrada.

Para arquivos de largura fixa:

**Nome** Nome do analisador. Sempre "fixed\_parser" para arquivos fixos. Requerido.

**config** Configurações do analisador.

**conjunto de caracteres**

A codificação de caracteres do arquivo. O nome de um conjunto de caracteres Java, como "UTF-8", "ISO-8859-2", "GB18030", etc. Opcional. O padrão é "UTF-8".

**invalidChars**

Determina como caracteres inválidos (sequências de bytes que não correspondem aos caracteres na codificação) devem ser manipulados. Um valor nulo indica que eles devem ser tratados como erros; uma sequência vazia indica que eles devem ser descartados silenciosamente; e uma sequência não vazia (geralmente um único caractere) indica que eles serão substituídos pelo conteúdo da sequência. Opcional. O padrão é nulo.

**lineOriented**

JSON verdadeiro ou falso. Indica quantos registros estão definidos. Se for verdadeiro, os registros serão definidos (delimitados) por novas linhas iniciando o arquivo ou terminando o arquivo. Se for falso, os registros serão definidos pela duração do registro em bytes. Opcional. O padrão é verdadeiro.

**recordLength**

Um valor positivo que é a duração de um registro em bytes. Esse valor será ignorado se lineOriented for verdadeiro. Ele será obrigatório se lineOriented for falso.

**initialRecordsToSkip**

O número de registros no início do arquivo a ser ignorado. Um número inteiro não negativo. Opcional. O padrão é zero.

**fieldLocations**

Uma matriz de dois elementos de números inteiros. Cada matriz de dois elementos especifica a coluna de início e o comprimento de um campo. As colunas são numeradas começando com zero. Comprimentos devem ser positivos. Requerido.

**fieldOrder**

Uma matriz de sequências. Os elementos da matriz são obtidos para serem os nomes dos campos na ordem em que aparecem no registro. Se vazia, a ordem dos campos no registro será assumida como sendo a ordem dos campos no modelo de dados. Opcional. O padrão é a matriz vazia.

```
{
  "parser" : {
    "name": "fixed_parser",
    "config" : {
      "charset" : "UTF-8",
      "invalidChars" : null,
      "lineOriented" : true,
      "initialRecordsToSkip" : 0,
      "fieldLocations" : [[0,2], [2,1], [3,3]],
      "fieldOrder" : [ "age", "education", "height" ],
    }
  }
}
```

Para arquivos Excel:

**Nome** Nome do analisador. Sempre "excel" para arquivos Microsoft Excel. Requerido.

**config** Configurações do analisador.

Seleciona uma planilha do Excel. Indexe ou nomeie a planilha. Opcional. O padrão é a primeira planilha.

**rangeFirstNonBlank**

Se for configurado como verdadeiro, o intervalo de células iniciará na primeira linha que não está em branco. Opcional. O padrão é verdadeiro. Se for configurado para falso, o parâmetro &quot;explicitRange&quot; deverá ser especificado.

**explicitRange**

Define um intervalo explícito para as células do Excel, por exemplo, &quot;A1:B14&quot;. O intervalo também pode ser definido usando notação do Excel &quot;RxCx&quot;, por exemplo, &quot;R1C2:R4C5&quot;.

### hasHeader

Se for configurado como verdadeiro, a primeira linha do intervalo de células selecionado conterá os nomes de campo. Opcional. O padrão é falso.

### fieldOrder

Uma matriz de sequências. Os elementos da matriz são obtidos para serem os nomes dos campos na ordem em que aparecem no registro. Se vazia, a ordem dos campos no registro será assumida como sendo a ordem dos campos no modelo de dados. Opcional. O padrão é a matriz vazia.

### stopOnBlankRow

Se for configurado como verdadeiro e mais de uma linha em branco for encontrada, parará a leitura. Caso contrário, continuará lendo todos os dados até o final da planilha, incluindo linhas em branco. Opcional. O padrão é falso.

```
{
  "parser" : {
    "name": "excel",
    "config" : {
      "worksheet" : 0,
      "hasHeader" : false,
      "rangeFirstNonBlank" : true,
      "stopOnBlankRow" : true,
      "fieldOrder" : [ "age", "education", "height" ]
    }
  }
}
```

## Guia Configurações

**Descrição de origem de dados** é um campo de texto editável que permite fornecer um texto explicativo sobre a origem de dados.

**Atributos customizados** permite definir atributos customizados para a origem de dados. Clique em qualquer célula na coluna Valor para editá-la.

- Um clique em **Incluir** incluir uma nova linha (atributo) na tabela com o nome que você fornece no diálogo Definir uma Chave Exclusiva. Nomes fazem distinção entre maiúsculas e minúsculas, ignoram espaços em branco iniciais e finais e são protegidos contra injeção de SQL.
- Um clique em **Excluir** remove o atributo selecionado.

## Guia Projetos

A lista Projetos Disponíveis exibe projetos no consumidor ativo que não estão associados atualmente a essa origem de dados.

- A digitação na área de janela de procura filtra os projetos com a sequência de procura em seu nome.
- Os projetos podem ser movidos para Projetos da Origem de Dados usando o botão mover.

Os usuários que são participantes ou visualizadores dos projetos selecionados (supondo que eles tenham a permissão "Gerenciar Projetos Próprios") poderão ler essa origem de dados a partir do Modelador.

## Guia Proprietários

Os proprietários definidos nesta guia, supondo que eles tenham a permissão "Gerenciar Origens de Dados Próprias", poderão modificar ou excluir essa origem de dados.

A lista Usuários e Grupos Disponíveis exibe usuários e grupos no consumidor ativo que não estão atualmente associados a essa origem de dados.

- A digitação na área de janela de procura filtra os usuários e os grupos com a sequência de procura em seu nome.
- A seleção do ícone usuários acima da lista mostra os usuários disponíveis. O cancelamento da seleção do ícone oculta os usuários. Esse ícone fica selecionado por padrão.
- A seleção do ícone grupos acima da lista mostra os grupos disponíveis. O cancelamento da seleção do ícone oculta os grupos. Esse ícone fica selecionado por padrão.
- Os usuários e os grupos podem ser movidos para Usuários e Grupos da Origem de Dados usando o botão mover.

*Nota:* Os usuários com a permissão "Gerenciar qualquer Origem de Dados" têm acesso de leitura e gravação a cada origem de dados, independentemente de estarem ou não especificamente listados como participante na guia Participantes.



---

# Índice Remissivo

## C

- Console Administrativo 5
  - consumidores 5
  - origens de dados 11
  - permissões 8
  - projetos 9
  - roles 7
- consumidores
  - no Console Admin 5

## D

- dados de exportação
  - exportação do Analytic Server 3

## E

- exportação do Analytic Server 3, 4

## N

- nós de exportação
  - exportação do Analytic Server 4
- nós de origem
  - origem do Analytic Server 3

## O

- origem do Analytic Server 3
- origens de dados
  - no Console Admin 11

## P

- permissões
  - no Console Admin 8
- projetos
  - no Console Admin 9

## R

- roles
  - no Console Admin 7







Impresso no Brasil