IBM SPSS Deployment Manager 8.5

User's Guide

Note

Before using this information and the product it supports, read the information in "Notices" on page 211.

Product Information

This edition applies to version 8, release 4, modification 0 of IBM® SPSS® Collaboration and Deployment Services and to all subsequent releases and modifications until otherwise indicated in new editions.

Contents

Chapter 1. Overview 1
IBM SPSS Collaboration and Deployment Services 1
Collaboration 1
Deployment 2
System architecture 2
IBM SPSS Collaboration and Deployment Services Repository 3
IBM SPSS Deployment Manager 3
IBM SPSS Collaboration and Deployment Services Deployment Portal 4
Execution servers 4
Scoring server 5

Chapter 2. What's new in this release 7
What's new for IBM SPSS Deployment Manager users 7

Chapter 3. Getting started 9
Launching IBM SPSS Deployment Manager 9
Navigating through the system 9
Using the mouse versus pressing Enter 9
Dragging and dropping items in the user interface 9
Accessing help 9
Entry field content assist 10
Field conventions 10
Naming conventions within the system 10
Exiting the system 11

Chapter 4. The Content Explorer 13
Content Explorer overview 13
What Is a content object? 13
Organization of the Content Explorer 13
Working with servers 13
Creating new Content Server connections 14
Logging in to a server 15
Logging off from a server 15
Changing server passwords 15
Working with files 16
Opening external files 16
Permissions inheritance 16
Working with the repository 16
Adding files to the repository 16
Downloading files from the repository 17
Deleting files from the repository 17
Searching 17
Accessing the search dialogs 17
Simple Search 18
Advanced Search 18
Viewing Search Results 20
Terms Not Included in a Search 21
Object Locking 21
Locking Objects 21
Viewing the Locked Objects Table 22

Unlocking Objects ... 22
Obtaining updates ... 22

Chapter 5. Properties ... 23
Working with object properties ... 23
Viewing object properties ... 23
Editing object properties ... 24
Working With Expiration Dates and Expired Files ... 29
Working with server properties and user preferences ... 31
Server properties ... 31
User preferences ... 33
Version properties ... 35
Working with custom properties ... 36
Verifying access privileges to create custom properties ... 36
Accessing the custom properties dialog box ... 36
Creating custom properties ... 36
Editing custom properties ... 38
Searching for custom properties ... 38
Deleting custom properties ... 39
Working with topics ... 39
Process overview ... 40
Verifying access privileges to create topic definitions ... 40
Working with topic definitions ... 40
Searching for topics ... 42
Bulk property updates ... 42
Updating general properties in bulk ... 43
Updating version properties in bulk ... 43
Updating permissions in bulk ... 44

Chapter 6. Resource definitions ... 45
Credentials ... 45
Adding new credentials ... 45
Server Process Credential ... 46
Data sources ... 46
Selecting a data source definition type ... 46
Specifying a DSN for an ODBC data source ... 47
Specifying a JDBC name and URL ... 47
Specifying a JNDI name for an application server data source ... 50
Specifying properties for data service data sources ... 50
Modifying data source definitions ... 51
Message domains ... 52
Creating a new message domain ... 52
Modifying message domain definitions ... 53
Promotion policies ... 53
Adding promotion policies ... 53
Modifying promotion policies ... 55
Deleting promotion policies ... 55
Server definitions ... 55
Adding new server definitions ... 55
Modifying server definitions ... 57
Server clusters ... 58
Creating a new server cluster ... 58
Modifying a server cluster ... 59
Importing resource definitions ... 59

Chapter 7. Export, import, and promotion ... 61
Overview ... 61

Job components that are migrated during export and import....61
Exporting external references....61
Export and import restrictions....62
Recommended import order....62
Security permissions when importing folders....62
Exporting folders....62
Importing folders....63
Resolving import conflicts....64
Promotion....66
Promoting objects....67
Promotion considerations....67

Chapter 8. Analytic data views....69
Creating analytic data views....70
Data access plans....71
Creating data access plans....72
Adding mapped tables to analytic data views....73
Modifying the data mapping for analytic data view tables....74
Mapping table attributes to stream fields....75
Overriding batch data access plan data sources....76
Overriding real-time data access plan data sources....77
Previewing data access plan data....79
Deleting data access plans....79
Data models....79
Specifying the analytic data view editor advanced mode....81
Adding unmapped tables to analytic data views....81
Adding unmapped attributes to data model tables....82
Defining relationships between analytic data view tables....83
Adding derived attributes to analytic data view tables....84
Modifying data model properties....86
Deleting data model components....87
Business object models....88
Exporting data models as business object model archives....89
Importing business object models as data models....89
Execution object models....90
Exporting data models as execution object model archives....91

Chapter 9. Scoring....93
Supported scoring functions and models....94
Scoring configurations....96
New scoring model configuration....97
Model specific settings....97
Data provider settings....97
Input data order....98
Input data returned settings....98
Output data returned settings....98
Logging settings....99
Advanced settings....101
Scoring configuration aliases....101
Creating scoring configuration aliases....102
Editing scoring configuration aliases....103
Deleting scoring configuration aliases....103
Scoring configuration association sets....104
Creating scoring configuration association sets....107
Editing scoring configuration association sets....107
Deleting scoring configuration association sets....108
A/B split testing for scoring configurations....108

Evaluating scoring models by using A/B testing ... 109
Scoring view ... 110
Filtering the Scoring view ... 110
Editing a scoring configuration ... 111
Suspending and resuming scoring configurations ... 111
Deleting a scoring configuration ... 111
Scoring Graph view ... 111

Chapter 10. Jobs ... 113
What is a job? ... 113
Versioning and labeling jobs ... 113
Components of a job ... 113
Prerequisite to running jobs ... 114
External file dependencies ... 114
Job process overview ... 114
Working with jobs in the Content Explorer ... 115
Creating new jobs ... 115
Opening an existing job ... 116
Viewing job properties ... 116
Working with the job editor ... 116
General information tab ... 117
Job variables ... 117
Adding steps to a job ... 118
Executing job steps concurrently ... 119
Specifying relationships in a job ... 119
Sequential connector ... 120
Pass connector ... 120
Fail connector ... 120
Conditional connector ... 120
Viewing general relationship properties and modifying relationships ... 121
Deleting relationships in a job ... 121
Saving jobs ... 122
Job step results ... 122
Output file location ... 123
Output file permissions ... 123
Output file metadata ... 124

Chapter 11. Running jobs ... 125
On-demand job execution ... 125
Specifying on-demand execution options ... 125
Scheduled job execution ... 125
Creating schedules ... 126
Editing schedules ... 128
Deleting schedules ... 129
Message-based processing example ... 129

Chapter 12. Monitoring status ... 131
Accessing status views ... 131
Selecting a server in status views ... 131
Opening a job in the Job Editor ... 131
Refreshing status views ... 132
Reordering items in status views ... 132
Deleting a job from status views ... 132
Job schedule view ... 132
Job history view ... 133
Working with the job history table ... 133
Model management views ... 134

Model evaluation view....134
Champion challenger view....135
Predictors view....136
Filters....137
Filters common to all status views....137
Job schedule filters....137
Job History Filters....138
Model management filters....139
Server status view....140
Setting the refresh rate....140

Chapter 13. Notifications and subscriptions....141
Notifications....141
Job success and failure notifications....141
Job step notifications....142
Content notifications....142
Notifications based on model evaluation return codes....143
Label event notifications....143
Notification settings....143
Subscriptions....146
Subscribing to files....147
Modifying file subscriptions and unsubscribing from files....147
Subscription email address....147
Managing subscriptions....147
Delivery failures....148

Chapter 14. Visualization report job steps....149
General properties for visualization report steps....149
Data source for visualization report steps....149
Type for visualization report steps....150
Parameters for visualization report steps....150
Results for visualization report steps....150
Clean-up for visualization report steps....151
Notifications for visualization report steps....151

Chapter 15. SAS® job steps....153
General properties for SAS steps....153
Additional arguments for SAS steps....154
Results for SAS steps....154
SAS result formats....155
Example SAS step....155

Chapter 16. General job steps....157
General properties for general job steps....157
Input files for general job steps....158
Input file source location....158
Input file target location....158
Output files for general job steps....158
Output file intermediate location....159
General job step examples....159
Storing IBM Corp. PMML files....159
Batch scoring using IBM SPSS Statistics PMML files....160
Event-based scheduling....162

Chapter 17. Message-based job steps....167
General properties of message-based job steps....167

Chapter 18. Notification job steps..... 169
Adding a notification job step to a job....169
General information....170
Notification....170
Updating notifications.... 170
Selecting a new template.... 170

Chapter 19. Champion challenger job steps..... 173
Champion challenger overview.... 173
Model evaluation metrics....173
Order dependency....173
General information....174
Challengers.... 174
Selecting challengers....175
Invalid challengers....175
Selecting challenger data sources....176
Champion....176
Testing the champion 176
Data files.... 178
Data view....178
ODBC data sources....178
Cognos import....179

Chapter 20. Submitted jobs....181
Restrictions within the Submitted Jobs folder.... 181
Submitted jobs and expiration dates.... 181

Chapter 21. Administration Consoles..... 183
Getting started....183
Administered servers....183
Adding new administered servers.... 183
Viewing administered server properties.... 185
Connecting to administered servers....185
Disconnecting administered servers.... 185
Deleting administered servers....185
IBM SPSS Statistics Server Administration.... 185
Introduction to IBM SPSS Statistics Server Administration.... 185
IBM SPSS Statistics Server Connections....186
Controlling the IBM SPSS Statistics Server....186
IBM SPSS Statistics Server Configuration....186
IBM SPSS Statistics Server User Profiles and Groups.... 191
Monitoring IBM SPSS Statistics Server Users.... 193
Accessibility with the Keyboard....194
IBM SPSS Modeler Server Administration....195
Starting Modeler Administration Console 195
Restarting the web service.... 195
Configuring Access with Modeler Administration Console 195
SPSS Modeler Server connections....196
SPSS Modeler Server Configuration.... 197
SPSS Modeler Server Monitoring.... 203
Using the options.cfg file....204
Closing unused database connections.... 204

Chapter 22. Accessibility features....207
Keyboard navigation.... 207
Navigating the content explorer.... 207

Navigating tables....208
Navigating the job history and job schedule views....208
Navigating the job editor....208
Navigating the help system....208
Accessibility for visually impaired users....209
Accessibility issues for blind users....209

Notices....211
Privacy policy considerations....212
Trademarks....212
Glossary....215
A....215
B....215
C....216
D....216
E....217
F....217
G....217
I....217
J....218
K....218
L....219
M....219
N....219
O....219
P....220
R....220
S....220
T....221
U....221
V....221
W....222
X....222

Index....223

Chapter 1. Overview

IBM SPSS Collaboration and Deployment Services

IBM SPSS Collaboration and Deployment Services is an enterprise-level application that enables widespread use and deployment of predictive analytics.

IBM SPSS Collaboration and Deployment Services provides centralized, secure, and auditable storage of analytical assets and advanced capabilities for management and control of predictive analytic processes, as well as sophisticated mechanisms for delivering the results of analytical processing to users. The benefits of IBM SPSS Collaboration and Deployment Services include:

- Safeguarding the value of analytical assets
- Ensuring compliance with regulatory requirements
- Improving the productivity of analysts
- Minimizing the IT costs of managing analytics

IBM SPSS Collaboration and Deployment Services allows you to securely manage diverse analytical assets and fosters greater collaboration among those developing and using them. Furthermore, the deployment facilities ensure that people get the information they need to take timely, appropriate action.

Collaboration

Collaboration refers to the ability to share and reuse analytic assets efficiently, and is the key to developing and implementing analytics across an enterprise.

Analysts need a location in which to place files that should be made available to other analysts or business users. That location needs a version control implementation for the files to manage the evolution of the analysis. Security is required to control access to and modification of the files. Finally, a backup and restore mechanism is needed to protect the business from losing these crucial assets.

To address these needs, IBM SPSS Collaboration and Deployment Services provides a repository for storing assets using a folder hierarchy similar to most file systems. Files stored in the IBM SPSS Collaboration and Deployment Services Repository are available to users throughout the enterprise, provided those users have the appropriate permissions for access. To assist users in finding assets, the repository offers a search facility.

Analysts can work with files in the repository from client applications that leverage the service interface of IBM SPSS Collaboration and Deployment Services. Products such as IBM SPSS Statistics and IBM SPSS Modeler allow direct interaction with files in the repository. An analyst can store a version of a file in development, retrieve that version at a later time, and continue to modify it until it is finalized and ready to be moved into a production process. These files can include custom interfaces that run analytical processes allowing business users to take advantage of an analyst's work.

The use of the repository protects the business by providing a central location for analytical assets that can be easily backed-up and restored. In addition, permissions at the user, file, and version label levels control access to individual assets. Version control and object version labels ensure the correct versions of assets are being used in production processes. Finally, logging features provide the ability to track file and system modifications.

Deployment

To realize the full benefit of predictive analytics, the analytic assets need to provide input for business decisions. Deployment bridges the gap between analytics and action by delivering results to people and processes on a schedule or in real time.

In IBM SPSS Collaboration and Deployment Services, individual files stored in the repository can be included in processing **jobs**. Jobs define an execution sequence for analytical artifacts and can be created with IBM SPSS Deployment Manager. The execution results can be stored in the repository, on a file system, or delivered to specified recipients. Results stored in the repository can be accessed by any user with sufficient permissions using the IBM SPSS Collaboration and Deployment Services Deployment Portal interface. The jobs themselves can be triggered according to a defined schedule or in response to system events.

In addition, the scoring service of IBM SPSS Collaboration and Deployment Services allows analytical results from deployed models to be delivered in real time when interacting with a customer. An analytical model configured for scoring can combine data collected from a current customer interaction with historical data to produce a score that determines the course of the interaction. The service itself can be leveraged by any client application, allowing the creation of custom interfaces for defining the process.

The deployment facilities of IBM SPSS Collaboration and Deployment Services are designed to easily integrate with your enterprise infrastructure. Single sign-on reduces the need to manually provide credentials at various stages of the process. Moreover, the system can be configured to be compliant with Federal Information Processing Standard Publication 140-2.

System architecture

In general, IBM SPSS Collaboration and Deployment Services consists of a single, centralized IBM SPSS Collaboration and Deployment Services Repository that serves a variety of clients, using execution servers to process analytical assets.

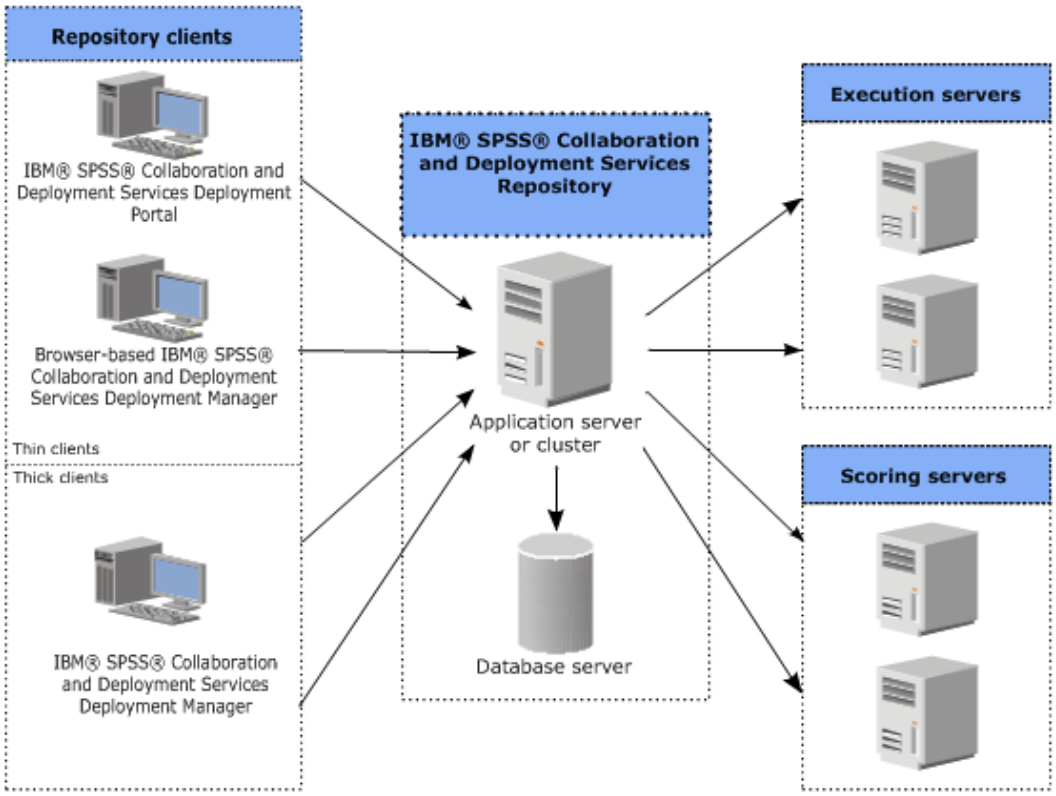

Figure 1. IBM SPSS Collaboration and Deployment Services Architecture

IBM SPSS Collaboration and Deployment Services consists of the following components:

- IBM SPSS Collaboration and Deployment Services Repository for analytical artifacts
- IBM SPSS Deployment Manager
- IBM SPSS Collaboration and Deployment Services Deployment Portal
- Browser-based IBM SPSS Deployment Manager

IBM SPSS Collaboration and Deployment Services Repository

The repository provides a centralized location for storing analytical assets, such as models and data. The repository requires an installation of a relational database, such as IBM Db2, Microsoft SQL Server, or Oracle.

The repository includes facilities for:

- Security
- Version control
- Searching
- Auditing

Configuration options for the repository are defined using the IBM SPSS Deployment Manager or the browser-based IBM SPSS Deployment Manager. The contents of the repository are managed with the Deployment Manager and accessed with the IBM SPSS Collaboration and Deployment Services Deployment Portal.

IBM SPSS Deployment Manager

IBM SPSS Deployment Manager is a client application for IBM SPSS Collaboration and Deployment Services Repository that enables users to schedule, automate, and execute analytical tasks, such as updating models or generating scores.

The client application allows a user to perform the following tasks:

- View any existing files within the system, including reports, SAS syntax files, and data files
- Import files into the repository
- Schedule jobs to be executed repeatedly using a specified recurrence pattern, such as quarterly or hourly
- Modify existing job properties
- Determine the status of a job
- Specify email notification of job status

In addition, the client application allows users to perform administrative tasks for IBM SPSS Collaboration and Deployment Services, including:

- Manage users
- Configure security providers
- Assign roles and actions

Browser-based IBM SPSS Deployment Manager

The browser-based IBM SPSS Deployment Manager is a thin-client interface for performing setup and system management tasks, including:

- Setting system configuration options
- Configuring security providers
- Managing MIME types

Non-administrative users can perform any of these tasks provided they have the appropriate actions associated with their login credentials. The actions are assigned by an administrator.

You typically access the browser-based IBM SPSS Deployment Manager at the following URL:

```
http://<host IP address>:<port>/security/login
```

Note: An IPv6 address must be enclosed in square brackets, such as `[3ffe:2a00:100:7031::1]`.

If your environment is configured to use a custom context path for server connections, include that path in the URL.

```
http://<host IP address>:<port>/<context path>/security/login
```

IBM SPSS Collaboration and Deployment Services Deployment Portal

IBM SPSS Collaboration and Deployment Services Deployment Portal is a thin-client interface for accessing the repository. Unlike the browser-based IBM SPSS Deployment Manager, which is intended for administrators, IBM SPSS Collaboration and Deployment Services Deployment Portal is a web portal serving a variety of users.

The web portal includes the following functionality:

- Browsing the repository content by folder
- Opening published content
- Running jobs and reports
- Generating scores using models stored in the repository
- Searching repository content
- Viewing content properties
- Accessing individual user preferences, such as email address and password, general options, subscriptions, and options for output file formats

You typically access the home page at the following URL:

```
http://<host IP address>:<port>/peb
```

Note: An IPv6 address must be enclosed in square brackets, such as `[3ffe:2a00:100:7031::1]`.

If your environment is configured to use a custom context path for server connections, include that path in the URL.

```
http://<host IP address>:<port>/<context path>/peb
```

Execution servers

Execution servers provide the ability to execute resources stored within the repository. When a resource is included in a job for execution, the job step definition includes the specification of the execution server used for processing the step. The execution server type depends on the resource.

Execution servers currently supported by IBM SPSS Collaboration and Deployment Services include:

- **Remote Process**. A remote process execution server allows processes to be initiated and monitored on remote servers. When the process completes, it returns a success or failure message. Any machine acting as a remote process server must have the necessary infrastructure installed for communicating with the repository.

  Note: The IBM SPSS Collaboration and Deployment Services Remote Process Server has a default thread pool core size of 16, which allows a maximum of 16 concurrent jobs to be executed on a single remote process server. Any concurrent jobs in excess of 16 must wait in the queue until the available thread pool has free resources. To manually configure the IBM SPSS Collaboration and Deployment Services Remote Process Server thread pool core size, add the following JVM option (with a user defined value) to the remote process server's startup script: `prms.thread.pool.coresize=<user defined value>`

For more information regarding the start-up script, see the "Starting and stopping the remote process server" section in the IBM SPSS Collaboration and Deployment Services Remote Process Server guide.

Execution servers that process other specific types of resources can be added to the system by installing the appropriate adapters. For information, consult the documentation for those resource types.

During job creation, assign an execution server to each step included in the job. When the job executes, the repository uses the specified execution servers to perform the corresponding analyses.

Scoring server

IBM SPSS Collaboration and Deployment Services Scoring Service is also available as a separately deployable application, the Scoring Server.

The Scoring Server improves deployment flexibility in several key areas:

- Scoring performance can be scaled independently from other services
- Scoring Server(s) can be independently configured to dedicate computing resources to one or any number IBM SPSS Collaboration and Deployment Services scoring configurations
- Scoring Server operating system and processor architecture does not need match the IBM SPSS Collaboration and Deployment Services Repository or other Scoring Servers
- Scoring Server application server does not need match the application server used for IBM SPSS Collaboration and Deployment Services Repository or other Scoring Servers

Chapter 2. What's new in this release

What's new for IBM SPSS Deployment Manager users

There are no new features for this version.

Chapter 3. Getting started

Launching IBM SPSS Deployment Manager

To launch the client:

1. From the Start menu, choose:

All Programs > **SPSS Inc.** > **IBM SPSS Collaboration and Deployment Services Deployment Manager**

The IBM SPSS Deployment Manager interface appears.

Navigating through the system

IBM SPSS Deployment Manager relies primarily on tab-based navigation.

The interface is divided into the following major sections:

Table 1. Sections in the IBM SPSS Deployment Manager

Section	Description	Location
Content Explorer	The Content Explorer is a tree that displays the contents of your repository. See the topic "Content Explorer overview" on page 13 for more information.	Left-hand-pane
Properties window	The Properties window displays the properties of the file that you select in the Content Explorer. Note that this Properties window is different from the individual properties that may appear as part of the Status tabs.	Lower-left-hand pane
Job Editor	The Job Editor allows you to drag-and-drop elements to create jobs.	Upper-right-hand pane

Using the mouse versus pressing Enter

The system is mouse-driven. It is not recommended that you use the **Enter** key to complete actions. Typically, pressing **Enter** will not submit your request.

Dragging and dropping items in the user interface

You can drag items in the user interface. For example, you can reorganize items within the Content Explorer, or you can drag files from the Content Explorer into the Job Editor.

The system adheres to the following guidelines for drag-and-drop behavior.

- The Content Repository root cannot be moved.
- You can move items from the Content Explorer into the Job Editor. However, you cannot drag items from the Job Editor to the Content Repository. You must work with items within the Job Editor. See the topic "What is a job?" on page 113 for more information.

Accessing help

Help is available via an online help system.

You can access help using any of the following methods:

Help menu. From the Help menu, choose **IBM SPSS Collaboration and Deployment Services Deployment Manager Help**.

Dialog-level help. To view online help in dialog boxes, click the **Help** button.

F1 help. In certain sections of the system context-sensitive help is available. To access context-sensitive help, press **F1**.

Entry field content assist

The content assist feature provides job variables and predefined system property variables that can be used to insert values into entry fields. Variables available through content assist vary from field to field and may include timestamps, object paths, object URLs, job/step start and end times, execution identifiers for jobs and job steps, completion codes, variables passed to a job step by another job step, and variables defined at the job level. Entry fields that allow content assist are marked with a light bulb icon.

To insert variable values into a field, type $. The list of available variables appears in a drop-down list. Click the variable name to display its description. Double-click the variable name to select it.

Date and timestamp variables allow the user to select the display format. To select a format after you have specified a variable, type . (period) after the variable name. The list of available formats appears in a drop-down list. Click the format name to display its description. Double-click the format name to select it.

Variables can be used to define file paths. However, a single backslash is an escape character for entry fields. Consequently, use a double backslash or a forward slash when specifying paths. For example, a path of

```
${JobVariable.Var1}\${JobVariable.Var2}
```

should be specified as

```
${JobVariable.Var1}\\${JobVariable.Var2}
```

or

```
${JobVariable.Var1}/${JobVariable.Var2}
```

Field conventions

In IBM SPSS Deployment Manager, when property values of multiple objects or object versions are compared, property fields that have varying values are empty and marked by alert sign icons.

Naming conventions within the system

At various points in the system, you are prompted to name items. For example, you will be asked to specify names for folders and jobs. All names within the system must be unique.

Note: Alphanumeric characters are recommended. The following symbols are prohibited:

- Quotation marks (single and double)
- Ampersands (&)
- Less-than (<) and greater-than (>) symbols
- Forward slash (/)
- Periods
- Commas
- Semicolons

Exiting the system

You can exit the IBM SPSS Deployment Manager using either of the following methods:

- From the File menu, choose **Exit**.
- Click the close button (**X**) in the title bar of the user interface.

Chapter 4. The Content Explorer

Content Explorer overview

The Content Explorer is your entry point into the IBM SPSS Collaboration and Deployment Services Repository.

The Content Explorer contains a tree hierarchy of objects in the repository. The objects that appear in the Content Explorer depend upon your permissions. For example, you can view only folders for which you have permissions. In the Content Explorer, you can perform the following tasks:

- Log in to and log off from a server
- Establish server and credential definitions
- View object properties
- Access and work with files in the Content Repository

What Is a content object?

A content object is any item that resides in the IBM SPSS Collaboration and Deployment Services Repository. Objects in the repository are stored in a relational database, such as Microsoft SQL Server or Oracle, in binary form.

You can add almost any type of file to the Content Repository. Examples of content objects include:

- IBM SPSS Modeler streams
- IBM SPSS Statistics syntax files
- SAS syntax files
- Jobs

Typically, you perform an action on a content object. For example, you may add a IBM SPSS Modeler stream to a job, or you may run a job. At a minimum, you can move, copy, and paste content objects within the Content Repository.

Organization of the Content Explorer

The server that hosts your IBM SPSS Collaboration and Deployment Services Repository is represented by a folder in the Content Explorer.

Within every server folder, the following items appear:

Content Repository. Contains all of your content objects. You can create subfolders within this folder.

Submitted Jobs. Displays the results of running reports using IBM SPSS Collaboration and Deployment Services Deployment Portal.

Resource Definitions. This folder contains server, credential, and data source definitions.

These items persist permanently at the root. You cannot move, copy, or delete these folders.

Working with servers

When working with servers, you can perform the following tasks in the Content Explorer:

- Create new server connections
- Log in to a server
- Log off from a server
- Delete a server

Creating new Content Server connections

Before you begin working, you must establish a connection to the server that houses your repository.

You need to create the server connection only once. After you have created the connection, the server folder appears in the Content Explorer, and you can simply log in to the server. See the topic "Logging in to a server" on page 15 for more information.

To create a server connection:

1. Launch IBM SPSS Deployment Manager .
2. From the File menu, choose:

   New > **Content Server Connection**

   The Create New Content Server Connection dialog box opens.
3. In the **Connection name** field, enter the name of your IBM SPSS Collaboration and Deployment Services Repository. This name is used in the Content Explorer at the root level.

   Note: Alphanumeric characters are recommended. The following symbols are prohibited:

   - Quotation marks (single and double)
   - Ampersands (&)
   - Less-than (<) and greater-than (>) symbols
   - Forward slash (/)
   - Periods
   - Commas
   - Semicolons
4. In the **Server URL** field, enter the complete connection URL for the server.

   The URL includes the following elements:

   - The connection scheme, or protocol, as either *http* for hypertext transfer protocol or *https* for hypertext transfer protocol with the secure socket layer (SSL)
   - The host server name or IP address

     Note: An IPv6 address must be enclosed in square brackets, such as `[3ffe:2a00:100:7031::1]`.
   - The port number. If the repository server is using the default port (port 80 for http or port 443 for https), the port number is optional.
   - An optional custom context path for the repository server

   Table 2. Example URL specifications . This table lists some example URL specifications for server connections.

URL	Scheme	Host	Port	Custom path
`http://myserver`	HTTP	*myserver*	default (80)	(none)
`https://9.30.86.11:443/spss`	HTTPS	9.30.86.11	443	*spss*
`http://[3ffe:2a00:100:7031::1]:9080/ibm/cds`	HTTP	`3ffe:2a00:100:7031::1`	9080	*ibm/cds*

   Contact your system administrator if you are unsure of the URL to use for your server.
5. Click **Finish**.

Logging in to a server

Before you begin working with the Content Explorer, you must log in to a server. The version of the content repository server must be equivalent to or greater than the IBM SPSS Deployment Manager version.

At a minimum, you must have one server connection defined in the Content Explorer. To log in to a server:

1. Double-click the server name. (Alternatively, you can expand the server folder by clicking on the **+** sign.) The Login to IBM SPSS Collaboration and Deployment Services Repository dialog box opens.
2. In the **User ID** field, enter a valid username for the server. Note that the username that you use determines what you see in the Content Explorer because different users may have different permission levels. See the topic "Modifying permissions" on page 26 for more information.
3. In the Password field, enter the password that corresponds to the username.
4. If multiple security providers have been configured for the server, use the Provider field to select the security provider against which to validate the user/password combination.
5. Click **OK**. The system opens the server folder and displays the expanded contents in the server directory.

Note: If single sign-on has been configured by the IBM SPSS Collaboration and Deployment Services administrator, the login page is bypassed and the repository is accessed without the user providing credentials. In that case, the user's Windows credentials are authenticated against an external directory service, for example, Windows Active Directory, which acts as a security provider for the system.

Logging off from a server

You can access multiple servers simultaneously; however, before logging off from any of them, be sure to save changes to files or jobs on each server because the system does not prompt you to do so.

To log off from a server:

1. In the Content Explorer, right-click the server from which you want to log off.
2. Select **Logoff**. You are then logged off from the server, and the contents of the server folder are collapsed.
3. To access stored information, you must log back in to the server. See the topic "Logging in to a server" on page 15 for more information.

Changing server passwords

Username and password information is required to log in to a server in the Content Explorer. The password can be changed at any time. However, availability of the change password option depends on the security provider associated with the credentials.

For example, passwords can be changed when using IBM SPSS Collaboration and Deployment Services native security or IBM i providers, but not when using Active Directory.

The password modification takes effect immediately. Logging off and logging back in to the system is not required.

To change a server password:

1. Right-click the server name and select **Change Password**. The Change Password dialog box opens. This dialog box displays the name of the server to which you are currently logged in and your username. Neither the server name nor the username can be modified in this dialog box.
2. In the Current Password field, enter your current password.
3. In the New Password field, enter your new password.
4. In the Confirm New Password field, enter your new password again. The **OK** button is disabled until all of the fields in the dialog box contain data. In addition, the button is disabled until the information in the New Password and Confirm New Password fields is identical.

5. Click **OK**. The Change Password dialog box opens showing the message **Password changed successfully.**
6. Click **OK**.

Working with files

In the Content Explorer, you can perform the following tasks:

- Open external files
- Add files to the repository
- Download files from the repository

Opening external files

From the Content Explorer, you can open and view files.

The method used to open and display the file depends on the file type. For example, if you open a text file, the text is displayed in the Job Editor. However, if you double-click a IBM SPSS Modeler stream, the system launches the IBM SPSS Modeler application.

Permissions inheritance

It is important to understand the relationship between IBM SPSS Collaboration and Deployment Services Repository resources, particularly when you copy or move resources.

The following guidelines apply:

- **Creating resources.** When a resource is added to the repository (for example, when a new job is created), it inherits the permissions of its parent folder; a user with *Write* permissions to a folder he will by default have the same permissions to any resources created in it.
- **Copying resources.** When you copy a resource to a new folder, the resource retains the permissions of the original folder. However, if the user copying the resource is not the owner of the original resource, then the system changes the ownership of the resource to the new user.
- **Moving resources.** When you move a resource from one folder to another, the resource retains the permissions of the original folder. When you cut and paste a resource, this is considered a move. Thus, the resource retains the permissions of the original folder.

Working with the repository

Files can be added to the repository or downloaded from the repository. Typically, individual files are added or downloaded.

It is important to note that adding or downloading files is different from importing or exporting files. See the topic "Overview" on page 61 for more information.

Adding files to the repository

You can add almost any type of file to the repository.

Using the File menu. To add a file to the repository:

1. In the Content Explorer, select the folder to which you want to add the file.
2. From the File menu, select **Add File to Repository**. The Add Files dialog box opens. If the **Add File to Repository** option is disabled, then you have clicked on an object in the Content Explorer and not on a folder.
3. Navigate to the file that you want to add to the repository.
4. Click **Open**.

Dragging and dropping files. Alternatively, you can drag files into the repository.

Downloading files from the repository

Using the Content Explorer, you can download files from the repository to another computer.

To download files:

1. In the Content Explorer, select the file that you want to download.
2. From the File menu, choose **Download File**. If your file has multiple versions, the Select File Version dialog box opens. See the topic "Selecting a version" on page 29 for more information. Otherwise, the Download Files dialog box opens.
3. Navigate to the folder in which you want to place the file.
4. Click **OK**. A copy of the file is saved in the folder that you specified.

Deleting files from the repository

Objects can be deleted from the repository individually or in bulk, provided the appropriate permissions are specified.

Searching

Contents of the IBM SPSS Collaboration and Deployment Services Repository are searchable. The results appear on the Search Results tab.

The following types of searches are available:

- Simple search.
- Advanced search.

It is important to note that the search function will search for full text matches only. Currently, partial text and wildcard searches are not supported. Finally, searches cannot be saved and do not persist.

Guidelines for searching

When searching for objects in the system, the following guidelines apply:

- In order to find an object, the name provided must match the repository object's name exactly.
- The search feature looks inside of IBM SPSS Modeler streams to find nodes that contain the search string.
- Although the following objects are returned if they match a search string exactly, the search feature does not look *within* the following types of objects: submitted jobs and resource definitions.
- The following characters are not allowed in search strings:
  - Single quotation marks (')
  - Double quotation marks (")
  - Parentheses (or)
- Certain terms are excluded from the search. See the topic "Terms Not Included in a Search" on page 21 for more information.

Accessing the search dialogs

To access the search dialogs:

1. Ensure that you are logged in to the server that contains the repository that you want to search.
2. In the Content Explorer, select any server folder. The search must be conducted at the server instance level. The search option is disabled at the subfolder level.
3. From the Edit menu, choose **Search**. By default, the simple search dialog appears. If the *Search* option is not enabled, then the appropriate folder level was not selected. The search option is enabled only at the server level. All subfolders are automatically searched.

4. To perform an advanced search, click **Advanced**. See the topic "Advanced Search" on page 18 for more information.

Simple Search

A simple search looks through the repository for objects that match the specified string and returns all matching results. The simple search does not distinguish among property types (e.g., author or title) or narrow results by time frame.

To search for a string:

1. In the **Search** field, enter a text string. Quotation marks are not necessary.
2. Click **Search**. The Search Results tab is populated with all of the objects that meet the search criteria. See the topic "Viewing Search Results" on page 20 for more information.

Advanced Search

The advanced search allows greater refinement of the search. For example, in addition to text strings, the advanced search option provides the ability to search by date range.

To perform an advanced search, the following options are available:

Property

The list of searchable properties depends on the types of files contained within the repository and may include:

- Author
- Description
- Keyword
- Label
- MIME type
- Object last modified by
- Parent URI
- Title
- URI
- Version created by
- Version URI

For any property selected, an exact value needs to be specified. A list of options will not appear for any of these property types. For URI properties, the URI must use the id and version marker specification. Paths and version labels cannot be used when searching by URI.

To search by property:

1. From the **Property** drop-down list, select a property type.
2. In the **Value** field, type a value that corresponds to the property type selected. Quotation marks are not necessary.
3. Click **Add**. The property type and the value appear in the Search Terms box. For the initial entry in the Search Terms box, no AND or OR grouping is specified. By default, any subsequent search terms are combined with the AND operator. The AND operator can be changed to OR at any time. See the topic "Switching Between AND and OR" on page 19 for more information.

Date Search

A date range can also be specified in an advanced search. Valid parameters for a search by date range include:

Date search. Valid values include: *Expiration Date*, *Last Modified*, and *Version Created On*.

Date range. A date range must be provided. The date in the **To** field must be equal to or greater than the date in the **From** field.

Time range. The time range is optional. However, to search by a time range a date range must be specified first.

Refining Search Terms

After parameters are specified in the property or date fields and added to the Search Terms list, the selections can be further refined.

Specifically, the following enhancements can be made:

- Switching between *AND* or *OR*.
- Grouping search terms
- Editing search term values.
- Reordering search terms.
- Deleting search terms.

Switching Between AND and OR

By default, search terms are joined by the Boolean operation *AND*. The Boolean operator can be changed from *AND* to *OR* and vice versa.

The button that is available depends upon the search term selected. If the search term is connected by *AND*, then the **OR** button appears. If the search term is connected by *OR*, then the **AND** button appears.

To change the connector:

1. Select the search term whose connector you want to change.
2. Click **AND** or **OR**.

Grouping and Ungrouping Search Terms

Search terms can be grouped to refine a search further. The ability to perform AND and OR searches allows searches to be organized across attributes. Grouping allows AND and OR searching to occur within the same attribute. Nested groups are supported.

For example, suppose that you want to search for an object modified in December 2007 that could have been created by one of two authors (Joe Author or Jane Author) and carries one of two labels (Test 1 or Test 2). By default, all search terms are joined by *AND*. A search for objects authored by Joe Author and Jane Author AND labeled Test 1 and Test 2 might return limited results. By changing the *AND* between authors and labels to *OR* and using grouping, the search could be refined so that the search results contain a list of objects authored by either Joe Author or Jane Author and labeled either Test 1 or Test 2 and last modified in December 2007.

Suppose the following property values were set and appeared in the Search Terms list:

```
'Author' = 'Joe Author'
OR 'Author' = 'Jane Author'
AND 'Label' = 'Test 1'
OR 'Label' = 'Test 2'
AND LastModified BETWEEN '12/1/07' AND '12/31/07'
```

When grouped the properties would be organized in the Search Terms list as follows:

```
('Author' = 'Joe Author'
OR 'Author' = 'Jane Author')
AND ('Label' = 'Test 1'
OR 'Label' = 'Test 2')
AND LastModified BETWEEN '12/1/07' AND '12/31/07'
```

To group search terms:

1. Select the search terms that should be evaluated as a group. To group selected items, the items must be adjacent to each other on the list of search terms. To select multiple terms, press **Ctrl** while selecting rows from the list.
2. Click **Group** or **Ungroup**. The *Ungroup* option only appears if the search terms selected were previously grouped together.

Editing Search Terms

The values for search terms can be edited. However, the property type cannot be edited. To add a new property type to the Search Terms list, a new property must be selected from the **Property** drop-down list and added to the Search Terms list.

To edit a previously specified property value:

1. Select the search term to edit.
2. Click **Edit**. The Edit Search Value dialog appears. The contents of the dialog may vary depending upon the search term selected. For example, the dialog for the *Author* property type contains a text field for the author's name; whereas, the dialog for the *Last Modified* property type contains a date range.
3. Specify a new value for the property type.
4. Click **OK**.

Reordering Search Terms

The order in which search terms appear on the list can be reordered. The order of search terms is only important when grouping terms together. Search terms must be adjacent to each other on the search terms list in order for grouping to be possible.

However, the order in which search terms (grouped or ungrouped) appear on the search terms list does not matter. The search feature will return results in order of relevancy. Thus, the search term that matches the largest number of repository objects will be listed first in the search results. The search term with the second largest number of hits will be listed second, and so forth.

To reorder items in the search terms list:

1. From the Search terms list, select the search term to move.
2. Click **Up** or **Down** until the search term appears in the target location. It is important to note that after rows are grouped, the movement of rows is limited. Movement that would potentially affect a grouping is not allowed.

Deleting Search Terms

Search terms can be deleted from the search terms list. After a search term is deleted from the list, it cannot be recovered. The search term would have to be re-created.

To delete a search term:

1. Select the search term to delete.
2. Click **Delete**. The search term is removed from the list.

Viewing Search Results

By default, the Search Results table is hidden. When a search is executed, the Search Results table appears.

To open the Search Results table independently, choose **Search Results** from the View menu.

The search results table contains the following information.

Table 3. Search results

Column	Description
Name	The name of the file that contains the search string.

Table 3. Search results (continued)	
Column	**Description**
Folder	The folder that contains the file.
Author	The user who created the file.
Type	The file type. External applications are prefaced with `application/x-vnd`. For example, the type for a IBM SPSS Modeler stream is `application/x-vnd.spss-clementine-stream`.
Last Modified	Date and time of the most recent modification to the returned item.
Version Count	The number of versions of the search object.

Terms Not Included in a Search

Stopwords are excluded from searches. The words in the following alphabetical list are not indexed:

```
a, all, am, an, and, any, are, as, at, be, but, by, can, could, did, do,
does, etc, for, from, goes, got, had, has, have, he, her, him, his, how,
if, in, is, it, let, me, more, much, must, my, nor, not, now, of, off,
on, or, our, own, see, set, shall, she, should, so, some, than, that,
the, them, then, there, these, this, those, though, to, too, us, was,
way, we, what, when, where, which, who, why, will, would, yes, yet, you
```

As a result, the system ignores them when searching for items. Currently, you cannot modify the terms in this list.

Object Locking

In the Content Explorer, objects can be locked to prevent other users from making changes.

For example, suppose that a file associated with a step within a job needs to be modified. Locking the file would prevent others users from making changes that could potentially conflict with the current user's changes. If another user opens a locked object, the new user receives a message indicating that the file is locked and can only be opened in read-only mode.

The following guidelines apply to locked objects:

- Most objects in the repository--including resource definitions--can be locked. However, folders cannot be locked. In addition, objects within the *Submitted Jobs* folder cannot be locked.
- When an object is locked, all versions of that object are locked. Thus, properties of that object cannot be edited. However, properties will be visible in read-only mode.
- Locks, which are denoted by the lock icon, persist across sessions.
- Objects remain locked until the lock is explicitly removed. Expiration dates cannot be associated with a lock.
- A locked object cannot be renamed, moved, or deleted. However, alocked object can be copied. The lock is not copied along with the object.
- Locks do not persist from one repository instance to the next. If a locked object is exported, the lock is not exported along with the object.
- Any user can lock an object. Users can also unlock their own objects. However, to unlock additional objects, users must have the action, *Manage Locks*.

Locking Objects

To lock an object:

1. In the Content Explorer, right-click the object and select **Lock**. The lock icon appears for the object.

Viewing the Locked Objects Table

The Locked Objects table contains a list of locked objects and corresponding details about those objects.

Specifically, the Locked Objects table contains the following information:

Name. The name and path of the object that is locked.

Locked By. The user who locked the object.

Locked Since. The date and time at which the object was initially locked.

To access the Locked Objects table:

1. From the View menu, select:

Show View > **Locked Objects**

Unlocking Objects

Objects can be unlocked using either of the following methods:

Content Explorer. To unlock an object, right-click the object and select **Unlock**. The object is unlocked, and the lock icon is removed.

Locked objects table. To unlock an object:

1. Navigate to the Locked Objects table.
2. Select the object(s) to unlock. To select multiple rows, press the **Ctrl** key and click simultaneously.

   Note: Only objects for which the user has appropriate permissions can be unlocked.
3. Click the **Unlock** icon. The selected objects are unlocked and removed from the Locked Objects table.

Obtaining updates

Periodically, updates may be available on the server. Each time IBM SPSS Deployment Manager is launched, the system checks the server for updates. Note that the update check is only performed once per client session.

If updates are available, the system automatically applies the updates and sends a message indicating that the updates have been applied. After updates have been applied, the IBM SPSS Deployment Manager needs to be restarted for the changes to take effect.

Chapter 5. Properties

Working with object properties

Properties are metadata associated with a content object, which describe the overall content object. Author and modification date are characteristics associated with a content object. Typically, more properties are available as you delve deeper into the content tree.

For example, for a folder at the root of the content tree, *title* is the only property available. However, when you view properties for a IBM SPSS Modeler stream, you see additional properties, such as author, description, version, and so on.

When working with properties, you can:

- View properties
- Edit properties

Viewing object properties

When you select an object in the Content Explorer, properties appear in the Properties pane.

The type and number of properties that are displayed depend upon the object and your location within the content tree. For example, if you click a server folder at the root level, you will see a title in the Properties pane.

Although you view object properties in the Properties pane, you edit them in the Properties dialog box.

Table 4. Object properties

Property	Description	Can you view this in the Properties pane?	Can you modify this in the Properties dialog box?
Author	Author refers to the creator of the job. Typically, the author corresponds to a user ID in the system and not an individual person. For example, the author of the job may be the administrator.	Yes	No
Description	The description displayed here corresponds to the description that you specified when you created the object.	Yes	Yes
Modification date	The modification date and time of the most recent modification to the object is displayed here.	Yes	No
Title	The title is the name of the object.	Yes	Yes. Title corresponds to name in the Properties dialog box.
Version	Identifier that distinguishes one version of an object from another. The identifier consists of the version number, a colon, the version modification date, and the version modification time.	Yes	Yes

Table 4. Object properties (continued)			
Property	**Description**	**Can you view this in the Properties pane?**	**Can you modify this in the Properties dialog box?**
Name	The name of your object.	Yes. Name corresponds to title in the Properties pane.	Yes
File type	The MIME type associated with the job.	No	Yes
Last modified by	The last user who made changes to the job.	No	No
Custom properties	Properties defined in the Custom Properties dialog box. These properties will vary as they are specific to each server instance.	No	Yes

Editing object properties

You may want to modify properties. For example, you may want to add keywords, which are used by the system to search for objects.

Properties are grouped into the following categories:

- General
- Labels
- Permissions
- Versions

To access the Properties dialog box:

1. In the Content Explorer, right-click the object whose properties you want to view.
2. Choose **Properties**. The Properties dialog box opens.

Editing general properties

To edit general object properties:

1. In the Content Explorer, right-click the object whose properties you want to view.
2. Choose **Properties**. The Properties dialog box opens. By default, the General properties dialog box opens.

The General properties dialog contains information about the object that you selected. Specifically, this dialog describes the following information, which is modifiable, unless otherwise noted.

Title. The name of the object.

Object URI. The uniform resource identifier (URI) for the object. This value is generated by the system and cannot be modified.

Author. The user ID under which the object was created.

Content type. The object type. A new object type can be specified for most objects in the repository. However, a new content type cannot be applied to the following objects:

- A job
- Any object in the *Resource Definitions* folder
- Any object in the *Submitted Jobs* folder

Note: Properties can be modified for an object in the *Submitted Jobs* folder once the object has been moved out of that folder into the Content Repository. See the topic Chapter 20, “Submitted jobs,” on page 181 for more information.

To change the file type that is associated with the object:

1. Click the ellipsis button. The **File Types** dialog box opens.
2. From the Type list, select the new file type.
3. Click **OK**. The new file type appears in the General properties dialog box.

Last modified by. The user ID under which the last modification was made. This value is generated by the system and cannot be modified.

Modification date. The date and time at which the last modification occurred. This value is generated by the system and cannot be modified.

Custom properties. If custom properties were defined for this server instance, custom property values can be assigned in this dialog box.

Topics. If topics were defined for this server instance, topic values can be assigned in this dialog box.

Setting values for custom properties

After a custom property has been created, you can assign values to the custom properties fields. Custom properties appear in the General properties dialog box.

It is important to note that you may not see custom properties for all objects on a server. Custom properties can be restricted to certain objects (for example, jobs) or certain file types.

To set values for an existing custom property:

1. Right-click on any object within the server and choose **Properties**. The General properties dialog box opens.
2. Specify the values for the custom properties that appear. The options that appear depend upon the types of custom properties that were created. See the topic “Creating custom properties” on page 36 for more information.
3. To save your changes, click **OK**.

Assigning topics

After you have created and saved a topic definition, you can assign the topic to a content object.

To assign a topic:

1. Navigate to the **General Properties** dialog box. The **Topics** table lists any previously assigned topics.
2. To assign a new topic, click **Add**. The **Add Topics** dialog box opens. This dialog box displays the topic hierarchy that was established in the **Topics Definition** dialog box.
3. Select a topic from the hierarchy.
4. Click **OK**. The **Properties** dialog box reopens, and the topic is added to the **Topics** table. Note the format of the topic reference. Each slash in the topic reference represents a folder in the topic hierarchy. The first slash represents the uppermost topics folder.

Removing topics

To remove a topic assignment:

1. Navigate to the General Properties dialog box. The **Topics** table lists any previously assigned topics.
2. Select a topic from the **Topics** table.
3. To remove a topic, click **Remove**. The topic is removed from the table.

Modifying permissions

Each object has a permission level associated with it. At times, you may want to modify permissions. For example, if you have added steps to a job and scheduled its execution for the next six months, you may want to restrict other users to read-only permissions to prevent changes to the job.

To access the permissions for the object:

1. In the Properties dialog box, click the **Permissions** tab. A list of users and their corresponding permissions appears.

You can perform the following tasks in this dialog box:

- Add a new user or group
- Modify permissions for an existing user or group
- Delete an existing user or group

Adding new users or groups

To add a new user or group and assign permissions:

1. Click **Add**. The Select User or Group dialog box opens.
2. From the **Select provider** drop-down list, select the entity that contains your user and group information. The default is *Local User Repository*.
3. In the Find field, type the first few letters of the user ID you want to add. To search for all available user IDs, leave this field blank.
4. Click **Search**. The list of users and groups that correspond to your search appears in the dialog box.
5. Select a user or group from the list.
6. Click **OK**. The user or group appears in the Principal list in the Permissions dialog box.

Modifying permissions for an existing user or group

To modify permissions:

1. Click in the Permissions cell that corresponds to the user or group whose permissions you want to modify.
2. From the drop-down list, select a new permissions level. You can select from the following options:

- Read
- Write
- Delete
- Modify Permissions

Cascading permissions

After permissions have been established in the Permissions table, the permissions changes can be cascaded, which means that the content of the current folder will have the same permissions as the parent folder. Permissions can be cascaded completely, partially, or not at all. By default, permissions are not cascaded.

Note: The ability to cascade permissions to an object is controlled by the Modify Permissions option. Permissions are cascaded to expired objects as well as active objects. If an object is not visible because it is expired and the user is not assigned the View Expired action but has Modify Permissions permission, permissions will still be cascaded. If the user does not have Modify Permissions permission, then cascading permissions will not be applied.

To cascade permissions:

1. Select the level of cascading from the drop-down list. Valid values include:
   - Do not cascade permission changes to child folders and content.

- Cascade only the updated permissions to all child folders and content.
- Cascade all permissions to all child folders and content.

2. Click **OK**.

Deleting a user or group from the permissions list

To delete a user or group:

1. Click in the Principal cell of the user or group that you want to remove.
2. Click **Delete**. The Delete Confirmation dialog box opens.
3. Click **OK**. The user is removed from the Permissions list.

Working with labels

Version labels allow a version of an object to be distinguished from another version of the same object, using a user-defined name instead of a system-generated identifier. You can use a label to reference the specific version without having to determine the version identifier.

For example, version labels are useful when scheduling jobs. Associating the schedule with a specific label ensures that only the version of the job with that label is executed when the schedule activates. Any other versions of the job, including those that were created after the labeled version, are ignored by the schedule.

Although any particular object version can be associated with multiple different labels, labels must be unique across the versions of an object. For example, there cannot be two versions of a particular job that are both labeled *Production*. This restriction allows a one-to-one mapping of label references to object versions.

Labels can be moved from one version to another version. When you move a label, every item that references the object with that label will use the new version of the object instead of the old. For example, if you move the *Production* label of a report from the first version to the seventh, any job that references the *Production* version of the report automatically uses the seventh version instead of the first.

The system automatically associates the internal label *LATEST* with the most recent version of an object. Any time a new version of an object is created, the *LATEST* label is moved to the new version automatically. Because this behavior is automatic, referencing a version by using the *LATEST* label is discouraged. Doing so often results in undesirable results. For example, if a job references the *LATEST* version of a report, the most recent version is executed. The most recent version may be a working version that is not yet finalized. However, if the job references the *Production* version, the job ignores every interim version created after the Production version. When a new version is ready for general use, the *Production* label can be moved and the job will execute that version.

Version labels persist until they are explicitly changed or removed. It is important to note that removing a version label does not remove the label from the repository. Removing a version label simply severs the connection between an object and the version label. The label is still available for use by other objects.

The following information appears on the Labels tab of the Properties dialog box:

- **Version**. The version identifier, which consists of the version number, suffixed with the date and time that the version was created. The version identifier cannot be modified.
- **Label**. The labels associated with the version. If a version has multiple labels, the labels appear in this column separated by commas. If the column width prevents all of the labels from being displayed, position the mouse cursor over the Labels cell to display the complete list. Apply and remove labels by using the **Edit Version Labels** dialog box. See the topic “Editing the version label” on page 28 for more information.
- **Creation Date**. The date and time on which the version was created. The creation date cannot be modified.
- **Expiration Date**. The date on which the version is set to expire. The expiration date is set on the Versions tab of the Properties dialog box. See the topic “Version properties” on page 35 for more information.

- **Created By**. The user who created the version. This value cannot be modified.

From the Labels tab, you can perform the following actions:

- Apply a new version label
- Edit an existing version label
- Apply a predefined version label
- Specify multiple version labels
- Remove a version label
- Delete a version

Version label recommendations

A label policy that defines a set of labels and outlines how they should be used is strongly recommended. Such a policy will help prevent inadvertently breaking automated processes as objects change.

In particular, this policy should consider the following:

- When creating new objects and object versions, apply specific version labels, such as *Test* or *Production*, that associate the version with a procedural stage. The label should be moved to a new version when that version is ready for general use. This ensures that references to the object will consistently be linked to the expected version and the references will not change without an explicit label move.
- Job versions should be identified with specific labels that allow execution of an intermediate test version of the job while the validated production version continues to operate. This allows new objects or object versions to be tested before being promoted to general use. The test version of the job can be promoted by moving the version label for production when the test job is ready for general use.
- Job steps should always reference object versions by using a specific version label instead of the system *LATEST* label to ensure that a specific object version will be used regardless of whether more recent versions of that object exist. This permits the job to run successfully while the referenced object is being revised. If the reference uses the *LATEST* version, the job will use the most recent object version, which may require modifications of the job properties. Without those changes, the job could fail. By referencing a specific version label and by labeling job versions, you can modify the job properties to accommodate changes to the referenced object version. Once those changes are validated, you can promote the new object version and the revised job for production use.
- Use label security, which allows for restricting the users allowed to assign or move specific version labels, to ensure that a proper review and validation of objects and jobs by a qualified individual occurs before new versions are promoted to production use.

Editing the version label

To edit the version label, select the object version from the table on the Labels tab of the Properties dialog box and click **Edit**.

To apply a new label, perform the following steps:

1. Type the name in the **New Label** field. Labels cannot contain commas, apostrophes, semicolons, double quotation marks, greater than symbols, or less than symbols.
2. Click the arrow button to add the label to the **Applied Labels** field.

To apply an existing label, perform the following steps:

1. Select the label(s) from the **Available Labels** list. To select multiple labels, press **Ctrl** and make multiple selections.
2. Click the arrow button to add the selected label(s) to the **Applied Labels** list.

To remove the association between a version and a label, perform the following steps:

1. Select the label from the **Applied Labels** list.
2. Click the arrow button to remove the label from the **Applied Labels** list.

When your modifications are complete, click **OK**. The Versions table displays the updated values.

Resolving duplicate version labels
Although multiple version labels can be applied to an object, each individual version label must be unique for that object. (However, a label with the same name can be used for another object.) In the event of a conflict, the Labels Already Exist dialog box describes the version label duplication.

To resolve a duplicate label conflict, the following options are available:

- **Apply the label.** To apply the label to the existing version and remove the label from the previous version, click **OK**.
- **Specify a new label.** To maintain the label with the previous version and specify a new label for the current version, click **Cancel**. The Edit Version Labels dialog reappears. Specify a new label in this dialog.

Deleting versions

To delete an IBM SPSS Collaboration and Deployment Services Repository object version:

1. On the Labels tab of the Properties dialog box, select the version.
2. Click **Delete**. The Delete Confirmation dialog box opens.
3. Click **OK**. The version is removed from the list.

Note: Once a version is deleted from IBM SPSS Collaboration and Deployment Services Repository, it cannot be recovered.

Selecting a version
At various points in the system, you may be asked to select a version of a file. For example, if you want to download a file that has multiple versions, the system prompts you to select a version.

To select a version:

1. If your file has multiple versions, the Select file version dialog box opens.
2. Select the version of the file on which you want to perform an action.
3. Click **OK**.

If the selected version of the file has multiple labels associated with it and the version opens in the IBM SPSS Deployment Manager environment, the first label appears with the name of the file in the tab title to assist in identifying the version.

Working With Expiration Dates and Expired Files

Working with Expiration Dates and Expired Files

An expiration date specifies the date after which a file is no longer actively in use. When an item expires, it is no longer visible to the general user. However, the file owner and administrators can see expired files.

Note that expiration does not equal deletion. An expired file is not removed from the repository. By default, files do not expire. An expiration date must be explicitly set. Expiration dates can be set, modified, or reactivated by any of the following:

- The owner of the file
- An administrator
- Any user with write permissions to the file

The file or job itself does not have an expiration date associated with it. Expiration dates are set at the version level. Version numbering continues sequentially. However, a gap in version numbers may appear on the versions list if some file versions have expired.

When a file version expires, any corresponding labels are unaffected. If an expired file version is opened, an error message appears, indicating that the file has expired. When the labeled version of a job expires, it cannot be executed after the expiration date. Thus, any schedules that contain the expired job version are affected.

Schedules are not deleted when the associated version expires. The system will still attempt to execute the scheduled job at the designated time. However, because the job version contained in the schedule has expired, the job will fail, generating an error message in the log. The failure of the scheduled job is recorded in the job history. See the topic "Job history view" on page 133 for more information.

Setting and Modifying Expiration Dates

Expiration dates are set in the Properties dialog box. In addition, an expiration date can also be specified when a job is initially created.

By default, objects do not expire. Expiration dates can be modified at any time. In addition, expiration dates can be retroactive. To set an expiration date on an object:

1. In the Content Explorer, right-click on the object and select **Properties**. The Properties dialog box opens.
2. Click the Versions tab and select the version on which to specify an expiration date.
3. Click the ellipsis button next to the **Expiration Date** field. The Expiration Date dialog box opens. By default, no expiration date is set.
4. Select one of the following options:
   - **No Expiration Date.** The version does not expire.
   - **Expire in.** The version expires at a time relative to today's date—for example, in 30 days. Valid relative time periods include days, weeks, months, or years. The maximum expiration date allowed is 10 years.
   - **Expire on.** The version expires on the date selected.
5. Click **OK**. The expiration date appears in the **Expiration Date** field. For both relative and specific dates, the exact calendar expiration date appears in this field. If no expiration date was specified, the field remains blank.

Alternatively, the expiration date can be specified when a new job is created. The process for accessing and supplying information for the expiration date dialog box is the same. However, the dialog box is accessed from the **New Job Information** dialog box instead of the Properties dialog box.

Importing and Exporting Expiration Dates

When files are imported into and exported out of the system, the expiration date attached to a file version is included as well.

Restrictions on Expiration Dates

Expiration dates can be applied to most files within the Content Explorer. However, expiration dates cannot be applied to objects in the Resource Definitions folder—for example, server, credential, and data source definitions.

Viewing Expired Files

Expired files are visible depending upon whether or not all versions of a file have expired.

Files with some versions expired. For files with some expired versions, the expired versions are displayed.

Files with all versions expired. The expired files are hidden in the Content Explorer. However, if the owner of the expired files or an administrator views the tree, the expired files are visible.

Searching for Expired Files

The process of searching for expired files is the same as the process of searching for other objects in the Content Explorer.

See the topic "Searching" on page 17 for more information.

To see expired files in the search results, a user must be one of the following:

- A user with the *view expired files* action
- The owner of the file

- An administrator

Reactivating Expired Files

Expired files can be reactivated. To reactivate an expired file version:

1. Navigate to the existing expiration date in the Properties dialog box.
2. Modify the expiration date to a future date.

Working with server properties and user preferences

Server properties are divided into the server properties and user preferences subcategories.

Both types of properties apply to the current server instance (or user) only. To access the Properties dialog box for a server instance, right-click on a server name in the Content Explorer and choose **Properties**. The Server Properties dialog box opens.

Server properties

Server properties are similar to content object properties. However, server properties apply to the overall server instance.

Users with the appropriate access privileges can view and modify the following:

- Custom properties. See the topic "Working with custom properties" on page 36 for more information.
- Server connection
- Label security
- Topic definitions. See the topic "Working with topic definitions" on page 40 for more information.

Server connection

Server connection properties define the information that is needed to connect to a IBM SPSS Collaboration and Deployment Services Repository. To access the connection properties, click the **+** to expand the server properties options in the Server Properties dialog box and select Server Connection.

Name
The name for the server connection.

Server URL
The complete connection URL for the server that hosts the IBM SPSS Collaboration and Deployment Services Repository.

The URL includes the following elements:

- The connection scheme, or protocol, as either *http* for hypertext transfer protocol or *https* for hypertext transfer protocol with the secure socket layer (SSL)
- The host server name or IP address

  Note: An IPv6 address must be enclosed in square brackets, such as `[3ffe:2a00:100:7031::1]`.
- The port number. If the repository server is using the default port (port 80 for http or port 443 for https), the port number is optional.
- An optional custom context path for the repository server

Table 5. Example URL specifications . This table lists some example URL specifications for server connections.

URL	Scheme	Host	Port	Custom path
`http://myserver`	HTTP	*myserver*	default (80)	(none)
`https://9.30.86.11:443/spss`	HTTPS	9.30.86.11	443	*spss*
`http://[3ffe:2a00:100:7031::1]:9080/ibm/cds`	HTTP	`3ffe:2a00:100:7031::1`	9080	*ibm/cds*

Contact your system administrator if you are unsure of the URL to use for your server.

Label security

Label security provides control over which principals can view or modify user-defined version labels in the system, allowing the specification of which users can move or delete a label. Although users may be able to view a version of a resource in the IBM SPSS Collaboration and Deployment Services Repository, only those users with permissions for the labels associated with the version can access the labels. Moreover, label security applies to any version of any resource using the label. For example, to control who can assign production versions of resources, an administrator may restrict access to the *Production* label to prevent other users from moving that label from one version to a newer version. The security defined for the *Production* label applies to every version of every resource using that label.

To access label security properties, click the **+** to expand the server properties options in the Server Properties dialog box and select Label Security.

The Labels table lists all labels in the system you can access. Select a label to view the associated principals and their permissions

For a selected label, the Principals table identifies label permissions for principals in the system. A principal can have one of the following permissions for a label:

- **Use versions.** Any principal with this permission can view the version of a resource associated with the label by referencing the label.
- **Manage label.** Principals with this permission can apply, move, and delete the label, as well as delete the version of a resource having the label. If a user does not have the *Manage label* permission for a label, that user cannot delete a resource or version of a resource to which the label is applied

To edit the security for a label, such as adding or deleting principals, you must have the *Manage label* permission for the label. Administrators have this permission for all labels in the system. The *Manage label* permission does allow you to alter your own permission for a label, making it possible for you to edit your permissions for a label in such a way that you will be unable to work with the label in the future. For example, you might change your permission to *Use versions*, which would subsequently prevent you from changing the security for a label. If situations such as this occur, contact your administrator to have your permissions adjusted as needed.

Adding principals for a label

To add a principal for a label:

1. Select the label from the Labels table.
2. Click **Add**. The Select User or Group dialog box appears.
3. Select the principal to add.
4. Click **OK**.

The principal appears in the Principals table with the default permission.

Modifying permissions for principals

To modify the label permissions for a principal:

1. In the Principals table, click the Permissions cell for the principal.
2. Select the permission.

Deleting principals for a label

To delete a principal for a label:

1. Select the label from the Labels table.
2. In the Principals table, select the principal to be deleted. To select multiple principals, hold down the Ctrl key while selecting principals.
3. Click **Delete**.

The principal is removed from the Principals table and will be unable to access the label.

User preferences

User preferences apply to the user ID that is currently logged in to the server.

Users with the appropriate access privileges can view and modify the following:

- Analytic data view editor advanced mode
- Default permissions
- Distribution channels
- Subscriptions recipient
- Version label

Default permissions

The default permissions option allows a user to set up a default permissions for new files and folders that are created. Thus, when a user creates a new object, such as a job, the object's permissions list defaults to the list of principals and permissions defined in the user's preferences for the instance of IBM SPSS Collaboration and Deployment Services.

In the Default Permissions dialog box, the following changes can be made:

- Add a new principal or delete an existing principal
- Modify permissions for a principal

To specify default permissions:

1. Navigate to the Properties dialog box for the server instance.
2. In the Properties tree, expand the User Preferences list and select **Default Permissions**.
3. Modify the list of principals as needed. Users or groups can be added to or deleted from the Principal list.

   Add a new principal. To add a new principal, click **Add**. The Select User or Group dialog box opens. Select a user or group from the list and click **OK**.

   Delete an existing principal. To delete an existing principal, select a principal from the list and click **Delete**. The Confirm Deletion dialog box opens. Click **OK**.
4. To modify the permissions associated with a principal, click the drop-down arrow in the Permissions column and select the permission level to associate with the principal. Valid values include:

   - Read

- Write
- Delete
- Modify Permissions

5. Click **OK**.

Important: For members of the *administrators* group, the permissions of the newly-created objects will default to Modify Permissions even if they are set to a different level in user preferences.

Distribution channels

The Distribution Channels option allows the user to specify how they receive notifications. For example, if your company uses IBM SPSS Collaboration and Deployment Services RSS feeds, you may choose to disable email notifications and only enable RSS feeds.

To select distribution channels for notifications and subscriptions:

1. Navigate to the Properties dialog box for the server instance.
2. In the Properties tree, expand the User Preferences list and select **Distribution Channels**.
3. To enable notifications via email, select **Enable Email**. To enable notifications via syndicated feeds, select **Enable RSS**.

   Note that the **Enable RSS** option is only available to users who are granted the **Access Syndicated Feeds** action by an administrator.

   Syndication feeds provide a format for delivering regularly changing content. IBM SPSS Collaboration and Deployment Services syndicates its notifications as RSS feeds or Atom feeds. An RSS aggregator (feed reader) that supports authenticated feeds is required for viewing notification feeds. For example, RSSBandit and Microsoft® Outlook® 2007 are common desktop-based feed readers, and Google Reader™ and Yahoo® Reader are common Web-based feed readers.

   Administrators can override your settings by disabling feeds for all users via the browser-based IBM SPSS Deployment Manager. Administrators also define the output format used for all syndication feeds (RSS or Atom), limit the total number of syndication feed entries served to the RSS/Atom subscribers, control which users can access feeds, etc.
4. Click **OK**.

Subscriptions recipient

The Subscriptions Recipient option allows the user to choose between the following options:

- Typing in an email address to be used as the default for subscriptions
- Using the email address associated with the user from a backing directory, such as active directory

The email address applies per server instance and per user, which means that a user can specify different email addresses for different servers.

To specify an email address for a subscriptions recipient:

1. Navigate to the Properties dialog box for the server instance.
2. In the Properties tree, expand the User Preferences list and select **Subscriptions Recipient**.
3. Select one of the following options:
   - In the **Email Address** field, type the email address to use as the default.
   - Select the **Use email address from directory** check box to use an email address from the directory.
4. Click **OK**.

Version label (user preferences)

To access server version label properties:

1. In the Content Explorer, right-click on a server name and choose **Properties**. The Server properties dialog box opens.
2. Click **+** to expand the user preference options.
3. Click **Version Label**. The Version Label dialog box opens.
4. From the **Default label used for files in Jobs** drop-down list, select the default version label to apply to files within a job.
5. Click **OK**.

Version properties

The IBM SPSS Collaboration and Deployment Services Repository allows you to maintain multiple versions of objects.

To access the version listing for an object:

1. In the Properties dialog box, click the Versions tab. A list of versions for the object appears.

The list displays the version number and the created timestamp of the version. The most recent version is labeled *LATEST*. *LATEST* is the version selected by default in the list. The list also displays the *All Versions* entry.

In addition to the general properties defined for IBM SPSS Collaboration and Deployment Services Repository objects, certain properties can also be defined on a version level.

Default version properties include:

- **Description.** A user-defined label for the version.
- **Keywords.** The metadata assigned to IBM SPSS Collaboration and Deployment Services Repository object versions for the purposes of content search.
- **Expiration Date.** The date after which IBM SPSS Collaboration and Deployment Services Repository object versions are no longer active. See the topic "Working with Expiration Dates and Expired Files" on page 29 for more information.

The dialog box also displays custom version properties defined for the server. See the topic "Creating custom properties" on page 36 for more information. The properties are displayed for a selected version in the right pane of the properties dialog box. They can be set for individual versions, for multiple versions, or for all existing versions. The properties allow the user to enhance the metadata for a particular version of an object. For example, the version description property could be used to clearly state its differences from other versions, while specifying keywords makes the search more precise.

To set or modify properties for a single version:

1. Select the version.
2. Modify the properties as necessary. The properties are set to the specified values for the selected version.

To set or modify properties for multiple versions:

1. Hold down the Shift key to select the versions. Any property fields that have different values between the selected versions are empty and marked by alert signs.
2. Modify the properties as necessary. The properties are set to the specified values for the selected versions.

To set or modify properties for all versions:

1. Select *All Versions* in the list. Any property fields that have different values between the existing versions are empty and marked by alert signs.
2. Modify the properties as necessary. The properties are set to the specified values for all versions.

Working with custom properties

Custom properties consist of user-defined metadata that apply to objects within the repository. You must have the appropriate access privileges to create and modify custom properties.

Custom properties are applied per server instance. Custom properties persist across sessions and remain associated with the server instance until they are removed.

Currently, you cannot copy custom properties from one server instance to another. In addition, you cannot make any of the custom properties fields required.

The process for working with custom properties consists of the following tasks:

- Ensuring that you have access privileges to create custom properties
- Creating custom properties
- Setting values for custom properties

Typically, an administrator creates custom properties, and non-administrative users set values for custom properties.

Although custom properties are created at the server level, values for custom properties are set at the content object level. Specifically, the custom properties that you create in the Custom Properties dialog box appear in the General properties dialog box for each content object. See the topic "Editing general properties" on page 24 for more information.

Verifying access privileges to create custom properties

Before you begin working with custom properties, you must have the appropriate access privileges to create and modify custom properties. If you log in to a server and you do not have access privileges to create custom properties, the custom properties option will not be visible.

A user with administrative privileges to the server instance automatically has permissions to create, edit, and delete custom properties.

Suppose that you log in to a repository and you do not have the appropriate access privileges to create custom properties. If your administrator assigns you access privileges while you are still logged in to the repository, you will need to log out of the repository and then log back into the repository before the updated access privileges take effect.

Accessing the custom properties dialog box

To access the Custom Properties dialog box:

1. Log in to the server for which you want to create custom properties. See the topic "Logging in to a server" on page 15 for more information.
2. Right-click on the server name and choose **Properties**. The Properties dialog box opens.

   Note: You must right-click on the server name to view custom properties. If you select any other object within the server folder, only general properties and permissions appear.
3. Click **Custom Properties**. The Custom Properties table appears. If you have previously defined custom properties, they appear in the table.

Creating custom properties

To create a new custom property:

1. Navigate to the **Custom Properties** dialog box. See the topic "Accessing the custom properties dialog box" on page 36 for more information.
2. To create a new custom property, click **Add**. The **Custom Property Parameters** dialog box opens.
3. In the **Custom Property Parameters** dialog box, you need to provide the following information:

- **Label**. The label is the name of the custom property as it will appear in the user interface. The label can be up to 128 characters in length and must be unique to the server instance. If you specify a duplicate name, the Duplicate Property dialog box opens indicating that the name is already in use. However, you may specify the same label for custom properties on two different servers.
- **Property Type.** The property type describes the input value for the custom property. You have the following options:

Table 6. Custom property types

Property type	Description	Representation in the General properties dialog box
String	Any valid string. String is the default property type.	Text field.
Number	Any numeric value.	Text field.
Single choice	A single selection from the list of options.	Drop-down list.
Multiple select	One or more selections from the list of options.	A series of check boxes.
Yes/No	Either yes or no. Selecting the check box indicates yes. A blank check box connotes no. By default, the check box is cleared. Thus, the default is no.	A single check box.
Date	A valid date in the format m/d/yyyy—for example, 5/18/2006.	A drop-down calendar that prompts the user to specify a date. Alternatively, the user can also type the date.

- **Selection Values.** For custom property types that require the user to choose from a list of values (for example, single select and multiple select), you need to define the values available to the user in the Selection Values list. Items in the list are line delimited. Thus, when you add selection values, each value must appear on a new line.
- **Property Applies to.** This field indicates the objects to which the custom property is applied. You must select at least one check box. You have the following options:

Property Applies to	Description
Folders	Applies to all folders in the repository.
Files	Applies to all files in the repository.
File types	Applies only to the file types you specify. If you select the **File types** option, the **Files** option is automatically selected. To specify a file type: a. Click the ellipsis button. The **File Types** dialog box opens. b. From the **File Types** list, select the file type you want to associate with the custom property. To select multiple file types, press the Ctrl key and select multiple rows. If you select one file type, that file type appears in the text box. If you select multiple file types, the first file type appears in the list followed by an ellipsis. c. Click **OK**.
Jobs	Applies only to jobs in the repository.

By default **Folders**, **Files**, and **Jobs** are selected.

- **Property value can be set.** This field indicates whether the property can be set only at the object level or the version level.

4. To save your changes, click **OK**.

Editing custom properties

After you have created and saved a custom property, you can edit its attributes at any time.

In this context, editing refers to modifying the values available to the user. Editing custom properties takes place at the server level. This process is separate from setting values for the custom property, which takes place at the content object level.

To edit a custom property:

1. Navigate to the Custom Properties dialog box. See the topic "Accessing the custom properties dialog box" on page 36 for more information.
2. In the Custom Properties table, select the custom property that you want to edit.
3. Click **Edit**. The Custom Property Parameters dialog box opens.
4. You can modify any of the parameters in this dialog box except for the following:
   - **Property Type.** When you edit a custom property, the **Property Type** field is disabled. To change the property type, you must create a new custom property. See the topic "Creating custom properties" on page 36 for more information.
   - **Property Value Can Be Set.** The level at which a custom property value is set cannot be modified. When you edit a custom property, this section of the dialog box is disabled.
5. After you have made changes, click **OK** to save them.

Editing selection values

For the single select and multiple select property types, you can modify the selection values available to users. It is important to note that if you add or delete any of the options in the Selection Values table, that change is propagated to all of the objects on the server associated with that custom property.

For example, suppose you have a multiple select property type with the following options:

- *Harry*
- *Ron*
- *Fred*
- *George*

If you remove *George* from the Selection Values list, the value *George* would be removed from any content objects to which it was previously assigned. In addition, any users that had previously selected *George* would no longer see *George* as an option.

Searching for custom properties

Custom properties are searchable and used in auditing reports, with the following exceptions:

- **Yes/no label values.** The label that you apply to any yes/no property types will not appear in a search.
- **Numeric values.** For any number property types, the numeric value that you type into the text field will not appear in a search.

See the topic "Searching" on page 17 for more information.

Deleting custom properties

Custom property definitions are stored in the repository. If you delete a custom property, the custom property and all of its attributes are removed from the repository. Once a custom property is deleted, it cannot be recovered.

It is important to note that when you delete a custom property, the deletion affects all of the objects in the repository that contained that custom property. Once deleted, a custom property will no longer be applied to any object. It is strongly recommended that you do not delete custom properties, unless absolutely necessary.

For example, suppose you have a custom property called **Reviewers** that applies to all of the files on your server. After you delete the custom property and navigate to the Properties dialog box, the **Reviewers** custom property will no longer appear for any file on the server.

To delete a custom property:

1. Navigate to the Custom Properties dialog box. See the topic "Accessing the custom properties dialog box" on page 36 for more information.
2. From the Custom Properties table, select the property that you want to delete. To select multiple custom properties, press the Ctrl key and select multiple rows.
3. Click **Delete**. The Confirm Deletion dialog box opens.
4. Click **OK**. The Custom Properties dialog box reopens, and the Custom Properties table is refreshed.

Custom Properties and Server Connections

When you delete a server connection, the system removes the connection between the client and the server that hosts your repository. This process does not modify any items in the repository itself. Thus, deleting a server connection has no effect on the custom properties associated with that repository.

Working with topics

Topics allow the definition of a classification system for the content stored in the IBM SPSS Collaboration and Deployment Services Repository, providing a hierarchical map to guide users to the resources they need. Topics function like a directory structure but differ from directories in that a single object can be listed under multiple topics.

For example, you might want to create a topic structure that mirrors your organization, with separate topics for marketing, finance, development, and so on. Users can then choose from the available topics when storing content. In addition, users can limit content searches to specific topics to accelerate the retrieval process. Because a given item can be listed under multiple topics, cross-indexing is also possible.

Alternatively, consider the following topic hierarchy, in which the model type is the basis for classifying resources:

```
Model
   association
     Apriori
     CARMA
     GRI
     Sequence
   clustering
     K-means
     Kohonen
     TwoStep
   decision tree
     C5
     CHAID
     C&RT
     QUEST
   neural network
   screening
     Anomoly detection
     Feature selection
   statistical
     factor PCA
     linear regression
     logistic regression
   text extraction
```

Any model in the repository could be assigned a topic in this hierarchy to assist in finding desired resources. For example, a user might want to find all association models that use a specific field. Alternatively, the search may be restricted to CARMA models only.

Process overview

Topics consist of user-defined metadata that apply to objects within the repository. You must have the appropriate access privileges to create and modify topics. Working with topics consists of a two-part process.

Creating topic definitions. Topic definitions and the topic hierarchy are established in the **Topic Definitions** dialog box. The topics that you create in the **Topic Definitions** dialog box appear in the **General Properties** dialog box for each content object. See the topic "Editing general properties" on page 24 for more information.

The following guidelines apply:

- Topic definitions are applied per server instance
- Topic definitions persist across sessions and remain associated with the server instance until they are removed
- Topic definitions cannot be copied from one server instance to another
- Topics that you define cannot be made into required fields

Assigning topic values. Although topic definitions are created at the server level, values for topics are set at the content object level. It is important to note that topics are assigned at the object level only. Topics cannot be assigned to folders. In addition, topics cannot be assigned to any items in the *Resource Definitions* folder of the Content Explorer.

Typically, an administrator creates topic definitions, and non-administrative users assign topics.

Verifying access privileges to create topic definitions

Before you begin working with topic definitions, you must have the appropriate access privileges to create and modify topic definitions. If you log in to a server and you do not have access privileges to create topic definitions, the topic definition option will not be visible.

A user with administrative privileges to the server instance automatically has permissions to create, edit, and delete topic definitions.

Suppose that you log in to a repository and you do not have the appropriate access privileges to create topic definitions. If your administrator assigns you access privileges while you are still logged in to the repository, you will need to log out of the repository and then log back into the repository before the updated access privileges take effect.

Working with topic definitions

The following tasks can be performed in the **Topic Definitions** dialog box:

- Create new topic definitions
- Rename topic definitions
- Move topic definitions within the topic hierarchy
- Delete topic definitions

Accessing the topic definitions dialog box

To access the **Topic Definitions** dialog box:

1. Log in to the server for which you want to create topics. See the topic "Logging in to a server" on page 15 for more information.
2. Right-click on the server name and choose **Properties**. The **Properties** dialog box opens.

Note: You must right-click on the server name to view topics. If you select any other object within the server folder, only general properties, permissions, and versions appear.

3. Click **Topic Definitions**. The topics hierarchy appears. If you have previously defined topics, they appear in the hierarchy. Otherwise, only the main topics folder appears.

Creating new topic definitions

To create a new topic definition:

1. Navigate to the **Topic Definitions** dialog box. See the topic "Accessing the topic definitions dialog box" on page 40 for more information.
2. Select the folder of the topic hierarchy in which to create the new topic definition.
3. To create a new topic definition, click **Add**. The **New Topic** dialog box opens.
4. In the **Topic Name** field, type a name for your topic.
5. Click **OK**. The **Topic Definitions** dialog box reopens and the topic hierarchy is refreshed.

Renaming topic definitions

Topic definitions can be renamed. After a topic definition is renamed, the change is propagated to all content objects to which the topic is assigned.

Any topic definition can be renamed. However, the root topics directory cannot be renamed. The **Rename** button is disabled when you select the **Topics** folder.

To rename a topic definition:

1. Navigate to the **Topic Definitions** dialog box. See the topic "Accessing the topic definitions dialog box" on page 40 for more information.
2. From the topic hierarchy, select the topic you want to rename.
3. Click **Rename**. The **Rename Topic** dialog box opens.
4. In the **Topic Name** field, type a name for your topic.
5. Click **OK**. The **Topic Definitions** dialog box reopens, and the topic hierarchy is refreshed.

Moving topic definitions

Topic definitions can be moved by cutting the topic definition and pasting it to a new location within the topic hierarchy. It is important to note that when a topic is cut from the topic hierarchy, all of topics within it are cut as well.

Any topic definition can be moved. The only exception is the root topics directory. This directory cannot be moved. The **Cut** and **Paste** options are disabled when you select the **Topics** folder.

To move a topic definition:

1. Navigate to the **Topic Definitions** dialog box. See the topic "Accessing the topic definitions dialog box" on page 40 for more information.
2. In the topic hierarchy, right-click on the topic you want to move and choose **Cut**.
3. Right-click on the folder into which you want to paste the topic definition and choose **Paste**. The topic definition is moved into its new location.

Deleting topic definitions

Topic definitions are stored in the repository. If you delete a topic, the topic and all of its attributes are removed from the repository. Once a topic definition is deleted, it cannot be recovered.

For example, suppose you have a topic definition called *Analysis*. After you delete the topic definition and navigate to the Properties dialog box, the *Analysis* topic will no longer appear for any content object on the server.

It is important to note the following about deleting topic definitions:

- When you delete a topic definition, the deletion affects all of the objects in the repository that contained that topic. Once deleted, a topic definition will no longer be applied to any object.
- When you delete a topic definition, all of the topic definitions within it are deleted as well.
- You cannot delete the root topics directory. The **Delete** button is disabled when you select the **Topics** folder.

To delete a topic definition:

1. Navigate to the **Topic Definitions** dialog box. See the topic "Accessing the topic definitions dialog box" on page 40 for more information.
2. From the topic hierarchy, select the topic that you want to delete.
3. Click **Delete**. The **Delete Topic Confirmation** dialog box opens.
4. Click **OK**. The **Topic Definitions** dialog box reopens, and the topic hierarchy is refreshed.

The effect of deleting server connections on topics

When you delete a server connection, the system removes the connection between the client and the server that hosts your repository.

This process does not modify any items in the repository itself. Thus, deleting a server connection has no effect on the topic definitions or assignments associated with that repository.

Deleting topic definitions versus removing topic assignments

Deleting a topic definition is different from removing a topic assignment.

Deleting a topic definition removes the definition from the system, preventing any content object from being associated with that topic definition. Thus, deleting a topic definition affects all content objects associated with that topic definition.

Removing a topic assignment severs the link between a content object and the topic property. However, removing a topic assignment from a content object does not affect any other content objects that were assigned that topic. See the topic "Removing topics" on page 25 for more information.

Searching for topics

Topics are searchable and used in auditing reports.

See the topic "Searching" on page 17 for more information.

Bulk property updates

You can update properties for multiple objects simultaneously (in bulk).

The properties that can be updated in bulk include:

- Description
- Keywords
- Topics
- Permissions
- Expiration date
- Custom properties
- Content language

To update properties in bulk:

1. In the Content Explorer, select the objects. Selected objects can include files, folders, and jobs.
2. Right-click and choose **Properties**. The Shared Properties dialog box opens. The dialog box title bar indicates how many objects have been selected.

3. Modify properties on the General, Versions, and Permissions tabs as necessary.

   Note: The Shared Properties dialog box does not display the Versions tab if any folders are selected for the bulk update.

4. Click **OK**.

Updating general properties in bulk

To access the general properties for the selected objects:

1. In the Shared Properties dialog box, click **General**.
   - If selected objects are files or jobs, Author, Content Type, Last Modified By, and Modification Date fields are displayed. The dialog box also displays object-level custom properties defined in the repository and the topics common to all objects.
   - If selected objects are folders, only Description, Last Modified By, and Modification Date fields are displayed. The dialog box also displays all custom properties defined for the repository and the topics common to all objects.
   - If selected objects include both files or jobs and folders, only Last Modified By and Modification Date fields are displayed. The dialog box also displays object-level custom properties defined for the repository. The dialog box does not display topics.
   - Any property fields that have different values between the selected objects are empty and marked by alert signs. Green boxes are displayed for check box custom properties when the state of the check box is not common for all selected objects.
2. Modify the property values as necessary. To change the content type of the objects, click the ellipsis button to open the **File Types** dialog box and select the type. Click **Add** to open the **Add Topic** dialog box and select topics. To remove a topic, select the topic in the list and click **Remove**. See the topic "Working with topics" on page 39 for more information.

Note: If there are jobs among selected objects, the Content Type field is read-only.

Updating version properties in bulk

To access the version properties for the selected objects:

1. In the Shared Properties dialog box, click **Versions**.

- **Description.** A user-defined label for the version.
- **Keywords.** The metadata assigned to IBM SPSS Collaboration and Deployment Services Repository object versions for the purposes of content search.
- **Expiration Date.** The date after which IBM SPSS Collaboration and Deployment Services Repository object versions are no longer active. See the topic "Working with Expiration Dates and Expired Files" on page 29 for more information.

The dialog box also displays custom version properties defined for the server. See the topic "Creating custom properties" on page 36 for more information. The properties are displayed either for the latest versions of the selected objects or for all versions. Any property fields that have different values between the selected objects are empty and marked by alert signs. Green boxes are displayed for check box custom properties when the state of the check box is not common for all selected objects.

To set or modify properties for the latest versions of selected objects:

1. Select *LATEST* in the list.
2. Modify the properties as necessary.

To set or modify properties for all versions:

1. Select *All Versions* in the list.
2. Modify the properties as necessary.

Updating permissions in bulk

To access the permissions for the selected objects:

1. In the Shared Properties dialog box, click **Permissions**.

Owner. The users who created any of the selected objects are marked with a green box in the Owner column.

Principal. The user ID and the number of the selected objects for which the users have permissions defined.

Permissions. The Permissions column displays the user's permissions for the selected objects. If the user is not defined in the permissions of all objects, an asterisk (*) is displayed. To set users permissions to the objects, click in the column and select a permissions level from the list.

Add. Add users to the list. See the topic “Adding new users or groups” on page 26 for more information.

Delete. Delete users from the list. See the topic “Deleting a user or group from the permissions list” on page 27 for more information.

Chapter 6. Resource definitions

In the Content Explorer, the Resource Definitions folder contains credential definitions, data source definitions, message domains, promotion policies, server definitions, and server cluster specifications. These resources are often required for running jobs .

For example, a IBM SPSS Modeler job step requires the definition of a IBM SPSS Modeler execution server to process the step.

To create new resource definitions, you must have the appropriate actions associated with your login credentials. In addition, you need *Write* permission to the folder that will contain the new definition. Contact your IBM SPSS Collaboration and Deployment Services administrator if your actions or permissions require modifications.

You cannot save any other items into the *Resource Definitions* folder. For example, jobs cannot be saved here. If you try to save or move an item into the *Resource Definitions* folder, an error message appears.

Credentials

Execution of some job steps requires the specification of user credentials, allowing credential validation before executing a job .

User credentials authenticated through an external directory system, such as Active Directory, can be associated with a domain. Write permissions to the Credentials folder are required to create and modify data credentials. See the topic "Modifying permissions for an existing user or group" on page 26 for more information. Also, the user must be assigned the *Define Credentials* action through the role.

Adding new credentials

To add a new credential:

1. In the Content Explorer, open the *Resource Definitions* folder.
2. Open the *Credential Definitions* folder.
3. Select the domain in which you want to create a new credential definition.
4. From the File menu, choose:

   New > **Credentials Definition**

   The Add New Credentials wizard opens.

   Note: Alternatively, the new credentials dialog box can be accessed by clicking **New** next to the login field on the General tab for certain steps (for example, IBM SPSS Modeler job steps).
5. In the Name field, enter a nickname for the credential. The credential destination dialog box opens.

Credential destination

If you are defining the credential under a domain, after you have specified the name, the credential destination dialog box opens.

1. Select the domain.
2. Click **Next**. The user and password credentials dialog box opens.

Specifying a user and password for credentials

After you have specified a name and destination for your credential and clicked **Next**, the User and password credentials dialog box opens.

1. In the user ID field, enter the username for the credential.
2. In the Password field, enter the password that corresponds to the username.

3. In the Confirm password field, reenter the password.
4. If the credential must be validated against a security provider, select the provider using the Security Provider drop-down list. Leave this field blank if provider validation is not necessary for the credential.
5. Click **Finish**. The updated information appears in the *Credentials* folder.

Creating a new domain

To create a new domain for your credentials:

1. In the Content Explorer, navigate to the **Resource Definitions** folder.
2. Right-click the **Credential Definitions** folder and select:

   New > **Domain**
3. A new domain appears in the **Credential Definitions** folder, called `New Domain`.
4. Assign a name to the new domain.

   Note: Domains are LDAP recognized.

Server Process Credential

Server Process Credential is the built-in credentials definition of the user profile under which the repository server is run. In Active Directory or OpenLDAP-based single sign-on environment, Server Process Credential can be used instead of regular repository user credentials to:

- Run reporting job steps and schedule time-based jobs
- Query a security provider for a list of user and group profiles

Important! Server Process Credential cannot be used in non-single sign-on environments and requires a number of additional configuration steps. For more information, see IBM SPSS Collaboration and Deployment Services Repository configuration documentation.

Server Process Credential cannot be opened, edited, copied, cut pasted, labeled, or versioned, but you can modify permissions to the credential if you have sufficient authority.

Data sources

Open database connectivity (ODBC) provides a mechanism for client programs to access databases or data sources. Similarly, Java database connectivity (JDBC) determines how Java applications access databases.

Data source definitions allow other components of the system to access the data sources used in the system. Data source definitions are created in the Content Explorer and stored in the IBM SPSS Collaboration and Deployment Services Repository. The appropriate permissions are required to create and modify data source definitions. If a user does not have permissions, the **Data Source Definitions** folder is not accessible. See the topic "Modifying permissions" on page 26 for more information. The process for adding a data source definition consists of the following steps:

1. Selecting a data source definition type.
2. Specifying data source definition parameters.

Data source definitions can be exported and imported.

Selecting a data source definition type

To create a new data source definition:

1. In the Content Explorer, expand the **Resource Definitions** folder. Right-click on the **Data Source Definitions** folder and select:

   New > **Data Source Definition**

The **Select Data Source Definition Type** dialog box opens.

2. In the Name field, enter the name for your data source definition. This name, which will appear in the **Data Source Definitions** folder, is used when other components of the system invoke the data source definition.
3. From the Type drop-down list, select the data source definition type. Valid values depend on the system configuration and may include *ODBC Data Source*, *JDBC Data Source*, *App Server Data Source*, and *Data Service Data Source*.
4. Click **Next**. The dialog box that appears depends on the data source type that was selected.

- **ODBC**. The **DSN** dialog box opens.
- **JDBC**. The **JDBC Name and URL** dialog box opens.
- **App Server Data Source**. The **JNDI Name** dialog box opens
- **Data Service Data Source**. The **Data Service Data Source Properties** dialog box opens.

Specifying a DSN for an ODBC data source

The data source name (DSN), which corresponds to the database associated with the data source, is specified in the DSN dialog box.

1. In the DSN field, enter the data source name. The name that you enter in this field must match your data source name exactly. The name is also case sensitive.
2. Click **Finish**. The new data source definition appears in the **Data Sources** folder.

Specifying a JDBC name and URL

The JDBC name and URL, which define how Java applications connect to a database, are specified in the JDBC Name and URL dialog box.

1. In the JDBC Driver Name field, enter the class name of your JDBC driver.
2. In the JDBC Driver URL field, enter JDBC driver URL.
3. Click **Finish**. The new data source definition appears in the **Data Sources** folder.

IBM SPSS Collaboration and Deployment Services is installed with a set of JDBC drivers for all major database systems. It also provides JDBC drivers for IBM SPSS Statistics file access.

For current supported database information, refer to the software product compatibility reports on the IBM Technical Support site.

If single sign-on is enabled for your system, include the optional `AuthenticationMethod` property in the driver URL string to use single sign-on for the database connection. Valid value for this property include *kerberos* and *ntlm* (NT LAN Manager).

Db2

Class name: *spssoem.jdbc.db2.DB2Driver*

URL string template:

```
jdbc:spssoem:db2://<host>:<port>;DatabaseName=<database>[;AuthenticationMethod=kerberos]
```

Greenplum

Class name: *org.postgresql.Driver*

URL string template:

```
jdbc:postgresql://<host>:<port>/<databasename>
```

Informix

Class name: *spssoem.jdbc.informix.InformixDriver*

URL string template:

```
jdbc:spssoem:informix://<host>:<port>;InformixServer=<Informix server>
```

MS SQL Server

Class name: *spssoem.jdbc.sqlserver.SQLServerDriver*

URL string template:

```
jdbc:spssoem:sqlserver://<host>:<port>;DatabaseName=<database>[;AuthenticationMethod=kerberos|ntlm]
```

MySQL

Class name: *spssoem.jdbc.mysql.MySQLDriver*

URL string template:

```
jdbc:spssoem:mysql://<host>:<port>;DatabaseName=<database>
```

Oracle

Class name: *spssoem.jdbc.oracle.OracleDriver*

URL string template:

```
jdbc:spssoem:oracle://<host>:<port>;SID=<database>[;AuthenticationMethod=kerberos]
```

SAP Hana

Class name: *com.sap.db.jdbc.Driver*

URL string template:

```
jdbc:sap://<host>:<port>
```

Sybase

Class name: *spssoem.jdbc.sybase.SybaseDriver)*

URL string template:

```
jdbc:spssoem:sybase://<host>:<port>;databaseName=<database>[;AuthenticationMethod=kerberos]
```

Teradata

Class name: *com.teradata.jdbc.TeraDriver*

URL string template:

```
jdbc:teradata://<host>/Database=<databasename>
```

IBM SPSS Statistics data files

Class name: *com.spss.statistics.datafile.jdbc.openaccess.OpenAccessDriver*

URL String template:

```
jdbc:spssstatistics://<hostname>:<port>;ServerDatasource=SAVDB;
CustomProperties=(<data_file_identifier>;UserMissingIsNull=<1|0>;MissingDoubleValueAsNAN=<1|0>)
```

Note:

- The *hostname* parameter is the host name or IP address of the computer on which the driver is running. The default value is `localhost`.
- The default is port is 18886.
- Replace *<data_file_identifier>* with a string identifying the data file. This may be specified in a variety of ways. Use `CONNECT_STRING=<path_to_file>` to specify the path to the file, where the path is either a repository URI or a path relative to the host on which the driver is running. The path cannot contain any equals sign or semicolon characters. Alternatively, for data files stored in the repository, use `REPOSITORY_ID=<file_id>;REPOSITORY_VERSION=<file_version>` to specify the repository file and version identifiers. The file version may be specified as a label or as a version marker, with spaces replaced by the standard escape character of `%20`.
- The *UserMissingIsNull* parameter is optional and specifies the treatment of user-defined missing values. `0` indicates that user-defined missing values are read as valid values. `1` indicates user-defined missing values are set to system-missing for numeric variables and blank for string variables. If *UserMissingIsNull* is not specified, it is set to a default value of `1`.
- The *MissingDoubleValueAsNAN* parameter is optional and specifies the treatment of missing numeric values. `0` indicates that user missing values are displayed with the original missing value in the data file. `1` indicates that user missing values are read as not a number (NaN). For JDBC, *UserMissingIsNull* should always be set to `1`.
- When testing a connection that specifies `AuthenticationMethod=ntlm`, use any credential available in the system. The selected credential is not actually passed to the data source due to the authentication method, but the test requires that a credential be specified.

System z data sources

IBM SPSS Collaboration and Deployment Services provides the capability to access the following System z data sources using standard DataDirect Db2 drivers:

- Db2 for z/OS
- Db2 LUW on Linux for System z
- Oracle on Linux for System z

Third-party JDBC drivers

If IBM SPSS Collaboration and Deployment Services does not include a driver for a needed database, you can update your environment to include a third-party driver for the database.

For example, if access to a Netezza or Teradata database is needed, obtain the appropriate driver from the vendor and update your system. To add a JDBC driver to IBM SPSS Deployment Manager:

1. Close IBM SPSS Deployment Manager if it is running.
2. Create a folder named *JDBC* at the root level of the IBM SPSS Deployment Manager installation.
3. Place the driver files in the *JDBC* folder.

Note that for Netezza, the version 5.0 driver should be used to access both version 4.5 and 5.0 databases.

After adding the driver files to your environment, the driver can be used in a data source definition. In the JDBC Name and URL dialog box, type the name and URL for the driver. Consult the vendor documentation for the driver to obtain the correct class name and URL format.

If the data source will be used by the IBM SPSS Collaboration and Deployment Services Repository server, such as in scheduled jobs or reports, the server must also be updated to include the driver files. Contact your administrator to perform the necessary modifications described in the *JDBC drivers* section

of the *Installation and Configuration* chapter in the IBM SPSS Collaboration and Deployment Services installation and configuration documentation.

Specifying a JNDI name for an application server data source

The Java Naming and Directory Interface (JNDI) name identifying the data source is specified in the JNDI dialog box.

1. In the **JNDI Name** field, enter the JNDI name. The name that you enter in this field must match your data source name exactly. The name is also case sensitive.
2. Click **Finish**. The new data source definition appears in the **Data Sources** folder.

Specifying properties for data service data sources

Data service data source definitions are used to retrieve the data required for scoring tasks. The definitions can specify the custom driver class and related properties. The custom driver enables non-standard data sources as input data passed with a scoring request in real time.

For example, when a score is requested for a customer based on credit rating and geocode, the credit score and geocode will use the custom driver retrieving data for the request.

Properties for data service data sources are specified in the **Data Service Data Source Properties** dialog.

To specify data service data source properties:

1. Specify the Driver Class name.
2. If necessary, add additional driver properties.
3. Click **Finish**. The new data source definition appears in the **Data Sources** folder.

For information about creating custom drivers, see the SPSS Collaboration and Deployment Services customization documentation.

Defining tables for data service data sources

Optionally, tables can be defined or imported for both the context data and custom driver options.

To define tables:

1. After the context data or custom driver parameters have been specified, click **Next**. The **Define Tables** dialog box appears.

   Add. Click to define a table. See the topic "Table properties" on page 50 for more information.

   Edit. Click to edit the selected tables. See the topic "Table properties" on page 50 for more information.

   Import. Click to import tables from previously exported data structure definition.

   Export. Click to export data structure definition.
2. After data structure information has been provided, click **OK**.
3. Click **Finish**. The new data source definition appears in the **Data Sources** folder.

Table properties

To define a table for a data service data source:

1. In the Define Tables dialog, click **Add** or select a table definition and click **Edit**. The Table Properties dialog appears.

   Name. Define the table name.

   Add Column. Click to add a column. The Add Column dialog appears. See the topic "Adding columns" on page 51 for more information.

Manage Keys. Click to define keys for the table. See the topic "Managing keys" on page 51 for more information.

Remove Column. Click to remove a selected column.

2. After the table information has been provided, click **OK**. The table appears on in the list in the Define Tables dialog.

Adding columns

To add a column to a table:

1. In the Table Properties dialog, click **Add Column**. The Add Column dialog appears.

   Name. Define the column name.

   Type. Select the column data type from the drop-down list, for example, Boolean, Date, Decimal, String, or Timestamp.

2. After the column information has been provided, click **OK**. The column appears on in the list in the Define Tables dialog.

Managing keys

The Manage Keys dialog box provides options for viewing and defining keys for each table. A **key** is a set of columns that can be used to identify or access a particular row or rows. The key is identified in the description of a table, index, or referential constraint. The same column can be part of more than one key.

Note: A key can be based on more than one column. Rows in two tables match if the values of all columns in the table keys match each other, in the same order.

A unique key is a key that is constrained so that no two of its values are equal. The columns of a unique key cannot contain NULL values. For example, an employee number column can be defined as a unique key, because each value in the column identifies only one employee. No two employees can have the same employee number.

Assigning keys to a table

1. Click **Manage Keys**. The Manage Keys dialog box opens.
2. Click **Add** to define a new key. The Key Properties dialog box displays, allowing you to define or edit an existing key.
3. Enter an appropriate key name in the **Name** field and click **Unique** if the key is unique. This name appears under the list of keys and is the name a consuming application will use to determine the key.
4. Select the appropriate columns from the Available Columns table and click the right-arrow button. The selected column names are displayed in the Key Columns list. Derived columns can be used as keys.

   You can remove existing columns from the Key Columns list by selecting the appropriate columns and clicking the left-arrow button.

5. Click the **Up** and **Down** buttons to rearrange the key column order.
6. Click **OK** to save the key definition. The Key Properties dialog box closes, and the key name is displayed in the Manage Keys dialog.
7. Click **Close**. The selected table displays the number of currently defined keys in the corresponding column.

Modifying data source definitions

After a data source definition has been established, definition properties can be modified. For example, a JDBC driver URL may need to be updated.

Although data source parameters can be modified, the data source definition type cannot be changed. An ODBC data source cannot be converted into a JDBC data source, and vice versa. To change the data source definition type, a new data source definition must be created.

To modify the properties of an existing data source definition:

1. Navigate to the **Data Source Definitions** folder.
2. Double-click on the data source definition to be modified. The dialog box that opens depends on the data source definition type.
   - **ODBC.** The DSN dialog box opens.
   - **JDBC.** The JDBC Name and URL dialog box opens.
3. Make the modifications.
4. Click **Finish**. The changes are applied to the data source definition.

Message domains

Message domains define JMS (Java Messaging Service) queues or topics used by IBM SPSS Collaboration and Deployment Services to communicate with third-party applications.

Messages domains are used to set up message-based schedules and message-based job steps for IBM SPSS Collaboration and Deployment Services jobs. Message domain definitions reference the messaging service which is set up outside IBM SPSS Collaboration and Deployment Services, usually using the JMS infrastructure of the underlying application server.

Write permissions to the Message Domains folder are required to create and modify data credentials. See the topic "Modifying permissions for an existing user or group" on page 26 for more information. Also, the user must be assigned the *Define Message Domains* action through the role.

IBM SPSS Deployment Manager provides the ability to create new message domains and modify and delete domain definitions.

Creating a new message domain

To add a message domain definition:

1. In the Content Explorer, open the *Resource Definitions* folder.
2. Open the *Message Domains* folder.
3. From the File menu, choose:

New > **Message Domain**

The Add New Message Domain wizard opens.

The wizard provides the capability to specify new message domain name and JMS properties.

Message domain name

The first step of adding a new message domain involves the definition of the domain name.

1. In the **Name** field, enter the name for the message domain.

   Note: Currently the only supported message domain type is topic.
2. Click **Next**. The Message Domain Page of the wizard opens.

Message domain properties

1. On Message Domain Properties dialog, specify the following JMS properties:
   - **Destination Name.** The name of the topic or queue.
   - **Credentials.** Optional credentials. The credentials will only be required in certain instances, depending upon the configuration of the JMS server and how we are required to connect to the JMS server (whether the JMS Message topic is secured or not).

   Note: If JBoss JMS messaging service is being used, message domain must be specified in the `topic/<topic name>` format.

2. Click **Finish**. The updated information appears in the Message Domains folder.

Modifying message domain definitions

To modify a message domain definition:

1. In the Content Explorer, open the *Resource Definitions* folder.
2. Open the *Message Domains* folder.
3. Double-click the message domain. The Edit Message Domain dialog opens.
4. Modify message domain properties as necessary.
   - **Destination Name.** The name of the topic or queue.
   - **Credentials.** Optional credentials. The credentials will only be required in certain instances, depending upon the configuration of the JMS server and how we are required to connect to the JMS server (whether the JMS Message topic is secured or not).
5. Click **Finish**. The updated information appears in the Message Domains folder.

Promotion policies

A promotion policy is a resource definition that specifies rules and properties to apply when promoting objects. Use of a promotion policy prevents having to redefine the set of rules each time an object is promoted. Instead, you apply the policy to the promotion request, and the policy rules are automatically enforced. A promotion policy is defined once, but can be applied to any number of promotion requests. Information specified in a promotion policy includes the following:

- promotion timing
- resource definition inclusion
- related resource exclusion

To create or modify a promotion policy, the role for the user must include the *Define Promotion Policies* action. In addition, the user must have write permissions to the *Promotion Policies* subfolder of the *Resource Definitions* folder.

IBM SPSS Deployment Manager provides the ability to add promotion, modify, and delete promotion policy definitions.

Adding promotion policies

To add a promotion policy:

1. In the Content Explorer, open the *Resource Definitions* folder.
2. Open the *Promotion Policies* folder.
3. From the File menu, choose:

New > **Promotion Policy**

The Add Promotion Policy wizard opens.

Promotion timing

Promotion timing determines when the promoted object becomes available in a new IBM SPSS Collaboration and Deployment Services Repository. In immediate promotion, the promoted object is added to the new repository server automatically as part of the promotion request. In this case, the promotion policy must specify the target repository server as well as valid credentials for accessing that server to enable the source and target repository servers to communicate with each other. In contrast, delayed promotion saves the promoted object to a specified file on the file system for subsequent transfer at a later time. In this case, the promotion policy includes no information about the target repository server. The promotion is completed by manually importing the resulting file into the target server.

The Promotion Timing page of the wizard allows you to specify whether the promotion should be immediate or delayed. The following options are available:

- **Immediate.** Select the option to immediately promote the objects to a different repository server.
- **Target Server.** If you selected the Immediate option, select the target repository server.
- **Target Credentials.** If you selected the Immediate option, select the credentials for accessing the target repository server. For information about credentials, see "Adding new credentials" on page 45 In addition, the creation of new items in the target server may require that the role associated with the target credentials include certain actions. See the topic "Promotion considerations" on page 67 for more information.
- **Delayed.** Select the option to save to an export file which can be later imported into the targeted server later.

1. Click **Next**. The Resource Definition Policy page of the wizard opens.

Promotion policy resource definitions

The resource definition portion of a promotion policy specifies how to handle the resource definitions referenced by a promoted object. For example, a job may include steps that refer to server and credential definitions. Some steps may rely on data source definitions. If these definitions do not exist in the target server, the promoted item will not function properly without manual intervention to redefine these properties. By including the necessary resource definitions in the promotion, the promoted item should function in the target server exactly as it does in the source server without any redefinition. However, if you are sure the necessary resource definitions exist in the target server, the definitions from the source server can be omitted from the promotion.

The Resource Definition Policy page of the wizard allow you to specify the rules for resource definitions handling.

1. Select one of the following options:
   - **Recommended .** Promote resource definitions where no identifier or name conflicts exist in the target repository.
   - **Exclude.** Do not promote resource definitions.
   - **Include.** Promote resource definitions.
2. If Include option is selected, you can specify which of the following resource types should be excluded from promotion:
   - Credentials
   - Data sources
   - Message domains
   - Servers
   - Topics
   - Custom properties
   - Notifications
   - Server clusters
   - Promotion policies
3. Click **Finish** to add the promotion policy or click **Next** to proceed to MIME Type Filter page.

MIME type filtering

If an object being promoted uses other IBM SPSS Collaboration and Deployment Services Repository items, the promotion policy should include information about which of those items should be excluded from the promotion. The promotion policy specifies the items to exclude by their MIME types.

The MIME Type Filter page of the wizard allow you to optionally specify the content types of resources used by the promoted item that should be excluded from the promotion.

1. Select the MIME type of objects to be excluded from promotion:
   - To add a file type, click add. The **File Types** dialog box opens. Click the file type entries to select or clear. Use the Ctrl or Shift keys to select multiple entries. After all file types have been selected, click **OK**.
   - To delete a file type, select the file type and click **Delete**.
2. Click **Finish**.

Modifying promotion policies

To modify a promotion policy:

1. In the Content Explorer, open the *Resource Definitions* folder.
2. Open the *Promotion Policies* folder.
3. Right-click a promotion policy and select **Open**.

The Edit Promotion Policy wizard opens. Specify the policy properties such as promotion timing, resource definition handling, and MIME types of used resources to be excluded from promotion.

Deleting promotion policies

To delete a promotion policy:

1. In the Content Explorer, open the *Resource Definitions* folder.
2. Open the *Promotion Policies* folder.
3. Right-click a promotion policy and select **Remove**.

Server definitions

Executing an IBM SPSS Collaboration and Deployment Services Repository resource as a job step requires the specification of an appropriate corresponding server to process the instructions contained in the job step. The connection information for such a server is specified within a server definition.

Server definitions can be classified as either execution servers or repository servers.

- Execution servers process the contents of an IBM SPSS Collaboration and Deployment Services Repository resource. The execution server type must correspond to the resource type being processed. A SAS job step, for example, requires a SAS server definition.
- A repository server corresponds to an IBM SPSS Collaboration and Deployment Services repository installation. A server of this type is typically used by job steps that need to return result artifacts to a repository.

Server definitions are contained in the *Resource Definitions* folder of the Content Explorer. Specifically, they are defined in the *Servers* subfolder.

Adding new server definitions

To add a new server:

1. In the Content Explorer, open the *Resource Definitions* folder.
2. Click the *Servers* folder.
3. From the File menu, choose:

New > **Server Definition**

The Add New Server Definition wizard opens. Alternatively, the new server definition dialog box can be accessed by clicking **New** next to a server field on the General tab for some steps. The process for defining new servers consists of:

1. Naming the server definition and specifying its type. Note that the server types available depend on which product adapters are installed to the repository.
2. Selecting a location in the *Servers* folder for the definition.
3. Specifying parameters for the server that define connection or execution information. The parameter set depends on the server type.

Server definition type

The Server Definition Type page specifies identification information for a server definition.

1. In the Name field, enter the name that you want to assign to your server definition.
2. From the Type drop-down list, select the server type. Note that the server types available depend on which product adapters are installed to the repository.

   Note: If the server is to be used to run IBM SPSS Statistics custom dialogs in IBM SPSS Collaboration and Deployment Services Deployment Portal, then select **Remote IBM SPSS Statistics Server** for the server type.

Click **Next** to define additional parameters.

Server destination

The Server Destination page specifies the location in the IBM SPSS Collaboration and Deployment Services Repository in which the definition is stored.

1. Navigate to the folder, and select it.
2. Click **Next**. The next dialog box that opens depends on the server type that you selected.

Content repository server definition

A Content Repository Server Definition specifies the connection parameters for a IBM SPSS Collaboration and Deployment Services Repository server.

1. In the **Server URL** field, enter the complete connection URL for the server.

   The URL includes the following elements:

   - The connection scheme, or protocol, as either *http* for hypertext transfer protocol or *https* for hypertext transfer protocol with the secure socket layer (SSL)
   - The host server name or IP address

     Note: An IPv6 address must be enclosed in square brackets, such as `[3ffe:2a00:100:7031::1]`.
   - The port number. If the repository server is using the default port (port 80 for http or port 443 for https), the port number is optional.
   - An optional custom context path for the repository server

Table 7. Example URL specifications . This table lists some example URL specifications for server connections.

URL	Scheme	Host	Port	Custom path
`http://myserver`	HTTP	*myserver*	default (80)	(none)
`https://9.30.86.11:443/spss`	HTTPS	9.30.86.11	443	*spss*

Table 7. Example URL specifications . This table lists some example URL specifications for server connections. *(continued)*

URL	Scheme	Host	Port	Custom path
`http://[3ffe:2a00:100:7031::1]:9080/ibm/cds`	HTTP	`3ffe:2a00:100:7031::1`	9080	*ibm/cds*

Contact your system administrator if you are unsure of the URL to use for your server.

2. Click **Finish**.

The new definition is included in the Servers folder.

SAS server parameters

A SAS Server Definition specifies the execution file used for processing SAS job steps.

The execution file must be installed on the same host as the IBM SPSS Collaboration and Deployment Services Repository. Alternatively, a remote process server could be installed on a remote machine to allow IBM SPSS Collaboration and Deployment Services to access an execution file installed on the remote machine.

In the Executable field, specify the full path to the *sas.exe* file to use as the execution server. If the system path includes the location of this file, the full path can be omitted and the default value of *sas* used.

On some systems, spaces in the path for the execution file may cause failures for jobs using the execution server. These problems can be eliminated by specifying the path using 8.3 notation. For example, instead of:

```
C:\Program Files\SAS Institute\SAS\V8\sas.exe
```

Specify:

```
C:\Progra~1\SASINS~1\sas\v8\sas.exe
```

After defining the executable path, click **Finish**. The new definition appears in the *Servers* folder.

Remote process server parameters

A Remote Process Server Definition specifies the connection parameters for a server configured for remote processing. A remote process server allows IBM SPSS Collaboration and Deployment Services to access functionality installed on the remote machine.

1. In the Host field, enter the name of the host where the remote process server resides.
2. In the Port field, enter the port number to be used to connect to the host.
3. To specify that Secure Socket Layer (SSL) is to be used for the server connection, select **This is a secure port**.
4. Click **Finish**. The new definition appears in the *Servers* folder.

Modifying server definitions

To modify a server definition:

1. In the Content Explorer, open the *Resource Definitions* folder.
2. Open the *Servers* folder.
3. Double-click the server to be modified. The Edit Server Definition dialog opens.
4. Modify the server definition parameters as necessary.
5. Click Finish to save your changes.

Server clusters

IBM SPSS Statistics, IBM SPSS Modeler, and remote process servers can be grouped into server clusters to allow load balancing across servers.

When a job step uses a server cluster for execution, IBM SPSS Collaboration and Deployment Services determines which managed server in the cluster is best suited to handle processing requests at that time and routes the request to that server.

The load balancing algorithm for routing server requests uses a weighted least-connection algorithm based on server scores and server loads. When a new connection to a cluster is requested, the system determines a score for each running server in the cluster by using the following formula:

$W_i*C_i/(N_i+1)$

The value of W_i is the weight associated with server *i*. The C_i value is the number of CPUs for server *i*. The value of N_i is the number of current and pending connections for server *i*.

Using the average server load, the system classifies each server as either *available* or *busy*. The new connection is assigned to the *available* server having the highest score. If there are no *available* servers, the connection is assigned to the *busy* server having the highest score. If multiple servers have the same score, the connection is assigned to the one that has the smallest server load.

If two servers in a cluster have the same number of CPUs and server load, the ratio of the number of connections for each will depend entirely on the server weights. For example, if server A has a weight that is twice the weight of server B, server A will handle twice the number of connections. Conversely, if two servers have the same weight and server load, the ratio of the number of connections for each will depend entirely on the number of CPUs for the servers. If server C has eight CPUs and server D has two, server C will handle four times the number of connections.

Notice that the server scores are based on both the number of current and pending connections. If many job steps initiate connections to a server cluster simultaneously, a server in the cluster may not be able to report new connections as current before another connection request is attempted. By including the pending connection count, the scores accurately reflect the impending server load allowing the algorithm to optimize the distribution of requests across all servers in the cluster. A configuration setting defines the time interval during which a connection is classified as pending. For information on modifying this value, consult the administrator documentation.

Server cluster definitions are created in the Content Explorer and stored in the IBM SPSS Collaboration and Deployment Services Repository. Write permissions to the Server Clusters folder are required to create and modify server clusters. See the topic "Modifying permissions for an existing user or group" on page 26 for more information. Also, the user's role must include the *Define Server Clusters* action.

IBM SPSS Deployment Manager provides the ability to create new server clusters and modify and delete cluster definitions.

Important:

Only IBM SPSS Statistics, IBM SPSS Modeler, and remote process rervers that are registered with the IBM SPSS Collaboration and Deployment Services Coordinator of Processes Service can be added to server clusters and managed by the load balancing algorithm. For details, see "Coordinator of Processes configuration" on page 201.

Creating a new server cluster

To add a server cluster definition:

1. In the Content Explorer, open the *Resource Definitions* folder.
2. Select the *Server Clusters* folder.
3. From the File menu, choose:

New > **Server Cluster Definition**

The Add New Server Cluster wizard opens. This wizard provides the capability to specify the cluster name and settings.

Server cluster definition name

The first step of adding a new server cluster involves the definition of the cluster name.

1. In the **Name** field, enter the name for the cluster.
2. Click **Next**. The Server Cluster Settings page of the wizard opens.

Server cluster settings

The settings for a server cluster define the servers included in the cluster and the weight associated with each server.

To add a server to the cluster, click the Add button. The Add Servers to Cluster dialog box appears. See the topic "Add servers to cluster" on page 59 for more information.

To remove a server from the cluster, select the server to be removed in the server list and click the Remove button. To select multiple servers for removal, press the Ctrl key when selecting servers.

To modify the weight for a server, click the weight to be modified and click the ellipsis button that appears in the cell. The Set Server Weight dialog box appears. See the topic "Set server weight" on page 59 for more information.

After defining the cluster settings, click **Finish**.

Add servers to cluster

The Add Servers to Cluster dialog box lists all servers on the network currently registered with IBM SPSS Collaboration and Deployment Services.

Select the server to be added to the cluster and click the OK button. To select multiple servers, press the Ctrl key when selecting servers.

Set server weight

The Set Server Weight dialog box allows the specification of a weight for the server.

Weights must be between 1 and 100, inclusive. Weights represent a relative value for workload capacity. For example, a server with a weight of 10 has 10 times the capacity of a server with a weight of 1. Click OK to apply the weight value to the server and return to the Server Cluster settings page.

Modifying a server cluster

To edit a server cluster definition:

1. In the Content Explorer, open the *Resource Definitions* folder.
2. Select the *Server Clusters* folder.
3. Right-click the cluster to be modified and select Open.

The Edit Server Cluster dialog box opens. This dialog box provides the capability to modify the cluster settings. See the topic "Server cluster settings" on page 59 for more information.

Importing resource definitions

Like other IBM SPSS Collaboration and Deployment Services Repository objects, resource definitions can be exported and imported.

The procedure for exporting resource definitions is similar to the procedure for exporting regular folders. See the topic "Exporting folders" on page 62 for more information. Note that all existing credential, data source, and server information is included in the export file.

Resource definitions in a IBM SPSS Collaboration and Deployment Services export file can be imported "wholesale" or separately. The procedure for importing all resource definitions from an export file is

similar to the procedure for regular folder import. See the topic “Importing folders” on page 63 for more information.

To import resource definitions separately:

1. In the Content Explorer, expand the *Resource Definitions* folder.
2. Right-click on a resource definition and select **Import**. The Import Folder dialog box opens.
3. Navigate to the `.pes` import file and select it.
4. Click **Open**. The progress dialog box appears. When the import is complete, the Import dialog box opens.
5. Click **OK**. The imported folder and its contents appear in the Content Explorer tree.

Note: The import process performs a check of execution server definitions. If one of the definitions is invalid, the entire process fails.

Chapter 7. Export, import, and promotion

Overview

If you maintain multiple IBM SPSS Collaboration and Deployment Services Repository instances, you may have a need to move items from one repository server to another.

For example, you may have a server dedicated to objects under current development. When an object is ready, you add it to a test server to evaluate its performance. If the object performance meets the testing criteria, you add the item to a production server for use within your enterprise. IBM SPSS Collaboration and Deployment Services provides the following ways to transfer repository objects:

- **Export** can be used to save the content of an entire repository folder to a compressed archive file with the extension `.pes`. The archive will include all subfolders, child objects, and any custom properties applied to the folder. It is also possible to include external references, that is, objects outside the exported folder, such as resources definitions, in the export. Objects can be selected for export based on label. When the export file is **imported** into the target repository, the directory structure of the source folder will be re-created at the root of the folder into which it is imported. See "Exporting folders" on page 62 for more information.
- **Promotion** allows you to transfer individual repository objects with their dependent resources using **promotion policies**. See "Promotion" on page 66 for more information.

Note: When you use promotion to transfer objects between repository servers, the servers must all be the same version and have the same adapters installed. Import/export can be used to transfer objects between different versions, but the adapters must be compatible. If you need to move an entire repository between repository servers that are of different versions, follow the migration process described in the repository installation documentation. You can only import from the same or a previous version (you can't import an export from a newer version).

Job components that are migrated during export and import

When importing and exporting jobs, the following components of a job are migrated:

- Versions
- Notifications
- Job schedules

The job history associated with a job is not migrated during the import or export process.

Exporting external references

External references are resources that are defined outside of the folder that is being exported. For example, external references may consist of IBM SPSS Modeler streams and other resources (e.g., syntax files) that defined in different folder (i.e., not in the folder selected for export).

By default, external references are included during the export process. However, the **Include external references** option can be disabled when exporting objects. See the topic "Exporting folders" on page 62 for more information.

It is important to note that selecting the **Include external references** check box has no effect in the following instances because external references are not applicable.

- When exporting from the root of the Content Repository tree
- When exporting the Resource Definitions folder

Export and import restrictions

When exporting and importing, the system enforces certain restrictions.

Export Restrictions

The content of the *Submitted Jobs* folder cannot be exported.

Import Restrictions

- All jobs containing streams must have the streams in the same relative location on the target server unless the streams are included in the `.pes` file.
- The environment of the target IBM SPSS Collaboration and Deployment Services Repository must include any adapters needed for working with the artifacts in the archive being imported. For example, if the archive contains IBM SPSS Modeler server definitions, the target environment must include the IBM SPSS Modeler adapters.
- When importing large archive files, it may be necessary to use a disk mapping implementation for identifier lookup. Use the IBM SPSS Deployment Manager to change the *Resource Transfer Lookup Table* configuration value to *DISK* before attempting the import. For more information, see the Administrator's Guide.

Recommended import order

When importing content into IBM SPSS Collaboration and Deployment Services Repository, it is recommended that the following order be followed:

- Import IBM SPSS Collaboration and Deployment Services Resource Definitions (usually performed once). See the topic "Importing resource definitions" on page 59 for more information.
- Import other objects that will be stored in the IBM SPSS Collaboration and Deployment Services Repository (usually performed multiple times).

Security permissions when importing folders

The following security restrictions apply when importing folders:

- The system accepts the security configuration (permissions) of the parent folder. See the topic "Modifying permissions" on page 26 for more information.
- If a job contains a server-related job step, the system tries to match the server and credential names between the source and target machines. If the server and credential definitions do not match, the import process will fail. See the topic "Adding new credentials" on page 45 for more information.

Exporting folders

The process for exporting files consists of the following tasks:

1. Specifying a location for the exported file.
2. Determining whether to export external references.

To export a folder:

1. In the Content Explorer, right-click on the folder to export, and select **Export**. The **Export** dialog box appears.
2. To select an export destination, click **Browse**. The **Browse for Folder** dialog box opens.
3. Select the location in which you want to store the exported folder. An existing folder can be selected or a new folder can be created.
   - **Using an existing folder.** Navigate to an existing folder within the tree, and select it.

- **Creating a new folder.** To create a new folder, click **Make New Folder**. A new folder is inserted into the tree with the name *New Folder*. Rename this folder.

4. Select one of the following options for including versions of child in the export.
   - **Export all labeled versions.** Include all labeled versions.
   - **Export all versions** Include all versions (labeled and unlabeled).
   - **Export labeled versions** Include versions with the specific labels.
5. If export labeled versions option is selected, specify the labels to be included in export.
   - To add a label to the list of export labels, select an available label and click **Add**.
   - To add all labels to the list of export labels, click **Add All**.
   - To remove a label from the list of export labels, select a label and click **Remove Label**.
   - To remove all labels from the list of export labels, click **Remove All**.
6. Click **Save**. The **Export** dialog reappears.
7. Determine whether to include external references in the exported file. By default, external references are included during the export process. To exclude external references, clear the **Include external references** check box. See the topic "Exporting external references" on page 61 for more information.
8. Click **OK**. The progress dialog box appears. When the export is complete, the **Export** dialog box opens.
9. Click **OK**.

Importing folders

The process for importing files consists of the following tasks:

1. Specifying a path.
2. Determining how the system should resolve import conflicts.

To import a file:

1. In the Content Explorer, right-click on the directory into which you want to import the folder and select **Import**. The Import dialog box opens.
2. Click **Browse** to navigate to the file that you want to import . The Browse for Folder dialog appears.
3. Select the file to import, and click **Open**. The Import dialog returns. The name of the files selected for import appears in the Import File field. The version product version number associated with the imported file appears in the Import file version field. The import file version is determined by the system and is not modifiable.
4. Specify whether the system should resolve import conflicts individually or globally by selecting the applicable option. By default, IBM SPSS Deployment Manager resolves conflicts individually. If global conflict resolution is selected more options are enabled in the Import dialog. Specify the settings for global conflict resolution. See the topic "Resolving import conflicts globally" on page 64 for more information.
5. Click **OK** to launch the import process.
6. If any conflicts arise during the import process, the Import Conflicts dialog appears. See the topic "Resolving import conflicts" on page 64 for more information.
7. If no conflicts occurred, the progress dialog box appears. When the import is complete, the Import complete dialog box opens.
8. Click **OK**. The imported folder and its contents appear in the Content Explorer tree.

Resolving import conflicts

At times, an object that is imported may conflict with an existing object in the repository. For example, the imported object may be a duplicate of an object that is currently in the repository. When conflicts occur, they must be resolved before the import process can continue.

Conflicts can be resolved in a global or individual basis. Global conflict resolution is done by the system. Individual conflict resolution requires a user-specified decision for each conflict.

The method used to resolve conflicts is specified in the Import dialog. By default, the system resolves conflicts individually. Any setting changes made in this dialog persist until they are modified in a future session.

Resolving import conflicts globally

If Resolve conflicts globally is selected in the **Import** dialog box, additional settings must be specified. These settings determine how the system resolves conflicts that arise during the import process. Unlike individual conflict resolution, a list of conflicts prompting user decisions for each conflict will not appear if global conflict resolution is selected. The system will automatically deal with the conflicts that arise based on the criteria specified.

The global resolution options are organized as follows:

Resolve Resource Conflicts. If duplicate IDs or names are encountered during the import process, the following options are available:

- *Keep target item*. The target item will be maintained. The source item with the duplicate ID, which is contained in the *.pes* file, will be ignored.
- *Add new version of target item or rename source item*. This option is typically used to resolve ID conflicts or naming conflicts. If a duplicate ID conflict occurs between the source object and the target object, then a new version of the object will be created in the target location. If a naming conflict occurs, the imported object will be renamed in the target location. Typically, renamed objects are appended with _1, _2, and so forth. In the event that two versions of an object have the same label, the system keeps one label and discards the duplicate label because no two versions of the same item can have the same label. Whether the source label or the target label is maintained depends upon the value specified in the *Use labels from* drop-down field. By default, the label in the source object will be maintained when a conflict occurs. To use labels from the target directory, select **Target** from the *Use labels from* drop-down list.

Lock Resolution. Locked resources can affect the import process. See the topic "Object Locking" on page 21 for more information. When setting import defaults for lock resolution, the following options are available:

- *Continue import, even if some objects cannot be imported due to locking conflicts*. Selecting this option may result in a partial import.
- *Abort import if all objects cannot be imported*. If any conflicts are encountered due to object locks, the import process will be terminated and fail.

Resolve Invalid Version Conflicts. If an invalid version is encountered during the import process, the following options are available:

- *Import*. The invalid version will be imported. This option is the default.
- *Discard*. The invalid version will be deleted.

Resource Definitions. The system imports resource definitions by using one of the following rules:

- **Recommended.** A resource definition is imported only if the identifier or name does not conflict with a target definition. Any resource definitions that have a conflict are not imported.
- **Exclude.** No resource definitions from the import files are imported. Imported objects may need to be modified to reference available resource definitions.
- **Include.** All resource definitions in the import file are imported. You can select one or more resource definition types to exclude from the import by selecting the corresponding check box.

Troubleshooting Import Failures Caused By Duplicate Name Conflicts

Occasionally, the import process may fail due to import configuration settings. If the import failure was caused by naming conflicts, it is possible that changing a setting in the Import dialog may resolve the issue. To resolve an import failure due to naming conflicts:

1. Repeat the import process. See the topic "Importing folders" on page 63 for more information.
2. In the *Resolve Resource Conflicts* section of the Import dialog, select the **Add new version of target item or rename source item** option.

Resolving import conflicts individually

If Resolve conflicts individually was selected in the *Import* dialog and conflicts arise during the process, the Import Conflicts dialog appears.

The process for resolving individual conflicts consists of the following tasks:

1. Specifying conflict resolution defaults.
2. Determining if defaults need to be overridden in the import conflicts table.

Individual conflict resolution defaults

Defaults for the following import conflicts need to be specified:

Resolve Resource Conflicts. If duplicate IDs or names are encountered during the import process, the following options are available:

- *Keep target item*. The target item will be maintained. The source item with the duplicate ID, which is contained in the *.pes* file, will be ignored.
- *Add new version of target item or rename source item*. This option is typically used to resolve ID conflicts or naming conflicts. If a duplicate ID conflict occurs between the source object and the target object, then a new version of the object will be created in the target location. If a naming conflict occurs, the imported object will be renamed in the target location. Typically, renamed objects are appended with _1, _2, and so forth. In the event that two versions of an object have the same label, the system keeps one label and discards the duplicate label because no two versions of the same item can have the same label. Whether the source label or the target label is maintained depends upon the value specified in the *Use labels from* drop-down field. By default, the label in the source object will be maintained when a conflict occurs. To use labels from the target directory, select **Target** from the *Use labels from* drop-down list.

Lock Resolution. Locked resources can affect the import process. See the topic "Object Locking" on page 21 for more information. When setting import defaults for lock resolution, the following options are available:

- *Continue import, even if some objects cannot be imported due to locking conflicts*. Selecting this option may result in a partial import.
- *Abort import if all objects cannot be imported*. If any conflicts are encountered due to object locks, the import process will be terminated and fail.

Resolve Invalid Version Conflicts. If an invalid version is encountered during the import process, the following options are available:

- *Import*. The invalid version will be imported. This option is the default.
- *Discard*. The invalid version will be deleted.

Resource Definitions. The system imports resource definitions by using one of the following rules:

- **Recommended.** A resource definition is imported only if the identifier or name does not conflict with a target definition. Any resource definitions that have a conflict are not imported.
- **Exclude.** No resource definitions from the import files are imported. Imported objects may need to be modified to reference available resource definitions.

- **Include.** All resource definitions in the import file are imported. You can select one or more resource definition types to exclude from the import by selecting the corresponding check box.

Troubleshooting Import Failures Caused By Duplicate Name Conflicts

Occasionally, the import process may fail due to import configuration settings. If the import failure was caused by naming conflicts, it is possible that changing a setting in the Import dialog may resolve the issue. To resolve an import failure due to naming conflicts:

1. Repeat the import process. See the topic "Importing folders" on page 63 for more information.
2. In the *Resolve Resource Conflicts* section of the Import dialog, select the **Add new version of target item or rename source item** option.

Working with the import conflicts table

Individual import conflicts are listed in the Import Conflicts table.

The Import Conflicts table contains the following information. With the exception of the Action column, none of the information in the table can be modified.

Source Path. The source location of the object to be imported.

Hierarchy. The location within the repository hierarchy--for example, folder.

Conflict. The nature of the import conflict. Examples include *Duplicate* and *Invalid version*.

Target. The destination location of the object to be imported. The Destination path field is populated only for duplicates.

Marker. The object identifier of the item. Conflicts are determined by this identifier. For example, the system searches for duplicate object identifiers.

Additional information. Any supplementary information about the import conflict. For example, *Missing IBM SPSS Modeler stream* might appear in this column, if that is the cause of an invalid version.

Action. The *Default* that appears in the Action column refers to the default actions specified in the *Conflict Resolution Defaults* section of this dialog box. In Action column, the default can be overridden on an individual conflict basis. The options that appear vary depending upon the type of conflict. For example:

- Actions for duplicates include *Use global default*, *Keep target item*, and *Use imported item*.
- Actions for invalid versions include *Default*, *Import*, and *Ignore*.

To modify the default Action setting:

1. Highlight the conflict whose default setting you want to modify.
2. Click the ellipsis button next to *Default* in the Action column. A dialog, which is specific to the type of conflict appears.
3. Select the options within the dialog box and click **OK**. The import conflicts table returns. For items whose default actions were modified, *Custom* appears in the Action column.
4. After all default actions have been modified in the table, click **OK** to continue the import process.

Promotion

Promotion provides a policy-based approach to transferring individual objects between IBM SPSS Collaboration and Deployment Services Repository instances. A promotion request requires the specification of the following two items:

- an object to promote
- a promotion policy governing the promotion process

A promotion policy identifies the rules used for a particular promotion. The policy determines which related repository objects, if any, are promoted along with the specified object. For example, when

promoting a job, you often want to include any files referenced by the job steps. In addition, you may to include execution server and credential definitions. The policy dictates how these items are handled.

Promotion is typically used in conjunction with label event notifications. For example, an analyst can set up the *Production* label to notify an administrator when the label is applied to a job so that the job can subsequently be promoted to production status.

Promotion re-creates the directory structure of the source repository server in the target repository server at the root (Content Repository folder) level. This allows location-based references between promoted objects to resolve correctly in the target server.

To submit a promotion request, the role for the user must include the *Promote Objects* action. In addition, the creation of new items in the target server may require that the role associated with the target credentials include certain actions. See the topic "Promotion considerations" on page 67 for more information.

Promoting objects

To promote an object:

1. In the Content Explorer, right-click on an object and select **Promote**. The **Promote Object** dialog box appears.
2. Specify the following information:
   - **Promote Version.** The version of the object to be promoted.
   - **Promote Policy.** The policy to be used for promotion.
3. If selected policy is for a delayed promotion, **Export Location** dialog box is displayed. Click **Browse** and specify the path of the export file.
4. Click **OK**. Depending on the specified policy, the object will be either promoted directly to the target repository or exported. If delayed promotion is performed, you must subsequently import the export file into the target repository. See the topic "Importing folders" on page 63 for more information.

Promotion considerations

For an object, resource definitions, and referenced files to all be promoted successfully, the role for the target server credentials must include all relevant actions for the items included in the promotion request.

The following table identifies the actions needed for a variety of items that may be included in the promotion set.

Table 8. Actions affecting promotion

Included in promotion set	**Action needed for target credentials**
Resource having notifications	*Define and Manage Notifications*
Credentials	*Define Credentials*
Resource having custom properties	*Define Custom Properties*
Data source	*Define Datasources*
Message domain	*Define Message Domains*
Server cluster definition	*Define Server Clusters*
Server definition	*Define Servers*
Resource having topics	*Define Topics*
Job	*Job Edit*
Resource having subscriptions for a different principal	*Create Subscriptions*

Table 8. Actions affecting promotion (continued)	
Included in promotion set	**Action needed for target credentials**
Resource having subscriptions for the target principal	*Manage Subscriptions*
Job having a schedule	*Schedules*
Resource version that is not the latest version in the target server	*Show All Versions*
Latest resource version	*Show LATEST*

In addition, the target credentials must have the *Manage label* permission for any label associated with a promoted object version.

Chapter 8. Analytic data views

An analytic data view defines a structure for accessing data that describes the entities used in predictive models and business rules. The view associates the data structure with physical data sources for the analysis.

Predictive analytics requires data organized in tables with each row corresponding to an entity for which predictions are made. Each column in a table represents a measurable attribute of the entity. Some attributes may be derived by aggregating over the values for another attribute. For example, the rows of a table could represent customers with columns corresponding to the customer name, gender, zip code, and the number of times the customer had a purchase over $500 in the past year. The last column is derived from the customer order history, which is typically stored in one or more related tables.

The predictive analytic process involves using different sets of data throughout the lifecycle of a model. During initial development of a predictive model, you use historic data that often has known outcomes for the event being predicted. To evaluate the model effectiveness and accuracy, you validate a candidate model against different data. After validating the model, you deploy it into production use to generate scores for multiple entities in a batch process or for single entities in a real-time process. If you combine the model with business rules in a decision management process, you use simulated data to validate the results of the combination. However, although the data that is used differs across the model development process stages, each data set must provide the same set of attributes for the model. The attribute set remains constant; the data records being analyzed change.

An analytic data view consists of the following components that address the specialized needs of predictive analytics:

- A data view schema, or data model, that defines a logical interface for accessing data as a set of attributes organized into related tables. Attributes in the model can be derived from other attributes.
- One or more data access plans that provide the data model attributes with physical values. You control the data available to the data model by specifying which data access plan is active for a particular application.

Important:

- Components of the analytic data view are defined by using IBM SPSS Modeler streams. Familiarity with IBM SPSS Modeler concepts and experience creating streams are necessary to work with analytic data views.
- Adapters for IBM SPSS Modeler must be installed for your IBM SPSS Collaboration and Deployment Services Repository to define analytic data view components. For more information about these adapters, see the IBM SPSS Modeler documentation.

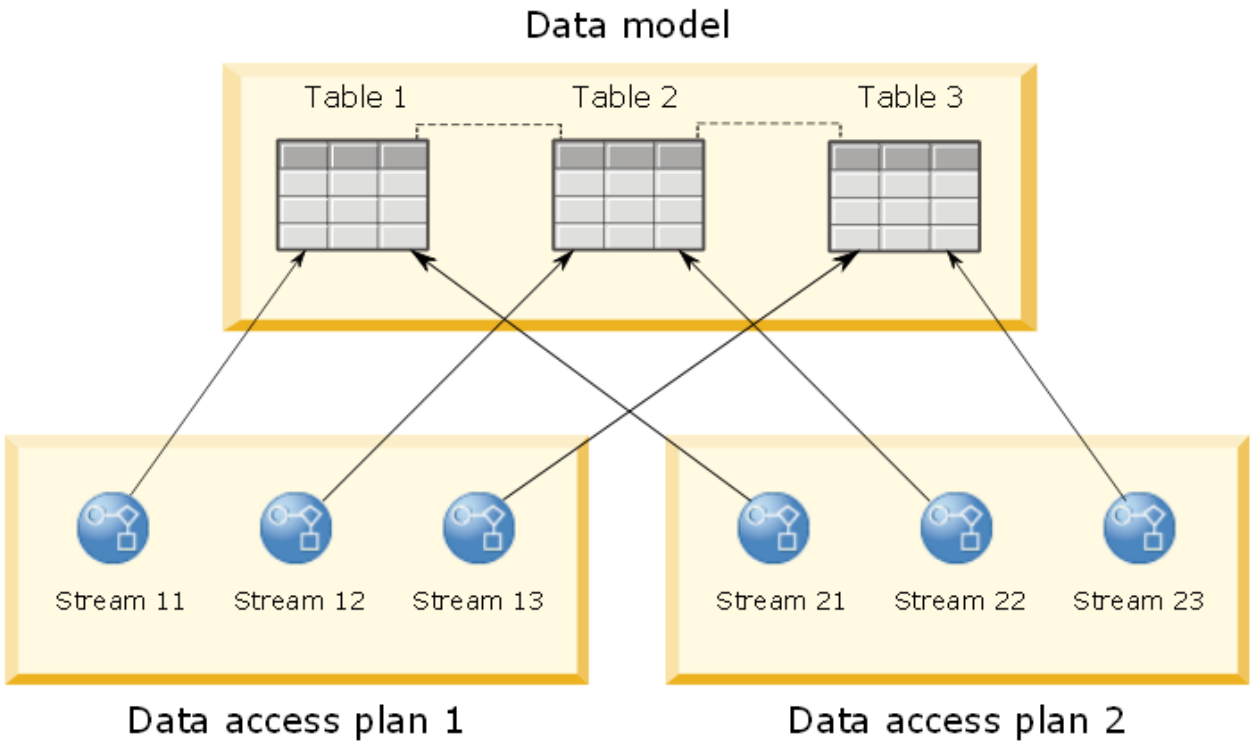

Figure 2. Analytic data view

Figure 2 on page 70 illustrates an analytic data view that contains two data access plans for a data model. The data model includes three tables, with relationships defined between tables 1 and 2 and between tables 2 and 3. Data access plan 1 associates each table with an IBM SPSS Modeler stream. Data access plan 2 associates the data model tables with three different streams. Under data access plan 1, the model retrieves data from terminal nodes in stream 11, stream 12, and stream 13. Under data access plan 2, the model retrieves data from terminal nodes in stream 21, stream 22, and stream 23. By changing the data access plan in use, you can switch the data available to the model.

Creating analytic data views

By creating an analytic data view, you define a container for a data model and its associated data access plans. This structure provides a user-defined interface for accessing data.

Procedure

1. In the **Content Explorer**, select the folder in which to save the analytic data view.
2. From the menus, select **File** > **New** > **Analytic Data View**.
3. Enter a name for the data view.
4. Optional: Select **Lock** to prevent other users from modifying the object settings.

   If you lock the object, you are the only user able to edit the analytic data view until you unlock the object.
5. Click **Finish**.

Results

The analytic data view is created in the selected folder and the analytic data view editor opens.

What to do next

Create a data model and data access plan for the analytic data view.

Data access plans

A data access plan associates the data model tables in an analytic data view with physical data sources. A data source corresponds to a terminal node from an IBM SPSS Modeler stream.

To be used as a source for a data model table, the data structure for the terminal node must correspond to the table structure. The measurement types for the terminal node fields must match the types for the data model tables. For example, if the data model table includes a field for gender that expects values of 0 and 1, the terminal node field that is mapped to this table field must be of the integer type.

A stream that is used as a data source can include most IBM SPSS Modeler nodes before the terminal node. For example, the stream can include multiple source nodes and use a **Merge** node to combine the data. You can use **Field Operations** nodes to create specific data structures.

A data access plan can use the source nodes as they are defined in the stream, or you can override them with alternative settings. The type of the data access plan determines how the source nodes are replaced. The following types are available:

- A batch data access plan specifies new values for source node parameters in the stream. For example, you can specify a different file for a **Variable File** node or a different table for a **Database** node.

  Restriction: A batch plan cannot replace a source node with a source node of another type. For example, you cannot replace a **Database** node with a **Variable File** node.
- A real-time data access plan replaces source nodes with data source definitions stored in the IBM SPSS Collaboration and Deployment Services Repository. For example, you can replace a source node with an application server data source.

An analytic data view typically contains multiple data access plans. You specify the plan to use when you access the analytic data view in your application. For example, you could use a batch plan when training a model and a real-time plan when generating scores based on the trained model. The same stream is used in both instances, but the two plans access different data. You could also create data access plans for an analytic data view that are based on entirely different streams, or on different terminal nodes in the same stream.

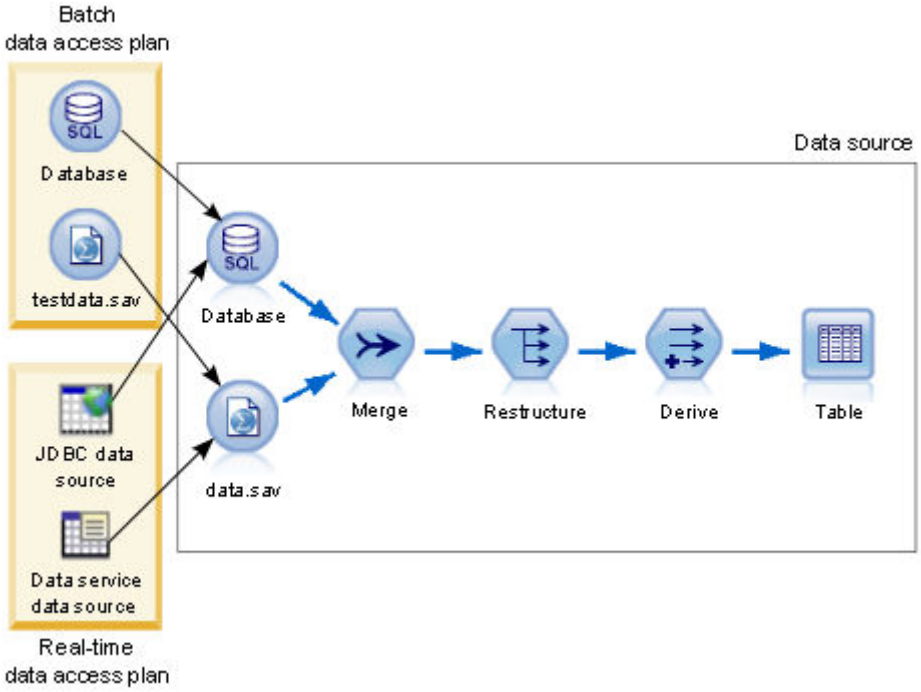

Figure 3. Data access plans for a data source

Figure 3 on page 71 illustrates two different data access plans for a stream that is used as a data source in an analytic data view. The stream reads data by using two source nodes, one for a database and one for the IBM SPSS Statistics data file `data.sav`. A series of subsequent nodes merges the data, restructures the content, and derives a new field. The terminal node presents the results as a table.

Two different data access plans specify settings for the source nodes in the stream. The batch plan uses the database settings as defined in the stream for the database node. For the data file, this plan substitutes a new data file, `testdata.sav`. In contrast, the real-time data access plan replaces the database node with a JDBC data source, and replaces the data file node with a data service data source.

Regardless of which data access plan is in use, the data structure that is defined by the stream must be satisfied. The `testdata.sav` file and the data service data source must both provide fields of the same type as the original `data.sav` file in the stream. Similarly, the JDBC data source must provide fields of the same type as the original database node. By maintaining the fields and their types, you ensure that the merge, restructure, and derive operations can successfully operate on the data to produce a table that contains the required fields of the correct type for an associated data model table.

Notes

- If a terminal node associated with an analytic data view table becomes unavailable, the mapping of the table to physical data is removed from any data access plans that use the terminal node as a data source. For example, if the version of the stream containing the terminal node expires or is removed from the IBM SPSS Collaboration and Deployment Services Repository, any tables associated with the stream need to be associated with a different data source before you can use the analytic data view.
- Refer to the `IBM SPSS Modeler Deployment Guide` for information regarding stream limitations and their use with Data Access Plans.

Creating data access plans

Associate analytic data view tables with physical data sources by using data access plans.

Before you begin

- Store any IBM SPSS Modeler streams to be used as data sources for analytic data view tables in the IBM SPSS Collaboration and Deployment Services Repository.
- From the **Content Explorer**, open the analytic data view to contain the data access plan .

About this task

An analytic data view typically contains multiple data access plans. You can change the data that is being accessed for the analytic data view tables by selecting a different data access plan for the view.

Procedure

1. In the analytic data view editor, select the **Data Access Plans** tab.
2. In the **Plan** field, click the down arrow and select **<Create new data access plan>**.
3. In the **New Data Access Plan** dialog box, enter a unique name for the data access plan in the **Name** field.
4. Select the plan type.
   - A **batch** data access plan uses the source nodes in the streams to access the data. You can use the node settings as specified in the streams or override the settings with new values. For example, if a stream contains a **Statistics File** node that references the file `data.sav`, you can change the referenced file to `testdata.sav` in the data access plan.
   - A **real-time** data access plan allows values to be specified interactively. You can override the source nodes in the data source streams with data providers to access the data. For example, if a stream contains a **Statistics File** node that references the file `data.sav`, you can override the entire node with a JDBC data source in the data access plan.

5. Click **OK**.
6. From the menus, select **File** > **Save** to save a version of the analytic data view that contains these settings.

   Note: Every time that you save the analytic data view, a new version is created. To limit the number of versions that are created, make all necessary changes to the view before saving.

Results

The analytic data view includes a data access plan having the specified name and type.

What to do next

- Add one or more tables to the analytic data view.
- Modify the data mapped to the analytic data view tables by the data access plan.
- Override the source nodes in the mapped IBM SPSS Modeler streams to access different data.

Adding mapped tables to analytic data views

To add a mapped table to an analytic data view, you select the terminal node of an IBM SPSS Modeler stream on which to base the table. The table attributes are generated from the fields available in the terminal node.

Before you begin

- Store any IBM SPSS Modeler streams to be used as data sources for analytic data view tables in the IBM SPSS Collaboration and Deployment Services Repository.
- From the **Content Explorer**, open an analytic data view.

About this task

An analytic data view table corresponds to the terminal node of an IBM SPSS Modeler stream. The node fields determine the attributes and attribute types included in the table.

For example, suppose that the terminal node of a stream has the following fields and storage types:

- CustomerID as an integer
- Gender as a string
- Income as a real

A table that is based on this terminal node has the attributes CustomerID, Gender, and Income. The attribute types correspond to the storage type as defined in the stream.

Procedure

1. In the analytic data view editor, select the **Data Access Plans** tab.
2. Select the data access plan for the mapping of data to the new table.
3. In the **Data View Table** panel, click **New**.
4. In the **New Data View Table** dialog box, click **Browse** and select the IBM SPSS Modeler stream from the IBM SPSS Collaboration and Deployment Services Repository that defines the fields to be mapped to the analytic data view table attributes.
5. Select the label that specifies the version of the selected stream to be used.
6. Select the terminal node that specifies the fields to be used for the data model table.
7. Click **OK**.

   A table with a name corresponding to the terminal node is added to the analytic data view. If the name is not unique, it is automatically modified to prevent conflicts with other tables.

8. Specify the amount of time (in seconds) to cache the table during data retrieval.

   Caching tables can improve overall performance by preventing retrieval of the same data repeatedly.
9. Optional: Right-click the table name in the analytic data view editor and select **Rename** to modify the table name.
10. From the menus, select **File** > **Save** to save a version of the analytic data view that contains these settings.

    Note: Every time that you save the analytic data view, a new version is created. To limit the number of versions that are created, make all necessary changes to the view before saving.

Results

The data model for the analytic data view includes the new table. The table contains attributes that correspond to the fields in the specified IBM SPSS Modeler stream. The physical data that is mapped to the table corresponds to the data from the stream. This data mapping is part of the current data access plan

What to do next

- Add more tables to the analytic data view.
- Define the relationships that exist between the tables in the data model.
- Add derived attributes to the tables.
- Create a data access plan to define other data mappings for the analytic data view tables.
- When the analytic data view contains all needed tables and data access plans, use the view in your application to access the data. For example, you can use the analytic data view as a data provider for scoring configurations. You can also reference the analytic data view by using a Data View node in an IBM SPSS Modeler stream.

Modifying the data mapping for analytic data view tables

You can change the IBM SPSS Modeler stream terminal node mapped to an analytic data view table to modify the data associated with the table or to update the overall table structure.

Before you begin

- Store any IBM SPSS Modeler streams to be used as data sources for analytic data view tables in the IBM SPSS Collaboration and Deployment Services Repository.
- From the **Content Explorer**, open an analytic data view that contains at least one table.

About this task

When you add a mapped table to an analytic data view, the current data access plan automatically associates the table attributes with the data from the selected terminal node. You can change the mapped stream to associate the table with different data. By changing the mapped stream, you can also change the overall data model structure. For example, you can add new attributes to existing tables.

If the analytic data view contains a table that is not mapped to any physical data, you must map the table to a stream terminal node before the analytic data view can be used to access data for attributes of that table.

Procedure

1. In the analytic data view editor, select the **Data Access Plans** tab.
2. Select the data access plan for the new mapping of data to the table.
3. Select the table to be mapped to a different terminal node.
4. Click **Map**.

5. In the **Map Stream to Data View Table** dialog box, click **Browse** and select the IBM SPSS Modeler stream from the IBM SPSS Collaboration and Deployment Services Repository that defines the fields to be mapped to the analytic data view table attributes.
6. Select the label that specifies the version of the selected stream to be used.
7. Select the terminal node that specifies the fields to be used for the data model table.
8. Optional: Click **Clear** to remove the current stream, label, and terminal node settings.

   The table is not mapped to any physical data source if you clear these settings. You will need to map the table to a data source before you can use it in any analyses.
9. Click **OK**.
10. From the menus, select **File** > **Save** to save a version of the analytic data view that contains these settings.

    Note: Every time that you save the analytic data view, a new version is created. To limit the number of versions that are created, make all necessary changes to the view before saving.

Results

The table attributes are mapped to fields from the selected terminal node that have the same name. If an attribute type does not match the type for the new mapped field, the Data View Schema displays a warning icon next to the attribute name to indicate that the mapping requires additional attention. Table attributes that do not correspond to fields in the new terminal node are not automatically mapped to any data in the current data access plan and need to be mapped manually.

What to do next

After mapping a table to a new terminal node, you should review the automatic attribute mappings and modify any as needed manually.

Mapping table attributes to stream fields

Manually map table attributes to the fields available in an IBM SPSS Modeler stream terminal node to manage the data that is associated with the attributes.

Before you begin

From the **Content Explorer**, open an analytic data view that contains at least one table.

About this task

When you map an analytic data view table to an IBM SPSS Modeler stream terminal node, the table attributes are automatically mapped to the node fields by using the attribute and field names. You can manually remap table attributes to associate the attributes with different data.

Procedure

1. In the analytic data view editor, select the **Data Access Plans** tab.
2. Select the data access plan.
3. Select the table containing the attributes to be mapped.
4. Expand the **Field Map** section to view the current attribute mapping.
5. Click the **Data View Attributes** cell for the stream field to be mapped.

   The cell displays a list that includes a blank value, **<Create new attribute>**, and all unmapped data view attributes in the analytic data view.
6. Specify the mapping for the stream field.
   - Select an unmapped attribute to map to the stream field.
   - Select the blank value to unmap the stream field.

- Select **<Create new attribute>** to create a table attribute for the field. Specify a unique name for the attribute in the **New Data View Attribute** dialog box. The attribute type corresponds to the field type.

7. From the menus, select **File** > **Save** to save a version of the analytic data view that contains these settings.

   Note: Every time that you save the analytic data view, a new version is created. To limit the number of versions that are created, make all necessary changes to the view before saving.

Results

The table attributes are mapped to the stream fields as defined in the **Field Map**.

What to do next

After mapping attributes to different fields, preview the data to verify that the correct values are being accessed.

Overriding batch data access plan data sources

You can specify settings for IBM SPSS Modeler stream source nodes that differ from the settings in the stream. The data that is accessed by the plan corresponds to the new source node settings.

Before you begin

From the **Content Explorer**, open an analytic data view that includes a batch data access plan.

Procedure

1. In the analytic data view editor, select the **Data Access Plans** tab.
2. Select the batch data access plan that contains the data source settings to be overridden.
3. Select the data model table to be associated with a different data source.
4. In the **Override Data Sources** section, select the source node to be overridden.

   The Source Nodes list contains every source node included in the data source stream.
5. Specify new parameters for the selected source node. The parameters available depend on the source node type.
6. Optional: Click **Preview Source Node** to view the data that is accessed for the specified parameters.
   a) In the **Modeler Information** dialog box, select the execution server to use to process the data source.

      The list of available servers includes all IBM SPSS Modeler servers that are defined as resources for the current IBM SPSS Collaboration and Deployment Services Repository.
   b) Select the credential to use for accessing the server.

      The list of available credentials includes all credentials that are defined as resources for the current IBM SPSS Collaboration and Deployment Services Repository.
   c) Click **Get Data**.
   d) In the **Preview Result** dialog box, review the data to identify any issues.
7. From the menus, select **File** > **Save** to save a version of the analytic data view that contains these settings.

   Note: Every time that you save the analytic data view, a new version is created. To limit the number of versions that are created, make all necessary changes to the view before saving.

Results

The data access plan contains the updated source node settings.

What to do next

Verify the data access by previewing the data that results from the current plan settings.

Overriding real-time data access plan data sources

For real-time data access plans, you can override the IBM SPSS Modeler stream source nodes with data providers. The data that is accessed by the plan corresponds to the data providers instead of the source node settings in the stream.

Before you begin

From the **Content Explorer**, open an analytic data view that includes a real-time data access plan.

About this task

Important: The fields in the data provider must match the fields defined in the data model. For example, if the data model includes a field named gender that contains string values, the data provider must provide string values for gender.

Procedure

1. In the analytic data view editor, select the **Data Access Plans** tab.
2. Select the real-time data access plan that contains the data source node to be overridden.
3. Select the data model table to be associated with a different data source.
4. In the **Override Data Sources** section, select the source node to be overridden.
   The Source Nodes list contains every source node included in the data source stream.
5. Select the data provider type and specify the data provider settings for the selected source node.
   The settings available depend on the data provider type.

Option	Description
Context Data	For a context data provider, the data set maps to a single table. The values in the table are provided in real-time. For example, when a score for a customer is based on credit rating and geocode, the credit score and geocode values are provided as the context data for the request. For this provider, specify the table to use as context data .
JDBC Data Source	For a JDBC data source provider, the data set maps to a single table or view within a database that is accessed by using JDBC. The following settings must be specified: • **JDBC Connection**. Select from the list of JDBC data sources defined in the system. To create a new JDBC data source, click **New**. • **Credential**. Select the credential for accessing the JDBC data source from the list of credentials defined in the system. To create a new credential, click **New**. • **Table**. Select this option to specify a table from the data provider. • **Query**. Select this option to specify a SQL query that extracts data from the data provider. Either type the SQL statements in the **Edit query** field or click **Load From File** to select a text file containing the SQL statements.
Application Server Data Source	For an application server data source provider, the data set maps to a single SQL table or view within a SQL database. You do not specify a credential when connecting to the data source; the credentials are defined in the application server. The following settings must be specified:

Option	Description
	• **JNDI Connection**. Select from the list of application server data sources defined in the system. To create a new application server data source, click **New**. • **Table**. Select this option to specify a table from the data provider. • **Query**. Select this option to specify a SQL query that extracts data from the data provider. Either type the SQL statements in the **Edit query** field or click **Load From File** to select a text file containing the SQL statements.
Data Service Data Source	For a data services data source provider, the data set maps to a single table that is defined in the data service. Select the source from the list of data service data sources defined in the system. To create a new data service data source, click **New**.

6. To limit the records being retrieved to those meeting specific criteria, select **Filters** to define filter conditions.
   a) Click **Add**.
   b) In the **Filter Information** dialog box, select the column to use for filtering.
   c) Select the filter type. When selecting a reference filter, the source node(s) and the terminal node are listed. Note that using the terminal node could impact performance. If the filtering field is available on a source node, we recommend using that instead.
   d) Specify the table and value to use as the filter criterion.
7. For **Record Count Limit**, specify the maximum number of records to retrieve from the data source.
   If this value is exceeded during data retrieval, an error occurs. If you do not specify a value, the data access plan uses the default configuration value that is defined for your system. For more information, see the Data Service configuration settings available in the browser-based IBM SPSS Deployment Manager.
8. Optional: Click **Preview Source Node** and in the **Preview Source Node Input Data** dialog, specify a value for the filter to view the data that is accessed for the specified data provider.
   a) Click **Get Data**.
   b) In the **Preview Result** dialog box, review the data to identify any issues.
9. Optional: In the **Preview Table** section, click **Preview Data View Table** and in the **Preview Data View Table Input Data** dialog, specify a value for the filter to view the data view schema table.
   a) Click **Get Data**.
   b) In the **Preview Result** dialog box, review the data to identify any issues.
10. From the menus, select **File** > **Save** to save a version of the analytic data view that contains these settings.

    Note: Every time that you save the analytic data view, a new version is created. To limit the number of versions that are created, make all necessary changes to the view before saving.

Results

The data access plan contains the updated data source settings.

What to do next

Verify the data access by previewing the data that results from the current plan settings.

Previewing data access plan data

You can verify that the correct data is being retrieved by a data access plan by previewing the data.

Before you begin

From the **Content Explorer**, open an analytic data view that includes a data access plan.

Procedure

1. Select the data access plan that contains the data source settings to be previewed.
2. Select the data view schema table to be previewed.
3. In the **Preview Table** section, click **Preview Data View Table**.
   a) In the **Modeler Information** dialog box, select the execution server to use to process the data source.

      The list of available servers includes all IBM SPSS Modeler servers that are defined as resources for the current IBM SPSS Collaboration and Deployment Services Repository.
   b) Select the credential to use for accessing the server.

      The list of available credentials includes all credentials that are defined as resources for the current IBM SPSS Collaboration and Deployment Services Repository.
   c) Click **Get Data**.
   d) In the **Preview Result** dialog box, review the data to identify any issues.

Results

Several values for the data model table fields are displayed.

Deleting data access plans

Remove a data access plan from an analytic data view if the set of stream and field mappings are no longer needed.

Before you begin

From the **Content Explorer**, open the analytic data view that contains the data access plan to be deleted.

Procedure

1. In the analytic data view editor, select the **Data Access Plans** tab.
2. In the **Plan** field, click the down arrow and select the data access plan to be deleted as the current data access plan

   Note: An analytic data view must contain at least one data access plan. If the view contains only one access plan, you cannot delete the plan.
3. Click **Delete**.
4. In the warning message dialog box, click **OK**.

Results

The data access plan is removed from the analytic data view.

Data models

The data model for an analytic data view is a schema that defines the interface through which data is processed as a set of one or more tables. Each table in the data model represents a concept or entity

involved in the predictive analytic process. Attributes for the tables correspond to attributes of the entities represented by the tables.

For example, if you are analyzing customer orders, your data model could include a table for customers and a table for orders. The customers table might have attributes for customer identifier, age, gender, marital status, and country of residence. The orders table might have attributes for the order identifier, number of items in the order, total cost, and the identifier for the customer who placed the order. The customer identifier attribute could be used to associate the customers in the customers table with their orders in the orders table.

Table attributes can be derived from attributes in related tables. For example, if you need the number of orders for each customer, you can include an order count attribute in the customers table. The value of this attribute is derived by aggregating over the orders table, counting the number of orders for each customer.

Relationships

A data model includes relationships between the model tables that describe how the entities represented by the tables are related. The relationship cardinality indicates how the rows of a table relate to the rows of another table, and determines how the data in the tables is combined.

The cardinality of a relationship between two tables is defined as one of the following types:

- *One-to-one*. One row of table A corresponds to at most one row of table B. Each row of table B corresponds to at most one row of table A.
- *One-to-many*. One row of table A corresponds to any number of rows of table B. Each row of table B corresponds to at most one row of table A.
- *Many-to-many*. One row of table A corresponds to multiple rows of table B. Each row of table B corresponds to multiple rows of table A.

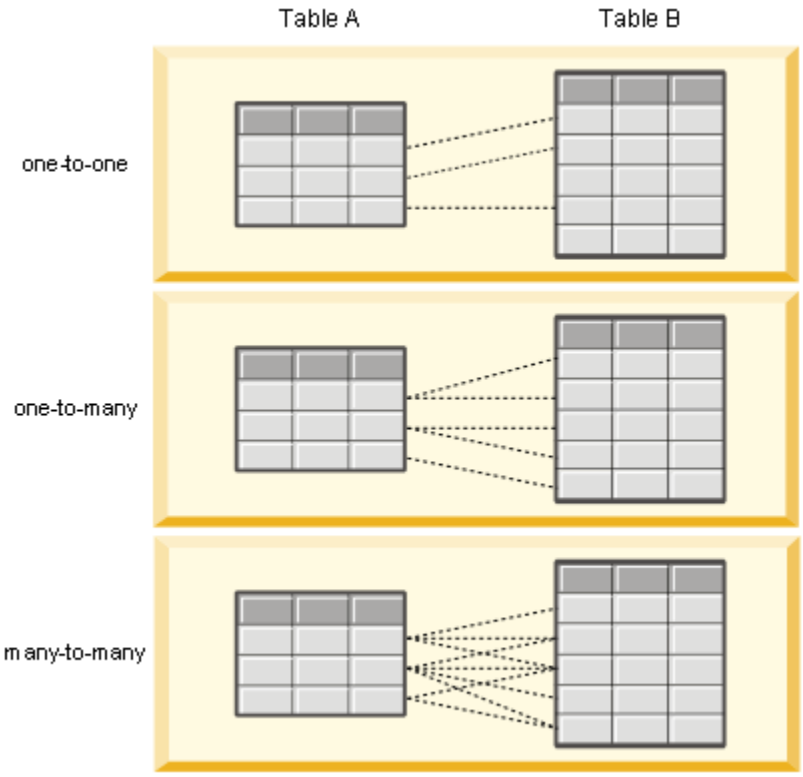

Figure 4. Table relationship types

Figure 4 on page 80 illustrates the three relationship types. In the one-to-one relationship, each row of Table A is connected to one row of Table B. For example, suppose Table A represents reward program

membership and Table B represents customers. A membership identifier in Table A is only associated with a single customer who participates in the program. Some of the customers in Table A do not participate in the rewards program, but the participating customers are associated with one membership.

In the one-to-many relationship, the first row of Table A is connected to two rows of Table B. The second row of Table A is also connected to two rows of Table B. The third row of Table A is connected to the last row of Table B. For Table B, each row is connected to only one row of Table A. For example, suppose Table A represents customers and Table B represents orders. Each customer in Table A can make multiple orders. However, each order in Table B can be associated only with the one customer who placed the order.

In the many-to-many relationship, The first row of Table A is connected to three rows of Table B. The second row of Table A is connected to four rows of Table B. The third row of Table A is connected to two rows of Table B. For Table B, the first row is connected to one row of Table A. The second row of Table B, however, is connected to two rows of Table A. Any row in either table can be connected to any number of rows in the other table. For example, suppose Table A represents orders and Table B represents products. Any order in Table A can include any number of products. Furthermore, any product in Table B can be included in multiple orders.

Note: All table relationships in an analytic data view have a cardinality of one-to-many.

Specifying the analytic data view editor advanced mode

To modify the data model for an analytic data view, you must enable the advanced mode.

About this task

Many analyses require an analytic data view that contains a data model that consists of a single table. If your data model requires a more complex structure, you must enable the advanced mode. In this mode, you can manually define tables and attributes, create derived attributes, and specify table relationships. You can import a data model corresponding to an IBM Operational Decision Manager Business Object Model (BOM), or export a data model as an archive file.

Procedure

1. In the **Content Explorer**, right-click the server name and select **Properties**.
2. In the **Properties** dialog box, open the **User Preferences** section and select **Analytic Data View Editor Advanced Mode**.
3. Select **Advanced Mode** to enable editing of the data model for analytic data views. Clear this option to disable data model editing.
4. Click **OK**.

Results

If you enable the advanced mode, the **Data View Configuration** tab is available when you open the analytic data view editor. If you disable this mode, the tab is not available.

Adding unmapped tables to analytic data views

You can add tables to analytic data views manually if they are needed by applications that access the data view.

Before you begin

- Enable the advanced mode for the analytic data view editor.
- From the **Content Explorer**, open an analytic data view for editing.

About this task

When you define mapped tables, the data model adds attributes to the tables based on the fields available in the associated stream. Alternatively, you can add unmapped tables manually.

Procedure

1. In the analytic data view editor, select the **Data View Configuration** tab.
2. Right-click the area under the **Data View Schema** and select **New** > **New Table**.
3. Right-click the new table and select **Rename** to enter a name for the table.

The following steps involve properties related to business object models. If you do not plan on exporting the data model to a BOM archive, the default values are sufficient.

4. In the Properties panel, define the table verbalization.

   The table verbalization determines the phrases available when creating business rules based on the table.
5. Optional: Select **Static** to make the table available at the class level in the object model.
6. Optional: Select **Final** to ensure that the table cannot be changed in the object model.
7. Optional: Select **Deprecated** to mark the table as deprecated.
8. From the menus, select **File** > **Save** to save a version of the analytic data view that contains these settings.

   Note: Every time that you save the analytic data view, a new version is created. To limit the number of versions that are created, make all necessary changes to the view before saving.

Results

The data model table includes the new table.

What to do next

After adding an unmapped table, add attributes to the table.

Adding unmapped attributes to data model tables

You can add attributes to data model tables manually if they are needed by applications that access the data view.

Before you begin

- Enable the advanced mode for the analytic data view editor.
- From the **Content Explorer**, open an analytic data view for editing.

About this task

When you define mapped tables, the data model adds attributes to the tables based on the fields available in the associated stream. Alternatively, you can add fields to the data model tables manually.

Procedure

1. In the analytic data view editor, select the **Data View Configuration** tab.
2. Select a table in the data view schema.
3. Right-click and select **New** > **New Attribute**.
4. Right-click the new attribute and select **Rename** to enter a name for the attribute.
5. In the Properties panel, specify the type of data associated with the attribute.
   - **String**. Sequences of alphanumeric characters.

- **Integer**. Whole numbers.
- **Real**. Fractional numeric values.
- **Date**. Values representing calendar dates.
- **Time**. Values representing time points.
- **Timestamp**. Values containing both date and time information.

The following steps involve properties related to business object models. If you do not plan on exporting the data model to a BOM archive, the default values are sufficient.

6. Specify the read/write properties for the field.

   The read/write property determines the methods included in the object model and the verbalization associated with the attribute.

   - **Read/Write**. The object model includes a get and a set method for the attribute.
   - **Read Only**. The object model includes only a get method for the attribute.
   - **Write Only**. The object model includes only a set method for the attribute.
7. Define the attribute verbalization.

   The attribute verbalization determines the phrases available when creating business rules based on the attribute.

   - **Read**. The verbalization specifies a navigation phrase for the attribute. The structure of a navigation phrase is typically `{`*`attribute_name`*`} of {this}`, where `{this}` represents the entity in the object model described by the attribute.
   - **Write**. The verbalization specifies an action phrase for the attribute. The structure of an action phrase is typically `set the` *`attribute_name`* `of {this} to {`*`attribute_name`*`}`, where `{this}` represents the entity in the object model described by the attribute.
8. Optional: Select **Static** to make the attribute available at the class level in the object model.
9. Optional: Select **Final** to ensure that the attribute cannot be changed in the object model.
10. Optional: Select **Deprecated** to mark the attribute as deprecated.
11. From the menus, select **File** > **Save** to save a version of the analytic data view that contains these settings.

    Note: Every time that you save the analytic data view, a new version is created. To limit the number of versions that are created, make all necessary changes to the view before saving.

Results

The data model table includes the new attribute.

What to do next

After adding attributes, map them to physical data by using a data access plan.

Defining relationships between analytic data view tables

Specify how the data records in multiple tables are combined by defining relationships between the tables in an analytic data view.

Before you begin

- From the **Content Explorer**, open an analytic data view that contains at least two tables.
- Enable the advanced mode for the analytic data view editor.

About this task

You define a relationship between tables by creating a collection attribute. This attribute specifies a table that contains data on which attributes of the original table can be based.

Procedure

1. In the analytic data view editor, select the **Data View Configuration** tab.
2. Select a table in the data model.
3. Click **Add a new collection attribute to the selected table**.
4. In the **Properties** pane, enter a name for the collection attribute.
5. For the type of collection, select the other data model table to include in the relationship with the current table.
6. Optional: If you plan on exporting the data model to a BOM archive, specify the read/write and verbalization properties.

   For information about specific properties, see “Modifying data model properties” on page 86
7. In the Relationship section, select an attribute in the **From Table** and an attribute in the **To Table** to be linked in the relationship.

   The two selected attributes must be of the same type.
8. Click **Map**.
9. Optional: To remove a mapping between two attributes, select one of the attributes and click **Unmap**.
10. Optional: To remove all mappings between attributes, click **Unmap All**.
11. From the menus, select **File** > **Save** to save a version of the analytic data view that contains these settings.

    Note: Every time that you save the analytic data view, a new version is created. To limit the number of versions that are created, make all necessary changes to the view before saving.

Results

One table contains a collection attribute that references a second table.

What to do next

- Create a derived attribute based on the relationship
- Define relationships for other table pairs in the data model

Adding derived attributes to analytic data view tables

You can add attributes to analytic data view tables that are computed from the values of other table attributes.

Before you begin

- Enable the advanced mode for the analytic data view editor.
- In an analytic data view, create a relationship between the table that will contain the derived attribute and the table containing the attributes on which the derived attribute will be based.

About this task

A derived attribute for a table has values that are computed from attributes in related tables. The definition of a derived attribute has the following structure:

```
<aggregation expression> where <condition expression>
```

The parameters **<aggregation expression>** and **<condition expression>** correspond to verbalization elements from the data model, aggregation operators, and conditional operators.

Procedure

1. In the analytic data view editor, select the **Data View Configuration** tab.
2. Select a table in the data model.
3. Click **Add a new derived attribute to the selected table**.
4. In the **Properties** pane, enter a name for the field.
5. In the Properties panel, specify the type of data associated with the attribute.
   - **String**. Sequences of alphanumeric characters.
   - **Integer**. Whole numbers.
   - **Real**. Fractional numeric values.
   - **Date**. Values representing calendar dates.
   - **Time**. Values representing time points.
   - **Timestamp**. Values containing both date and time information.
6. Define the attribute verbalization.

   The verbalization determines the phrases available when creating business rules based on the attribute. The verbalization is also used as the description of the attribute in some applications that use the analytic data view, such as IBM SPSS Modeler.
7. Specify the definition for the derived attribute.

   Either type the complete expression in the **Definition** field or press the spacebar to interactively build the expression by using context assist.
   a) Press the spacebar and select `<aggr>`.
   b) Select the aggregation function and specify the function arguments.
   c) Optional: To include a conditional expression after the aggregation expression in the definition, select `where <condition>` and construct the conditional expression using the operators and verbalization elements.
8. Optional: Select **Static** to make the table or attribute available at the class level in the object model.
9. Optional: Select **Final** to ensure that the table or attribute cannot be changed in the object model.
10. Optional: Select **Deprecated** to mark the table or attribute as deprecated.
11. From the menus, select **File** > **Save** to save a version of the analytic data view that contains these settings.

    Note: Every time that you save the analytic data view, a new version is created. To limit the number of versions that are created, make all necessary changes to the view before saving.

Results

The analytic data view table includes the new derived field.

Example

Suppose a data model includes a table that contains customer information and a table that contains order information. A derived attribute that indicates the order total for each customer for the third quarter could have the following definition:

```
the total amount of the orders of this customer
where the order date of each order is after 7/1/2013 and before 9/30/2013
```

What to do next

Test the derived attribute by previewing the data associated with the table.

Modifying data model properties

You can export an analytic data view for use in IBM Operational Decision Manager. Properties of the analytic data view determine how the model is processed.

Before you begin

- Enable the advanced mode for the analytic data view editor.
- From the **Content Explorer**, open an analytic data view for editing.

About this task

The analytic data view includes metadata that describes the tables and attributes in the model. The default property values will suffice for many applications. However, if you will be using the analytic data view tables in IBM Operational Decision Manager, you can modify the metadata to control how the model is processed.

Procedure

1. In the analytic data view editor, select the **Data View Configuration** tab.
2. Select either a table or an attribute in the data model and specify its properties.

   The properties available depend on your selection.

Option	Description
Table	For a table, specify the following properties: • **Name**. The name of the table, which corresponds to the class name in the object model. • **Superclass**. The class in the object model from which the current class is derived, if any. • **Interface**. The interface implemented by the class, if any. • **Verbalization**. The table verbalization determines the phrases available when creating business rules based on the table.
Attribute	For an attribute, specify the following properties: • **Name**. The name of the attribute in the object model. • **Type**. The type of data associated with the attribute. – **String**. Sequences of alphanumeric characters. – **Integer**. Whole numbers. – **Real**. Fractional numeric values. – **Date**. Values representing calendar dates. – **Time**. Values representing time points. – **Timestamp**. Values containing both date and time information. • **Read/Write**. The read/write property determines the methods included in the object model and the verbalization associated with the attribute. – **Read/Write**. The object model includes a get and a set method for the attribute. – **Read Only**. The object model includes only a get method for the attribute. – **Write Only**. The object model includes only a set method for the attribute.

Option	Description
	• **Verbalization**. The attribute verbalization determines the phrases available when creating business rules based on the attribute. – **Read**. The verbalization specifies a navigation phrase for the attribute. The structure of a navigation phrase is typically `{attribute_name} of {this}`, where `{this}` represents the entity in the object model described by the attribute. – **Write**. The verbalization specifies an action phrase for the attribute. The structure of an action phrase is typically `set the attribute_name of {this} to {attribute_name}`, where `{this}` represents the entity in the object model described by the attribute.

3. Optional: Select **Static** to make the table or attribute available at the class level in the object model.
4. Optional: Select **Final** to ensure that the table or attribute cannot be changed in the object model.
5. Optional: Select **Deprecated** to mark the table or attribute as deprecated.
6. From the menus, select **File** > **Save** to save a version of the analytic data view that contains these settings.

   Note: Every time that you save the analytic data view, a new version is created. To limit the number of versions that are created, make all necessary changes to the view before saving.

Results

The data model elements have the new properties.

What to do next

After you have specified the properties for all elements of the analytic data view, export the data model as a business object model or execution object model for subsequent use in IBM Operational Decision Manager.

Deleting data model components

Remove tables or attributes from the data model if the items are no longer needed.

Before you begin

- From the **Content Explorer**, open an analytic data view with at least one table or attribute to be deleted.
- Enable the advanced mode for the analytic data view editor.

Procedure

1. In the analytic data view editor, select the **Data View Configuration** tab.
2. Select the item to be deleted in the data model.
3. Right-click and select **Delete**.
4. From the menus, select **File** > **Save** to save a version of the analytic data view that contains these settings.

   Note: Every time that you save the analytic data view, a new version is created. To limit the number of versions that are created, make all necessary changes to the view before saving.

Results

The table or attribute is no longer included in the data model.

Business object models

A data model in an analytic data view has a one-to-one correspondence with a business object model (BOM) in IBM Operational Decision Manager. A BOM specifies the business elements and vocabulary used to define business rules. The vocabulary allows users to create business rules without any knowledge of how the data referenced in the rules is structured.

A business rule has the form "if **`criteria`** then **`outcome`**". For example, a decision on whether or not to reject a loan application based on the credit score of the applicant corresponds to the following business rule:

```
If the credit score of the applicant is less than 300 then reject the application.
```

The applicant is the business entity on which this rule is based. The credit score of the applicant is the entity attribute on which the decision is made. The business object model must include an element representing the applicant and that element must include an attribute corresponding to the credit score. The model also specifies the vocabulary used to reference the model elements in the rules. In this case the vocabulary includes "credit score of the applicant".

Structurally, the business object model is similar to the Java object model. Business entities correspond to a classes and can be grouped into packages. Classes can be nested in other classes. The entity attributes correspond to class attributes that have a type that indicates the type of data values allowed for the attribute. For the example rule, the BOM includes an applicant class that has a credit score attribute of the integer type.

Elements of the data model for an analytic data view correspond to elements in the business object model. A data model table corresponds to a BOM class. A field in the data model table corresponds to an attribute for the class.

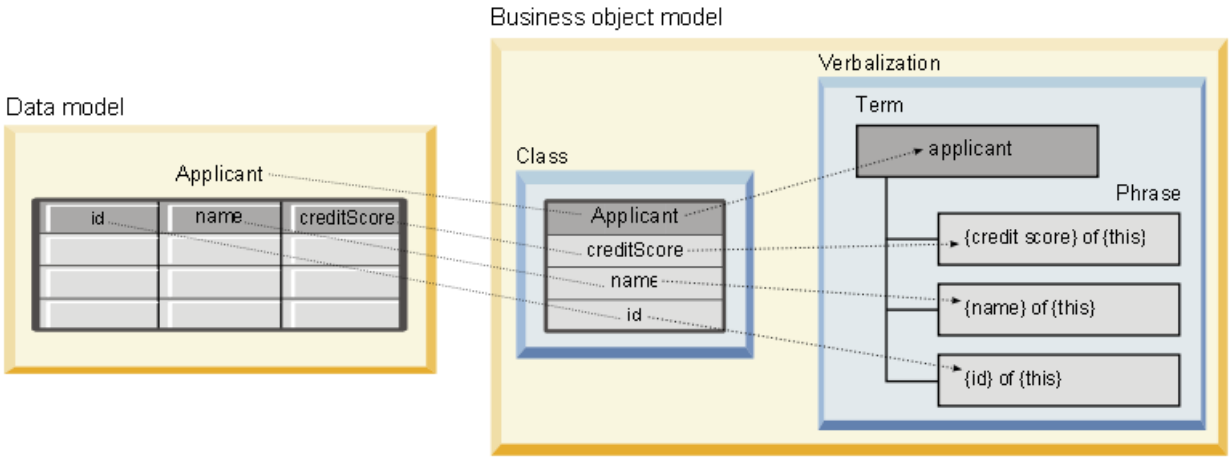

Figure 5. Relationships between the data model and the business object model

Figure 5 on page 88 illustrates the relationships between the data model elements and the business object model elements for a data model that includes a table named Applicant. That table corresponds to a BOM class that has the same name. The verbalization of the class includes the term applicant for the class. The data model table fields of id, name, and creditScore correspond to attributes of the BOM class. Each attribute of the class yields a verbalization phrase of the form `{`**`attribute`**`} of {this}`. The variable `{this}` refers to the parent term applicant in the verbalization.

You can use the correspondence of the data model and the BOM to combine the results from different analytical approaches. The output from predictive models that are based on an analytic data view can be combined with business rules based on a BOM if the data model fields match the BOM elements. The common data interface allows scores from predictive models to be integrated with business rule criteria.

Exporting data models as business object model archives

You can export a data model from an analytic data view as an IBM Operational Decision Manager business object model archive file. The data interface can then be used by any application that can process the business object model.

Before you begin

- Enable the advanced mode for the analytic data view editor.
- From the **Content Explorer**, open an analytic data view that contains a data model with at least one table.

About this task

The elements of a data model exported as a BOM archive can be used in IBM Operational Decision Manager to create business rules. Business rules that are based on this data interface use the same inputs as predictive models that are based on the data model. This correspondence facilitates combining predictive models and business rules.

A BOM archive file contains the following items:

- A `.bom` file that describes the business object model structure
- A `.voc` file that describes the verbalization for the business object model

Note: The archive file does not include a `.b2x` file that maps the business object model to an execution object model. You can create this mapping for your execution object model by using IBM Operational Decision Manager.

Procedure

1. In the analytic data view editor, select the **Data View Configuration** tab.
2. Right-click the data model and select **Export** > **Export BOM**.
3. In the **Save As** dialog box, specify the name and location of the archive file to contain the exported BOM.
4. Click **Save**.

Results

An archive that contains BOM artifacts corresponding to the data model is available at the specified location.

What to do next

Use the archive in IBM Operational Decision Manager to create business rules.

Importing business object models as data models

If the data model for your analytic data view is empty, you can populate the model with the structure defined in an IBM Operational Decision Manager business object model.

Before you begin

- Enable the advanced mode for the analytic data view editor.
- From the **Content Explorer**, open an analytic data view containing an empty data model for editing.

About this task

You can import a BOM archive file to create a data model that uses tables and fields that correspond to the elements in the BOM. Predictive models that are based on this data model use the same elements as business rules based on the BOM. This correspondence facilitates combining predictive models and business rules.

A BOM archive file contains the following items:

- A `.bom` file that describes the business object model structure
- A `.voc` file that describes the verbalization for the business object model

Procedure

1. In the analytic data view editor, select the **Data View Configuration** tab.
2. Right-click the data model and select **Import**.
3. In the **Open** dialog box, select the archive that contains the BOM to be imported.
4. Click **Open**.

Results

The data model includes the tables and fields that are defined in the BOM.

What to do next

Associate the data model tables with IBM SPSS Modeler streams by using a data access plan.

Execution object models

A data model in an analytic data view has a one-to-one correspondence with an execution object model (XOM) in IBM Operational Decision Manager. A XOM specifies the objects against which business rules are executed.

A XOM is derived from XML Schema Definitions. A XOM archive contains an `.xsd` schema file that contains definitions of the field types in the data model. Data model tables correspond to `complexType` entries. Each `complexType` entry contains a sequence of elements that correspond to the fields in that table. The `type` attribute for the elements designates the field type.

For example, suppose the data model for an analytic data view contains two tables. One table represents customers and contains fields for customer identifier, name, gender, birth date, zip code, a collection relationship, and the total of all orders. The other table represents orders and contains fields for order identifier, customer identifier, date, number of items, and order total. The schema file for this model follows:

```
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema"
attributeFormDefault="unqualified" elementFormDefault="unqualified"
targetNamespace="http://" xmlns="http://www.w3.org/2001/XMLSchema"
xmlns:xom="http://">
  <xsd:complexType name="Orders">
    <xsd:sequence>
      <xsd:element name="orderid" type="xom:int" />
      <xsd:element name="custid" type="xom:int" />
      <xsd:element name="date" type="xom:java.sql.Date" />
      <xsd:element name="numitems" type="xom:int" />
      <xsd:element name="total" type="xsd:double" />
    </xsd:sequence>
  </xsd:complexType>
  <xsd:complexType name="Customers">
    <xsd:sequence>
      <xsd:element name="custid" type="xom:int" />
      <xsd:element name="name" type="xsd:string" />
      <xsd:element name="gender" type="xsd:string" />
      <xsd:element name="birthdate" type="xom:java.sql.Date" />
      <xsd:element name="zipcode" type="xom:int" />
      <xsd:element maxOccurs="unbounded" minOccurs="0"
```

```
         name="ordersCollection" type="xom:Orders" />
      <xsd:element name="ordertotal" type="xsd:double" />
    </xsd:sequence>
  </xsd:complexType>
</xsd:schema>
```

The attributes `minOccurs` and `maxOccurs` for the field elements denote the minimum and maximum number of possible occurrences for the fields. If the attributes are not specified, the default value for both attributes is 1 and the field is required by any business rules that are based on the `.xsd` file.

Exporting data models as execution object model archives

You can export a data model from an analytic data view as an IBM Operational Decision Manager execution object model archive file. The data interface can then be used by any application that can process the execution object model.

Before you begin

- Enable the advanced mode for the analytic data view editor.
- From the **Content Explorer**, open an analytic data view that contains a data model with at least one table.

About this task

The elements of a data model exported as a XOM archive can be used in IBM Operational Decision Manager to execute business rules. Business rules that are based on this data interface use the same inputs as predictive models that are based on the data model. This correspondence facilitates combining predictive models and business rules.

Procedure

1. In the analytic data view editor, select the **Data View Configuration** tab.
2. Right-click the data model and select **Export** > **Export XOM**.
3. In the **Save As** dialog box, specify the name and location of the archive file to contain the exported XOM.
4. Click **Save**.

Results

An archive that contains XOM artifacts corresponding to the data model is available at the specified location.

What to do next

Use the archive in IBM Operational Decision Manager to create business rules.

Chapter 9. Scoring

Scoring is the process of generating real-time values by supplying predictive models with input data. A scoring model is any artifact that can be used to produce output values given input data, such as a PMML file from IBM SPSS Statistics.

In general, to use a model for generating scores:

1. Select a model to use for scoring from the IBM SPSS Collaboration and Deployment Services Repository.
2. Define a scoring configuration for the model.
3. Supply the configured model with data and generate scores.

The predictive model used for scoring can be defined using IBM SPSS Modeler streams or PMML that is generated from IBM Corp. products. IBM SPSS Collaboration and Deployment Services supports PMML 4.2 and earlier versions. In addition, legacy markup (e.g. SPSS-ML) from older products may also be used for scoring. Note that some nodes in streams, such as the ADP node, must be trained before they can be used for scoring. For more information, see the IBM SPSS Modeler documentation.

Important: To use a particular model type for scoring, a scoring adapter for that model type must be installed on the IBM SPSS Collaboration and Deployment Services Repository server. For example, to generate scores based on PMML files, the IBM SPSS Collaboration and Deployment Services Scoring Adapter for PMML must be installed. To use IBM SPSS Modeler files for scoring, the IBM SPSS Collaboration and Deployment Services Scoring Adapter for IBM SPSS Modeler must be installed.

Use the IBM SPSS Deployment Manager to define model scoring configurations and to monitor the model scoring performance. Results generated from scoring can be viewed in the IBM SPSS Collaboration and Deployment Services Deployment Portal or in custom client applications.

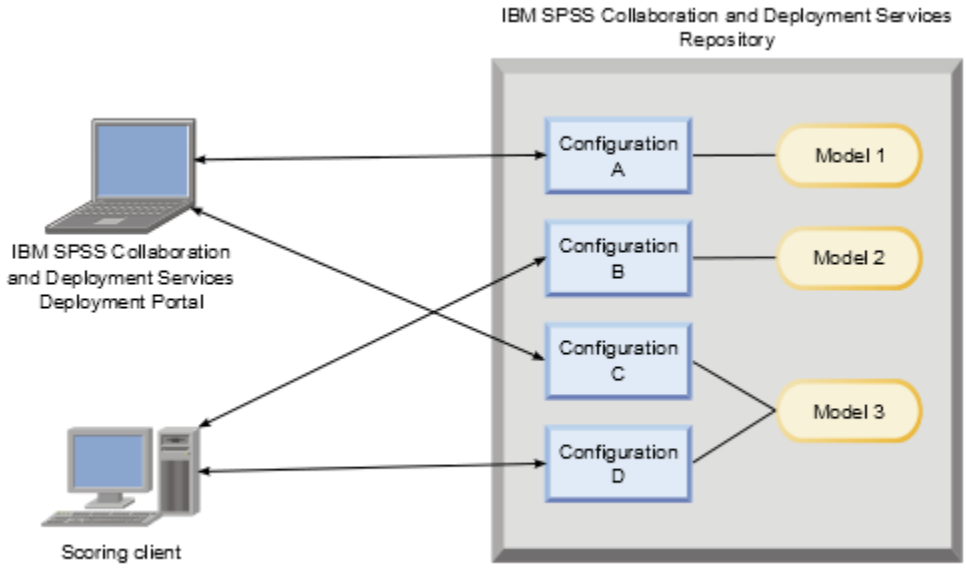

Figure 6. General scoring architecture

Figure 6 on page 93 illustrates the general architecture for scoring. In this example, the IBM SPSS Collaboration and Deployment Services Repository contains three predictive models. Configuration A defines the scoring settings for Model 1. Configuration B defines the scoring settings for Model 2. Configurations C and D define different scoring settings for Model 3.

IBM SPSS Collaboration and Deployment Services Deployment Portal users send separate scoring requests to models 1 and 3 that use the settings in configurations A and C. The scoring service generates the scores and returns the values to the users.

As an alternative to using IBM SPSS Collaboration and Deployment Services Deployment Portal, you can create a custom scoring client to send scoring requests and to receive results. In this example, a custom client sends separate scoring requests to models 2 and 3 that use the settings in configurations B and D. The scoring service generates the scores and returns the values to the client.

Supported scoring functions and models

Applying a predictive model to a set of data can produce a variety of scores, such as predicted values, predicted probabilities, and other values based on that model. The type of score produced is referred to as the **scoring function**. The following scoring functions are available:

Scoring function	Description
PREDICT	*Returns the predicted value of the target variable.*
STDDEV	*Standard deviation.*
PROBABILITY	*Probability associated with a particular category of a target variable.* Applies only to categorical variables. In the absence of the optional third parameter, *category*, this is the probability that the predicted category is the correct one for the target variable. If a particular category is specified, then this is the probability that the specified category is the correct one for the target variable.
CONFIDENCE	*A probability measure associated with the predicted value of a categorical target variable.* Applies only to categorical variables.
NODEID	*The terminal node number.* Applies only to tree models.
CUMHAZARD	*Cumulative hazard value.* Applies only to Cox regression models.
NEIGHBOR	*The ID of the kth nearest neighbor.* Applies only to nearest neighbor models. In the absence of the optional third parameter, k, this is the ID of the nearest neighbor. The ID is the value of the case labels variable, if supplied, and otherwise the case number.
DISTANCE	*The distance to the kth nearest neighbor.* Applies only to nearest neighbor models. In the absence of the optional third parameter, k, this is the distance to the nearest neighbor. Depending on the model, either Euclidean or City Block distance will be used.

The following table lists the set of scoring functions available for each type of model that supports scoring. The function type denoted as `PROBABILITY (category)` refers to specification of a particular category (the optional third parameter) for the `PROBABILITY` function.

Table 9. Supported functions by model type

Model type	Supported functions
Tree (categorical target)	`PREDICT`, `PROBABILITY`, `PROBABILITY (category)`, `CONFIDENCE`, `NODEID`
Tree (scale target)	`PREDICT`, `NODEID`, `STDDEV`
Boosted Tree (C5.0)	`PREDICT`, `CONFIDENCE`
Linear Regression	`PREDICT`, `STDDEV`

Table 9. Supported functions by model type (continued)

Model type	Supported functions
Automatic Linear Models	PREDICT
Binary Logistic Regression	PREDICT, PROBABILITY, PROBABILITY (category), CONFIDENCE
Conditional Logistic Regression	PREDICT
Multinomial Logistic Regression	PREDICT, PROBABILITY, PROBABILITY (category), CONFIDENCE
General Linear Model	PREDICT, STDDEV
Discriminant	PREDICT, PROBABILITY, PROBABILITY (category)
TwoStep Cluster	PREDICT
K-Means Cluster	PREDICT
Kohonen	PREDICT
Neural Net (categorical target)	PREDICT, PROBABILITY, PROBABILITY (category), CONFIDENCE
Neural Net (scale target)	PREDICT
Naive Bayes	PREDICT, PROBABILITY, PROBABILITY (category), CONFIDENCE
Anomaly Detection	PREDICT
Ruleset	PREDICT, CONFIDENCE
Generalized Linear Model (categorical target)	PREDICT, PROBABILITY, PROBABILITY (category), CONFIDENCE
Generalized Linear Model (scale target)	PREDICT, STDDEV
Generalized Linear Mixed Model (categorical target)	PREDICT, PROBABILITY, PROBABILITY (category), CONFIDENCE
Generalized Linear Mixed Model (scale target)	PREDICT
Ordinal Multinomial Regression	PREDICT, PROBABILITY, PROBABILITY (category), CONFIDENCE
Cox Regression	PREDICT, CUMHAZARD
Nearest Neighbor (scale target)	PREDICT, NEIGHBOR, NEIGHBOR(K), DISTANCE, DISTANCE(K)

Table 9. Supported functions by model type (continued)	
Model type	**Supported functions**
Nearest Neighbor (categorical target)	PREDICT, PROBABILITY, PROBABILITY (category), CONFIDENCE,NEIGHBOR, NEIGHBOR(K),DISTANCE, DISTANCE(K)
Bayesian Network	PREDICT, PROBABILITY, PROBABILITY (category), CONFIDENCE
Support Vector Machine (categorical target)	PREDICT, PROBABILITY, PROBABILITY (category), CONFIDENCE
Support Vector Machine (scale target)	PREDICT, STDDEV

- For the Binary Logistic Regression, Multinomial Logistic Regression, and Naive Bayes models, the value returned by the CONFIDENCE function is identical to that returned by the PROBABILITY function.
- For the K-Means model, the value returned by the CONFIDENCE function is the least distance.
- For tree and ruleset models, the confidence can be interpreted as an adjusted probability of the predicted category and is always less than the value given by PROBABILITY. For these models, the confidence value is more reliable than the value given by PROBABILITY.
- For neural network models, the confidence provides a measure of whether the predicted category is much more likely than the second-best predicted category.
- For Ordinal Multinomial Regression and Generalized Linear Model, the PROBABILITY function is supported when the target variable is binary.
- For nearest neighbor models without a target variable, the available functions are NEIGHBOR and DISTANCE.

Scoring configurations

Before a model can be used for scoring, supplemental information must be defined. For example, an IBM SPSS Modeler stream containing a Data View source node needs tables from a data access plan specified to retrieve required data.

This additional information constitutes a scoring configuration for the model, and defines scoring parameters such as:

- Identification information for the configuration itself
- Identification information for the model used for scoring
- The data provider for the input
- Settings for logging
- The order of the input attributes
- Cache size used for scoring models

A single model may be used in a variety of scoring situations that require different scoring parameters. For example, scores may be based on a test data provider for internal purposes and on a different data provider for production usage. Alternatively, the information being logged as result of scoring may depend on the scoring situation. To allow a model to be used in differing scoring circumstances, any model may be associated with multiple scoring configurations.

Scoring configurations can be suspended to temporarily prevent the processing of score requests. A suspended configuration must be reactivated before it can be used to generate scores.

To create a scoring configuration for a model, right-click the model to be configured for scoring in the Content Explorer and select **Configure Scoring**. The Scoring Configuration wizard appears.

Important: If the **Configure Scoring** option is not available for your model type, verify with your system administrator that the scoring adapter for that model type is installed for the IBM SPSS Collaboration and Deployment Services Repository server. When you connect to a repository server that includes scoring adapters, your environment updates to include the scoring functionality.

New scoring model configuration

The Add New Scoring Model Configuration panel of the Scoring Configuration wizard identifies the model being used for scoring.

Name. The name for the scoring configuration. A model may have many scoring configurations, each identified by its unique name.

Model File. The name of the model file associated with the configuration.

Label. The label identifying the version of the model file configured for scoring.

Click **Next** to specify additional settings.

Model specific settings

The Model Specific Settings panel of the Scoring Configuration wizard defines settings parameters for the specific model being configured. The settings parameters available vary across different models.

Specify values for the specific settings offered for the model. Click **Next** to specify additional settings.

Data provider settings

To generate a score, you must provide the predictive model with values for the input fields. This data can be supplied manually by providing a value for each field or retrieved automatically by using a data source.

For example, suppose a model requires values for the input fields age, income, and gender. To generate a score for a customer, you can enter the value for that person for each field manually. Alternatively, you can use an identifier for the customer to automatically retrieve the field values from a customer database.

The **Data Provider Settings** panel of the **Scoring Configuration** wizard identifies the data provider to use for automatic retrieval of input values. When you use a data provider, the provider must be one of the following objects that are defined in the IBM SPSS Collaboration and Deployment Services Repository:

- A data access plan for an Analytic Data View

After you specify the data provider settings, click **Next** to specify more settings.

Analytic data view

Analytic Data View. Select the data view to use from the list of all available analytic data view definitions that are currently specified within the system.

Label. Specify the data view version by selecting the corresponding label from the list of all labels available for the selected analytic data view.

Data Access Plan. Select the data access plan to use from the list of plans available in the specified version of the analytic data view.

Model Input Table. This list identifies the tables that are used by the scoring model. When you use a data access plan, at least one input table must be associated with a table from the plan. A boldface font for a table name indicates that table is associated with a data access plan table. The values for any input tables that are not associated with a plan table must be supplied manually for each score request.

Analytic Data View Table. Defines the table from the data access plan that supplies the values for a selected input table.

Associate an input table with a data access plan table, by performing the following steps:

1. Select the table from the **Model Input Table** list.

2. Select the analytic data view table from the **Table** list.

Repeat these steps for each input table that retrieves data from the data provider.

Input data order

The Input Data Order panel of the Scoring Configuration wizard allows control over the order of the input fields for scoring.

The list displays the input fields organized into tables used by the predictive model. Click the plus sign preceding a table name to reveal the input fields for that table. Click the minus sign preceding a table name to hide the input fields for that table.

If a client scoring request omits field names, the input values being passed are assumed to be in the order defined by this panel. Omitting the field names may streamline client and server communications for configurations involving a high volume of scores. In addition, if input fields are included in the scoring response, the order of the returned fields corresponds to the order defined here. See the topic "Input data returned settings" on page 98 for more information.

To change the order of the input fields within a table, select a field and click the **Up** or **Down** buttons to move the field to the target position. You cannot move a field from one table to another. However, you can reorder the tables by selecting the table name and clicking the **Up** or **Down** buttons.

Click **Next** to specify additional settings.

Input data returned settings

The Input Data Returned Settings panel of the Scoring Configuration wizard allows control over the contents of the scoring response for inputs.

Return request inputs in response. If selected, the scoring response contains values for specified request inputs.

The list displays the input fields organized into tables used by the predictive model. Click the plus sign preceding a table name to reveal the input fields for that table. Click the minus sign preceding a table name to hide the input fields for that table.

The box before each input field and table name indicates whether or not the field or table are included in the response. If the box contains a check mark, the field is included. To add a field to the scoring response, click the empty box before the field name. Click the empty box before a table to add all fields in the table to the response. Clicking a box that contains a check mark removes the corresponding field or table from the response.

Click **Next** to specify additional settings.

Output data returned settings

The Output Data Returned Settings panel of the Scoring Configuration wizard allows control over the contents of the scoring response for outputs.

Scoring requests can produce a variety of scoring outputs, with the type of output being dependent on the model used for scoring. The scoring configuration allows the specification of a subset of outputs to be returned in response to a scoring request. For models that offer many scoring outputs, restricting the returned values to a small set can help to optimize scoring performance.

The model outputs that can be included in the response appear in the list. For information about specific items for PMML files, see "Supported scoring functions and models" on page 94.

Select the outputs to include for this particular scoring configuration. At least one output item must be selected. Click **Next** to specify additional settings.

Logging settings

The Logging Settings panel of the Scoring Configuration wizard defines the content of audit logs for the scoring process.

To capture logging information, select **Enable Logging**. Measures that can be included in the logs for scoring appears in the list.

Request Inputs. The values of the input fields used for scoring.

Context Data. The values of the context fields used for scoring.

Model Outputs. Output values from the model. The list of available outputs depends on the type of model being configured for scoring. See the topic "Supported scoring functions and models" on page 94 for more information.

Important:

- For inputs to be logged, logging must be explicitly enabled for each input, including request inputs, and context inputs. The default is to not log input values.
- Request inputs are logged separately from context inputs. A request input can have the same name as a context input, but they will appear separately in the log.
- If the user does not explicitly provide the input and the data service does not calculate a value for the input, it will be logged as a null value (empty tag).
- The data service will not provide the input unless required by the model. If logging is enabled for an input that is not required, a null value will be logged.

Scoring Engine Metrics. Measurements that reflect the performance of the scoring configuration. Available metrics include:

- **Score Elapsed Time.** Amount of time, measured in milliseconds, between the request for a score and the generation of the score.
- **Score Data Initialization Time.** Amount of time, measured in milliseconds, that the data service takes to be initialized for a score request.
- **Score Data Access Time.** Amount of time, measured in milliseconds, that the data service takes to access its data.
- **Score Computation Wait Time.** Amount of time, measured in milliseconds, that the score provider worker spends waiting for the data service.
- **Score Computation Time.** Amount of time, measured in milliseconds, that the score provider worker spends computing the score.
- **Average Latency.** Average amount of time, measured in milliseconds, between the request for a score and the generation of the score.
- **Minimum Latency.** Smallest amount of time, measured in milliseconds, between the request for a score and the generation of the score.
- **Maximum Latency.** Largest amount of time, measured in milliseconds, between the request for a score and the generation of the score.
- **Score Data Initialization Time.** Amount of time, measured in milliseconds, that the data service takes to be initialized for a score request.
- **Average Data Initialization Time.** Average amount of time, measured in milliseconds, that the data service takes to be initialized for a score request.
- **Minimum Data Initialization Time.** Smallest amount of time, measured in milliseconds, that the data service takes to be initialized for a score request.
- **Maximum Data Initialization Time.** Largest amount of time, measured in milliseconds, that the data service takes to be initialized for a score request.
- **Score Data Access Time.** Amount of time, measured in milliseconds, that the data service takes to access the data.

- **Average Data Access Time.** Average amount of time, measured in milliseconds, that the data service takes to access the data.
- **Minimum Data Access Time.** Smallest amount of time, measured in milliseconds, that the data service takes to access the data.
- **Maximum Data Access Time.** Largest amount of time, measured in milliseconds, that the data service takes to access the data.
- **Score Computation Wait Time.** Amount of time, measured in milliseconds, that the score provider worker spends waiting for the data service.
- **Average Computation Wait Time.** Average amount of time, measured in milliseconds, that the score provider worker spends waiting for the data service.
- **Minimum Computation Wait Time.** Smallest amount of time, measured in milliseconds, that the score provider worker spends waiting for the data service.
- **Maximum Computation Wait Time.** Largest amount of time, measured in milliseconds, that the score provider worker spends waiting for the data service.
- **Score Computation Time.** Amount of time, measured in milliseconds, that the score provider worker spends computing the score.
- **Average Computation Time.** Average amount of time, measured in milliseconds, that the score provider worker spends computing the score.
- **Minimum Computation Time.** Smallest amount of time, measured in milliseconds, that the score provider worker spends computing the score.
- **Maximum Computation Time.** Largest amount of time, measured in milliseconds, that the score provider worker spends computing the score.
- **Average Log Serialization Time.** Average amount of time, measured in milliseconds, to create a log entry in XML format.
- **Minimum Log Serialization Time.** Smallest amount of time, measured in milliseconds, to create a log entry in XML format.
- **Maximum Log Serialization Time.** Largest amount of time, measured in milliseconds, to create a log entry in XML format.
- **Average Log Queue Time.** Average amount of time, measured in milliseconds, to place the XML log data on to the JMS queue.
- **Minimum Log Queue Time.** Smallest amount of time, measured in milliseconds, to place the XML log data on to the JMS queue.
- **Maximum Log Queue Time.** Largest amount of time, measured in milliseconds, to place the XML log data on to the JMS queue.
- **Configuration Scores.** Total number of scores produced by a specific scoring configuration.
- **Score Elapsed Time.** Amount of time, measured in milliseconds, since the previous score generation.
- **Configuration Uptime.** Amount of time, measured in seconds, that the scoring configuration has been available for scoring.
- **Cache Hits.** Number of successful attempts to retrieve data from the memory cache for a scoring configuration.
- **Cache Misses.** Number of failed attempts to retrieve data from the memory cache for a scoring configuration. Each failed attempt results in a new service call to retrieve the necessary data.

Scoring Engine Properties. Characteristics of the scoring configuration itself. Entries include:

- **Model Path.** IBM SPSS Collaboration and Deployment Services Repository path for the model file associated with the configuration.
- **Model MIME type.** MIME type of the model file associated with the configuration.
- **Analytic Data View Data Access Plan.** The name of data access plan that is used when accessing the analytic data view.

- **Model Version Marker.** Marker identifying the version of the model file associated with the configuration.
- **Model Version Label.** Label identifying the version of the model file associated with the configuration.
- **Scoring Configuration Name.** The name of the scoring configuration.
- **Model ID.** IBM SPSS Collaboration and Deployment Services Repository identifier for the model file associated with the configuration.
- **Scoring Configuration Serial.** Unique identifier of the scoring request for the configuration.

Select the items to include in the log for this particular scoring configuration. For each score request, values for all selected items will be entered into the scoring log.

Click **Next** to specify additional settings.

Advanced settings

The Advanced Settings panel of the Scoring Configuration wizard offers optional settings used to optimize the scoring process.

Usable in Batch Scoring. If selected, the scoring model configuration can be used in batch scoring requests.

Model Cache Size. The number of concurrent scoring requests for the configuration. To optimize performance, keep the cache size as low as possible.

Log Destination. By default, IBM SPSS Collaboration and Deployment Services uses the Java Message Service (JMS) for logging scoring information to a queue. If your environment is configured to use a custom message driven bean for logging, specify the log destination for that bean. Contact your administrator for the correct destination for your environment.

Click **Finish** to create a scoring configuration with the specified settings.

Scoring configuration aliases

A scoring configuration alias provides a fixed name to which scoring clients can send scoring requests. The alias passes the requests to an assigned scoring configuration. The scoring configuration that is assigned to the alias can be changed at any time, changing the model that is used for scoring.

Client applications reference the alias name for scoring requests instead of referencing a scoring configuration name. When you assign a different configuration to the alias, subsequent requests are routed to the new configuration. As a result, you can change the configuration that is used for scoring without modifying your client application in any way.

Consider the environment that is illustrated in Figure 7 on page 102. The IBM SPSS Collaboration and Deployment Services Repository contains two scoring models. *Model 1* is configured for scoring according to the settings that are specified in *Configuration A*. *Model 2* is configured for scoring according to the settings that are specified in *Configuration B*. *Configuration A* is assigned to a configuration alias. The scoring client sends a scoring request to the scoring configuration alias, which sends the request to *Configuration A*. *Model 1* generates the score and sends the result back to the client.

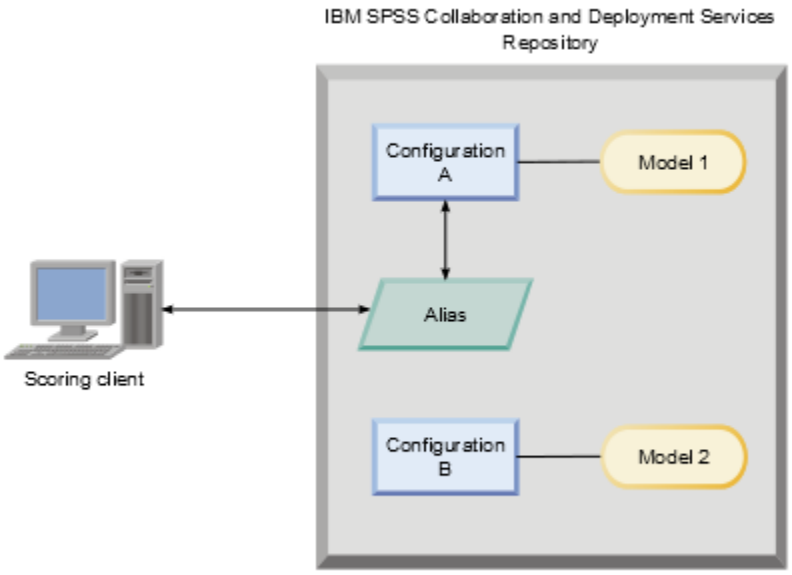

Figure 7. Example alias for scoring configurations

Suppose that you change the configuration that is assigned to the alias to *Configuration B*. In this case, the scoring client still sends the scoring request to the same alias, but the request is then sent to *Configuration B*. *Model 2* generates the score, which is then returned to the client. No change was made to the client, but an entirely different model produced the score. If you did not use the alias, the client would need to be changed to have *Model 2* used for scoring.

Scoring configuration aliases are typically used in A/B testing of scoring models, when you need the ability to replace a production model with a better-performing model without impacting any current production processes.

Creating scoring configuration aliases

Create an alias for a scoring configuration if you need the ability to change the configuration that is used by scoring clients without modifying the clients.

Before you begin

Open the Scoring view to access all available scoring configurations and configuration aliases. There must be at least one scoring configuration in the system to create an alias.

About this task

A configuration alias provides an extra layer between a scoring client and a scoring configuration. You can change the configuration that handles scoring requests without needing to update the client.

Procedure

1. Select a scoring configuration in the Scoring view.
2. Click **Create Alias**.
3. In the **Scoring Configuration Alias** dialog box, enter a name for the alias.
   Scoring tasks reference this name when requests are submitted for scoring.
4. Select the scoring configuration to assign to this alias.
   This configuration processes all scoring requests that are submitted to the alias.

Restriction: An alias cannot reference another alias. Aliases can reference scoring configurations only.

5. Click **OK**.

Results

The Scoring view updates to include an entry for the alias. The configuration alias is available for scoring requests. The alias passes all submitted scoring requests to the specified scoring configuration.

Editing scoring configuration aliases

Change the scoring configuration that is referenced by an alias to use a different configuration for generating scores for requests that are submitted to the alias.

Before you begin

Open the Scoring view to access all available scoring configurations and configuration aliases.

About this task

When you use a configuration alias to route scoring requests, you can change the scoring configuration that is used to generate scores by modifying the alias. This approach is typically done when a new scoring model outperforms another model. Update the alias to reference a scoring configuration for the new model to have all subsequent scoring requests that are submitted to the alias passed to the new scoring configuration.

When you assign a new configuration to an alias, the model that is used by the new configuration must be compatible with the model that is used by the old configuration. The inputs for the new model must be a subset of the inputs for the old model. The new model outputs must be a superset of the old model outputs. In addition, the input field types must match between the two models. For example, if the old model uses the fields *age*, *gender*, and *income* to produce predicted target values, probabilities, and confidence values, the new model must use at least one of the same input fields to produce output that contains these three prediction values at a minimum. If *gender* has a type of string in the old model, it must also be a string type in the new model. These compatibility restrictions ensure that the input from the client score request can be processed by both models and the results can be handled properly by the client. If the models are not compatible, changes to the scoring client are required when a new configuration is assigned to an alias.

Procedure

1. Select a scoring configuration alias in the Scoring view.
2. Click **Edit Alias**.
3. Select a new scoring configuration to assign to the alias.
4. Click **OK**.

Results

The Scoring view updates to show the new configuration that is assigned to the alias. The modified alias is available for scoring requests.

Deleting scoring configuration aliases

Delete a configuration alias if it is no longer needed for scoring tasks.

Before you begin

Open the Scoring view to access all available scoring configurations and configuration aliases.

About this task

If an alias is no longer needed, delete the alias from the system. The scoring configuration that is used by an alias is not deleted when the alias is deleted. If the configuration is no longer needed, delete the configuration separately.

Important: Verify that no client applications are currently using a configuration alias before you delete the alias. If you delete an alias that is used by a client application, that application can no longer generate scores.

Procedure

1. Select the alias in the Scoring view.
2. Click **Delete Configuration** or press the Delete key.

Results

The alias is removed from the system.

Scoring configuration association sets

A scoring configuration association set consists of a primary scoring configuration and one or more alternative configurations for generating scores. Scoring requests that are submitted to the primary configuration are assigned to one of the configurations in the set according to specified distribution percentages.

Use a scoring configuration association set to compare the performance of one or more models with a current production model. Scoring requests that are submitted to the production model are automatically assigned to one of the alternative models for processing. Over time, you can compare the scores that are generated by all of the models in the set to determine which one performs best.

You can stop testing of all alternative models by deleting the association set. When the set is deleted, the primary configuration handles all scoring requests. Alternatively, you can stop testing individual alternative models by removing them from the set and adjusting the distribution percentages.

Consider the environment that is illustrated in Figure 8 on page 105. The IBM SPSS Collaboration and Deployment Services Repository contains an association set that includes three scoring configurations. *Configuration B* is the primary configuration and contains the scoring settings for *Model 2*. *Configuration A* and *Configuration C* define the scoring settings for *Model 1* and *Model 3*. The association set sends 15% of score requests to *Configuration A* and 15% of score requests to *Configuration C*. The remaining 70% or score requests are processed by *Configuration B*.

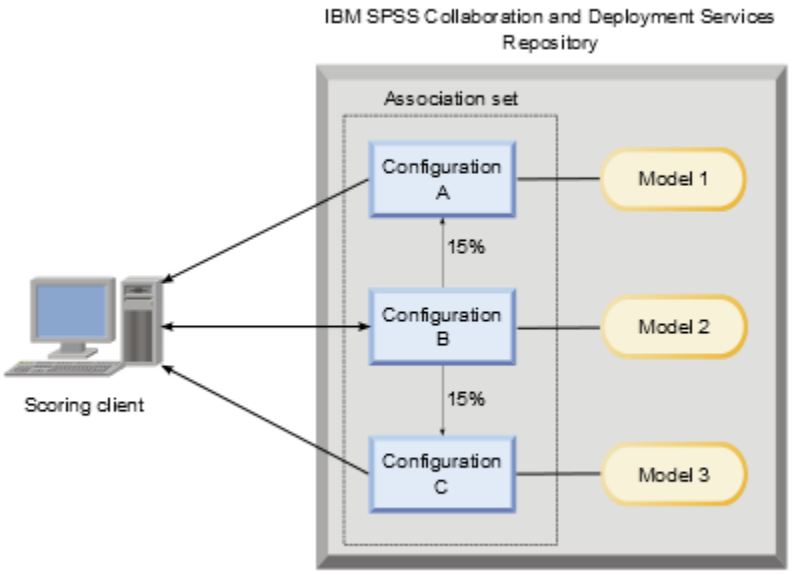

Figure 8. Example association set for scoring configurations

When a scoring client sends a scoring request to *Configuration B*, it is routed to one of the configurations in the set. That configuration generates the score and returns the result to the client.

If the configuration that is assigned to a scoring request is deleted from the system, is not running, or is experiencing an error that makes it unavailable, the request is rerouted to the primary configuration for the association set.

Configuration compatibility

To be included in an association set, a scoring configuration must be compatible with the primary configuration for the set. Compatibility ensures that score requests sent to the primary configuration can be processed by any associated configurations. In addition, the outputs that are returned by associated configurations can be processed by the client application that sends the score request.

To be compatible, the following criteria must be satisfied:

- All input table names in an associated configuration must have corresponding table names in the primary configuration
- The inputs for an associated scoring configuration must be a subset of the primary configuration inputs
- The outputs for an associated scoring configuration must be a superset of the primary configuration outputs

For the associated scoring configuration inputs to be a subset of the primary configuration inputs, the following criteria must be satisfied:

- Names of fields that are required in the input tables for an associated configuration must be included in the corresponding tables of the primary configuration. The fields must be required in the primary configuration.
- Names of fields that are optional in the input tables for an associated configuration do not need to be included in the tables of the primary configuration. However, if those fields are included in the primary configuration, they can be either required or optional.
- The types for the input fields in the associated configuration must match the types for the input fields in the primary configuration

Figure 9 on page 106 illustrates both compatible and incompatible scoring configurations for a primary configuration.

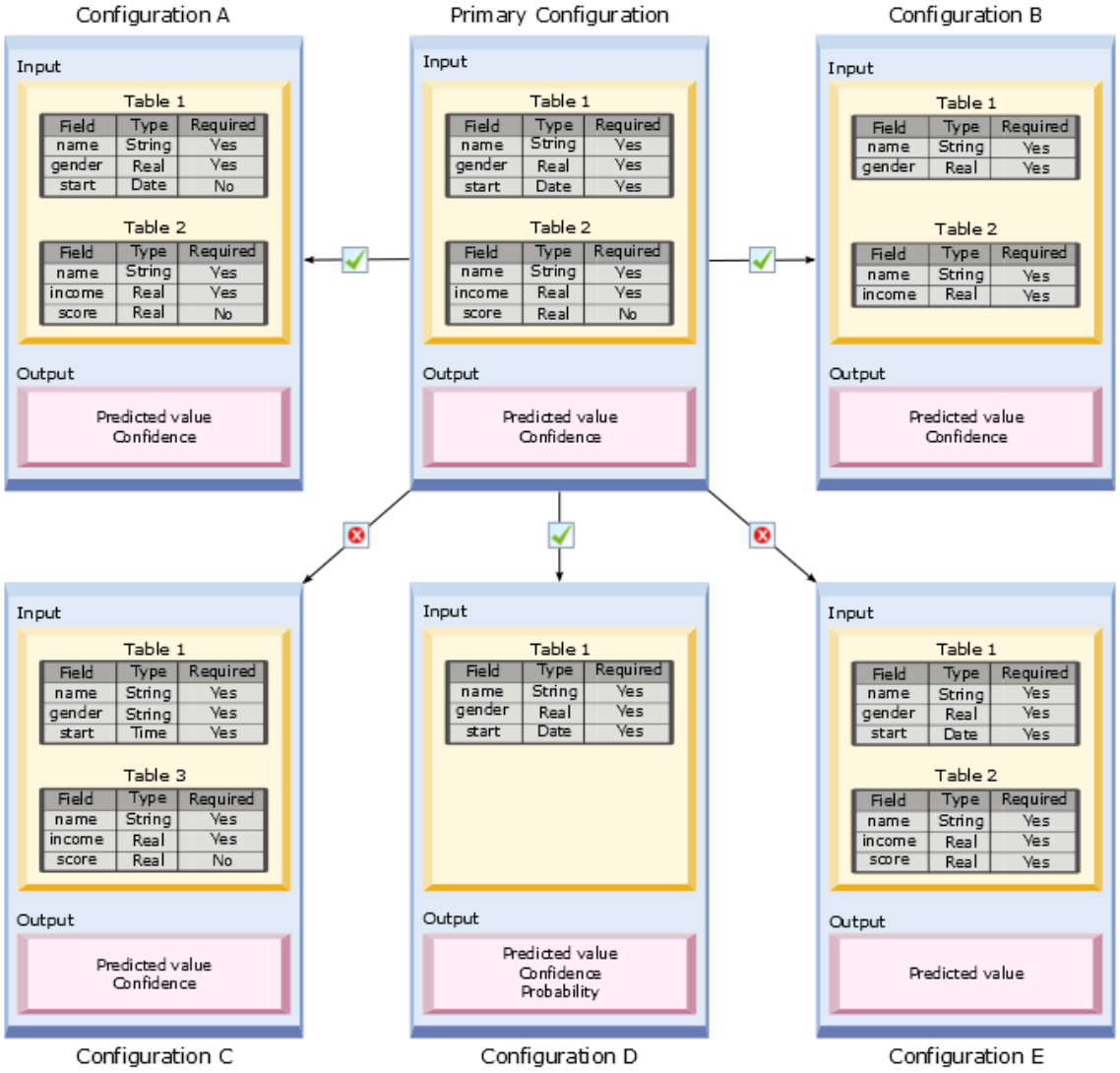

Figure 9. Scoring configuration compatibility

Configuration A and *Configuration B* are both compatible with the primary configuration. These configurations have two tables that have the same names as the tables in the primary configuration. All of the fields in the configurations have the same names and types as the fields in the primary configuration. All of the fields that are required for the associated configurations are also required in the primary configuration. Finally, all of the configurations return the same set of outputs.

Configuration D is also compatible with the primary configuration. The one input table in *Configuration D* has a corresponding table in the primary configuration. All of the fields in the associated configuration have the same names and types as the fields in *Table 1* of the primary configuration. The outputs of *Configuration D* are also a superset of the outputs for the primary configuration.

In contrast, *Configuration C* is not compatible with the primary configuration for two reasons. First, the primary configuration does not include a table that is named *Table 3*. A score request that is sent to the primary configuration would not include any data for this table. Second, the *start* field is of the `Time` type for *Configuration C* but has a type of `Date` in the primary configuration. The data for *start* is of different types in the two configurations.

Configuration E is also not compatible with the primary configuration for two reasons. First, the *score* field is required for the associated configuration but it is optional for the primary configuration. So, a score request that is sent to the primary configuration might not include data that is required for the associated configuration. Second, *Configuration E* returns only the predicted value while the primary configuration returns both the predicted value and the confidence. The associated configuration outputs are not a superset of the primary configuration outputs.

Creating scoring configuration association sets

Create a scoring configuration association set to distribute score requests across a set of configurations.

Before you begin

Open the Scoring view to access all available scoring configurations and configuration aliases. There must be at least two scoring configurations in the system to create an association set.

Procedure

1. Select a scoring configuration in the Scoring view.

   This configuration is the primary configuration for the association set.
2. Click **Create Associations**.
3. In the **Scoring Configuration Association** dialog box, select one or more scoring configurations to associate with the primary configuration.

   The list includes only the configurations that are compatible with the primary configuration.

   Press the **Ctrl** key to select multiple configurations. To select a contiguous set of configurations, press the **Shift** key and select the first and last configurations in the set.
4. Click **Next**.
5. Enter the percent of scoring requests to submit to each configuration by typing the value in the **Percentage** column.

   The specified percent values must sum to 100%.
6. Click **Finish**.

Results

The Scoring view updates to show the new configuration association set. Scoring requests sent to the primary scoring configuration in the association set are distributed across the configurations in the set in accordance with the specified percentages.

Editing scoring configuration association sets

Modify a scoring configuration association set to change the configurations included in the set or to change the percentage of requests that are sent to each configuration in the set.

Before you begin

Open the Scoring view to access all available scoring configurations and configuration aliases.

Procedure

1. In the Scoring view, select the primary configuration for an association set.
2. Click **Edit Associations**.
3. In the **Scoring Configuration Association** dialog box, select one or more scoring configurations to associate with the primary configuration.

   The list includes only the configurations that are compatible with the primary configuration.

   The configurations currently included in the association set are selected by default.

Press the **Ctrl** key to select multiple configurations. To select a contiguous set of configurations, press the **Shift** key and select the first and last configurations in the set.

4. Click **Next**.
5. Enter the percent of scoring requests to submit to each configuration by typing the value in the **Percentage** column.

   The specified percent values must sum to 100%.
6. Click **Finish**.

Results

The Scoring view updates to show the modified configuration association set. Scoring requests sent to the primary scoring configuration in the association set are distributed across the configurations in the set in accordance with the specified percentages.

Deleting scoring configuration association sets

Remove the associations from a scoring configuration to prevent scoring requests that are submitted to the configuration from being processed by any associated configurations.

Before you begin

Open the Scoring view to access all available scoring configurations and configuration aliases.

Procedure

1. In the Scoring view, select the primary scoring configuration in an association set whose associations you want to remove.
2. Click **Remove Associations**.

Results

The associations between the selected scoring configuration and other configurations are removed from the system. All subsequent scoring requests that are submitted to the selected configuration are processed by that configuration.

A/B split testing for scoring configurations

Use A/B split testing to compare different scoring configurations in terms of overall performance. You can add the best configuration to your existing production tasks to improve their predictive capabilities.

When you want to use a model to generate scores, you define a scoring configuration that is based on a particular version of the model in the IBM SPSS Collaboration and Deployment Services Repository. Over time, you might need to update the model to take into account extra considerations, such as new data availability or better predictive algorithms. However, before you replace a model that is used in production tasks, you might want to evaluate the performance of the new model to determine whether it is better than the current model.

In A/B split testing, you keep the existing scoring configuration deployed but reroute some of the incoming score requests from the production configuration to the potential replacement configuration. As more requests are processed, you can review the model results to evaluate the overall performance. If you determine that the replacement configuration is working better than the existing configuration, replace instances of the existing configuration in your production tasks with the new configuration.

The A/B split testing approach incorporates both association sets and configuration aliases. The association set defines the scoring configurations that are being evaluated and the number of requests that are sent to each. The alias controls the configuration that is used in your production tasks.

To illustrate this approach, suppose that you generate scores from Model 2 by using Configuration B. You develop two possible alternative models for generating scores, Model 1 and Model 3, and create a

scoring configuration for each. You want to determine whether either alternative model performs better than Model 2.

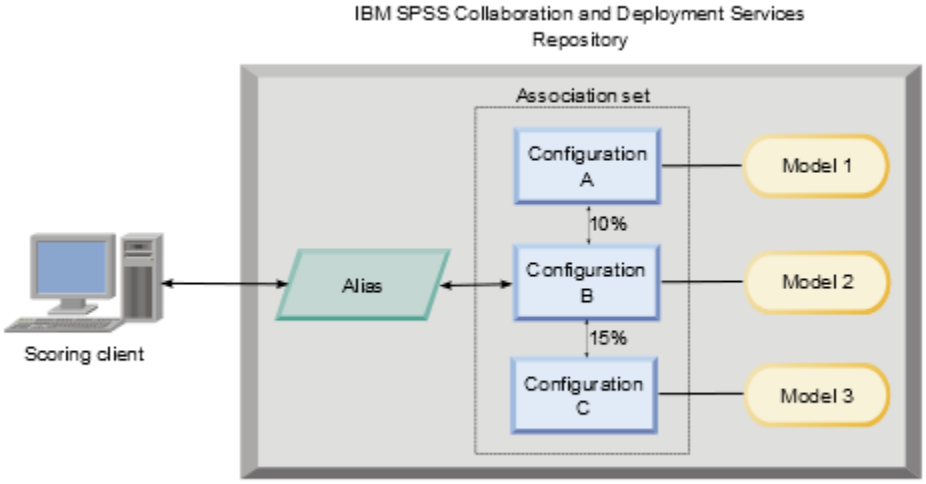

Figure 10. A/B testing example

Figure 10 on page 109 displays an A/B split testing environment for evaluating the configurations. A client application sends a score request to a configuration alias, which passes the request to Configuration B. Configuration B is the primary configuration in an association set that sends 10% of requests to Configuration A and 15% of requests to Configuration C. The remaining 75% of requests are processed by Configuration B. The configuration that is assigned to the request generates and returns the score output by using the associated model.

The scoring log table contains the results for all of the scoring configurations. After you examine the results, suppose that you determine that Configuration C produces better overall scores than Configuration B. To deploy Configuration C in your production tasks, edit the alias to reference this configuration. All subsequent requests that are sent to the alias are routed to the new configuration.

If none of the alternative configurations produce better results, remove the association set to have the existing configuration continue to produce scores in your production tasks.

Evaluating scoring models by using A/B testing

For A/B split testing, use a scoring configuration association set to define the set of configurations that are included in the tests. Use a configuration alias for the primary configuration in the set to manage the configuration that is used for production-level scoring tasks.

Procedure

1. Create an alias for the primary scoring configuration.

   Client applications send score requests to this alias.
2. Create an association set for the primary scoring configuration that includes the other configurations to be tested. Assign the percentages of score requests to be handled by each configuration.

   Scoring requests that are passed to the association set by the alias are distributed across the configurations in the set.
3. Send score requests to the alias and review the results.

   Use the information in the scoring log table to evaluate the performance of the configurations that are being tested.
4. If the results indicate that a scoring configuration produces better results than the primary configuration, update the alias to reference the new configuration.

Client applications that send requests to the alias automatically use the new configuration to generate scores.

5. Delete the association set that was used for testing.

Scoring view

The Scoring view lists all of the scoring configurations, configuration aliases, and configuration association sets currently defined in the system. From this view, aliases and associations can be created. In addition, configurations, aliases, and associations can be modified or deleted from the system.

To access the Scoring view from the menus, select:

View > **Show View** > **Scoring**

By default, configurations on the current server appear in the view. To access other servers, select the server from the **Server** list. If you are not logged in to the selected server, the **Login** dialog box opens, prompting you to log in to the server.

Each row in the view corresponds to a scoring configuration or configuration alias for the server. Properties that are shown for each entry include the following values:

Type
An indicator of the type of model that is used by the configuration.

Status
The status for the element.

Configuration Name
The name of the scoring configuration or alias. For an alias, the assigned scoring configuration is displayed in parentheses after the alias name.

Model
The name of the model file that is associated with the configuration.

Label
The label that identifies the version of the model file that is configured for scoring.

Average Score Time
The average time for a single score to be calculated.

Scores/Second
The number of scores that are produced per second at the time the scoring value is retrieved.

Total Scores
The total number of scores that are generated by using the configuration.

If a configuration is the primary configuration in an association set, the row begins with a triangle icon that indicates the entry can be expanded or collapsed. When you click the triangle, the row expands to show all of the configurations included in the set. The percentage of scoring requests that are assigned to each configuration is displayed in parentheses after the name. If you click the triangle for an expanded row, the row collapses to hide the association set details.

The configuration list can be refreshed to update the scoring statistics. To refresh the view, click the **Refresh** icon.

The list can be sorted by any column in the view. Click a column heading to sort by that column. Click the heading again to reverse the sort direction.

Filtering the Scoring view

The Scoring view can be filtered to reduce the number of results that appear in the list.

Once filtering is enabled, the specified filters persist throughout the IBM SPSS Deployment Manager session until changed. In addition, filter settings persist across server connections. For example, if filtering is enabled and the server selection is changed from server A to server B, the filter settings

established for server A persist for server B. To access view filters, click the Filters button. The Scoring Filters dialog box opens.

Enable Filtering. If selected, the specified filters are applied to the view.

The view can be filtered by the configuration name. Select the configurations to view in the Scoring view.

Editing a scoring configuration

To edit a scoring configuration:

1. In the Scoring View, right-click the configuration to be modified.
2. Select **Edit** from the pop-up menu. The Scoring Configuration wizard appears.
3. Modify the configuration settings as needed.
4. Click **Finish**.

Suspending and resuming scoring configurations

At times, it may be necessary to temporarily suspend a scoring configuration from processing requests. Suspending a configuration frees associated system resources and allows various maintenance tasks to be performed without affecting incoming score requests. Resuming a suspended configuration allows the configuration to begin processing scoring requests again.

To suspend a scoring configuration:

1. In the Scoring View, select the configuration to be suspended. To select multiple configurations, press the Ctrl key while selecting.
2. Right-click a selected configuration and select **Suspend Configuration(s)** from the pop-up menu. Alternatively, press the **Suspend Configuration(s)** button on the scoring view.

The status for the selected configurations updates accordingly.

To enable a suspended configuration to accept scoring requests, the configuration must be resumed. To resume a scoring configuration:

1. In the Scoring View, select the configuration to be resumed. To select multiple configurations, press the Ctrl key while selecting.
2. Right-click a selected configuration and select **Resume Configuration(s)** from the pop-up menu. Alternatively, press the **Resume Configuration(s)** button on the scoring view.

The status for the selected configurations updates, indicating the configurations are active.

Deleting a scoring configuration

To delete a scoring configuration:

1. In the Scoring View, select the configuration to be deleted. To select multiple items, hold down the Ctrl key and select additional rows.
2. Click the **Delete** button. The Delete Confirmation dialog box appears.
3. Click **OK**. The configuration is removed from the system.

Alternatively, right-click the configuration and select **Delete** from the pop-up menu.

Deleting a scoring configuration does not remove the model associated with the configuration from the system. Furthermore, any other configurations using the model remain in the system until specifically deleted.

Scoring Graph view

The Scoring Graph view displays the scoring throughput for a selected scoring model.

To access this view:

1. In the Scoring View, select the configuration to be viewed as a graph.
2. Click the **Graphic View** button.

Alternatively, right-click the configuration and select **Graphic View** from the pop-up menu. The Scoring Graph view for the configuration opens.

The appearance of the graph is defined by the following controls:

Content Shown. Select the scoring metric appearing in the graph. Available metrics include the following:

- **Service Scores.** Total number of scores produced by all configurations.
- **Service Uptime.** Amount of time, measured in seconds, that the scoring service has been available for scoring.

For a description of other available measurements, see the Scoring Engine Metrics section of "Logging settings" on page 99.

Show Latest Period (Hour). Limits the graph to the most recent time period. Select the size of the period as the number of hours to show. For example, to show the past half-hour, specify a value of *0.5*. To show the past two hours, specify a value of *2*.

The graph in the view is regenerated at regular intervals, with the most recent information taking precedence over older results.

The Scoring Graph view displays the performance for a single selected configuration. To monitor multiple configurations, open multiple graph views.

Chapter 10. Jobs

What is a job?

A job is a container for a set of steps. Each step has parameters associated with it. Before you execute a step, you must embed it within a job. In order to generate results, a job must contain at least one step. (Although an empty job will run, it does not produce any results.) Jobs are scheduled and executed by an application server.

Job steps can be executed sequentially or conditionally. For example, you can make the second step in a job contingent upon the outcome of the first step in a job. Objects stored in the repository can provide input for a job step, and job results can be stored in the repository. For example, a job might include a data preparation step, which reads data from a data mart, followed by a step in which a IBM SPSS Modeler stream reads the prepared data and computes a propensity score based on the data.

Jobs are created in the Content Explorer and modified in the Job Editor. Jobs are stored in the database associated with the repository.

You can move jobs anywhere within the Content Repository. However, unlike other files, jobs are native and unique to the system. You must create and work with the job in the IBM SPSS Deployment Manager.

Jobs are executed on a server. The content of the job determines the type of execution server required. For example, if a job contains a IBM SPSS Modeler stream, a IBM SPSS Modeler server is required. The server can reside on the same system as the IBM SPSS Collaboration and Deployment Services server. Alternatively, certain jobs can be executed on remote servers.

Versioning and labeling jobs

Just like any other repository object, jobs can be versioned. In addition, a job can have many schedules associated with it. However, although a job can have many schedules, a schedule can contain only one job.

The following guidelines apply:

- Each time that changes are saved to a job, a new version of the job is created.
- Only a labeled version of a job can be scheduled.
- The system applies the *LATEST* label to the most recent version of a job. If a label is not explicitly supplied, the only option when scheduling a job is the version labeled *LATEST*.

Components of a job

A job can contain any combination of the following components:

- Visualization report files. See the topic Chapter 14, “Visualization report job steps,” on page 149 for more information.
- SAS syntax files. See the topic Chapter 15, “SAS® job steps,” on page 153 for more information.
- General job steps. See the topic Chapter 16, “General job steps,” on page 157 for more information.
- Message-based job steps. See the topic Chapter 17, “Message-based job steps,” on page 167 for more information.
- Notification job steps. See the topic Chapter 18, “Notification job steps,” on page 169 for more information.

In addition, products that collaborate with IBM SPSS Collaboration and Deployment Services may offer additional components that can be included in a job, including:

- IBM SPSS Modeler streams.

- Champion challenger job steps.
- IBM SPSS Statistics syntax files.

To include these components, install the necessary adapters and plug-ins as described in the documentation for the product offering the additional functionality.

Prerequisite to running jobs

For certain types of jobs, you must establish server and credential definitions.

- Server definitions are required for IBM SPSS Modeler streams, IBM SPSS Statistics syntax, and SAS syntax job steps.
- Credential definitions are required for IBM SPSS Modeler stream job steps.

See the topic Chapter 6, “Resource definitions,” on page 45 for more information.

External file dependencies

Job steps often refer to external sources, such as data files or databases. For those jobs to complete successfully, the external sources must be accessible to the IBM SPSS Collaboration and Deployment Services Repository.

For example, consider the following IBM SPSS Statistics syntax that references the data file *Employee data.sav*:

```
GET FILE='C:\Program Files\data\Employee data.sav'.
GRAPH
  /BAR(SIMPLE)=MEAN(salary) BY jobcat .
```

If this syntax is used as an IBM SPSS Statistics job step, the data file must exist in the *C:\Program Files\data* directory on the IBM SPSS Collaboration and Deployment Services Repository computer. If not, the job will fail.

An alternative approach for external files involves storing them on a network location that the IBM SPSS Collaboration and Deployment Services Repository can access. This could involve mapping to a shared drive on a remote machine or using a UNC file reference.

Another method for dealing with external files involves storing the files in the repository. Use a General job step to extract them to a location necessary for subsequent steps. For example, if the repository contains *Employee data.sav*, a General job step can place it in the *C:\Program Files\data* directory for a subsequent IBM SPSS Statistics step.

Finally, if steps make use of databases, an ODBC source for the database must be defined on the IBM SPSS Collaboration and Deployment Services Repository computer.

Job process overview

Although the components of a job vary, the basic process for working with jobs consists of the following tasks:

1. Create a new job or open an existing job.
2. Add steps to the job.
3. Specify the relationship between job steps, if applicable.
4. Save the job.
5. Test the job by running it immediately (optional).
6. Schedule the job for execution or run the job immediately.
7. Specify email notifications (optional).
8. Save changes to the job.
9. View job status (optional).
10. View output results from the job (optional).

Working with jobs in the Content Explorer

Typically, the Content Explorer is the starting point and the end point for jobs in the system. When you first work with a job, you either create or select the job in the Content Explorer. After you have created and saved the job, you can access it again in the Content Explorer.

Creating new jobs

Before you can schedule any step for execution, you must create a job. You can create an empty job and add steps to it later; however, to execute a job, your job must contain at least one step.

To make jobs easier to find, it is recommended that you have a designated place for storing them. For example, you can create a separate folder for jobs, or you can store jobs in the same folder as their respective model files. You must save a job within a folder. You cannot save a job within another object.

As you create jobs, it is important to note the following:

- By default, when you create a job, it is automatically saved in the repository. You need to explicitly save the job again after you add steps to it.
- If the Job option is disabled when you try to create a job, this means that you selected an item in the Content Explorer that is not a folder. For example, if you click a IBM SPSS Modeler stream file, the **New > Job** option is disabled because you cannot store a job within a IBM SPSS Modeler stream.
- Although you can house jobs almost anywhere within the Content Explorer, you cannot store jobs within the Resource Definitions folder. See the topic Chapter 6, “Resource definitions,” on page 45 for more information.

To create a new job:

1. In the Content Explorer, create a folder to house your job, if it does not already exist.
2. Select the folder in which you want to store your job.
3. From the File menu, choose:

   New > **Job**

   The **New Job Information** dialog box opens.
4. In the **New Job Information** dialog box, provide the following information:

Table 10. New job information

Field	Description	Required or optional
Name	You must specify a unique name for your job. The name must be unique within the folder. See the topic “Naming conventions within the system” on page 10 for more information.	Required
Description	This field contains a description of your job. Although this field is optional, it is recommended that you provide a detailed description for your job because the description can be searched. For example, if your description contains the phrase *propensity analysis*, and you search for that term, this job appears in the search results.	Optional
Keywords	The keywords that you specify here are used by the search facility. See the topic “Searching” on page 17 for more information.	Optional

Table 10. New job information (continued)		
Field	**Description**	**Required or optional**
Expiration date	An expiration date can be specified for the job. See the topic "Working with Expiration Dates and Expired Files" on page 29 for more information.	Optional

5. After you have specified the job details, click **Finish**. The new job appears in the folder that you selected previously in the Content Explorer and as a tab in the Job Editor.

You are now ready to add steps to your job. See the topic "Adding steps to a job" on page 118 for more information.

Opening an existing job

To open an existing job:

1. In the Content Explorer, navigate to the job and double-click on it. The steps within the job appear in the Job Editor.

Viewing job properties

Properties for jobs are similar to properties for other objects in the Content Explorer. See the topic "Viewing object properties" on page 23 for more information.

Working with the job editor

The Job Editor is where you modify jobs. You can perform the following tasks:

- Add steps to a job.
- Specify parameters for a job step.
- Establish relationships between steps.
- Schedule a job for execution.
- Specify notifications.
- View job status.

The Job Editor is divided into the following sections:

Job canvas. The job canvas provides a visual representation of your job. You add steps to your job in the job canvas. You can drag files from the Content Explorer and select job step types from the Job Steps list. In addition, you can establish relationships between items within a job.

Job palette. The job palette is divided into two main sections:

- **Relationships.** Using the tools in the Relationships section, you can establish relationships between the steps in your job. See the topic "Specifying relationships in a job" on page 119 for more information.
- **Job Steps.** In the Job Steps section, you can add several types of job steps to your job--for example, champion challenger, general, message-based, or notification job steps.

Job properties section. This section is divided into the following tabs:

- **General information.** Describes general attributes associated with the job. See the topic "General information tab" on page 117 for more information.
- **Job Variables.** Defines variables for the job whose values can be passed to steps within the job. See the topic "Job variables" on page 117 for more information.
- **Notifications.** Allows you to specify email notifications for job failure and job success. See the topic "Job success and failure notifications" on page 141 for more information.

The General information, Job Variables, and Notifications tabs apply to the entire job. Each step within a job has its own tabs associated with it. When you click a job step, the job processing tabs change accordingly.

General information tab

When you first open a job, the **General Information** tab appears by default. You cannot modify any items within the **General Information** tab. The information on this tab is generated and updated by the application, based on information that is provided at other points in the job creation process.

The **General Information** tab contains the following details about the job:

Job Path. This field specifies the directory location for the job.

Job Label. This field contains the label that has been applied to the job. If no label has been applied, this field remains blank.

Schedule Status. This field describes whether or not the job has already been scheduled. Valid values are *Scheduled* or *Not scheduled*.

Last Run Status. This field describes whether the job ran successfully or failed the last time it was run.

Job variables

Job variables define parameters whose values can be passed to any step within the job. Using variables, any job can be used as an iterative consumer, in which values external to the job can be used to control job processing. Values for the variables can be defined:

- When initiating the job
- In schedules associated with the job
- In other jobs executing before the job

The Job Variables tab for a job displays a table identifying variables defined for a job.

Variable Name. Lists the names of variables defined for the job.

Default Value. Identifies the default value for each job variable. If a variable has no specified default value and a value has not been assigned in some other fashion, the user is prompted for a value during job execution.

Description. Lists informative text about each variable typically used to aid in identifying the variables.

Job variables can be used in any job step field that supports entry field content assist. In addition to system properties, the list of available variables appearing when typing a $ in those fields includes all variables defined at the job level.

Adding variables to a job

The Add Job Variable dialog box creates new job variables. To define a new variable for a job, click **Add** on the Job Variables tab for an open job.

1. Type a unique name for the variable. No two variables for the same job can have the same name. Job variable name can contain English letters and digits only, and the first character must be a letter.
2. Type the value of the variable to use as the default.
3. Type a description for the variable.
4. Click **OK**.

The new variable appears in the variable list for the job.

Editing existing job variables

To modify an existing variable for a job:

1. Open the job.
2. Select the Job Variables tab.
3. In the variables table, select the cell containing the value to be changed. The variable name cannot be modified.
4. Modify the value.
5. Press the Enter key.

The updated information appears for the variable in the list.

Removing variables from a job

To remove an existing variable from a job:

1. Open the job.
2. Select the Job Variables tab.
3. Select the variable to remove.
4. Click **Remove**.

The variable is removed from the job and no longer appears in the variable list.

Adding steps to a job

After you have created or selected a job, you are ready to add steps to it. See the topic "Components of a job" on page 113 for more information.

To add steps to a job:

1. Open the job. The job appears in the Job Editor.
2. Select the item that you want to add to the job. You have the following options:
   - **Add a file.** From the Content Explorer, click the file you want to add, and then drag the file into the job canvas. Alternatively, you can right-click on the file and select **Add to Job**.
   - **Add a job step.** From the Job Steps section of the job palette, select the type of job step that you want to add. The option you select will be highlighted. Then click anywhere in the job canvas. By default, the system numbers steps in the order in which they appear. Thus, subsequent job steps appear as Event 2, Event 3, and so on. You can rename a step at any time. The number has no effect upon the order in which steps are executed.
   - **Copy and paste an existing job step.** Right-click an existing job step from any job and select **Copy**. Right-click the canvas for the target job and select **Paste**. Note that job step notifications are not preserved when copying steps.
3. If you want to add additional steps, return to step 2.

If you have two or more steps in your job, you can either:

- Run job steps concurrently. For more information, see "Executing job steps concurrently" on page 119.
- Establish relationships between the steps. For more information, see "Specifying relationships in a job" on page 119.

Many job steps reference IBM SPSS Collaboration and Deployment Services Repository resources that may have multiple versions. These steps reference a particular version of the resource by label. You must have adequate authority to the label selected for the job step to execute the job. If your authority does not permit referencing the job step resource by the selected label when the job runs, the system will be unable to find the resource and the execution will fail. See the topic "Label security" on page 32 for more information.

Executing job steps concurrently

If you create a job with multiple steps but you do not specify a relationship between the steps, the system will execute the steps concurrently. Thus, the steps are not contingent upon one another.

Specifying relationships in a job

When you add steps to a job, you have the option of establishing a relationship between the steps. The relationships that you specify determine the sequence of steps in the job and the conditions under which each step is executed.

Relationships are established by relationship connectors. When you connect two steps, there is an antecedent step and a consequent step. The antecedent step is the starting point for the relationship. This step, as its name suggests, is executed first. The consequent step is contingent upon the completion of the antecedent step. It is important to note that completion does not always mean successful completion. For example, if you specify a fail relationship between two steps, the consequent step is executed only if the antecedent step fails to complete successfully.

For example, suppose you have four steps in your job: A, B, C, and D. A simple scenario for the four steps would be to execute them sequentially--that is, execute A first, then B, then C, and then D. In another case, you might want to execute step A and then execute step B only if step A succeeds. If step A fails, you might want to execute step C. In both cases, you would want to generate a report in step D.

If you cancel a job, the relationships that you establish will not be followed. For example, suppose that you have a job with steps A and B with a sequential relationship between them. If you run the job and cancel it while step A is running, step B will not execute. This behavior holds true for all relationship types, including a fail relationship.

When you create a job, you can establish the following relationships between steps in the job:

Table 11. Relationship connectors

Relationship connector	**Conditional expressions allowed?**
Sequential	No
Pass	No
Fail	No
Conditional	Yes. In order to use a conditional connector, you must specify at least one argument.

Although the relationships that you can establish between steps vary, the process for establishing a relationship is the same.

To specify a relationship between two steps in a job:

1. Open an existing job. The job appears in the Job Editor.
2. From the palette, select the relationship you want to establish. The relationship is highlighted on the palette.
3. In the job canvas, click the antecedent step in the sequence and then click the consequent step in the sequence. When the mouse is released, a directional arrow connects the two steps. The icon in the center of the arrow describes the relationship.
4. To establish another relationship within the job, return to step 2.
5. Save your changes.

Sequential connector

If you connect two steps using the sequential connector, the consequent step is executed immediately after the antecedent step is completed.

The sequential connector is always executed. Unlike pass, fail, or conditional relationships, there are no conditions associated with a sequential connector. The only criterion for the execution of the consequent step is the completion of the antecedent step.

Pass connector

The pass connector is executed if the antecedent step is executed successfully. If the antecedent step is completed successfully, the system proceeds to the consequent step. If the antecedent step fails to complete successfully, the system will not execute the consequent step.

Fail connector

The fail connector is executed if the antecedent step does not execute successfully. If the antecedent step is executed successfully, the consequent step in a fail relationship is not executed. For example, the consequent step in a fail relationship may generate a report indicating that the antecedent step failed.

Conditional connector

The conditional connector is executed if the antecedent step satisfies the condition(s) you specify. For conditional relationships, you must specify an expression that contains the condition. This is the only relationship connector in which you can pass arguments to the system.

The Conditional Connector Expression can be any script the system can evaluate, but there are two special parameters that relate to the consequent step:

Completion_code. The `completion_code` is an integer value. When the `completion_code` expression is used, it must be lowercase. The job step interprets the meaning of the completion code. For the General job step, the completion code is the executable return code. For example, for most Windows commands, `File Not Found = 2`.

Success. The `success` parameter is a Boolean, where `true` indicates that the job step execution was successful. When the `success` expression is used, it must be lowercase.

You can compare and combine parameters, using standard scripting operators. For example,

- `&&`
- `||`
- `==`
- `!=`
- `<`
- `>`
- `!`
- `()`

The following are examples of expressions for a conditional connector:

```
completion_code == 0 && success == true
completion_code == 1 || success == false
completion_code >= 10 || success == false
(completion_code >= 1 &&  completion_code <= 5)  && success == true
```

Warning expressions in conditional connectors

A special case of using conditional job step connectors is specifying a relationship based on the warning code returned by the antecedent step.

The following job step types support warning codes:

- General
- IBM SPSS Modeler
- IBM SPSS Statistics
- SAS

To specify a warning-code-based relationship:

1. Connect the steps with a conditional connector.
2. Enter `warning==true` as the conditional expression. When the `warning` expression is used, it must be lowercase.
3. Open the parent job step and override the default warning code on the General tab if necessary.

Once the antecedent step is completed, the system will evaluate the conditional expression and proceed to the next step if `true` is returned.

Viewing general relationship properties and modifying relationships

Each object in the job canvas has properties associated with it. The relationship between steps is described on the General tab. At times, you may want to modify the relationship between steps in a job.

Viewing relationship properties. To view the properties for a relationship:

1. In the Job Editor, click the relationship icon. Two small black boxes appear at each end of the arrow to indicate that the connector has been selected. The General tab appears in the job properties section.

Modifying relationship properties on the General tab. To modify a relationship:

1. Select a new relationship from the Relationship drop-down list. The connector icon changes to reflect the new relationship.
2. If the new relationship requires that you provide parameters, you must specify the parameters. For example, if you change a pass connector to a conditional one, you must specify the parameters that describe the condition. See the topic "Conditional connector" on page 120 for more information.
3. Save your changes.

Modifying relationship properties in the job canvas. Alternatively, you can modify relationships within the job canvas. To do so:

1. In the Job Editor, right-click the arrow connecting the two steps whose relationship you want to modify. Two small black boxes appear at each end of the arrow to indicate that the connector has been selected.
2. Select **Relationship**.
3. Choose the new relationship you want to establish.
4. Save your changes.

Deleting relationships in a job

At times, you may want to delete relationships between steps in a job. If you delete a step from a job, any relationships that were established between the deleted step and other steps in the job are automatically deleted.

If you simply want to modify a relationship, you do not need to delete it. See the topic "Viewing general relationship properties and modifying relationships" on page 121 for more information.

To delete a relationship between two steps:

1. In the Job Editor, right-click on the connector icon. Two small black boxes appear at each end of the arrow to indicate that the connector has been selected.
2. Select **Delete**. The relationship is deleted.
3. Save your changes.

Saving jobs

It is recommended that you save changes to your jobs regularly. Changes that you make while working with jobs are not committed to the system until you save the job. In the Job Editor, a job containing unsaved changes is preceded by an asterisk in its name.

If you attempt to close a job before you have saved your changes, the system prompts you to save changes. The dialog box that opens depends on how you chose to close the job.

Closing a job in the Job Editor. If you attempt to close a job with unsaved changes in the Job Editor, a Save Resource dialog box opens.

You have the following options in this dialog box:

1. To save your changes and close the job, click **Yes**.
2. To disregard your changes and close the job, click **No**.
3. To exit the dialog box and return to the job, click **Cancel**. The Job Editor reopens.

Exiting the system. If you attempt to exit the system and you have unsaved jobs in the Job Editor, a Save Resources dialog box opens. This dialog box provides a list of the jobs with unsaved changes in them.

To save the changes:

1. Select the check boxes next to the jobs with changes you want to save. To expedite the process, click **Select All** to save changes in all of the jobs in the list, or click **Deselect All** to discard changes in all of the jobs in the list.
2. Click **OK**.

To exit the system without saving changes to any of the jobs in the list, click **Cancel**.

It is important to note that if any changes are made to the files in a job--for example, an IBM SPSS Modeler stream (`.str`)--any job that contains the file is affected. When changes are made to the file, a new version of the file is saved to the repository. However, the job that contains the file is not automatically updated with the modified file. To incorporate the file updates into the affected job:

1. Reopen the job. When the job is reopened, an asterisk appears with the job name in the job canvas, indicating that the job contains unsaved changes.
2. Resave the job.

Job step results

Most types of IBM SPSS Deployment Manager job steps generate results in the form of output files. The files can be saved in the IBM SPSS Collaboration and Deployment Services Repository or in the file system running the application that generated them (for example, a system running IBM SPSS Statistics server). Certain types of job steps can generate multiple output files. Most types of job steps allow you to select output formats, such as HTML, text, or PDF.

Job step output settings are displayed on the Results tab of the Job Editor for IBM SPSS Statistics, SAS, IBM SPSS Modeler, and reporting job steps and on the Output Files tab for general job steps. The Results tab and available options are different for different job step types.

Regardless of the job step type, the following output file settings can be modified:

- Output file location
- Output file permissions

- Output file metadata

Output file location

Click the ellipsis button in a selected Location cell to define the location for the results.

Save to repository. Saves output files to the repository in the specified folder. Click **Browse** to select the folder in which to save the output. If the files already exists, new versions are saved.

Save to server file system. Saves output files to the file system where the IBM SPSS Collaboration and Deployment Services Repository Server is running. In the Folder field, specify the name of the target directory in which to save the file. The target directory can be your local file system or a network sharing directory in which proper access permission must be granted.

Discard. Job output is discarded.

Output file permissions

When a job is run, any output generated by the job is owned by the user that ran the job. Output file permissions for additional users or groups can be set in the Output Permissions dialog box. This dialog can be used to:

- Add a new principal. See the topic "Adding new users or groups" on page 26 for more information.
- Add a principal using a variable. See the topic "Add principal by variable" on page 123 for more information.
- Delete an existing principal. See the topic "Deleting a user or group from the permissions list" on page 27 for more information.
- Modify permissions for an existing principal.

Accessing the Output Permissions dialog box varies by the job step type.

- **IBM SPSS Statistics, SAS, and IBM SPSS Modeler job step output files.** In the Results tab, click the ellipsis in the Permissions column next to the file entry.
- **Reporting job step output files.** Click **Browse** next to the Permissions field.
- **General job step output files.** On the Output Files tab, click the ellipsis in the Permissions column next to the file entry.

To modify permissions for an existing user or group:

1. Select the principal from the table.
2. In the Permissions column, click the drop-down arrow and select a new permissions level from the list.
3. Click **OK**.

Add principal by variable

Job step output file permissions can be assigned dynamically by adding users and groups to the permissions list as iterative variable values. In such cases, the variables defined within the step are used to retrieve the recipient address value every time the step is executed.

To add a user or group to the permissions list, in the Output Permissions dialog box click **Add Variable**. The Add Principal By Variable dialog box opens.

1. Select the security provider for authenticating the principal defined by the iterative variable.
2. Select **User** or **Group** as the principal type.
3. Enter the variable in the Principal Variable field. Use field content assist if necessary. See the topic "Entry field content assist" on page 10 for more information.

Output file metadata

IBM SPSS Deployment Manager provides the capability to specify metadata properties for IBM SPSS Statistics, IBM SPSS Modeler, SAS, and reporting job step output files similar to the way properties are specified for IBM SPSS Collaboration and Deployment Services Repository objects.

To specify properties for a job step output file:

1. Click the step in an open job.
2. Click the Results tab.
3. In the list of output files, click in the Properties column next to the file. The ellipsis button is displayed. Disregard this step for reporting job steps.
4. Click the ellipsis button. The Output Properties dialog box opens. For reporting steps, click **Browse** next to the **Metadata** field.
5. Provide the following information.
   - **Description.** A user-defined description of the output file.
   - **Keywords.** The metadata assigned to the output file for the purposes of content search.
   - **Author.** A string identifying the author of the output file.

   Note: The dialog box may also display custom properties if such properties have been defined for the server. See the topic "Creating custom properties" on page 36 for more information.

   - **Expiration Date.** The date after which the file is no longer active. By default, no expiration date is set for the job output file. You can specify a time interval from the current date for file expiration (for example, one year) or enter a specific date.
   - **Version Labels.** A user-defined label for the file. By default, no label is assigned to the file. To specify a new label or select from a list of existing labels, click **Browse**. The Edit Version Labels dialog appears. See the topic "Editing the version label" on page 28 for more information. If your label permissions change after defining output labels, labels may appear in the list that can no longer be applied. Job execution will fail due to an inability to assign a label. The label that cannot be assigned must be removed from the output properties or your permissions to the label will need to be changed. See the topic "Label security" on page 32 for more information.
   - **Topics.** The topics assigned to the file. Click **Add** to open the **Add Topic** dialog box and select topics. To remove a topic, select the topic in the list and click **Remove**. See the topic "Working with topics" on page 39 for more information.
6. To save your changes, click **OK**.

Note: Fields marked with a light bulb icon support the insertion of variables to the property at run time. For example, you can append the date to an output file/folder name at run time. For burst reporting job step types, you can also insert burst variables. See "Entry field content assist" on page 10.

Chapter 11. Running jobs

Running a job results in all of its job steps executing according to their relationships defined in the job. Job execution can be initiated on-demand in response to a user request or scheduled to occur in the future in response to a specified event.

The job status and job history are displayed in the history view for jobs run through schedules and jobs run immediately. See the topic “Job history view” on page 133 for more information. Existing time-based and message-based schedules are listed in the job schedule view. See the topic “Job schedule view” on page 132 for more information.

On-demand job execution

At times, you may want to run a job immediately after you have created or modified it. This process, which is typically used to test a job, is launched manually.

To run a job immediately:

1. In the Content Explorer, select an existing job.
2. From the **Tools** menu, select **Run Job Now**. The **Job Execution** dialog box opens. Alternatively, you can select the **Run Job Now** button from the toolbar.
3. Click **OK**.

The execution uses the specified default values for any variables in the job. In addition, any notifications associated with the job and its steps will be processed and delivered. To specify alternative variable values or disable notifications for the execution, use **Run Job with Options**.

It is not necessary to open a job in order to run it. You only need to select the job. If you want to view the contents of the job, double-click the job to open the Job Editor. Job execution can also be initiated from the Job Editor and the Job History view. See the topic “Job history view” on page 133 for more information.

Specifying on-demand execution options

The default variable values and notification treatment can be overridden by specifying execution options. For example, it may be desirable to turn off notifications when testing jobs.

To specify options when running a job on-demand:

1. In the Content Explorer, select an existing job.See the topic “Opening an existing job” on page 116 for more information.
2. From the **Tools** menu, select **Run Job with Options**. The **Run Job with Options** dialog box opens.
3. Modify variable values as needed by clicking the value to be changed and typing the new value.
4. Specify whether or not the job and job step notifications should be delivered as a result of the execution.
5. Click **OK**.

The job executes with the specified options.

Scheduled job execution

A job often needs to execute multiple times in response to future events. For example, a report created by a job may be needed every Friday. To define the execution pattern, the job must be associated with a schedule specifying the execution settings. A schedule can be either time-based or message-based.

- A time-based schedule relies on the occurrence of a specified time or date to initiate job execution. For example, a time-based schedule may run a job at 5:00 PM every Thursday.

- A message-based schedule relies on the receipt of a JMS message to trigger job execution. In this case, the job executes every time a specified message domain receives a message.

Given that a scheduled job executes at some point in the future, the schedule must use credentials defined in the system. Jobs cannot be scheduled using single sign-on (SSO) credentials because a user might not be logged in when the job executes.

Creating schedules

To create a new schedule for a job, right-click on a job in the Content Explorer and select:

New Schedule > **Time-Based**

or

New Schedule > **Message-Based**

The Job Schedule wizard opens. Schedules for jobs must specify the following information:

- Settings for the job being scheduled
- Settings for the schedule that depend on whether the schedule is time-based or message-based
- Values for job variables

Job information

The **Job Information** page of the **Schedule** wizard identifies the version of the job associated with the schedule and the credentials used for job execution.

Job. The name and IBM SPSS Collaboration and Deployment Services Repository path for the job associated with the schedule.

Credentials. Credentials, which specify permission levels, determine the user under which the scheduled job runs. Specifically, the system verifies credentials when writing output and saving files to the IBM SPSS Collaboration and Deployment Services Repository. Click **Browse** to select a credential from those defined in the system. To create a new credential for the scheduled job, click **New**. The specification of a credential for a job schedule is required. In Active Directory or OpenLDAP-based single sign-on environment, Server Process Credential can be used instead of regular user credential. See the topic "Server Process Credential" on page 46 for more information.

Note: If the job step is run under Active Directory user credentials, the credentials definition must be associated with the corresponding Active Directory domain. See the topic "Credential destination" on page 45 for more information.

Label. The label for the version of the job being scheduled. Select the value from the list of existing labels. The user under which the job runs must have adequate authority to the selected label when the schedule triggers job execution. If the user authority does not permit referencing the job by the selected label when the schedule triggers, the system will be unable to find the job and the execution will fail. See the topic "Label security" on page 32 for more information.

Time-based schedule settings

For time-based schedules, the Schedule Time and Recurrence page of the Schedule wizard defines the time of day and frequency at which the job runs, as well as the duration of the schedule.

Start Time. Select the time of day at which you want to run the job from the drop-down list. To schedule a job at a time that is not available from the drop-down list (for example, 5:45 AM), type the value in the **Start time** field.

Recurrence Pattern. The recurrence pattern defines how often the job runs. Select the pattern and frequency from the following options:

- **Once.** The job runs one time only. From the **Date** drop-down list, select the date on which you want the job to run.

- **Minutely.** Runs a job on a minute basis. The recurrence pattern indicates the number of minutes between runs. For example, set the interval to 10 to run the job every 10 minutes.
- **Hourly.** The job runs at hourly intervals. Specify the schedule frequency using the **Recur every <value> hours** field. For example, a value of 2 results in the job running every two hours.
- **Daily.** The job runs regularly at daily intervals. Specify the schedule frequency using the **Recur every <value> days** field. For example, a value of 3 results in the job running on every third day.
- **Weekly.** The job runs at weekly intervals on the specified day. Specify the schedule frequency using the **Recur every <value> weeks** field. For example, a value of 4 results in the job running on every four weeks on the designated day.
- **Monthly.** The job runs at monthly intervals on the specified day of the month. Specify the schedule frequency using the **Recur every <value> months** field. For example, a value of 4 results in the job running every fourth month on the designated day.

Range of recurrence. The range of recurrence is the duration for which the job runs and consists of three parts:

- **Start date.** The date for the initial run of the job. The earliest that a daily schedule can begin is the day after it is created.
- **Start time.** The time for the initial run of the job.
- **End date.** The date for the final run of the job. For daily schedules, specify a date one day after the last run day to allow the schedule to run at the required time on the final day.
- **End time.** The time for the final run of the job. If you specify no end date and time, the job will run indefinitely according to the schedule settings you specified.
- **Weekday.** The days of the week to run the job.

Click **Next** to specify values for variables used in the scheduled job.

Message-based schedule settings

A message-based schedule is triggered by an outside event signaled by a JMS (Java Messaging Service) message. For example, when IBM SPSS Collaboration and Deployment Services job relies on the input from a third-party application, the application must send IBM SPSS Collaboration and Deployment Services a JMS message when the input file is ready for processing. For message-based schedules, the Message-Based page of the Schedule wizard defines the message domain and filter for the schedule.

Message Domain. The message domain identifies the JMS topic to which to subscribe. Select the domain from the list or click **New** to create a new message domain. See the topic "Message domains" on page 52 for more information.

Message Filter. Optional values that must be satisfied by a message for schedule activation to occur. Filters can be based on the text and the header of the message.

- **Message Text.** For JMS text messages, text that must be included in the message to activate the schedule.
- **Message Selector.** Optional selector text of contents of the message header. For example, you can indicate that the message header must contain the ID of a specific repository resource (`ResourceID=<resource ID>`) or specific custom properties (`NewsType='Sports'` or `NewsType='Business'`).

Use Durable Subscriptions. The option to save the messages specified for the schedule while IBM SPSS Collaboration and Deployment Services is not actively listening on the messaging service, so that the saved message can be retrieved later when the system starts listening.

Note: Job processing may depend on several different external events. The initial event that triggers the start of the execution is specified in the message-based schedule. If processing requires subsequent events to take place, they must be specified by message-based job steps. See the topic Chapter 17, "Message-based job steps," on page 167 for more information.

Click **Next** to specify values for variables used in the scheduled job.

Job variables for schedules

The Job Variables page for the Schedule wizard specifies values for the variables defined for the job associated with the schedule.

The variables table contains the following columns:

- **Name.** The names of the existing job variables.
- **Value.** The value currently assigned to each variable. These are typically the default values defined for the variables.
- **Description.** Informative text about each variable.

When scheduling a job, you may want to use alternative values for one or more variables. To change a variable value:

1. Click the value to be changed.
2. Type the new value for the variable.
3. Press the Enter key.

The existing value updates to the new value. To return the variables to their default values, click **Restore Defaults**.

Click **Finish** to create the schedule for the job.

Mapping JMS header properties to job variables

For message-based schedules, the JMS message may contain variables in the header that can be mapped to the values for job variables. This mapping may be manual or automatic, as defined by the Job Variables page of the Schedule wizard.

In manual mapping, the header properties from the message are assigned to specific job variables. Message header variables are referenced using the following syntax:

```
${JMSHeader.propertyName}
```

The value of *propertyName* corresponds to the name of the property in the message header. For example, suppose the header for the message associated with the schedule includes the property *SalesRegion* and the job being scheduled includes the job variable *region*. The value of *SalesRegion* is assigned to *region* by specifying *${JMSHeader.SalesRegion}* as the value for *region*.

In contrast, for automatic mapping, the values for header properties in the JMS message are automatically used for any job variables that have the same name as the header property. Values defined in the schedule for the matching job variables are replaced with values from the message. Job variables that do not match a header property use the values specified in the schedule. If any of those values reference a header property that is unavailable, the name of the value is used. To use an empty string instead of the value name, use **quiet reference notation** by inserting an exclamation point between the dollar sign and the property name:

```
$!{JMSHeader.propertyName}
```

Header properties in the message that do not match a job variable are unused by the schedule. To enable automatic mapping, select the **Automatically map JMS header properties to job variables** option.

Editing schedules

A list of schedules defined in the system can be accessed using the Job Schedule view. See the topic "Job schedule view" on page 132 for more information.

To edit an existing job schedule, right-click on the schedule in the Job Schedule view and select **Edit Schedule**. Alternatively, you can click the *Edit Schedule* icon.

The Schedule wizard opens, showing the settings for the selected schedule. Modify the values as needed and click **Finish** to save the updated schedule.

Reactivating a dormant schedule

If a version label of a job is removed, or if that version of the job is deleted, the schedule associated with the labeled version becomes dormant. Thus, the dormant schedule cannot be used until it is reassociated with a valid labeled job version.

If a schedule becomes dormant, the following message appears:

`The scheduled label, <label name>, for this job does not exist anymore. You must select a different label, or re-apply the originally scheduled label to the job.`

`<label name>` represents the version label that was removed or the job version that was deleted.

The resolution for a dormant schedule depends upon how the schedule became dormant.

- **Deleted job version.** If the job version was deleted, a different job version must be selected for scheduling.
- **Deleted version label.** If the label for the job version was removed, the label needs to be reapplied to the job version.

Deleting schedules

To delete an existing job schedule, right-click on the schedule in the Job Schedule view and select **Delete Schedule(s)**. Alternatively, you can click the *Delete Schedule(s)* icon.

The job associated with the schedule remains in the system. However, the job will no longer execute in accordance with the deleted schedule.

Message-based processing example

Message-based scheduling functionality of IBM SPSS Collaboration and Deployment Services can be used to trigger processing by repository events as well as by third-party applications. For example, a job can be configured to be rerun when the IBM SPSS Modeler stream used in one of the job steps is updated.

The procedure involves the following steps:

1. Using IBM SPSS Deployment Manager, create a JMS message domain.
2. Set up a message-based schedule for the job using the message domain. Note that the JMS message selector must indicate the resource ID of the IBM SPSS Modeler stream as in the following example:

   ```
   ResourceID=<resource ID>
   ```

   The repository resource ID of the IBM SPSS Modeler stream can be found in the object properties.
3. Set up a notification for the IBM SPSS Modeler stream based on the JMS subscriber you have defined.
4. To test the message-based schedule, the stream must be opened in IBM SPSS Modeler, modified, and stored in the repository. If everything has been set up correctly, the schedule will trigger the job.

Chapter 12. Monitoring status

In IBM SPSS Deployment Manager, the status of a job can be analyzed via a number of job summary views. Information is organized in a table and is designed to provide an at-a-glance summary for jobs in the repository.

The following views are available:

- Job Schedule
- Job History
- Model Management
- Predictors

Although individual tables contain specialized information, all job summary views are accessed in a similar way.

In addition, the status of servers and server clusters can also be monitored. See the topic "Server status view" on page 140 for more information.

Accessing status views

In the following section, the **<View Type>** refers to the specialized view. Examples include:

- *Job Schedule*
- *Job History*
- *Model Management*
- *Server Status*

Job summary views can be accessed using any of the following methods:

Toolbar. This option launches a blank job summary view. To access the job schedule view from the toolbar:

1. From the **View** menu, select:

   Show View > **<View Type>**

   Content Explorer. This option launches a populated job summary view. To access the job summary view from the Content Explorer:
2. Right-click on a job and select **Show <View Type>**.

Selecting a server in status views

By default, jobs on the current server appear in the job list.

To view jobs for other servers:

1. From the **Server** drop-down list, select a server.

The job list is updated with jobs on the selected server.

Note: If you are not logged into the server that was selected, the *Login to IBM SPSS Collaboration and Deployment Services Repository* dialog box opens, prompting you to log into the server.

Opening a job in the Job Editor

At times, when a job summary table is open (e.g., the job schedule or the job history), it may be useful to open the job in the Job Editor and view the job contents in conjunction with the job table.

To open a job in the Job Editor:

1. Select the job in the table.
2. Click the **Open job in editor** icon. The job opens in the canvas of the Job Editor.

Refreshing status views

When a job is run, the corresponding job tables (e.g., job schedule and job history) are not refreshed automatically. You must refresh job tables manually.

To refresh the job table:

1. Click the Refresh icon. The job table displays the updated status.

Reordering items in status views

By default, items in the job summary tables (e.g., job schedule and job history) are organized in chronological order. The order of items in the job summary tables can by reorganized by column.

To reorder items:

1. Click in the header row of the column that you want to reorder.
2. Click on the arrow in the row title. The table is reordered.

Deleting a job from status views

To delete an item from a job summary table (e.g., a job schedule or a job history):

1. Select the item in the table. To select multiple items, hold down the **Ctrl** key and select additional rows.
2. Click the **Delete** icon. The Delete Confirmation dialog box opens.
3. Click **OK**. The item is removed from the job schedule table.

Job schedule view

A scheduled job is run automatically at a designated date and time or when a JMS message is received from an outside application. Job schedules are described in the Job Schedule table, which contains the following summary information:

Job Name. The name of the job.

Version Label. The label that has been applied to the job. If no label has been applied to the job, this field is blank.

Schedule Summary. The frequency with which the job is scheduled to run. Valid values are *Once*, *Minutely*, *Hourly*, *Daily*, *Weekly*, and *Monthly*.

Next Start. The next date and time at which the job is scheduled to run.

Last Run. The date and time when the job was last run.

Last Run Status. The status of the job when it was last run. Valid values include *SUCCESS* or *FAILED*. A canceled job is recorded as a failure.

Credential. The credential under which the scheduled job will run.

The job schedule type (message-based or time-based) is specified by the icon in the column before the job name.

The first column indicates any problems with a schedule, such as a reference to a message domain that has been deleted, by displaying a warning icon. Right-click a problematic schedule and select **Show Error Messages** to view any errors associated with the schedule.

Job history view

Each time a job is run, the action is recorded in the job history, which provides status information about the job and its corresponding steps.

Specifically, the *Name* column describes the overall job. When the **+** is expanded, information about the individual job steps within a job appears.

Information in the job history table cannot be modified because this table reflects information obtained from other components in the system. Job information is provided on a per server basis. See the topic "Selecting a server in status views" on page 131 for more information.

In addition, filters can be applied to the job history to reduce the number of jobs that appear in the list. See the topic "Job History Filters" on page 138 for more information.

Working with the job history table

The job history table contains the following information:

Name. The name of the job or step.

Version. The version label that has been applied to the job. If no version label was explicitly applied in the job schedule, then the *LATEST* label is used by default.

Status. The current status of the job or step. Valid values include *Success*, *Running* or *Failed*. A canceled job is recorded as a failure. The status of the individual job steps that comprise each job, as well as any corresponding logs, appear underneath each job. To expand the job history list for a specific job, click the **+** next to the job.

Start date. The date and time at which the job or step was started.

Run time. The time taken to run the job or step. It is important to note that a value in this field does not necessarily imply success. The *Status* column describes whether or not the job ran successfully.

User. The user who last scheduled the job.

Blank cells in the job history view

If no label has been applied to a job, the corresponding **Label** field is blank. For the **Status**, **Start Time**, and **Run Time** fields, a blank cell indicates that the job has not been run yet.

Canceling a job

You can cancel a job while it is running. It is important to note that canceling a job is not the same as deleting a job. Canceling only stops the job from running. Canceling does not remove a job from the content repository.

Important: You must be assigned the *Schedule* action to be able to cancel jobs.

If you cancel a job that contains relationships, the relationships will not be followed. For example, suppose your job contains step A and step B, which are joined by a sequential connector. If you cancel the job while step A is running, step B will not be executed. This process applies to all relationship connectors, including the fail connector.

To cancel a job:

1. In the Job History table, select the job that you want to cancel. The job status of the selected job should be *Running*.
2. Click the **Stop selected running job** icon.

Note: On Windows, if you cancel a General job step that executes an external program or script (such as a Python script), the script execution will not be canceled. This is a limitation of JDK on Windows. You may need to stop the script manually outside of IBM SPSS Collaboration and Deployment Services.

Viewing job results

If results are available, separate rows appear under the job steps in the table. Each row contains the path in which the results are stored. For certain applications, such as IBM SPSS Modeler, double-clicking on the results path launches the IBM SPSS Modeler client and displays the results.

The system can open any result file that is written back to the IBM SPSS Collaboration and Deployment Services Repository. If the file type has a corresponding editor in the IBM SPSS Deployment Manager (e.g., HTML, text, image files), the system opens the file within the IBM SPSS Deployment Manager. If a file path is preceded by the machine name, the IBM SPSS Deployment Manager does not have a editor to display the file and will call upon the operating system to open the file.

Viewing job logs

The job history table includes logs for the overall job as well as for steps within the job. These logs are system-generated logs and cannot be modified. Note that some types of steps do not produce logs.

In the table, the overall log appears under a subheading corresponding to the name of the job. This log contains information relevant to the entire job, such as whether or not residual artifacts have been deleted. The values used for any job variables also appear in this log.

If a log is available for a job step, it too can be accessed from the job history table. Job step logs appear under a subheading corresponding to the job step name and pertain to the corresponding step only. If the job contains multiple steps that produce logs, the logs appear under each step in the table. These are not logs for the entire job, and they are not logs for the system.

To view a log, double-click **Log** in the Name column for the cell that corresponds to the log that you want to view. The log appears in a separate editor.

Model management views

A summary of model results is provided in the model management views. Similar to the job history and job schedule views, the model management tables provide information about model analyses.

The following views are available:

- Model evaluation
- Champion challenger

The information provided in each view, which may vary by model type, is not modifiable. In addition, filters can be used to narrow the results of a model management view even further. See the topic “Model management filters” on page 139 for more information.

To select a specific model management view:

1. From the **View** menu select:

   Show View > **Model Management**

   The Model Management tab appears.
2. From the **Server** drop-down list, select the server name.
3. From the **Type** drop-down list, select the view type. Options include model evaluation or champion challenger.

Note: The model management views present information about IBM SPSS Modeler files stored in the repository. To store these files in the repository, the IBM SPSS Collaboration and Deployment Services environment must include IBM SPSS Modeler adapters. For information on installing the adapters, see the IBM SPSS Modeler documentation.

Model evaluation view

The model evaluation view describes results for model evaluation scoring branches. This view is designed to provide an at-a-glance summary of the overall performance of the branch. For example, this view

describes whether the branch is trending upward, downward, or has remained the same. Jobs may appear multiple times in the model evaluation view. One entry appears for each labeled version of a job that contains a model evaluation scoring branch.

The Model Evaluation view contains the following information.

- **Scoring Branch.** The branch containing the scoring node.
- **File.** The file containing the scoring branch.
- **Version.** The version of the file used to generate the results.
- **Index.** A percentage value. The colored circle next to the index value corresponds to the thresholds for performance (e.g., good, better, bad) that were specified in the General tab for the job step. For example, a red circle indicates that the index value falls into the lowest-performing range.
- **Trend.** The percentage change in the model. An arrow indicates whether the trend is up or down. If the trend is 0.00, a horizontal bar appears, which represents no change. The first time the file is run, the trend field is blank.
- **Author.** The author of the file.
- **Type.** The type of analysis. Examples include *gains* and *accuracy*.
- **Data.** The actual data used for the source node. The value depends on the source node type. For example, for an ODBC node, the value is a DSN name. For a Variable File node, the value is a file name.
- **Job.** The job that references the file.
- **Job Version.** The version of the job used to analyze the model.
- **Last successful run.** The most recent date and time at which the job was run successfully. All of the model evaluation views only display information for the most recent job that was run successfully. A full job history, including any previously run jobs or failed jobs is available in the job history view. See the topic "Job history view" on page 133 for more information.

Champion challenger view

The Champion Challenger view describes files that were compared to one another within a Champion Challenger job step. The file containing the scoring branch that was deemed to be the most effective is designated as the champion. The champion file appears as the uppermost row in the Champion Challenger view. To see all of the branches analyzed in the Champion Challenger job step, simply expand the tree.

It is important to note that one of the scoring branches in the list was designated the champion. The champion file is then renamed in accordance with the parameters specified on the Champion tab of the champion-challenger job step. Thus, the renamed file appears in the uppermost row.

The Champion Challenger table contains the following information.

- **Champion.** Champion-challenger files are organized around the champion. Thus, the first line introducing a set of champion-challenger files contains the name and path information of the champion file. The files underneath the champions are challenger files.
- **Is Copy.** This column indicates whether or not the champion file listed in the first row of the current set of champion-challenger files is a copy of the winning file. (The winning file is designated by a star in the Index column.) If *Yes* appears in this field, then a copy of the champion file is displayed in the first row. By default, IBM SPSS Deployment Manager makes a copy of the winning champion file. This setting can be disabled when creating or editing a champion-challenger job step. If the option to copy the champion file was cleared, then the winning champion file appears in the first row under the stream name.
- **File.** The file containing the scoring branch.
- **Version.** The version of the file used to generate the results.
- **Index.** A percentage value. If a star appears next in the Index column, this file was the winning file in the champion-challenger analysis.

- **Trend.** The percentage change in the model. An arrow indicates whether the trend is up or down. If the trend is 0.00, a horizontal bar appears, which represents no change. The first time the file is run, the trend field is blank.
- **Author.** The author of the file.
- **Type.** The type of analysis. Examples include *gains* and *accuracy*.
- **Source.** The name of the source node that supplies the data for the scoring branch. This may not be the original source node defined in the file. For Champion Challenger steps, a source node from any of the challengers may be used. The column values include the name of the file containing the source node as a prefix before the node name to identify the location of the source node used.
- **Data.** The actual data used for the source node. The value depends on the source node type. For example, for an ODBC node, the value is a DSN name. For a Variable File node, the value is a file name.
- **Job.** The job that references the file.
- **Job Version.** The version of the job used to analyze the model.
- **Job step.** The name of the job step that references the file.
- **Last successful run.** The most recent date and time at which the job was run successfully. All of the model evaluation views only display information for the most recent job that was run successfully. A full job history, including any previously run jobs or failed jobs is available in the job history view. See the topic "Job history view" on page 133 for more information.

Predictors view

Both champion challenger and model evaluation files use predictors to generate results. A predictor is a variable that is used as an input for a model. Typically, a model contains multiple predictors. Predictors are evaluated and then ranked based on their relevance to the results.

To view the Predictors table:

1. In the Content Explorer, select an object.
2. From the **View** menu select:

Show View > **Predictors**

The Predictors table appears.

The predictor effectiveness table contains the following information. Some parameters apply at the predictor level. Other columns in the table apply to the entire file.

Predictor-level information

Name. The predictor name. Predictors are organized by file. Thus, the first line introducing a set of predictors contains the file name and path information.

Source. The data source used to obtain predictors.

Value. The value of the predictor.

Rank. The rank of the predictor. Predictors are listed in order of descending rank. The rank of a predictor corresponds to its importance within the model. For example, if the a predictor for household income is listed first on the list, then for this model, the level of household income was most highly correlated with a positive response.

File-level information

Version. The version of the file used to generate the results.

Author. The author of the file.

Job. The job that references the file.

Last successful run. The most recent date and time at which the job was run successfully. All of the model evaluation views only display information for the most recent job that was run successfully. A full job history, including any previously run jobs or failed jobs is available in the job history view. See the topic "Job history view" on page 133 for more information.

Note: The Predictors view presents information about IBM SPSS Modeler files stored in the repository. To store these files in the repository, the IBM SPSS Collaboration and Deployment Services environment must include IBM SPSS Modeler adapters. For information on installing the adapters, see the IBM SPSS Modeler documentation.

Filters

Any of the available views can be filtered to reduce the number of results that appear in the view table. Although some filtering options are common to all views, specific filtering parameters may vary by the view type. Multiple filters can be used simultaneously. For example, the job filter is often used in conjunction with other filters.

By default, filtering is disabled. Once filtering is enabled, the ability to filter the job history persists across IBM SPSS Deployment Manager sessions. In addition, filter settings persist across server connections. For example, if filtering is enabled and the server selection is changed from server A to server B, the filter settings established for server A persist for server B.

Filters common to all status views

The following filters are common to all views:

Job. The jobs that appear in the view table can be limited to the following:

- **Selected job in job editor.** Filters the table, so that only information for the currently selected job appears.
- **User-defined job.** Used to search for a job schedule by name. Selecting **Browse** allows a search of the entire content repository.

Version Label. Limits the list to objects containing the specified version label. Typically, this option is used in conjunction with the another filter. For example, if the version label *LATEST* were chosen without a corresponding job filter, every job in repository would appear in the job history because the current version of any job is implicitly labeled *LATEST*. For the model management filters, the results can be limited by job version label or file version label.

Job schedule filters

To access job schedule filters, click the **Filters** button in the job schedule view.

Enable Filtering. If filtering has not been previously enabled, select the **Enable Filtering** check box.

The job schedule can be filtered based on any of the following criteria:

Job. The jobs that appear in the job schedule can be limited to the following:

- **Selected job in job editor.** Filters the job schedule table, so that only the history for the currently selected job appears.
- **User-defined job.** Used to search for a job schedule by name. Selecting **Browse** allows a search of the entire content repository.

Version Label. Narrows the list to jobs containing the specified version label. Typically, this option is used in conjunction with the job filter. For example, if the version label *LATEST* were chosen without a corresponding job filter, every job in the repository would appear in the job schedule because the current version of any job is implicitly labeled *LATEST*.

Last Run Status. Narrows the list to jobs containing the selected status. The following status options can be specified:

- Success

- Failure
- Canceled

Schedule type Enables filtering by schedule type (message-based or time-based).

- **Time-Based Fields**. Selecting the option allows you to specify the criteria for filtering time-based schedules. This limits the view to schedules having a next start time within the specified time interval.

Table 12. Time-based filtering options

Field	Description
Relative Time Period	A finite time period relative to the current date. The following options are available: – Current day – Current week – Current month – Current quarter – Current year – Most recent 30 days – Most recent 3 months – Most recent 6 months – Most recent 9 months
Absolute Time Period	A specific date-based range of time. A valid range must be provided--for example, the end date must be equal to or greater than the start date.

- **Message-Based Fields**. Selecting the option allows you to specify the following criteria for filtering message-based schedules:

Table 13. Message-based filtering options

Field	Description
Message Domain	Filters schedules based on the JMS topic of the corresponding messages.
Message Text	Filters schedules based on the text of the corresponding JMS text messages.
Message Selector	Filters schedules based on the contents of the message headers, for example `NewsType='Sports'` or `NewsType='Business'`

Job History Filters

To access job history filters, click the **Filters** button in the job history view.

Enable filtering. If filtering has not been previously enabled, select the **Enable Filtering** check box.

The job history can be filtered based on any of the following criteria:

Job. The jobs that appear in the job history can be limited to the following:

- **The selected job in job editor.** Filters the job history table, so that only the history for the currently selected job appears.
- **A user-defined job.** Used to search for a job history by name. Selecting **Browse** allows a search of the entire content repository.

Version Label. Narrows the list to jobs containing the specified version label. Typically, this option is used in conjunction with the job filter. For example, if the version label *LATEST* were chosen without a corresponding job filter, every job in repository would appear in the job history because the current version of any job is implicitly labeled *LATEST*.

Last Run Status. Narrows the list to jobs containing the selected status. The following status options can be specified:

- Success
- Failure
- Canceled

Mode of execution. Narrows the list by mode of execution. Valid values include:

- Scheduled—Jobs that were previously scheduled and ran automatically at the specified date and time.
- Manual—Jobs that were executed manually using **Run Job Now**.
- Submitted—Jobs that were submitted for execution by an external source.

Start Date. Narrows the job history list based on the date the job was started. A relative or absolute time period can be specified.

- **Relative time period.** A finite time period relative to the current date. The following options are available:
- Current day
- Current week
- Current month
- Current quarter
- Current year
- Most recent 30 days
- Most recent 3 months
- Most recent 6 months
- Most recent 9 months
- **Absolute time period.** A specific date-based range of time. A valid range must be provided--for example, the end date must be equal or greater than the start date.

Model management filters

The results that appear in the model management view can be filtered based upon the following criteria.

Trend. The trends that appear in model management views can be limited to the following:

- **Direction.** Shows results having the same trend direction. For example, you can limit the view to all results trending up.
- **Values.** The range of the percentage change. A minimum and maximum percentage must be specified.

File. Limits the results to scoring branches contained in a specific file.

Index. A percentage value. A minimum and maximum percentage must be specified.

Evaluation type. Valid values include *Gains*, *Accuracy*, and *Accreditation*.

Data. The data used for the source node

Model Execution Date. A range of execution dates for scoring branches. Specify the start and end dates for the range.

Server status view

The status of servers and server clusters is available in the Server Status view. This view provides metadata about servers and server clusters for which you have permissions.

Information cannot be modified in the server status table. Any modifications to server or server cluster configurations must be done in the *Resource Definitions* folder. See the topic Chapter 6, "Resource definitions," on page 45 for more information.

To view the status for a server, from the **Server** drop-down list, select a server. A list of servers appears.

The server status table contains the following information:

Host. The host name for the server.

Port. The port number for the server.

Type. The type of server; examples include *IBM SPSS Statistics server*, *IBM SPSS Modeler server*, *Remote process server*, and *Cluster*.

Up Since. The last date and time at which the server was last started.

State. The status of the server--for example, *running*.

Connections. The number of connections into the server.

Weight. The weight of the server within a server cluster.

Viewing Servers by Cluster. By default, the server status table contains a list of servers. To view servers organized by cluster, click the show clusters icon. For server clusters, the host cell is blank. To view the servers within a cluster, click the **+** sign to expand the list. The host name for each server within the cluster appears in the table.

Setting the refresh rate

By default, the server status view is not automatically refreshed. To modify the refresh rate:

1. From the **Refresh Rate** drop-down list, select the refresh rate increment--for example, *15 minutes*.

Chapter 13. Notifications and subscriptions

IBM SPSS Collaboration and Deployment Services provides the mechanisms of notifications and subscriptions for keeping the users informed about changes to IBM SPSS Collaboration and Deployment Services Repository objects and job processing results.

Notifications and subscriptions generate email messages, RSS feeds, or JMS (Java Messaging Service) messages when corresponding events occur. Notifications can be defined for multiple recipients while subscriptions to IBM SPSS Collaboration and Deployment Services Repository objects are defined by the system users only for themselves.

Notifications and subscriptions are retained when IBM SPSS Collaboration and Deployment Services Repository content is exported and imported.

Notifications

Notifications can be defined for changes to the IBM SPSS Collaboration and Deployment Services Repository content and processing events.

File notifications are triggered by the creation of new file and job versions. Folder notifications are triggered by folder content changes, such as addition of files, creation of new file and job versions, and creation of subfolders. Already existing notifications can be modified or removed for a single object or for multiple repository objects. See the topic “Content notifications” on page 142 for more information.

A special kind of notifications are the notifications for content label events. These notifications are triggered when certain labels are applied to IBM SPSS Collaboration and Deployment Services Repository objects. See the topic “Label event notifications” on page 143 for more information.

Notifications can also be defined for job processing events such as job and job step success and failure. See “Job success and failure notifications” on page 141, “Job step notifications” on page 142 and “Notifications based on model evaluation return codes” on page 143 for more information.

A user must be assigned the corresponding action through roles to define notifications.

Job success and failure notifications

IBM SPSS Collaboration and Deployment Services Repository provides the capability to send email notification messages for a job success or failure.

To display job notifications:

1. Open the job.
2. Click the **Notifications** tab. The Notifications for Success and Notifications for Failure fields display the number of recipients defined for each type of notification.

To display the list of notification recipients, place the mouse pointer over the Update button.

To specify or update job notifications:

1. On the **Notifications** tab, click **Update** next to the notification type. The Notifications dialog box appears.
2. Specify the notification sender and recipients, and customize the message as necessary. See the topic “Notification settings” on page 143 for more information.
3. Save the job.

Note: Notifications can also be defined for jobs as content repository objects. Such notifications will be triggered by the creation of a new version of a job. See the topic “Content notifications” on page 142 for more information. For example, defining job success or failure notifications for a job or a job step creates a new version of the job and subsequently triggers notification for the change.

Job step notifications

IBM SPSS Collaboration and Deployment Services Repository provides the capability to set up notifications for the success and failure of individual job steps, as well as jobs.

A special case of job step notifications are notifications of success and failure of individual iterations; such notifications can be defined only for iterative job steps.

To display job step notifications:

1. Open the job.
2. Select the job step.
3. Click the **Notifications** tab. The Notifications for Success and Notifications for Failure fields display the number of recipients defined for each type of notification.

To display the list of notification recipients, place the mouse pointer over the Update button.

To specify or update job step notifications:

1. On the **Notifications** tab, click **Update** next to the notification type. The Notifications dialog box appears.
2. Specify the notification sender and recipients, and customize the message as necessary. See the topic "Notification settings" on page 143 for more information.
3. Save the job.

Content notifications

Content notifications keep users informed about changes to IBM SPSS Collaboration and Deployment Services Repository objects, such as the creation of new versions of files and jobs, and folder structures, such as the creation of new subfolders. Content notifications can be defined for a single object or multiple objects selected simultaneously.

To define or modify file or job notifications:

1. Right-click the file or job name and choose **Notifications**. To select multiple objects, hold down `Shift` or `Ctrl` key. The Notifications dialog box appears.
2. In the Notifications dialog box, specify the sender and recipients, customize the message as necessary, and, for folder notifications, indicate folder options. See the topic "Notification settings" on page 143 for more information.

To define or modify folder notifications:

1. Right-click the folder name and choose **Notifications**, then click **Folder Events** to define notifications for folder structure events or **Folder Content Events** to define notifications for changes to the objects in the folder.
2. In the Notifications dialog box, specify the sender, recipients, customize the message as necessary and, for folder notifications, indicate folder options. See the topic "Notification settings" on page 143 for more information.

To remove content notifications:

1. Right-click the file, job, or folder name and choose **Notifications**. The Notifications dialog box appears. Content notifications can also be removed for multiple objects selected simultaneously. To select multiple object, hold down `Shift` or `Ctrl` key.
2. In the Notifications dialog box, click **Remove Notification**. See the topic "Notification settings" on page 143 for more information.

Notes:

- The **Notifications** option is not available if objects of different kinds are selected simultaneously (for example, files and folders, folders and jobs).

- IBM SPSS Collaboration and Deployment Services jobs cannot be opened outside IBM SPSS Deployment Manager. If content notifications are defined for jobs, it is strongly recommended to modify the notification settings to remove the hyperlink to the job because it cannot be opened in a Web browser. See "Customizing the notification message" on page 146 for more information.
- If you import files and folders that were previously exported, no notifications are sent for any import events.

Notifications based on model evaluation return codes

Notifications can be set up to monitor specific model evaluation outcomes. The following are examples of when notifications might be based on model evaluation return codes:

- Notifications can be set up based on a specific outcome of the model evaluation. For example, if the result returned by a model evaluation analysis is less than 0.85, notifications could be sent to a predefined list of recipients.
- Notifications can be set up based on the traffic light level of a model evaluation job. For example, if the result is red, a notification could be sent. In addition, multiple notification rules could be set up with different email recipients for each traffic light level.

Label event notifications

IBM SPSS Collaboration and Deployment Services Repository provides the capability to set up notifications for label events.

For example, an analyst can set up the *Production* label to notify the repository administrator when the label is applied to a job so that the job can subsequently be promoted to production status. Label event notifications can be limited to certain content types, for example, repository jobs or SPSS Statistics files.

To access label event notifications, in IBM SPSS Deployment Manager, right-click the Content Repository (root) folder and choose

Notifications > **Label Events...**

The **Label Events** dialog box appears. The dialog allows you to add, change, and delete the notifications for label events.

To add a label event notification, click the **Add a label event** button. Specify the notification sender and recipients, the label and content types, and customize the subject and body of the message as necessary. See the topic "Notification settings" on page 143 for more information.

To change a label event notification, select a label event notifications entry and click the **Change selected label event** button. Specify the notification sender and recipients, the label and content types, and customize the subject and body of the message as necessary. See the topic "Notification settings" on page 143 for more information.

To delete a label event notifications, select a label event notifications entry and click the **Delete selected label event** button.

Notification settings

Notification settings, including the notification sender, recipients, message customizations, and folder options, are specified in the Notifications dialog box.

From. The address of the sender notification message. The field is prepopulated with the default email address in repository configuration options.

To. The list of notification recipients. The addresses can be entered, selected from a directory listing of a supported email application such as Lotus Notes, or selected from available IBM SPSS Collaboration and Deployment Services users. To edit an address, click the ellipsis button next to it in the list. The list also specifies whether the modified object is to be attached to the message for each individual recipient. See

the topic “Notification message attachments” on page 146 for more information. To remove a recipient, click **Delete**.

Subject. The subject of the notification message. By default, the field is populated by the default template associated with the notification event. Modify the message subject, if necessary. See the topic “Customizing the notification message” on page 146 for more information.

Message. The body text of the notification message. The field is populated by the default template associated with the notification event. Modify the message body text, if necessary. See the topic “Customizing the notification message” on page 146 for more information.

Remove Notification. Click to remove notification. When multiple objects are selected, clicking this button will remove all notifications for the objects.

Apply to subfolders. For folder notifications, applies the subscription to subfolders.

Label. For label event notifications, the label triggering the notification.

Notify when changes are made to these file types only. For folder and label event notifications, limits notification to selected file types. For example, notifications can be set up only for repository jobs or IBM SPSS Statistics files. Click the ellipsis button to open the **File Types** dialog box. The **File Types** dialog opens. Click the file type entries to select or clear. Use the Ctrl or Shift keys to select multiple entries. After all file types have been selected, click **OK**.

Preview. Click to preview the notification message. See the topic “Notification message preview ” on page 146 for more information.

Note:

- When notifications are modified for multiple objects, the list of recipients also displays a number next to the email address indicating how many of the selected objects have the recipient in their notification settings. For multiple objects with different notification settings, the fields with varying values are blank and marked with alert signs. When notifications are modified for multiple folders with varying folder option settings, the options are marked with green squares.
- If you use the Security Subscriber option to select notification recipients, email notifications are not sent to members of IBM SPSS Collaboration and Deployment Services groups. You must select individual users instead of groups, or define a distribution list associated with the IBM SPSS Collaboration and Deployment Services group (and IBM SPSS Collaboration and Deployment Services security must be capable of retrieving the email address defined for the group).

Entering notification recipient email addresses

To enter a recipient email address, on the Notification screen, click the ellipsis button next to the To field or click the down arrow next and then choose **Email Subscriber**. The Set email dialog box appears.

Email Address. The email address of the notification recipient. The address string must contain a `@` character followed by a period.

1. Enter the email address. Only one address can be entered at a time.
2. Click **OK**. The address will appear in the To list of the Notifications dialog box.

Selecting notification recipients from Lotus Notes

To add a notification email address from Lotus, in the Notifications dialog box, click the down arrow next to the To field and then choose **Lotus Subscriber**. You will be prompted for the Lotus Domino Server credentials, and then the Lotus Subscribers dialog box appears.

1. To select a recipient, highlight a directory entry in the list and click **To ->**. The address appears in the To-> field. To remove a recipient, select the entry in the To-> field and click `Remove`. Click `Remove All` to clear all recipients.
2. After all recipients have been selected, click **OK**. The email addresses of the selected recipients appear in the To list of the Notification for Job Success dialog box.

Selecting notification recipients from security subscribers

To send notifications to IBM SPSS Collaboration and Deployment Services users or groups, in the Notifications dialog box, click the down arrow next to the To field and then choose **Security Subscriber**. The Select User or Group dialog box appears.

1. From the **Select provider** drop-down list, select the entity that contains your user and group information.
2. In the Search field, type the first few letters of the user ID or group you want to add. To search for all available user IDs and groups, leave this field blank.
3. Click **Search**. The users and groups that correspond to your search appear in the dialog box.
4. Select one or more users or groups from the list.
5. Click **OK**. The users or groups appear in the To field of the Notifications dialog box.

To receive email notifications, users must specify a valid email address in their user preferences or on the supported system (LDAP, for example).

Note: If you use the Security Subscriber option to select notification recipients, email notifications are not sent to members of IBM SPSS Collaboration and Deployment Services groups. You must select individual users instead of groups, or define a distribution list associated with the IBM SPSS Collaboration and Deployment Services group (and IBM SPSS Collaboration and Deployment Services security must be capable of retrieving the email address defined for the group).

Users can also receive syndicated feeds (RSS/Atom feeds) if they have been granted the *Access Syndicated Feeds* action by an administrator. The feeds will be aggregated individually for the given security principal, based on individual subscriptions (such as file subscriptions) and notifications (such as job completion events and folder content events) created by administrators for this principal or any security group the principal belongs to. The user must be authenticated to access the feed. If authentication is successful and the user has been granted the *Access Syndicated Feeds* action, IBM SPSS Collaboration and Deployment Services will aggregate the feed based on the security principal identifier and security groups.

For more information about RSS feeds, see the IBM SPSS Collaboration and Deployment Services Deployment Portal help.

JMS subscriber

Notification events can generate JMS (Java Messaging Service) messages that trigger internal IBM SPSS Collaboration and Deployment Services processing as well as external applications.

For example, a IBM SPSS Collaboration and Deployment Services user can set up JMS notification for a IBM SPSS Modeler stream to rerun a job which includes the stream when a new version of the stream is created. For that, a JMS message domain must first be created. Then a message-based schedule must be set up for the job specifying the resource ID of the stream in the message selector. See the topic "Message-based schedule settings" on page 127 for more information. And finally, a notification must be set up for the IBM SPSS Modeler stream based on the JMS subscriber to be triggered when a new version of the stream is created.

To specify a JMS subscriber, in the Notifications dialog box click the down arrow next to the To field and then choose **JMS Subscriber**. The JMS Subscriber dialog box appears.

Message Domain. The JMS message domain for the subscriber. See the topic "Message domains" on page 52 for more information.

1. Select the message domain. A message domain can only be selected once for a given notification. To define a new message domain, click **New ...**.
2. Click **OK**. The JMS subscriber will appear in the To list of the Notifications dialog box.

Notification message attachments

Content notification messages can include file attachments. If an attachment is specified for a file or a job notification, the new version of the file or the job is included in the message. If an attachment is specified for a folder notification, the object that triggers folder notification (new file or job, or the new version of an existing file or job) is included in the message. Attachments are specified for individual notification recipients separately.

To include a file attachment in a notification message:

1. In the To list, click in the Attachment column next to the notification recipient and choose **Yes** or **No** from the drop-down.

Customizing the notification message

Notification message can be customized for individual notifications. The subject line and body text of the notification email message are defined by the default message template assigned to the corresponding event type. The Subject and Message fields of the Notification screen contain the default text, variable property values, and for HTML templates, formatting markup. If multiple objects are selected with varying sender, subject, and body text field values, the fields are displayed as blank. Modifying the values of these fields will apply the changes to all selected objects.

1. To change the subject or message body text or formatting, type the changes in the corresponding fields.

   If you make changes to the HTML formatting, you must make sure that the changes are valid. Invalid template may cause notification to fail.

2. You can use content assist to insert system property variable values into the From field, message subject, or body. See the topic “Entry field content assist” on page 10 for more information.
3. Click **Preview** to display the message. If the template cannot be parsed, an error message is displayed. See the topic “Notification message preview ” on page 146 for more information.

The customized message is saved for a specific notification. Customizing messages for individual notifications does not alter the default templates.

Note: The appearance and content of the notification message can also be customized by modifying the default template associated the notification event. Contact your IBM SPSS Collaboration and Deployment Services administrator to request global template changes.

Notification message preview

To preview notification on the notification screen, click **Preview**. The Notification Preview dialog box appears.

The preview displays the email message resulting from merging the template with sample values. The preview includes the specified sender and recipient addresses and customizations made to the message subject, text, and formatting. See the topic “Customizing the notification message” on page 146 for more information. If the customized message cannot be previewed due to invalid formatting, an error message appears.

Subscriptions

IBM SPSS Collaboration and Deployment Services Repository enables the user to subscribe to IBM SPSS Collaboration and Deployment Services files and jobs. Unlike notifications, subscriptions are defined by individual users for themselves.

A subscriber receives email messages when changes occur to the contents of the object and a new version is created. The email message contains a link to the URL of the changed repository object or a file attachment. The Content Explorer also provides the ability to manage subscriptions to individual objects. This may be necessary, for example, when an administrator must remove a subscription for a user who is no longer employed with the company. A user must be assigned corresponding actions through roles to subscribe to content and manage subscriptions.

Subscribing to files

You can subscribe to a file and specify whether a link to the file URL or the file attachment will be included in the notification message when the file is updated.

To create a subscription to a file, right-click the file name and choose **Subscriptions**. The **Create Subscription for Selected File** dialog box appears.

1. If the default email address has not been set for the user, click **Update**. The **Set User Preferences** dialog box appears. Update the address. See the topic "Subscription email address" on page 147 for more information.
2. Specify whether a file link or attachment is included in the email message.
3. Click **OK**.

Modifying file subscriptions and unsubscribing from files

Use the Modify Subscription for Selected File dialog to modify the subscription settings or to unsubscribe.

To modify a subscription for a file or a IBM SPSS Collaboration and Deployment Services Repository job, or to unsubscribe, right-click the file name and choose **Subscriptions**. The **Modify Subscription for Selected File** dialog box appears.

Email. The email address of the subscriber.

Link. Include a link to the URL of the file in the notification message.

Attachment. Attach the file to the message.

Modify subscription settings as necessary and click **OK**, or click **Unsubscribe** to remove the subscription.

Subscription email address

Use the Set email Address dialog to change the default email address for subscriptions.

To change the default email address for subscriptions, in the Create/Modify Subscriptions for Selected File dialog box, click **Update**. The **Set Email Address** dialog box appears.

Email Address. The email address of the user. The address string must contain a `@` character and a period. The default address preference for the user will be modified and all existing subscriptions will use the new address. See the topic "Subscriptions recipient" on page 34 for more information.

Note: The **Use email address from directory** option is available only for users authenticated by Active Directory and Active Directory with Local Override. The email address from the directory is used by default.

1. Enter the email address or select the **Use email address from directory** option.
2. Click **OK**. The address will appear in the To list of the Notifications dialog box.

Managing subscriptions

To manage subscriptions to a repository object, right-click the object and choose **Manage Subscriptions**. The Subscriptions for Selected Resource dialog box appears.

The list of subscriptions displays the name of the user and the subscription delivery type (link or attachment). Existing subscriptions can be removed.

1. Select the subscription. Use the **Ctrl** and **Shift** keys to select multiple entries.
2. Click the **X** button.
3. Click **OK**.

Delivery failures

Notification Message Delivery Failures

If a notifications email messages cannot be delivered to a specified recipient, a warning message similar to the following will be generated.

```
Your message did not reach some or all of the intended recipients.

      Subject:    IBM SPSS Collaboration and Deployment Services: Job ChurnAnalysis failed
      Sent:    4/5/2010 2:21 PM

The following recipient(s) could not be reached:

      jsmith@mycompany.com on 4/5/2010 2:21 PM

 The email account does not exist at the organization this message was sent to.
 Check the email address, or contact the recipient directly to find out the correct address.
 <smtp.mycompany.com #5.1.1>
```

The message is sent to the address indicated in the **From** field of the Notifications dialog box. If the address is invalid, the message is sent to the default IBM SPSS Collaboration and Deployment Services administrator's address. To correct the problem with the notifications delivery, verify that notifications recipients are specified correctly. Notification delivery problems can also occur due to email server setup, network configuration, and so on. If notifications delivery problems persist, contact your system administrator.

Subscription Message Delivery Failures

If you do not receive messages from subscriptions that have been defined, verify that your default subscriptions address is specified correctly. See the topic “Subscriptions recipient” on page 34 for more information. If subscription message delivery problems persist, contact your IBM SPSS Collaboration and Deployment Services administrator. The administrator is also notified of any delivery failures by system-generated messages.

Chapter 14. Visualization report job steps

A visualization report generates a visual representation of the information contained in a data source. The report is typically created using the IBM SPSS Visualization Designer, which provides an advanced visualization environment enabling users to create graphs ranging from basic business charts to rich, interactive representations using data from a variety of sources. Within the IBM SPSS Visualization Designer, users can access and explore data and define both the structure and the style of graphs. The application provides a range of deployment options for the visualization specifications and resulting graphs; for example, they can be stored in IBM SPSS Collaboration and Deployment Services Repository. Once in the repository, graph specifications can be associated with data as needed, rendered on demand, rendered on a recurring basis, and served to applications and web pages as required.

To add a visualization report step to a job, drag a `.viz` file from the Content Explorer to the job canvas. However, before the step can successfully run, several properties of the visualization reporting step must be defined. These properties can be classified into several categories:

- General properties, such as the job step name and version of the report.
- Data source properties, including the data source used by the report and its credentials.
- Type properties, such as whether the report type is a single run or an iterative run.
- Parameter properties, such as prompt variable values.
- Result properties, such as output file format and location.
- Clean-up properties, for automatically moving, deleting, or expiring existing output when a job runs.
- Notification properties, for setting up email notifications.

General properties for visualization report steps

The general properties for a visualization report step identify what runs when the job executes. To define general properties for the step, click the step in an open job. Click the **General** tab to view or edit the step properties.

Job Step Name. Type a name for the step. The default name appends a *_step* suffix to the file name.

Report Definition. Displays the location of the report used for the step.

Version Label. Select the version of the report submitted for processing using the drop-down list.

Data source for visualization report steps

Visualization reports support a variety of input data sources, including:

- Delimited text files
- IBM SPSS Statistics data files (`.sav`)
- Dimensions Data Models
- JDBC sources

Data sources based on files can access files stored on file systems or in the IBM SPSS Collaboration and Deployment Services Repository. For data sources based on reports or JDBC, you need to specify the data source login credentials to access the data. To define data source login credentials, click the step in an open job. Click the **Datasource** tab to view or edit the data source credentials.

Type for visualization report steps

To define report type properties, click a visualization report step in an open job. Click the **Type** tab to view or edit the step properties.

On the **Type** tab, select the report type for the reporting step.

Single. Runs once and returns a single output file. If the report contains variables, you can specify values for the prompt variables on the **Parameters** tab (see "Parameters for visualization report steps" on page 150 for more information). At run time, the report generates output based on the specified variable values.

Parameters for visualization report steps

Report data sources may contain variables, or prompts, that need an assigned value for proper processing. At run time, the report generates output based on the specified prompt variable values. The number of prompt variables available for selection is dependent on the number of prompts in the source file and the options used to define those prompts when the file was created. You can only assign prompt variables if your source file contains prompts.

To define prompt variables, click a reporting step in an open job. Click the **Parameters** tab to view or edit the step properties.

With your cursor in the Assignment column, click the button on the edge of the column. The Prompt Variable Values dialog box will be displayed. Select the variable value to be used for the report. Select **N** if no value is required for this prompt variable at run time, and you don't want to restrict the report results based on this prompt value (often referred to as a "null" value).

Results for visualization report steps

Execution of a reporting step yields report output. To define properties for this output, click a reporting step in an open job. Click the **Results** tab to view or edit the step properties.

File Name. Define the file name for the results. You can insert variables into the file output name, and various other fields. To insert variables, position your cursor in the target location in the *File Name* field, type **$**, and select one or more values from the drop-down. These variables can help users distinguish between result output. You can insert variables for information such as the date or time. When the report runs as scheduled, the variable information is inserted into the output file name(s). For example, if you have a scheduled report running daily, a user may have trouble determining the information they are viewing. Adding a date variable will insert the current date into the file name at run time, indicating the date the report ran. For more information, see "Entry field content assist" on page 10.

Location. Define the location in which to save the results by typing a path or clicking **Browse**. The Results Location dialog box will appear. If the specified folder does not exist, it will be created automatically at run time. To insert variables to append the date or time to the folder name at run time, see "Entry field content assist" on page 10.

Format. Select the file format for reporting step output. For example, select PDF to produce Portable Document Format (`*.pdf`) output.

Permissions. Define access permissions for the results by clicking **Browse**. The Output Permissions dialog box will appear. See the topic "Modifying permissions" on page 26 for more information.

Metadata. Define output properties for the results by clicking **Browse**. The Output Properties dialog box will appear. See "Output file metadata" on page 124.

Dimensions. Define the size of the resulting visualization output by specifying values for the output height and width. Alternatively, select **Use Default Dimensions** to accept the default output size.

Clean-up for visualization report steps

To define options for cleaning up existing output when a job runs, select a job step in an open job. Click the **Clean-up** tab to view or edit the step's clean-up options.

You should specify clean-up options every time you create a new job step, and then disable or enable clean-up later as necessary.

Important: Clean-up options are only applied if saving to the repository.

Choose from the following options.

- **No clean-up**. Output files from the previous run of the current job remain intact.
- **Clean-up Options**
  - **Delete**. Output files from the previous run of the current job will be deleted (even if they have since been moved to a different folder in the repository).
  - **Expire**. Output files from the previous run of the current job will be expired. See the topic "Working with Expiration Dates and Expired Files" on page 29 for more information.
  - **Move to**. Output files from the previous run of the current job will be moved to a different folder. Click **Browse** or type a target folder for old output files. If a file with the same name already exists in the **Move to** folder, it will be overwritten. If the specified folder does not exist, it will be created automatically during the clean-up process. To insert variables to append the date yyor time to the file name at run time, see "Entry field content assist" on page 10.
- **Label applied to cleaned-up versions**
  - **No label**. Select this option to specify no label is applied to the moved output.
  - **Specify label**. Select this option to specify a new label to apply to the moved output.
  - **Select label**. Select this option to select an existing label to apply to the moved output.
- **On Clean-up Error**
  - **Stop clean-up, do not run the job, mark as failed**. Select this option if you want the job to fail immediately if it encounters errors while attempting clean-up.
  - **Continue clean-up and run the job**. Select this option if you want the job to keep running if it encounters errors while attempting clean-up. If the job runs successfully, the clean-up error will not trigger a failed job history message.
  - **Continue clean-up, run the job, mark as failed**. Select this option if you want the job to keep running if it encounters errors while attempting clean-up. If the job runs successfully, the clean-up error will trigger a failed job history message.

Notifications for visualization report steps

To define email notification properties, click a reporting step in an open job. Click the **Notifications** tab to view or edit the step properties.

For details about notifications, see "Job step notifications" on page 142.

Chapter 15. SAS® job steps

A SAS step submits SAS syntax to a SAS execution server. This server corresponds to running SAS from the command line, using invocation options to control processing.

Step output consists of a log file detailing the processing history and a text listing file containing the procedure results. The Output Delivery System (ODS) can be used within the syntax to generate HTML results.

To add a SAS step to a job, drag a SAS syntax file from the Content Explorer to the job canvas. However, before the step can be successfully run, several properties of the SAS step must be defined. These properties can be classified into two categories:

- General properties, such as the execution server to process the step
- Result properties, such as output format and location

General properties for SAS steps

The general properties for a SAS step specify what gets run and which server processes the step when the job is run. To define general properties for a SAS step, click the step in an open job. Click the **General** tab to view or edit the step properties.

Job step name. Type a name for the step. The default name appends a *_step* suffix to the file name.

Object version. Select the version of the file submitted to the execution server using the drop-down list.

Additional arguments. Define optional system options that get passed to the SAS execution server when the job is run.

Warning expression. Define warnings for job steps connected by a Conditional connector. The warning expression (e.g., `completion_code`, `warning`, or `success`) must be lowercase.

To use warning expressions:

1. Connect two job steps with a Conditional connector. In the **Expression** field for the conditional connector, type `warning==true`.
2. Navigate to the General tab of the parent job step.
3. In the **Warning expression** field, specify a warning code--for example, `completion_code==18`. This expression overrides the default warning code, if any.

When the job is run, the system will execute the parent job step. Then the system will evaluate the condition for `warning==true`. If true, the system will look at the warning expression specified and determine if the condition was met. If the condition specified in the warning expression was met, the system proceeds to the next job step.

For SAS job steps, the default completion code is `completion_code==1`.

Note: If a SAS job step fails because SAS execution server is not defined correctly and the default warning code of 1 is used, the job step status will still be displayed as *Success* in Job History table. This is due to the conflict between Windows warning code 1 (an executable path cannot be found) and SAS default warning code value.

SAS server. Use the drop-down list to select a SAS server to process the step. The list contains all servers currently configured to execute SAS steps. To add a new server to the list, click **New**. See the topic "Creating new Content Server connections" on page 14 for more information.

Remote Process Server. If SAS is not installed on the machine hosting IBM SPSS Collaboration and Deployment Services Repository, use the drop-down list to select a IBM SPSS Collaboration and Deployment Services Remote Process Server to process the step. The list contains all servers currently configured as remote process servers. The machine hosting SAS needs to be configured as a remote process server and defined as such in IBM SPSS Collaboration and Deployment Services. To add a new

server to the list, click **New**. See the topic "Creating new Content Server connections" on page 14 for more information.

Additional arguments for SAS steps

The additional arguments list for SAS job steps corresponds to the system options that can be supplied to the executable file when invoking SAS.

By default, the additional arguments list includes the two options shown in the following table.

Table 14. Default arguments

Option	Description
-NOSPLASH	Suppresses the splash screen when invoking the SAS execution server.
-$NOSTATUSWIN	Suppresses the status window when invoking the SAS execution server.

Additional options can be added to this list for any SAS job step. Each option must be preceded by a hyphen. See the SAS options documentation at http://v8doc.sas.com/sashtml/lrcon/z0928597.htm for a comprehensive list of available arguments.

Controlling graph output

SAS steps that create graphs without the use of *ODS* to save them with HTML output require the specification of a device driver.

Table 15. DEVICE argument

Option	Description
-DEVICE devicename	Defines `devicename` as the SAS/GRAPH® output device driver.

For example,

```
-DEVICE Jpeg
```

defines the JPEG driver for creating graphs. In addition to `.jpeg`, devices commonly used for exporting graphs include `.bmp` (bitmaps), `.gif` (gif files), and `.wmf` (Windows metafiles). To see a list of devices available for your system, consult the *Devices* catalog for the SAS execution server associated with the step.

Results for SAS steps

Execution of a SAS step yields two general types of results: a processing log and an output listing file. To define properties for these results, click a SAS step in an open job. Click the **Results** tab to view or edit the step properties.

The **Name** column identifies the different result types for SAS steps. Properties for each result type include:

File name. Specify a name for the results file by clicking the cell for the target and typing the new name.

Location. Define the location in which to save the results by clicking the Location cell for the target. Click the resulting ellipsis button to specify the location. Defining result locations for SAS results is identical to doing so for IBM SPSS Statistics results.

Permissions. Define access permissions for the results file by clicking the Permissions cell for the output target. Click the resulting ellipsis button to specify the permissions. See the topic "Modifying permissions" on page 26 for more information.

The contents of the output listing file correspond to the output that would be produced in the Output window by executing the step syntax directly within the SAS environment. Although the syntax may produce additional files, such as exported data files or PMML model files, you cannot define properties for

those files using a SAS step. If control over the files is needed, use a General job step to execute the SAS syntax.

SAS result formats

Both the log and listing output are saved as text files. The listing can be saved as HTML by including ODS statements within the SAS program file.

When using ODS to control output format, if the results are saved to the repository, all output files generated by the step are saved to a single compressed file. The name of the file corresponds to the name specified for the Output Target.

SAS/Graph® results

If the SAS command file associated with a step does not use *ODS* for saving graphs with HTML output, the file must include the proper syntax for exporting the graph output.

1. Use `FILENAME` to create a *fileref* corresponding to the output directory.
2. Assign the *fileref* to the graphics stream using the `GSFNAME` graphic option.

For example, the syntax

```
FILENAME grout 'c:\graph_output';
GOPTIONS GSFNAME=grout;
```

uses the *c:\graph_output* directory for the graphics stream. To use the current working directory, specify:

```
FILENAME grout '.';
GOPTIONS GSFNAME=grout;
```

The format of the graph files depends on the `-DEVICE` option specified as an additional argument for the step. See the topic "Controlling graph output" on page 154 for more information.

To save the step results to the repository, use the current working directory as the *fileref* for the graphics stream. IBM SPSS Deployment Manager stores the graph output within a single compressed file in the repository. To save the graph files individually, use a General job step to process the SAS syntax.

Example SAS step

To illustrate the use of a SAS step in a job, this example uses SAS to compute descriptive output for the data introduced in the IBM SPSS Statistics example.

To enable the SAS step to use the data, we transform it to an IBM SPSS Statistics portable file using the following IBM SPSS Statistics syntax:

```
GET FILE='C:\Program Files\data\Tutorial\sample_files\customers_model.sav'.
EXPORT OUTFILE='C:data\customers_model.por'.
```

Create an IBM SPSS Statistics command file (`sav2por.sps`) containing this syntax and save it to the Content Repository. Create a new job and drag the file to the job canvas.

Click the step to set its properties. On the General tab, specify the IBM SPSS Statistics execution server to process the step. On the Results tab, click the Location cell for the Output Target, click the resulting ellipsis button, and specify a location of *Discard*.

This step is designed to create a new file; the IBM SPSS Statistics output is not of interest.

If the IBM SPSS Statistics step runs successfully, a file named *customers_model.por* will be created in the *C:\data* directory on the server. We will use the file as input to a SAS step. Consider the following SAS syntax:

```
FILENAME CUST DISK 'C:\data\customers_model.por';
LIBNAME CUST SPSS;
PROC FREQ DATA=CUST._FIRST_;
TABLES response*gender;
RUN;
```

These commands read the IBM SPSS Statistics portable file into a SAS data view and create a crosstabulation of *response* and *gender*. Create a SAS command file containing this syntax, save it under the name *frequencies.sas*, and save it to the Content Repository. Drag the file to the job canvas containing the IBM SPSS Statistics step.

Execution of the SAS step depends on the successful completion of the IBM SPSS Statistics step. To enforce this dependency, we create a *Pass* relationship from the IBM SPSS Statistics step to the SAS step. Select the Pass tool from the job palette, click the IBM SPSS Statistics step, and click the SAS step. With this relationship in place, the `sav2por.sps` step must succeed for the *frequencies.sas* step to be invoked.

Define properties for the SAS step by clicking it. On the General tab, select the SAS server for processing the step. On the Results tab, click the Location cell for the Output Target to define the Location. Click the ellipsis button in the cell to access the Location dialog box. For this example, we save the results to the *Trees* folder in the repository. Click **OK** to return to the Results tab.

Save the job and click the **Run Job Now** button to execute it. When it completes, press the F5 key to refresh the Content Explorer. The output and log files appear in the *Trees* folder.

Suppose we wanted to include graphs in the output. Consider the modified syntax below:

```
FILENAME CUST DISK 'C:\data\customers_model.por';
LIBNAME CUST SPSS;
PROC FREQ DATA=CUST._FIRST_;
TABLES response*gender;
RUN;

FILENAME grout '.';
GOPTIONS GSFNAME=grout;

AXIS1
    label=('Frequency')
    MINOR=NONE
    ;
AXIS2
    label=('Response')
    MINOR=NONE
    ;
TITLE;
PROC gchart DATA=CUST._FIRST_;
    vbar3d response /
    GROUP=GENDER
    DISCRETE
    RAXIS=AXIS1
    MAXIS=AXIS2
    AUTOREF
    SHAPE=Block
    PATTERNID=GROUP
    COUTLINE=BLACK
    FRAME
    ;
RUN;QUIT;

PROC gchart DATA=CUST._FIRST_;
    pie3d response /
    GROUP=GENDER
    NOLEGEND
    SLICE=OUTSIDE
    PERCENT=ARROW
    VALUE=NONE
    ACROSS=2
    COUTLINE=BLACK
    ;
RUN;QUIT;
```

These commands add bar and pie charts to the output, exporting them to the working folder as defined by the fileref *grout*. Create a new file named *freq_chart.sas* containing this new syntax and add it to the Content Repository. Delete the *frequencies.sas* file from the job canvas and add *freq_chart.sas*. Add a Pass relationship from the IBM SPSS Statistics step to the new SAS step.

The inclusion of graphs requires the addition of a system option to define the format. Click the SAS step to modify the properties and click the General tab.

Add the `-DEVICE JPEG` option to save the graphs in the JPG format. Save and run the job. When the job completes, press the F5 key to refresh the Content Explorer.

The *Trees* folder contains the results of the run in the file *freq_chart.zip*. The listing file contains the results of the FREQ procedure. The JPG files *gchart.jpg* and *gchart1.jpg* contain the two graphs. To save the three files individually, use a General job step to process the SAS syntax.

Chapter 16. General job steps

General job steps provide a method for scheduling generic executable processes using IBM SPSS Deployment Manager. These processes may involve execution files or batch files and are often controlled using command line options or switches.

General job steps can access files stored within the repository. As a result, stored data or model files can serve as input to processes. In addition, General job steps provide control over individual output files. For example, individual graphs can be stored in the repository by these steps.

To add a General job step to a job, select the General job tool from the job palette and click the job canvas. However, before the job can run successfully, several properties of the step must be defined. These properties can be classified into three categories:

- General properties, such as the command to run
- Input file properties, such as the location of files used as input to the command
- Output file properties, such as the location of files generated by the command

Note: On Windows, if you cancel a General job step that executes a Python script, the Python script execution will not be canceled.

General properties for general job steps

The general properties for a General job step specify the job to run and the directory in which to place intermediate files.

To define General properties for a General job step:

1. Click a General job step in an open job.
2. Click the General tab to view or edit the step properties.

- **Job step name.** Type a name for the step.
- **Iterative Variable List.** If the job step is connected to an iterative producer job step, select the name of the iterative producer from the drop-down list.
- **Command to run.** Type the command to be executed by the step, including the path to the file. Paths containing spaces must be enclosed in quotes or specified using short names, such as *PROGRA~1* for *Program Files*. For diagnostic or troubleshooting purposes, it may be useful to include standard output or standard error redirection in the command specification. Any redirection syntax must match the operating system that executes the command.
- **Working Directory.** Specify the working directory the step should use when dealing with files. Select **Automatic** to use a custom internal directory, or select **Directory** to specify a particular directory.
- **Warning expression.** The warning code expression for the job step. The warning expression (e.g., `completion_code`, `warning`, or `success`) must be lowercase. For example, specifying the warning expression `completion_code==26` instructs the system to treat the return code 26 as a success of the job step execution, although it may be different from the default success code.

Remote Process Server. If the execution file providing the functionality invoked by the job step is not installed on the machine hosting IBM SPSS Collaboration and Deployment Services Repository, use the drop-down list to select a remote process server to process the step. The list contains all servers currently configured as remote process servers. The machine hosting the execution file needs to be configured as a remote process server and defined as such in IBM SPSS Collaboration and Deployment Services. To add a new server to the list, click **New**. See the topic "Creating new Content Server connections" on page 14 for more information.

Input files for general job steps

The Input Files tab identifies files used as input to the General job step. The step copies an input file from its source location to a specified target location for subsequent processing when running the command.

To define Input File properties:

1. Click a General job step in an open job.
2. Click the Input Files tab to view or edit the properties for input files.
3. Click the **Add** button to add an input file to the step.

- **File name.** Type the name for the input file. This name should correspond to the name of the file used when the command is run. By default, a new input file is named *inputFile*.
- **Source Location.** Specify the location of the file. Define this location by clicking the cell and clicking the resulting ellipsis button. See the topic "Input file source location" on page 158 for more information.
- **Source Version.** For input files contained within the repository, specify the version of the file to be copied to the target location.
- **Target Location.** Specify the location to which the step should copy the file. This value should correspond to the location of the file that is accessed when the command is run. Define this location by clicking the cell and clicking the resulting ellipsis button. See the topic "Input file target location" on page 158 for more information.

Input file source location

Click the ellipsis button in a selected Source Location cell to define the location of an input file.

- **Location.** Define the location of the input file as either **Content Repository** or **File System**.
- **Path.** Specify the path to the input file. For input files in the Content Repository, click the **Browse** button to navigate to the path.

Input file target location

Click the ellipsis button in a selected Target Location cell to define the location to which the input file should be copied.

- **Copy to working directory.** Copies the input file from the source location to the working directory defined for the step. See the topic "General properties for general job steps" on page 157 for more information.
- **Copy to another folder.** Copies the input file to a specified folder.

Output files for general job steps

The Output Files tab identifies files generated by the General job step. When the command run completes successfully, the step copies output files from their intermediate locations to specified target locations.

To define Output Files properties:

1. Click a General job step in an open job.
2. Click the Output Files tab to view or edit the properties for output files.
3. Click the **Add** button to add an output file to the step.

- **File name.** The name of the output file. This name should correspond to the name of the file generated by running the command. By default, a newly added output file is named *outputFile*. To change the name, click in the **File Name** column and type the new name.
- **Intermediate Output Location.** Specify the location of the file generated by running the command. Define this location by clicking the cell and clicking the resulting ellipsis button. See the topic "Input file source location" on page 158 for more information.

- **Target Location.** The location of the file. To define location, open the Results Location dialog box by clicking in the column and then clicking the resulting ellipsis button. See the topic "Input file target location" on page 158 for more information.
- **Permissions.** Access permissions for the file if saved to the repository. To define permissions, open the Output Permissions dialog box by clicking in the Permissions column and then clicking the resulting ellipsis button. See the topic "Modifying permissions" on page 26 for more information.
- **Properties.** The properties (metadata) of the file. To define properties, open the Output Properties dialog box by clicking in the Properties column and then clicking the resulting ellipsis button. See the topic "Output file metadata" on page 124 for more information.

Output file intermediate location

Click the ellipsis button in a selected Intermediate Output Location cell to define the location of the output file created by running the command.

The command can create an intermediate file in one of two locations:

- **Working directory.** Identifies the working directory defined for the step as the location for the intermediate file. See the topic "General properties for general job steps" on page 157 for more information.
- **Write to another folder.** Identifies a specific folder as the location for the intermediate file.

General job step examples

Applications of the General job step are numerous and varied.

The following examples illustrate some of the possible uses.

Storing IBM Corp. PMML files

The IBM SPSS Statistics syntax in other examples created a PMML file for the generated tree model, but that file could not be stored in the repository by the IBM SPSS Statistics step. A General job step can execute the IBM SPSS Statistics syntax and save all results to the repository.

To create a General job step for storing:

1. Open a new job.
2. Select the General job tool from the job palette.
3. Click in the job canvas.

Defining General properties

To define General properties for a General job step:

1. Click the step to open the properties dialog box.
2. On the General tab, type a name of *Tree PMML* for the job step.
3. For **Command to run**, type `STATISTICSB -f tree_model.sps -type text -out tree_model.txt`. The switches instruct STATISTICSB to process the syntax file *tree_model.sps*, saving the results as text in the file *tree_model.txt*.
4. Select **Automatic** for the **Working Directory**.

Defining Input File properties

To define Input File properties for a General job step:

1. Click the Input Files tab to define the input for the step. Running the command uses a single input, the syntax file.

2. Add an input file for the step by clicking **Add**. Change the name of this input file to match the name of the syntax file stored within the repository, *tree_model.sps*.
3. Click the Source Location cell and click the resulting ellipsis button to define the location of the file. The syntax file is stored within the repository, so specify **Content Repository** as the Location.
4. Click **Browse** to navigate to the *Trees* folder to define the path to the file.
5. Click **OK** to return to the Input Files tab.

   Note: The syntax file itself references a data file, *customers_model.sav*, stored within the file system on the server. The step does not need to relocate the file so it does not need to be listed as an input. If the data file were stored in the repository or needed to be moved to another system folder for the syntax to access it, it would be listed as an input file for the step.
6. Click the Target Location cell for the input file and click the resulting ellipsis button to define a location to which the step should copy the file. The switch passed to the execution file does not include a path to the syntax file, so select **Copy to working directory** in the Target Location dialog box.
7. Click **OK** to return to the Input Files tab.

Defining Output File properties

To define Output File properties for a General job step:

1. Click the Output File tab to define the output for the step. Running the command creates a text output file, but we are currently not interested in that file. Our focus is on the PMML file generated by the syntax.
2. Add an output file for the step by clicking **Add**. Change the name of this output file to match the name of the file created by the syntax, *tree_model.xml*.
3. Click the Intermediate Output Location cell and click the resulting ellipsis button to define the location of the file. The syntax file creates the PMML file in the *C:\Program Files\SPSS* folder, so select **Write to another folder** in the Intermediate Location dialog box and type the path *C:\temp*.
4. Click **OK** to return to the Output File tab.
5. Click the Target Location cell and click the resulting ellipsis button to define the location for the output file.
6. Select **Save to repository** in the Results Location dialog box and click **Browse** to specify the *Trees* folder.
7. Click **OK** to return to the Output File tab.
8. Save the job and click the **Run Job Now** button to execute the job. IBM SPSS Deployment Manager runs the STATISTICSB command and stores the PMML file in the repository. Press the F5 key to refresh the Content Explorer and view the XML file containing the model.

Batch scoring using IBM SPSS Statistics PMML files

Once a predictive model has been built and the model specifications have been saved as an XML file, the model can be used to score data.

Consider the following syntax:

```
GET FILE='C:\data\customers_new.sav'.
RECODE Age
     ( MISSING = COPY )
     ( LO THRU 37 =1 )
     ( LO THRU 43 =2 )
     ( LO THRU 49 =3 )
     ( LO THRU HI = 4 ) INTO Age_Group.

IF MISSING(Age) Age_Group = -9.
COMPUTE Log_Amount = ln(Amount).

MODEL HANDLE NAME=tree FILE='C:\models\tree_model.xml'
      /MAP VARIABLES=Income_Grp MODELVARIABLES=Income_Group.

COMPUTE PredCatTree = ApplyModel(tree,'predict').

SAVE OUTFILE='C:\scores\scoring_results.sav'.
```

This syntax uses the PMML file *tree_model.xml* to generate scores for the data in *customers_new.sav*. The results are saved to *scoring_results.sav*. Save the syntax into a file named *scoring.sps* and add it to a folder named *Trees* in the Content Repository.

To create a General job step for scoring:

1. Open a new job.
2. Select the General job tool from the job palette
3. Click in the job canvas.

Defining General properties

To define General properties for a General job step:

1. Click the step to open the properties dialog box.
2. On the General tab, type a name of *Scoring* for the job step.
3. For **Command to run**, type `STATISTICSB -f scoring.sps -type text -out scoring.txt`. The switches instruct STATISTICSB to process the syntax file *scoring.sps*, saving the results as text in the file *scoring.txt*.
4. Select **Automatic** for the **Working Directory**.

Defining Input File properties

To define Input File properties for a General job step:

1. Click the Input Files tab to define the input for the step. This step requires two input files:

   - Syntax file
   - PMML file

   Both files are stored in the *Trees* folder of the Content Repository. To add the files, click **Add** to add an input file for the syntax. Change the name of the input file to match the name of the syntax file stored within the repository, *scoring.sps*.
2. Click the Source Location cell and click the resulting ellipsis button to define the location of the file. The syntax file is stored within the repository, so specify **Content Repository** as the Location.
3. Click **Browse** to navigate to the *Trees* folder to define the path to the file. Click **OK** to return to the Input Files tab.
4. Click the Target Location cell and click the resulting ellipsis button to define a location to which the step should copy the file. The switch passed to the execution file does not include a path to the syntax file, so select **Copy to working directory**. Click **OK** to return to the Input Files tab.
5. Click **Add** to add an input file for the PMML file. Change the name of the input file to match the name of the file stored within the repository, *tree_model.xml*.
6. Click the Source Location cell and click the resulting ellipsis button to define the location of the file. The syntax file is stored within the repository, so specify **Content Repository** as the Location.
7. Click **Browse** to navigate to the *Trees* folder to define the path to the file. Click **OK** to return to the Input Files tab
8. Click the Target Location cell and click the resulting ellipsis button to define a location to which the step should copy the file. The syntax requires the model file to be in the *C:\models* folder, so select **Copy to another folder**. Type the path *C:\models* and click **OK**.
9. Click **OK** to return to the Input Files tab.

Defining Output File properties

To define Output File properties for a General job step:

1. Click the Output File tab to define the output for the step. Running the command creates a text output file and the syntax creates a data file containing the scores.
2. Add an output file for the step by clicking **Add**. Change the name of this output file to match the name of the output file, *scores.txt*.
3. Click the Intermediate Output Location cell and click the resulting ellipsis button to define the location of the file. The command creates the file in the working directory, so select **Working directory** and click **OK** to return to the Output File tab.
4. Click the Target Location cell and click the resulting ellipsis button to define the location for the output file.
5. Select **Save to repository** and click **Browse** to specify the *Trees* folder. Click **OK** to return to the Output File tab.
6. Running the command creates a text output file and the syntax creates a data file containing the scores. To add the files to the Content Repository, add another output file for the step by clicking **Add**. Change the name of this output file to match the name of the data file containing the scores, *scoring_results.sav*.
7. Click the Intermediate Output Location cell and click the resulting ellipsis button to define the location of the file. The command creates the file in the *C:\scores* folder, so select **Write to another folder** and type the path *C:\scores*. Click **OK** to return to the Output File tab.
8. Click the Target Location cell and click the resulting ellipsis button to define the location for the output file. Select **Save to repository** and click **Browse** to specify the *Trees* folder. Click **OK** to return to the Output File tab.
9. Save the job and click the **Run Job Now** button to execute the job. IBM SPSS Deployment Manager runs the STATISTICSB command and scores the new data, storing the results in the repository. Press the F5 key to refresh the Content Explorer and view the file containing the scores.

Event-based scheduling

Typically, jobs are scheduled for execution based on a particular time-based recurrence pattern--for example, weekly or monthly. In addition, you can also create event-based (or triggered) steps within a job. In this case, your job would contain an event that must be completed successfully before the next step begins. Essentially, an event-based step is a special type of conditional step.

To use an event-based trigger, you create a job step that monitors the existence or the contents of a file. If the file exists or its contents indicate that a new external event has occurred, the job step executes successfully. The completed job step then passes control to the next step in the job. For example, the first step in your job could be a shell script or a batch file that checks to see if a particular file has been updated. If the file was updated, the step completes successfully. The system then proceeds to the next step in the job.

If you use a file in your event-based step, you may need to know the name of the file and specify the file name as a parameter in the script. For example, you may want to use the event name as the file name.

If the event-based step involves a time component, such as a polling interval, it is important to ensure that the polling interval does not conflict with the scheduling recurrence pattern. For example, if your polling interval is one minute and you have scheduled the job to run every hour, then you need to ensure that the monitoring step is complete before the job is scheduled to run—that is, the event-based step has to either succeed or fail before the next job is executed. In this instance, it would be very useful to set up a timeout value in the monitoring script that is less than the scheduling recurrence pattern--for example, a timeout value of 50 minutes.

Event-based scheduling process overview

Although the individual parameters may vary, the process for implementing an event-based step consists of the following tasks:

1. Create a script that monitors the contents of a file.
2. Create a new job. See the topic “Creating new jobs” on page 115 for more information.

3. Add a General job step to the job that executes the script. You may also need to specify parameters for the script—for example, file name and polling interval. This is the antecedent step. See the topic Chapter 16, “General job steps,” on page 157 for more information.
4. Add another step to the job that depends on the event-based step. This is the consequent step. See the topic “Adding steps to a job” on page 118 for more information.
5. Connect the antecedent step to the consequent step with a pass connector. See the topic “Pass connector” on page 120 for more information.
6. Schedule the job for execution.

Sample event-based scheduling script

The following is a sample shell script that monitors the contents of a predefined file. This script:

- Assumes that a separate process writes a timestamp to the event file when the given event is completed
- Uses this timestamp to determine whether or not the event has been processed
- Returns an exit code of 0 by default, if the script completes successfully

Script arguments

The script takes the following command line arguments, which you can specify in the IBM SPSS Deployment Manager when you create the General job step:

- **CONTROL_FILE**. The script monitors this file for changes. An external process writes events to this file. Periodically, the script checks the file for new events and processes them when found. After a new event has been processed, the script exits successfully. If the script exits successfully, the General job step proceeds to the next step in the job.
- **SLEEP_TIME**. This argument refers to the amount of time, in seconds, that the script sleeps between checks of the *CONTROL_FILE*.
- **TIMEOUT_MINS**. This argument controls the amount of time for which the script should run. If an event has not been processed after the number of minutes specified, script execution stops, and the script returns an exit code of 1. In this example, an exit code of 1 causes the job to fail because the event-based step was not successful. As a result, the job cannot proceed with the remaining steps.

Script Execution

The script executes the following tasks:

1. The script checks the last row of the *CONTROL_FILE* file for an event.
2. If the *RESULT_FILE* exists, the last event has already been processed. The script waits for the designated SLEEP_TIME and then returns to the *CONTROL_FILE* and repeats the process.
3. If the *RESULT_FILE* does not exist, then the script creates a new *RESULT_FILE* for the last event. The script considers the creation of the new *RESULT_FILE* a success and returns an exit code of 0. The General job step then proceeds to the next step in the job.

```
#!/bin/bash
#-------------------------------------------------------------------------
#
# Copyright (c) 2005 SPSS, Inc.
# All Rights Reserved
#
# This software is the confidential and proprietary information of
# SPSS, Inc. ("Confidential Information").  You shall not disclose
# such Confidential Information and shall use it only in accordance
# with the terms of the license agreement you entered into with SPSS.



#
#
#-------------------------------------------------------------------------


# Control file is written to by an external process.
# This is the file we are monitoring.
CONTROL_FILE=$1
```

```
# The duration in seconds we should sleep between
# checks on the control file
SLEEP_TIME=$2

# The timeout (in minutes)
TIMEOUT_MINS=$3

#-------------------------------------------------------------------------
# usage: Print usage information then exit
#-------------------------------------------------------------------------
function usage
{
    echo "usage watch.sh <control file> <sleep duration> <timeout in minutes>"
    exit -1
}

#-------------------------------------------------------------------------
# check_args: Validates the command line arguments
#-------------------------------------------------------------------------
function check_args
{
    if [ -z $CONTROL_FILE ]; then
        usage;
    elif [ -z $SLEEP_TIME ] ; then
        usage;
    elif [ -z $TIMEOUT_MINS ] ; then
        usage
    fi
}


#-------------------------------------------------------------------------
# check_file: Watches $CONTROL_FILE for new events.  If a new event occurs,
# touch an event file.  If not, sleep for $SLEEP_TIME seconds and try again.
#-------------------------------------------------------------------------
function check_file
{
        if [ -f $CONTROL_FILE ] ; then
            # Get the last event from the file
            LAST_EVENT=`tail -1 $CONTROL_FILE`;
            echo "Last event: $LAST_EVENT"
        else
            "$CONTROL_FILE does not exist.  Exiting."
            exit 1
        fi

        # Check to see if we've processed this event
        RESULT_FILE=${0}.$LAST_EVENT;
        echo "Checking for event $LAST_EVENT with file $RESULT_FILE";

        if [ -e $RESULT_FILE ] ; then
                # We've  processed it, sleep and
                # check again later for a new event
                echo "Sleeping for $SLEEP_TIME seconds"
                sleep $SLEEP_TIME;
        else
                # We haven't processed the latest event
                echo "A new event $LAST_EVENT has been fired.  Processing.";
                touch $RESULT_FILE;
                FINISHED=finished;
        fi
}

check_args;


# How long we should run before failing
DURATION=$[$TIMEOUT_MINS*60]

# Get the current time and add our duration to it
NOW=`date +%s`
END_TIME=$[$DURATION + $NOW]

# Finished flag - when set to "finished" we exit
FINISHED=not

echo "Running for a maximum of $DURATION seconds"

while [ $FINISHED != "finished" ] ; do
    check_file;
    NOW=`date +%s`
    if [ $NOW -gt $END_TIME ] ; then
        "Timed out after $DURATION seconds"
        exit 1
    fi
done

# If we've got this far we were successful
# so just return a success code
exit 0
```

Script failure

Common reasons for a script failure are:

- A new *RESULT_FILE* is not created before the script times out
- The *CONTROL_FILE* cannot be found
- Improper specification of command line arguments

If the script fails, it returns an exit code of 1. In this example, this failure will cause the job to fail because the next step in the process cannot be executed.

Chapter 17. Message-based job steps

Message-Based job steps provide a method for triggering processing within an IBM SPSS Collaboration and Deployment Services job with external events signaled by JMS (Java Messaging Service) messages.

Message-based job steps allow to specify message text and message selector as well as the timeout (a time to wait for the message). After the specified message is received, the system proceeds to the next step in the job. If the timeout is exceeded, the job is marked as failed.

To add a message-based step to a job, select the **Message-driven Work** tool from the job palette and click the job canvas. Before the job step can run successfully, its properties must be defined.

General properties of message-based job steps

The general properties for a message-based job step specify the message domain, message text, message selector, and the step timeout.

To define the properties for a message-based job step, click a message-driven step in an open job. Click the General tab to view or edit the step properties.

1. The properties include:

- **Job step name.** Type a name for the step.
- **Message Domain.** The message domain identifies the JMS topic to listen on. See the topic "Message domains" on page 52 for more information.
- **Message Text.** For text message, the text of the expected message of to trigger the job step execution.
- **Message Selector** The selector text to filter messages by contents of the message header.
- **Message Timeout.** The time (in days, hours, minutes, and seconds) to wait for the message triggering the job step. The timeout is effective from the start of the job step after predecessor steps have been completed.

Chapter 18. Notification job steps

Email notifications can be set up in IBM SPSS Deployment Manager using either of the following methods:

- Notifications tab. See the topic "Job success and failure notifications" on page 141 for more information.
- Notification job step.

A notification job step is a specialized method for sending notifications. Typically, the notification job step is used in conjunction with model evaluation.

Notifications can be set up to monitor specific model evaluation outcomes. The following are examples of when notifications might be based on model evaluation return codes:

- Notifications can be set up based on a specific outcome of the model evaluation. For example, if the result returned by a model evaluation analysis is less than 0.85, notifications could be sent to a predefined list of recipients.
- Notifications can be set up based on the traffic light level of a model evaluation job. For example, if the result is red, a notification could be sent. In addition, multiple notification rules could be set up with different email recipients for each traffic light level.

Using a notification job step is similar to specifying notifications as part of another job step. However, the following key differences apply:

- Using notification job steps, different notification templates can be selected for different outcomes. In addition, notification rules can be established for multiple sets of notification recipients.
- In notification job steps, more specific criteria can be established for sending notifications, using conditional connectors. Notifications sent via the Notifications tab apply only for success or failure.
- With notification job steps, all email messages are sent as links. The ability to send an email message as an attachment is not available.

Adding a notification job step to a job

Typically, a notification job step is used in conjunction with another job step.

To add a notification job step to a job:

1. Create a new job or open an existing job.
2. From the job palette, click **Notification**.
3. Click anywhere in the job canvas. The notification job step is added to the job.
4. Because a notification job step is typically connected to another job step, at this stage relationships would be established between the notification job step and other steps within the job. See the topic "Specifying relationships in a job" on page 119 for more information.

For example, if you want to send a notification to a set of recipients if a model evaluation index is less than 80, connect the model evaluation step to a notification step by using a conditional connector. If the evaluation step is successful, the value returned is the index rounded to the nearest integer. Specify the conditional expression for the connector as `success==true && completion_code<80`. If you want to send different emails to different recipients based on index ranges, connect the model evaluation step to multiple notification steps using conditional connectors and specify the index range for each. For example, for three connectors, you could specify the following conditional expressions:

- `success==true && completion_code>=100`
- `success==true && completion_code==200`
- `success==true && completion_code<=300`

After the notification job step has been added to the job, its general information and notification recipients must be specified on the corresponding tabs.

General information

The General tab contains information that pertains to the overall notification job step.

The General tab contains the following information:

- **Job step name.** Type a name for the step.
- **Iterative consumer.** Specifies whether or not the notification job step is an iterative consumer of another job step. Select the **Iterative Consumer** check box if this job step is an iterative consumer of another job step. If this check box is selected, one of the available iterative variables must be selected from the drop-down list.

Notification

The Notification tab is used to configure the recipients of notifications for the job step.

The tab contains the following information:

- **Recipients for notifications.** The number of recipients that receive notifications. This number is generated by the system based on the number of recipients specified in the Notifications dialog.

Updating notifications

To update notification parameters:

1. In the General tab, click **Update**. The Notifications dialog box appears.
2. In this dialog box, the following parameters can be modified:
   - **From.** The address of the sender notification message. The field is pre-populated with the default email address in repository configuration options.
   - **To.** The list of notification recipients. The addresses can be entered, or selected from a directory listing of a supported email application such as Lotus Notes. To edit an address, click the ellipsis button next to it in the list. To remove a recipient, highlight the recipient's email address and click **Delete**.
   - **Subject.** The subject of the notification message. By default, the field is populated by the default template associated with the notification event. Modify the message subject, if necessary. See the topic “Customizing the notification message” on page 146 for more information.
   - **Message.** The body text of the notification message. The field is populated by the default template associated with the notification event. Modify the message body text, if necessary. See the topic “Customizing the notification message” on page 146 for more information.
3. Preview the notification message. See the topic “Notification message preview ” on page 146 for more information.
4. Click **OK**.

Selecting a new template

When a notification job step is initially created, the template fields are blank, and a template needs to be selected from the list of available templates. The template selected applies to the current notification job step only and persists until a new template is selected.

In this dialog box, templates can only be modified. New templates cannot be created. To select a new default template:

1. In the Notifications dialog box, click **Select Template**. The Notification Template Selection dialog box appears.

2. Select a template from the list. A preview of the template appears in the Notification Template Selection dialog box. The template can only be viewed in this dialog box. Modifications can only be made in the main Notifications dialog box.
3. Click **OK**. The Notifications dialog box reappears.
4. Make any modifications necessary to the notification message. See the topic "Updating notifications" on page 170 for more information.
5. Click **OK**.

Chapter 19. Champion challenger job steps

Champion challenger overview

By using IBM SPSS Deployment Manager, it is possible to compare model files generated by IBM SPSS Modeler to determine which file contains the most effective predictive model. The champion challenger job step evaluates a model and compares it to one or more challengers.

After the system compares the results, the best model becomes the new champion.

Champion. The champion corresponds to the most effective model. For the initial execution of the champion challenger job step, there is no champion--only the first challenger and the corresponding list of challengers. For subsequent executions of the job step, the system determines the champion.

Challenger. Challengers are compared against each other. The challenger that generates the best results then becomes the new champion.

Champion Selection Process

The champion challenger comparison process consists of the following tasks:

1. Scoring each of the competing models.
2. Evaluating the resulting scores.
3. Comparing the evaluation results and determining which challenger is the champion.
4. Saving the new champion to the repository (optional).

Adding Champion Challenger Work to a Job

To add a champion challenger job step to a job, select the Champion Challenger tool from the job palette and click the job canvas.

Model evaluation metrics

Model evaluation and comparison can focus on accuracy, gains, or accreditation.

- **Accuracy.** The accuracy of a model reflects the percentage of target responses that are predicted correctly. Models having a high percentage of correct predictions are preferred to those having a low percentage.
- **Gains.** The gains statistic is an indicator of the performance of a model. This measure compares the results from a model to the results obtained without using a model. The improvement in the results when using the model is referred to as the gains. When comparing two models, the model having the higher gains value at a specified percentile is preferred.
- **Accreditation.** Model accreditation reflects the credibility of a model. This approach examines the similarity between new data and the training data on which a model is based. Accreditation values vary from 0 to 1, with high values indicating greater similarity between the predictors in the two data sets. When comparing two models, the model having the higher accreditation value is based on training data that is more similar to the new data, making it more credible and preferred.

Order dependency

Unlike other types of job steps, the tabs in the champion challenger job step are order dependent.

For example, a challenger must be selected in the Challengers tab before information can be modified in the Champion tab. In addition, the information that appears on some tabs is contingent upon the challengers selected in the Challengers table.

The process for executing a champion challenger comparison consists of the following steps:

1. Providing general job information.
2. Identifying challengers.
3. Specifying champion information.
4. Viewing parameter information.
5. Specifying notifications.

General information

The General tab contains information that pertains to the overall champion challenger job step.

IBM SPSS Modeler server and login information is required to execute a champion challenger job step. Content repository server and login information is required to execute the job and save new champions to the IBM SPSS Collaboration and Deployment Services Repository. (Content repository server and login information is required, even if you discard the results of the analysis.) Credentials are based upon the user currently logged in to the system.

Job step name. The name of the job step. By default, the name of the first job step is *Event 1*. Subsequent job steps are named *Event 2*, *Event 3*, and so on. The name specified here appears in the job history table after the job step is executed.

IBM SPSS Modeler server. The IBM SPSS Modeler server or server cluster on which the stream will be executed. The list contains all servers and server clusters currently configured to execute IBM SPSS Modeler steps. To change the server, select from the **IBM SPSS Modeler Server** drop-down list. To create a new server definition, click **New** to launch the server definition wizard.

IBM SPSS Modeler login. The credential information used to access the IBM SPSS Modeler server or server cluster. To change the credentials, select a credential definition from the **IBM SPSS Modeler Login** drop-down list. To define new credentials, click **New** to launch the credentials definition wizard.

Content Repository Server. The content repository server allows a job to save files to an IBM SPSS Collaboration and Deployment Services Repository. Typically, the content repository server is specified when refreshing models using IBM SPSS Modeler. To specify a content repository server, select a server from the **Content Repository Server** drop-down list. To create a new server definition, click **New** to launch the server definition wizard. To generate a content repository server definition based on the current server information, click **Generate**. A server definition is created and automatically populated in the *Content Repository Server* field.

Content Repository Login. The login information for the content repository server. To specify a content repository login, select a credential from the **Content Repository Login** drop-down list. To create a new login, click **New** to launch the content repository login wizard. If single sign-on is not used to connect to the IBM SPSS Collaboration and Deployment Services Repository, click **Generate** to generate a content repository server login based on existing security settings. A content repository login is created and automatically populated in the *Content Repository Login* field. Login generation is not available when using single sign-on.

Challengers

At least one first challenger must be selected to execute a champion challenger job step. It is important to note that the first challenger selected does not imply an order of comparison or any priority in the evaluation process. The first challenger is simply the baseline.

The data source and labels used to determine subsequent challengers are established by the first challenger. After the first challenger is selected or updated, the remaining fields in this tab are updated with information that corresponds to the first challenger.

First challenger. Name of the first challenger. To browse through the repository, click **Browse**.

First challenger label. Label associated with the model file containing the first challenger. Specify this value when selecting the first challenger.

Data Source Challenger. The challenger supplying the data source node used for the job step. Click **Browse** to choose this challenger from the list of entries selected in the Challengers table.

Metric. The measurement criteria by which the challengers are compared. Valid values include *accreditation*, *accuracy* and *gains*. If *gains* is selected, then a percentile needs to be specified as well. See the topic "Model evaluation metrics" on page 173 for more information.

Challengers Table

The Challengers table lists the default score branches for challengers that match the data source and the label associated with the first challenger. Only the challengers selected from the table will be compared to the first challenger when the job step is executed. Selecting (or clearing) a challenger from the list will cause the system to update the corresponding information on the other job step tabs accordingly.

Each time a job that contains a previously saved champion challenger step is opened, the list of challengers is automatically updated. New challengers may be added to the list if they match the data source and label criteria of the first challenger. Conversely, challengers that no longer meet those criteria may be removed from the list of challengers. If a selected challenger has been removed from the repository, the system will generate a message indicating that the challenger is no longer available.

Although challengers can be selected and cleared for comparison, information in the challengers table is not modifiable. Specifically, the Challengers table contains the following information.

Name. Name of the challenger.

Label. Label associated with the challenger.

Description. A description of the challenger.

Modifications to the first challenger

Changes made to the first challenger after the job is saved may affect the champion challenger analysis. For example, suppose that the first challenger is removed from the repository or the label associated with the first challenger is removed. When the Challengers tab is accessed, the system will generate a message, indicating that the first challenger is no longer available for use. In this case, a new first challenger needs to be specified.

Selecting challengers

To select challenger models for inclusion in champion challenger analyses, perform the following steps:

1. On the Challenger tab of a Champion Challenger step, click **Browse** for the First Challenger. If you are adding challengers manually, click **Add** for the Challengers table.
2. Select the model file by clicking **Browse**. The model file is a IBM SPSS Modeler stream that contains a default score branch with a valid model nugget.
3. Select the label designating the version of the selected model file to use.
4. In the Challengers table, select the score branch to use.
5. Click **OK**.

Invalid challengers

To be compared, model files must have scoring branches that use a common data structure.

Data characteristics that must match across challengers include the following items:

- Data sources must have the same number of fields.
- The field names must be identical across data sources.
- The field measurement levels must be identical across data sources.

If the system cannot find challengers comparable to the first challenger selected, the Invalid Challenger dialog appears. To select a new challenger:

1. Click **OK** to return to the Challengers tab.
2. Select a new challenger.

Selecting challenger data sources

To select the data source used in the champion challenger analyses, perform the following steps:

1. On the Challenger tab of a Champion Challenger step, click **Browse** for the data source challenger.
2. From the list of challengers included in the analysis, select the score branch that includes the data source to use.
3. Click **OK**.

Champion

Before information can be specified for a champion, at least one challenger must be selected. If the Champion tab is accessed before a challenger has been selected, the First Challenger Not Selected dialog appears, indicating that a challenger must be selected.

Do not create a new version of the champion. Select this option to prevent the creation of a new version of the champion. In this case, the selected labeled version of the champion will be modified. Clear this option to create a new version of the champion instead of modifying the labeled version.

File name. The name to use for the copy of the challenger identified as the champion.

Location. The location in which the copy of the champion file is stored.

Permission. The permissions associated with the copy of the champion.

Metadata. Properties associated with the copy of the champion. Specifying metadata for the champion output is the same as specifying metadata for other job output.

Using the champion in other jobs

After the champion challenger job has been run, the resulting champion can be used in other jobs. To include the champion in another job, the following information is required:

- The name of the champion.
- The location of champion.

When the champion is used in another job, the *LATEST* label is applied. This label cannot be modified.

Testing the champion

By default, the system creates a new copy of the champion each time the champion challenger job step is run, stores the copy to the output location specified, and writes the results to the job history log.

However, there may be instances in which saving a copy of the champion is not wanted. For example, suppose that you simply want to test the champion challenger job step.

To disable the creation of a copy, select the **Do not create a new version of the champion** check box. If this check box is selected, the remaining options in the tab are disabled. The system will use the same information applied to the current champion.

The system will execute the champion challenger job step and determine a new champion. However, a new version of the champion will not be created or saved to the repository. Instead, the results will only be written to the job history log, indicating which challenger would have been chosen as the champion.

For example, suppose a champion challenger job is run and the creator of the job elected not to create a new version of the champion. The resulting job history log might look like this:

```
Stream execution started
500 500
1000 1000
```

```
1500 1500
2000 2000
2500 2500
Stream execution complete, Elapsed=26.22 sec, CPU=18.97 sec
Stream execution started
1000 0
2000 0
Field 'Correct_Sum' has only one value
Field 'Count' has only one value
Field 'Traffic Light Result' has only one value
2855 145
2855 1145
2855 2145
Field 'campaign' has only one value
Field 'gold_card' has only one value
Field 'response' has only one value
Stream execution complete, Elapsed=0.39 sec, CPU=0.2 sec
Stream execution started
500 500
1000 1000
1500 1500
2000 2000
2500 2500
Stream execution complete, Elapsed=26.06 sec, CPU=17.75 sec
Stream execution started
1000 0
2000 0
Field 'Correct_Sum' has only one value
Field 'Count' has only one value
Field 'Traffic Light Result' has only one value
2855 145
2855 1145
2855 2145
Field 'campaign' has only one value
Field 'gold_card' has only one value
Field 'response' has only one value
Stream execution complete, Elapsed=0.48 sec, CPU=0.19 sec
Stream execution started
500 500
1000 1000
1500 1500
2000 2000
2500 2500
Stream execution complete, Elapsed=21.48 sec, CPU=17.34 sec
Stream execution started
1000 0
2000 0
Field 'Correct_Sum' has only one value
Field 'Count' has only one value
Field 'Traffic Light Result' has only one value
2855 145
2855 1145
2855 2145
Field 'campaign' has only one value
Field 'gold_card' has only one value
Field 'response' has only one value
Stream execution complete, Elapsed=0.39 sec, CPU=0.17 sec
The result for challenger cc_cartresponse.str is 98.809.
The result for challenger cc_neuralnetresponse.str is 98.844.
The result for challenger cc_c51response.str is 98.809.
The declared Champion is cc_neuralnetresponse.str.
```

Note the last line in the log file:

```
The declared Champion is cc_neuralnetresponse.str.
```

This line indicates that the *cc_neuralnetresponse* stream would have been the champion. However, a copy of this stream was not saved to the repository because the system did not create a new version of the stream. If the stream had been saved to the repository, the log would contain an additional line, indicating that the stream was saved to the repository--for example:

```
Adding artifact spsstr:/PMDemo/ModelManagement/cc_neuralnetresponse.str.
```

Data files

Data files information appears for the challengers that were selected from the list of challengers in the Challengers tab.

Any changes made in this tab apply to the champion challenger job step only. Modifications made to the data file information are not propagated back to the challenger saved in the IBM SPSS Collaboration and Deployment Services Repository. The data files table contains the following information.

Node Name. The name of the input node that contains the data used by the stream. The node name is not modifiable.

Node Type. The node type as defined in the stream. The node type is not modifiable.

File name. The name of the input data file. To change the name, click in the file name cell and change the name.

Format. The format of the output file--for example, a comma delimited file. To modify the file format type, click in the Format cell. A drop-down arrow appears. Select the format type.

Location. The location of the input data files. To modify the location, click in the column and then click the resulting ellipsis button. The Input File Location dialog box opens. Change the location as necessary.

Data view

Analytic data view information appears for the challengers that were selected from the list of challengers in the Challengers tab.

Any changes made in this tab apply to the champion challenger job step only. Modifications made to the data view information are not propagated back to the challenger saved in the IBM SPSS Collaboration and Deployment Services Repository. The data view table contains the following information.

Node Name. The name of the data view node that contains the data used by the stream. The node name is not modifiable.

Analytic Data View. The analytic data view referenced by the data view node.

Label. Label identifying the version of the analytic data view used.

Table Name. The table containing the input data fields.

Data Access Plan. The plan supplying the data records for the input data fields. To change the data access plan used for a node, select the cell containing the access plan and click the resulting ellipsis (...) button.

ODBC data sources

ODBC data source information appears for the challengers that were selected from the list of challengers in the Challengers tab.

Any changes made in this tab apply to the champion challenger job step only. Modifications made to the ODBC data source information are not propagated back to the challenger saved in the IBM SPSS

Collaboration and Deployment Services Repository. The ODBC data sources table contains the following information.

Node Name. The name of the input node that contains the data used by the stream. The name is prefixed by the names of any supernodes containing the node separated by slashes. For example, if the node *MyNode* is within a supernode named *Supernode1*, the name appears as */Supernode1/MyNode*.

Node Type. The node type as defined in the stream.

ODBC Data Sources. The current ODBC data source name (DSN). To change to a different ODBC data source, click the cell containing current data source name, then click the "..." button that is displayed. Doing so displays a dialog box from where you can choose an existing DSN or create a new one.

Credentials. To change the database username and password when changing the ODBC data source, click the cell containing the current credentials, then click the "..." button that is displayed. Doing so displays a dialog box from where you can choose an existing credential definition or create a new one.

Database Table. The database table that corresponds to the node.

Nodes within locked supernodes are not accessible. They cannot be viewed or modified.

Cognos import

If the model files contain any IBM Cognos BI source nodes, the Cognos connection details are displayed here.

Node Name. The name of the Cognos source node.

Connection URL. The URL of the Cognos server to which the connection is made.

Package Name. The name of the Cognos package from which the metadata is imported.

Anonymous. Contains **Anonymous** if anonymous login is used for the Cognos server connection, or **Credential** if a specific Cognos username and password are used.

Credentials. The username and password (if required) on the Cognos server.

Note: Cognos credentials must be created in a domain, which represents the Cognos namespace ID.

Chapter 20. Submitted jobs

The Submitted Jobs folder is a staging area in IBM SPSS Collaboration and Deployment Services Repository that displays the results of running reports using IBM SPSS Collaboration and Deployment Services Deployment Portal. From the reports, jobs and additional output are generated and appear in the Submitted Jobs folder.

The Submitted Jobs folder is at the same level as the Content Repository in the Content Explorer. The Submitted Jobs folder appears in the Content Explorer, whether or not content is available within it.

When a user submits a report to IBM SPSS Collaboration and Deployment Services, a folder is created for the user within the Submitted Jobs folder. The folder name corresponds to the user's login ID. Subsequently, output is placed into the folder. Each report is contained in a separate folder, whose name contains a timestamp. Each report folder contains a corresponding job and any additional artifacts related to the report.

The following guidelines apply to objects in the Submitted Jobs folder:

- The objects within the Submitted Jobs folder are saved to the IBM SPSS Collaboration and Deployment Services Repository.
- The contents of the Submitted Jobs folder are searchable. See the topic “Searching” on page 17 for more information.
- Objects, such as files and jobs, can be opened but only in read-only mode.

Restrictions within the Submitted Jobs folder

Objects in the Submitted Jobs folder cannot be modified—for example, names cannot be changed, jobs cannot be scheduled, and versioning is disabled. To modify an object, it must be moved from the Submitted Jobs folder to the Content Repository. Once an object is moved into the Content Repository, full functionality is restored.

Although objects can be moved from the Submitted Jobs folder to the Content Repository, the reverse is not possible. Objects from the Content Repository cannot be moved into the Submitted Jobs folder.

The only exception to these restrictions is permissions. Permissions can be changed for objects within the Submitted Jobs folder. All of the standard permissions are available; however, only the owner of an object can modify its permissions. For example, the owner can grant permissions to modify an object. By default the user who created the report is the owner. See the topic “Modifying permissions” on page 26 for more information.

Submitted jobs and expiration dates

By default, objects within the Submitted Jobs folder automatically expire after a set period of time. To prevent an object from expiring, it must be moved out of the Submitted Jobs folder and into the Content Repository.

When the object is moved into the Content Repository, the expiration date attached to the object is not automatically removed. To prevent the object from expiring, the expiration date must be explicitly removed or, if needed, a new expiration date must be set.

The expiration time period in the Submitted Jobs folder is different from setting a specific expiration date in the Content Repository. Unlike expired objects in the Content Repository, expired objects in the Submitted Jobs folder are automatically deleted from the repository. The expiration time period for the contents of Submitted Jobs folder is a configuration setting, which can be modified by an administrator. The default value is five days. Expiration dates in the Content Repository are user-defined and modifiable. See the topic “Working with Expiration Dates and Expired Files” on page 29 for more information.

Chapter 21. Administration Consoles

Administration consoles for IBM SPSS Statistics, IBM SPSS Modeler, and IBM SPSS Modeler Text Analytics are included in IBM SPSS Deployment Manager. This provides a single interface for server administration tasks.

Getting started

Administered servers

Server administration in IBM SPSS Deployment Manager involves:

1. Adding the server to be administered to the system.
2. Logging in to the server being administered.
3. Performing administrative tasks for the server as needed.
4. Logging off from the server being administered.

The Server Administration tab offers access to this functionality. This tab lists the servers currently available to be administered. This list persists across IBM SPSS Deployment Manager sessions, facilitating access to those servers.

From the menus choose:

Tools > **Server Administration**

The administered server list may include a variety of server types, including IBM SPSS Collaboration and Deployment Services Repository servers, IBM SPSS Modeler servers, and IBM SPSS Statistics servers. The actual administrative functionality available for a server depends on the server type. For example, security providers can be configured and enabled for repository servers but not for IBM SPSS Modeler servers.

Adding new administered servers

Before performing administrative tasks, a connection to the administered server must be established.

From the menus choose:

File > **New** > **Administered Server Connection**

The **Add New Administered Server** dialog box opens. Adding a new connection requires the specification of the administered server type and the administered security server information.

Selecting the administered server name and type

The first step of adding a new administered server to the system involves the definition of the name and type for the server.

Name. A label used to identify the server on the Server Administration tab. Including the port number in the name, such as *my_server:8080*, may help to identify the server in the administered server list.

Note: Alphanumeric characters are recommended. The following symbols are prohibited:

- Quotation marks (single and double)
- Ampersands (&)
- Less-than (<) and greater-than (>) symbols
- Forward slash (/)
- Periods

- Commas
- Semicolons

Type. The type of server being added. The list of possible server types depends on the system configuration and may include:

- IBM SPSS Collaboration and Deployment Services Repository Server
- Administered IBM SPSS Modeler Server
- Administered IBM SPSS Statistics Server
- Administered IBM SPSS Modeler Text Analytics Server

Selecting an administered server type

In the **Select Administered Server Type** dialog box:

1. Enter a name for the server.
2. Select the server type.
3. Click **Next**. The **Administered Server Information** dialog box opens.

Administered server information

The second step of adding a new administered server to the system involves the definition of the server properties.

For an IBM SPSS Collaboration and Deployment Services Repository server, you can specify the server URL.

The URL includes the following elements:

- The connection scheme, or protocol, as either *http* for hypertext transfer protocol or *https* for hypertext transfer protocol with the secure socket layer (SSL)
- The host server name or IP address

  Note: An IPv6 address must be enclosed in square brackets, such as `[3ffe:2a00:100:7031::1]`.
- The port number. If the repository server is using the default port (port 80 for http or port 443 for https), the port number is optional.
- An optional custom context path for the repository server

Table 16. Example URL specifications . This table lists some example URL specifications for server connections.

URL	Scheme	Host	Port	Custom path
`http://myserver`	HTTP	*myserver*	default (80)	(none)
`https://9.30.86.11:443/spss`	HTTPS	9.30.86.11	443	*spss*
`http://[3ffe:2a00:100:7031::1]:9080/ibm/cds`	HTTP	`3ffe:2a00:100:7031::1`	9080	*ibm/cds*

Contact your system administrator if you are unsure of the URL to use for your server.

For other server types, available properties include the following items:

Host
The name or IP address of the server.

Note: Alphanumeric characters are recommended. The following symbols are prohibited:

- Quotation marks (single and double)
- Ampersands (&)
- Less-than (<) and greater-than (>) symbols
- Forward slash (/)
- Periods
- Commas
- Semicolons

Port
The port number that is used for the server connection.

This is a secure port.
Enables or disables the use of a Secure Sockets Layer (SSL) for the server connection. This option is not offered for all types of administered servers.

After you define the properties, the new server is included in the administered server list on the Server Administration tab.

Viewing administered server properties

To view the properties of an existing administered server, right-click the server on the Server Administration tab and select **Properties** from the drop-down menu.

The displayed properties depend on the type of server selected.

Connecting to administered servers

For most servers, you must connect to a server in the administered server list to perform administrative tasks. From the Server Administration tab, double-click the server to administer.

Disconnecting administered servers

After completing your administrative tasks, log off from the server.

1. On the Server Administration tab, right-click the server.
2. Select **Logoff**.

To administer the server, you must log in again.

Deleting administered servers

A server appears in the administered server list until it is deleted from the list.

1. On the Server Administration tab, select the server to delete.
2. From the menus choose:

   Edit > **Delete**

Alternatively, right-click the server and select **Delete** from the drop-down menu.

If further administrative tasks for the server are needed in the future, the server will need to be added to the system again.

IBM SPSS Statistics Server Administration

Introduction to IBM SPSS Statistics Server Administration

The IBM SPSS Statistics Administration Console provides a user interface to monitor and configure your IBM SPSS Statistics Server installations. The IBM SPSS Statistics Administration Console uses the framework from the IBM SPSS Deployment Manager, so some of the documentation refers to the later

product. The IBM SPSS Statistics Administration Console can be installed only on Windows computers; however, it can administer IBM SPSS Statistics Server installed on any supported platform.

To start IBM SPSS Statistics Administration Console

1. From the Windows Start menu, choose:

[All] Programs > **IBM SPSS Collaboration and Deployment Services** > **IBM SPSS Collaboration and Deployment Services Deployment Manager**

IBM SPSS Statistics Server Connections

You must specify a connection to each IBM SPSS Statistics Server on your network that you want to administer. You must then log in to each server that you want to configure or monitor.

After you log in to your IBM SPSS Statistics Server, options are shown beneath the server name: Configuration and Monitoring Current Users. Double-click one of these options to display the associated pane.

Controlling the IBM SPSS Statistics Server

The IBM SPSS Statistics Administration Console allows you to pause, shut down, and restart the IBM SPSS Statistics Server.

1. In the Server Administration pane, select the IBM SPSS Statistics Server.
2. From the menus choose:

   Tools > **IBM SPSS Statistics Server**
3. Select an option:

Pause. Pause a server when want to stabilize usage while you diagnose a problem, or when you want to temporarily limit the number of connections. Pausing the server prevents end users from connecting. It doesn't affect end users that are connected to the server when it is paused. Choose **Resume** when you want the server to accept connections again.

Shutdown. Shut down a server when you need to terminate the server software. Once you have shutdown, you must go to the computer where the server software is installed to restart it. Avoid shutting down servers when end users are connected. Shutting down will disconnect the users, and they may lose work.

Restart. Some configuration changes require that you restart the server before they take effect. These are indicated by an asterisk (*) in the Configuration pane. Restarting a server temporarily stops the server software, and then restarts it. You must disconnect all end users before you can restart the server. Because end users can lose work when you disconnect them, please ask them to logoff from the client application. Restarting the server may take a few seconds.

IBM SPSS Statistics Server Configuration

The Configuration pane shows configuration options for IBM SPSS Statistics Server. To activate the pane, double-click the Configuration node beneath the desired server in the Server Administrator pane. To save any configuration changes, choose **Save** from the File menu. Some settings, indicated by an asterisk (*), require the server to be restarted before taking effect.

Connections

Host. The host name—an alphanumeric name or an IP address—is the network designation for the computer where the server software is running. Change the host name when you move the server software to a different computer, when the network designation of the computer changes, or when the computer has more than one IP address and the server software is using the wrong one. End users must use the same name when they connect to the server software. If you make a change, distribute the new name.

Port Number. The port number is the port that the server software uses for communications. Change the port number only if another application on the server computer is using the same port. Enter an integer between 1025 and 32767. If you make a change, be sure to distribute the new port number to end users. The default port number is 30<version>, where *<version>* is the major version of IBM SPSS Statistics Server. For example, the default port number for version 17 is 3017.

Maximum Number of Users. End users connect to the server software by logging in from a client application. When a connection is made, the server software launches a process to handle the end user. Increase or decrease the maximum number of users to adjust the load on the server computer. Enter an integer between 1 and 2000.

Secure Sockets Layer. If you use a Secure Sockets Layer (SSL) to encrypt the communications that occur when end users to connect to the server software, you can configure the server software to accept SSL connections from end users. If you select this option, you must specify the location of the public and private key files. See the topic "File Locations" on page 187 for more information.

Automatic Reconnection Timeout. If end users are disconnected from the server because of a network failure, the client can automatically reconnect to the server when the connection is restored. The automatic reconnection timeout specifies the maximum number of minutes during which the client will attempt to reconnect. Enter an integer between 0 and 999. If you enter 0, this feature is disabled and no clients will attempt to automatically reconnect to the server.

Group Authorization Service URL. End users are typically authorized by the operating system. If you want them to be authorized by a remote server, you can configure group authorization. Enter the URL of the IBM SPSS Collaboration and Deployment Services server that will authorize the users. Be sure to include the port number (for example, `http://myserver.mydomain.com:9080`).

File Locations

Configuration File. The server software stores the information that it needs in a configuration file. The server software must locate this file on startup, and it needs to write to the file periodically. Since the file is writable, you may want to change its location.

Server Temp Files Directory. When the server software accesses and processes data, it often has to keep a temporary copy of that data on disk. The amount of disk space that will be used for temporary files depends on the size of the data file that the end user is analyzing and the type of analysis that he or she is doing. It varies from the size of the data file to three times the size of the data file. Because the temporary files are writable and can get quite large, you may want to change their location. If you do, the location that you define is a root directory. When an end user connects to the server software, it creates a unique, temporary subdirectory in the root location that you specify. You can specify multiple locations separated by commas. Using multiple temporary file locations can improve performance when the locations are on different spindles. The temporary file location is a global setting for all users and can be overridden by the user profile or group setting. See the topic " IBM SPSS Statistics Server User Profiles and Groups" on page 191 for more information.

Background Output Directory. When an end user chooses to run a production job in a separate session on a remote server, the output generated from the production job is stored in the background output directory on the server machine until the user retrieves it. You can change this directory as needed, but be sure that users have read and write permissions to the directory. The background output directory location is a global setting for all users.

User Profile File. The profile file defines the user profile and group settings. It is recommended that you move the profile file from the default location. This ensures that subsequent installations of the server software do not overwrite your profile file. The settings in the profile file are modified within the administration application (IBM SPSS Statistics Administration Console). See the topic " IBM SPSS Statistics Server User Profiles and Groups" on page 191 for more information.

SSL Public Key File/SSL Private Key File. If you selected **Secure Sockets Layer** in the Connections group, you must specify the full path to the SSL public and private key files. If the public and private keys are stored in the same file, specify the same file path in both fields.

Python Home Directory/Python3 Home Directory. These settings specify the installations of Python 2.7 and 3.8 that are used by the IBM SPSS Statistics - Integration Plug-in for Python on this server. By default, the distributions of Python 2.7 and 3.8 that are installed with this IBM SPSS Statistics Server (as part of IBM SPSS Statistics - Essentials for Python) are used. They are in the `Python` (contains Python 2.7) and `Python3` directories under the directory where IBM SPSS Statistics Server is installed. To use a different installation of Python 2.7 or Python 3.8 on this server computer, specify the path to the root directory of that installation of Python.

Users

Admin Group. When a user connects to the IBM SPSS Statistics Administration Console, the server software checks if the user is a member of the administrator group. The administrator group corresponds to a system account group on the machine on which the server software is installed. Any member of the administrator group has adminstrative-level rights to the server software. This means that a member of this group can use the IBM SPSS Statistics Administration Console to configure the server. You can specify any system account group to be the administrator group. However, if the server software cannot find the group, it will revert to the default group, which is the administrator group for the computer on which the server software is installed (for example, the Administrators group on an English Windows system).

Client Data Access. When end users connect to the server software and open a data file, they are able to see the actual data in the file. Always displaying the data in the client application can negatively affect performance and increase network traffic. Therefore, you may choose to prevent the data from displaying in the client application. The client will still display the data dictionary, allowing the end user to check which variables appear in the data. The client data access setting is a global setting for all users and can be overridden by the user profile or group setting. See the topic “ IBM SPSS Statistics Server User Profiles and Groups” on page 191 for more information.

Sort. When end users connect to the server software and run a sort procedure, the server software uses the specified sort engine. The engine must be installed and configured on the server software computer.

Umask. You can adjust the umask that is used when users connect to the server software running on a UNIX machine. This three-digit octal number is a global setting for all users and can be overridden by the user profile or group setting. See the topic “ IBM SPSS Statistics Server User Profiles and Groups” on page 191 for more information. Note that this setting is not available when the server software is running under Windows.

Maximum Threads. You can specify a limit to the number of threads that the server software can use. You may do this to limit the system resources available for multithreaded procedures. The maximum threads setting is a global setting for all users and can be overridden by the user profile or group setting. See the topic “ IBM SPSS Statistics Server User Profiles and Groups” on page 191 for more information.

Cache Compression. When end users explicitly issue a `CACHE` command or run a procedure that creates scratch files automatically, the server software can compress the scratch (temporary) files using zlib. You should turn cache compression on if your users typically read large data files. A compressed scratch file reduces the amount of disk I/O compared to an uncompressed scratch file but requires more processing by the server software. To minimize the effect of the extra CPU time on job elapsed time, compression and decompression are performed using separate threads. That means that one or more processors can be handling these chores while others are executing commands. Because the extra overhead is greater when writing a file compared to reading it, let your users know that they should use a higher value for `SET CACHE`. Doing so will reduce the frequency of writing scratch files and increase the number of times each scratch file is read. The cache compression setting determines the default behavior for cache compression. You can also set the compression override setting to allow users to override this default. The cache compression setting is a global setting for all users and can be overridden by the user profile or group setting. See the topic “ IBM SPSS Statistics Server User Profiles and Groups” on page 191 for more information.

Compression Override. You can specify whether end users can override the cache compression setting with `SET ZCOMPRESSION` syntax. The compression override setting is a global setting for all users and can be overridden by the user profile or group setting. See the topic “ IBM SPSS Statistics Server User Profiles and Groups” on page 191 for more information.

Maximum JVM Memory. When end users connect to the server software, memory is allocated for the Java virtual machine (JVM) on the server computer. Memory is allocated for each user. You can change the amount of allocated memory by specifying a number in megabytes (MB). You may need to this if users have jobs that require a large amount of memory, for example when exporting very large custom tables. The JVM memory setting is a global setting for all users and can be overridden by the user profile or group setting. See the topic " IBM SPSS Statistics Server User Profiles and Groups" on page 191 for more information.

Note: At this time, compiled transformations are supported only for IBM SPSS Statistics Server running on Windows. UNIX/Linux is not supported.

Compiled Transformations. When a syntax job includes transformation commands, the server software converts the transformations to bytecode when the syntax job is run. This bytecode is then "interpreted" while running the syntax job. The server software can compile the transformations to machine code to improve performance when there are many cases in the data and several transformations in the syntax job. If **Compiled Transformations** is set to `Yes` and a user issues the `SET CMPTRANS=YES` command, the server software will:

- Convert the transformations to bytecode as usual.
- Convert the bytecode to C++ code.
- Compile the C++ code to machine code.
- Execute the machine code.

Compiler Path. The full path to the C++ compiler that is used for compiling transformations. You must use MinGW, which is a Windows port of the GNU Compiler Collection (GCC). Be sure to include the executable name in the path specification, for example `C:\MinGW\bin\g++.exe`.

Restrict Database Access. You can restrict database access to specific database sources. To restrict sources to a specific list of DSNs, set this to `Yes` and identify the permitted sources with the **Permitted Database Sources** setting. If this is set to `No`, end users can access any database that is configured on the server.

Permitted Database Sources. If **Restrict Database Access** is set to `Yes`, specify the DSNs to which users have access. Use semicolons to delimit DSNs (for example, `Fraud - Analytic;Fraud - Operational`).

COP Configuration

If your site is running IBM SPSS Collaboration and Deployment Services 3.5 or later, you can configure one or more servers to communicate with the Coordinator of Processes (COP). The COP registers the server and allows client computers to connect the server without knowing the server's host name. In other words, the COP allows client computers to search for available servers.

COP Status. The current status of the COP. You cannot change this value.

COP Host. The host name—an alphanumeric name or an IP address—is the network designation for the IBM SPSS Collaboration and Deployment Services server on which the COP is running.

COP Port Number. The port number is the port on the IBM SPSS Collaboration and Deployment Services server that the COP uses for communications.

COP Login Name. The login name for connecting to the IBM SPSS Collaboration and Deployment Services server on which the COP is running.

COP Password. The password for the login name.

COP Provider. The security provider against which to validate the user/password combination.

COP Enabled. Making a server COP enabled allows it to communicate with the COP. When an end user searches for available servers, the COP will display the server as one of the results.

Server Name. The COP displays this as the name of the server.

Description. The COP displays this description for the server.

Update Interval. The server software sends an update to the COP after the specified interval (in minutes) has elapsed. The update lets the COP know whether the server is running.

Weight. When a server is part of a cluster, the COP uses the weight to determine how many client computers can connect to an individual server. For example, a weight of 5 indicates that 5 times as many client computers can connect to the server compared to a server with a weight of 1.

Logging

The server software records session information in two types of log files:

- The active log file where session information is currently being recorded
- The backup log files

When the active log file reaches maximum size, it is moved to a backup log file, and the active log file starts over. The log file can also record performance information when the **Log Sample Interval** is set to a number greater than 0

Active Log(s). The name of the active log file.

Number of Backup Log Files. You can change the number of backup log files that the server keeps. Increase or decrease the number to accommodate the amount of information that you want to keep.

Performance Log Interval. By default, the log file does not record performance information. You can configure the log file to record this information if you are trying to determine problematic areas. You do this by setting the performance log interval to a number of seconds between 1 and 600. It is not recommended that you set the interval to a small number for a long period of time. Doing so may result in the log file being filled up quickly. It may also affect performance. To disable performance samples, set this to 0. Note that you can view performance log information directly in a grid. See the topic "Viewing Performance Log Information" on page 190 for more information.

Maximum Log File Size. When the active log file reaches maximum size, it is moved to the backup log file, and the active log file starts over. Increase or decrease the size to accommodate the number sessions that you want to record.

Viewing the Log Files

You can view the active log or the backup logs for a specific server.

To View a Log File

1. Select the IBM SPSS Statistics Server in the Server Administration pane.
2. From the menus choose:

   Tools > **IBM SPSS Statistics Server** > **View Active Log**
3. In the View Log Selection for IBM SPSS Statistics Server dialog box, select the log file that you want to view.
4. Click **OK**.

The file opens in a new pane. If you are looking for specific information in the file, press Ctrl+F to open the Find/Replace dialog box.

Viewing Performance Log Information

If the **Performance Log Interval** is set to a number greater than 0, performance log information is written to the log files. See the topic "Logging" on page 190 for more information. You can see the performance log information by viewing the log files. However, the performance log information is interspersed with all the other log information and may be difficult to read. The IBM SPSS Statistics Administration Console offers an alternate interface for reading the performance log directly in a grid.

To View the Performance Log Information Directly

1. Double-click the Performance Logs node beneath the desired IBM SPSS Statistics Server in the Server Administration pane.

Performance log information appears in the Performance Logs pane. Above the grid is the date/time range of the displayed performance log information. By default, performance log information from the last hour is displayed. You can change this date/time range. You can also subsequently export the displayed log information to a comma-separated values (CSV) file.

To Change the Performance Log Date/Time Range

1. Click **Change...** next to the displayed date/time range.
2. In the Select Date Range dialog box, enter a date and time for the start and end of the date/time range.
3. Click **OK**.

To Export the Performance Log Information to a CSV File

1. Click **Export...** next to the displayed date/time range.
2. In the Select File to Export dialog box, browse to the location where you want to save the CSV file.
3. Enter a name for the file. The file extension is added automatically, so you don't need to add it.
4. Click **Save**.

IBM SPSS Statistics Server User Profiles and Groups

The server software provides the ability to create profiles of individual users and groups of users. The profile specifies the temporary files directory, the UNIX `umask` setting, the CPU process priority, client data access, and maximum threads for the user or group of users. These settings override the associated global, default settings. For information about creating user profiles and groups, see “Creating and Editing IBM SPSS Statistics Server User Profiles and Groups” on page 191 .

Profile File

User profile and group settings are stored in the profile file, *UserSettings.xml*. By default, this file is located in the `config` subdirectory of the IBM SPSS Statistics Server installation directory. Before making changes to the user profiles and groups, it is recommended that you move the default file and specify the new location in the administration application (IBM SPSS Statistics Administration Console). See the topic “File Locations” on page 187 for more information.

How the Server Applies Settings

The server follows these steps when determining the temporary files directory, the umask setting, the CPU priority, client data access, and maximum number of threads for a particular user.

1. Look for a user profile for the user name and domain of the user connecting to the server. Select the first match found. On UNIX, the server ignores the domain, and case matters for the user name. On Windows, the case of the user name or domain does not matter. Also, if a user logged onto the server without specifying a domain, the server looks for a user name match with a blank domain. If it can't find one, it uses a match on only the user name.
2. If the user profile defines settings, apply these settings to the server process.
3. If the user profile does not define settings or does not define all settings, apply the user's group settings to the server process.
4. If some settings are still not defined or there was no matching user profile, apply the default umask, the temporary file directory, client data access setting, and maximum number of threads defined by the administration application (IBM SPSS Statistics Administration Console). No default CPU priority processing is used. For information about the configuring the default settings, see “Users” on page 188 .

Creating and Editing IBM SPSS Statistics Server User Profiles and Groups

User profiles and groups are managed in the Manage Users and Groups pane.

To Display the Manage Users and Groups Pane

1. Double-click the User Profiles and Groups node beneath the desired IBM SPSS Statistics Server in the Server Administration pane.

The Manage Users and Groups pane displays the currently defined user profiles and groups in the User Profiles and Groups grid. Select a user profile or groups from this grid. The grid on the right then displays the groups to which the selected user belongs or the users in the selected group, depending on what is selected. Both grids also display the current settings for the user profiles and groups.

To Create a New User Profile

1. In the Manage Users and Groups pane, click **New User Profile**.
2. In the Create New User Profile dialog box, enter the name of the user for whom you are creating the profile. If you want to limit the profile to a user with a specific domain, be sure to include the domain name (for example, `domain\user`). Note that you are not creating a user at this time. You are associating a profile to an *existing* user.
3. If necessary, define any of the available settings. If you are only creating a user profile to assign a user to a group, you don't have to define any settings. For descriptions of all the available settings, see “Available User Profile and Group Settings” on page 192 .
4. In the Manage Groups the User Belongs to area, identify the group(s) to which the user belongs. The grid on the left shows all defined groups, and the grid on the right shows the groups to which the user belongs. Select a defined group and click **>>>>** to add the user to the selected group. You can also click **Add All** to add the user to all groups. To remove a user from a group, select a group to which the user belongs and click **<<<<**. You can also click **Remove All** to remove the user from all groups.

To Create a New User Profile Group

1. In the Manage Users and Groups pane, click **New Group**.
2. In the Create New Group dialog box, enter a name for the group. The name of the group is arbitrary and does not correspond to a group on your system.
3. Define any of the available settings. For descriptions of all the available settings, see “Available User Profile and Group Settings” on page 192 .
4. In the Manage Users in Group area, identify the user(s) that belong to the group. The grid on the left shows all users with defined user profiles, and the grid on the right shows the users that belong to the group. Select a user and click **>>>>** to add the selected user to the group. You can also click **Add All** to add all users with defined user profiles to the group. To remove a user from a group, select a user in the group and click **<<<<**. You can also click **Remove All** to remove all users from the group.

To Edit or Delete an Existing User Profile or Group

1. In the Manage Users and Groups pane, select a user profile or group. If deleting, you can select multiple user profiles or groups. Be careful because there is no option to undo the deletion.
2. Click **Edit** to edit the user profile or group, or click **Delete** to delete the user profile or group.

Available User Profile and Group Settings

Following are the available settings that you can assign to user profiles and groups.

Priority
: The CPU priority for the server process. Use a negative or positive integer. On **UNIX**, this corresponds to the `nice` command, which has a range of -20 to 19, with -20 yielding more favorable scheduling and 19 yielding less favorable scheduling. `nice` settings can be viewed in the *NI* column of the `ps` command output. On **Windows**, a negative value yields a base priority of ABOVE_NORMAL (more favorable scheduling), a positive value yields a base priority of BELOW_NORMAL (less favorable scheduling), and `0` yields NORMAL priority scheduling. You can view base priority settings in the Task Manager.

Temp file location
The directory to which the server process writes temporary files. The user must have read/write access to this directory. To gain any performance benefit, you must specify a different physical drive for each user. You can use a different directory on the same drive for multiple users, but there will be no performance gain. You might do this so that users have temporary directories to which only they have access. You might also create different sized partitions and specify a partitioned drive for each user, thereby controlling the temporary files space allocated to each user. You can also specify multiple locations separated by commas. You can also use the administration application to specify this setting globally for all users. See the topic "Users" on page 188 for more information.

Client data access
Access to the data from the client. Select an option to indicate whether the user sees the data in the client Data Editor. This option may increase client/server performance because data passes can be slow over the network. **No Data Editor Data Access** indicates that the user does see the data in the Data Editor. Although the user is able to see all the cases in the data, this option may degrade performance in some circumstances. You can also use the administration application to specify this setting globally for all users. See the topic "Users" on page 188 for more information.

Maximum threads
The number of threads for the server process. This setting limits the number of threads that the server software can use for multithreaded procedures run by the user. You may do this to limit the system resources available for multithreaded procedures. You can also use the administration application to specify this setting globally for all users. See the topic "Users" on page 188 for more information.

Cache Compression
When end users explicitly issue a CACHE command or run a procedure that creates scratch files automatically, the server software can compress the scratch (temporary) files using zlib. You should turn cache compression on if an end user or group of users typically reads large data files. You can also use the administration application to specify this setting globally for all users. See "Users" on page 188 for more information about cache compression and the global setting.

Compression override
You can specify whether an end user or group of users can override the cache compression setting with `SET ZCOMPRESSION` syntax. You can also use the administration application to specify this setting globally for all users. See the topic "Users" on page 188 for more information.

Maximum JVM memory
When end users connect to the server software, memory is allocated for the Java virtual machine (JVM) on the server computer. Memory is allocated for each user. You can change the amount of allocated memory for a specific end user or group of users by specifying a number in megabytes (MB). You may need to this if users have jobs that require a large amount of memory, for example when exporting very large custom tables. You can also use the administration application to specify this setting globally for all users. See the topic "Users" on page 188 for more information.

Umask
The umask for the server process. Use an octal three-digit number. You can also use the administration application to specify this setting globally for all users. See the topic "Users" on page 188 for more information.

Monitoring IBM SPSS Statistics Server Users

The Monitor Current Users pane of the IBM SPSS Statistics Administration Console shows users that are connected to the IBM SPSS Statistics Server computer. To activate the pane, double-click the Monitor Current Users node beneath the desired server in the Server Administrator pane. This populates the pane with a list of connected users. The list refreshes at the rate shown. To refresh the list manually, from the menus choose:

View > **Refresh**

The monitoring pane displays the following information.

- **User.** The ID with which the user connected to the server software.
- **Client.** The name of the client application that the user is running.

- **Version.** The version of the client application that the user is running.
- **Connection ID.** An arbitrary number assigned to a user when the user connects to the server software. A lower number indicates that a user logged in earlier than another user with a higher number.
- **Authentication.** The access rights of the user.

Note: If the name or version is listed as**<Unknown>**, the user is connected with client software that the IBM SPSS Statistics Administration Console cannot recognize.

You can perform user-specific actions in this pane. Right-click to see a list of choices:

- **Disconnect.** Disconnect one or more users from the server software. See the topic "Disconnecting Users" on page 194 for more information.
- **Broadcast.** Send a message to one or more users. See the topic "Broadcasting a message to users" on page 194 for more information.

Disconnecting Users

Disconnect users when you need to limit drain on the server computer, or when you notice a user using the system inappropriately. Avoid disconnecting all users since it is possible that they may lose work. Your ID also appears in the user list. The IBM SPSS Statistics Administration Console will prevent you from disconnecting yourself. If all users—including yourself—need to be disconnected from the server computer, disconnect all users and then disconnect from the server.

1. On the Monitor Current User pane, click the Refresh button to ensure the user list is current.
2. Select the user(s) you want to disconnect.
3. Right-click and choose **Disconnect User(s)** from the menu.

Broadcasting a message to users

You can broadcast a message to a specific user or all users who are connected to the server software. You may do this if you need to restart the server software and want to warn users before doing so.

1. On the Monitor Current User pane, click the Refresh button to ensure the user list is current.
2. Select the users to whom you want to broadcast a message. If you want to broadcast the same message to all users, you don't need to select specific one.
3. Right-click and choose **Broadcast [All] Users** from the menu.
4. In the Broadcast dialog box, type the content of the message. You can add line breaks if needed.
5. Click **OK** to broadcast the message.

Accessibility with the Keyboard

- Use the Tab and arrow keys to move between controls in each pane.
- Use the arrow keys to move between items in the Server Administration pane. Press Enter to open and move to the pane for the selected item.
- Use the following key combinations to select the menus.
  - File menu: Alt-F
  - Edit menu: Alt-E
  - View menu: Alt-V
  - Tools menu: Alt-T
  - Help menu: Alt-H
- Use **View** > **Navigation** to move between panes.

IBM SPSS Modeler Server Administration

The Modeler Administration Console in IBM SPSS Deployment Manager provides a console user interface to monitor and configure your SPSS Modeler Server installations, and is available free-of-charge to current SPSS Modeler Server customers. The application can be installed only on Windows computers; however, it can administer a server installed on any supported platform.

Many of the options available through the Modeler Administration Console can also be specified in the `options.cfg` file, which is located in the SPSS Modeler Server installation directory under `/config`. However, the Modeler Administration Console provides a shared graphical interface that allows you to connect, configure, and monitor multiple servers.

Starting Modeler Administration Console

From the Windows Start menu, choose **[All] Programs**, then **IBM SPSS Collaboration and Deployment Services**, then **Deployment Manager**.

When you first run the application, you see empty Server Administration and Properties panes (unless you already have Deployment Manager installed with an IBM SPSS Collaboration and Deployment Services server connection already set up). After you configure Modeler Administration Console, the Server Administrator pane on the left displays a node for each SPSS Modeler Server that you want to administer. The right-hand pane shows the configuration options for the selected server. You must first set up a connection for each server that you want to administer.

Restarting the web service

Whenever you make changes to an IBM SPSS Modeler Server in the Administration Console, you must restart the web service.

To restart the web service on Microsoft Windows:

1. On the computer where you installed IBM SPSS Modeler, select **Services** from Administrative Tools on the Control Panel.
2. Locate the server in the list and restart it.
3. Click **OK** to close the dialog box.

To restart the web service on UNIX:

On UNIX, you must restart the IBM SPSS Modeler Server by running the **modelersrv.sh** script in the IBM SPSS Modeler Server installation directory.

1. Change to the IBM SPSS Modeler Server installation directory. For example, at a UNIX command prompt, type:

   `cd /usr/<modelersrv>`, where `modelersrv` is the IBM SPSS Modeler Server installation directory.
2. To stop the server, at the command prompt, type

   `./modelersrv.sh stop`
3. To restart the server, at the command prompt, type

   `./modelersrv.sh start`

Configuring Access with Modeler Administration Console

Administrator access to SPSS Modeler Server through the Modeler Administration Console included with IBM SPSS Deployment Manager is controlled with the `administrators` line in the `options.cfg` file, located in the SPSS Modeler Server installation directory under `/config`. This line is commented out by

default, so you must edit this line to allow access to specific people, or use * to allow access to all users, as shown in the following examples:

```
administrators, "*"
administrators, "jsmith,mjones,achavez"
```

- The line must begin with `administrators,` and the entries must be contained in quotation marks. Entries are case sensitive.
- Separate multiple user IDs with commas.
- For Windows accounts, do not use domain names.
- Use the asterisk with care. It allows anyone with a valid user account for IBM SPSS Modeler Server (which, in most cases, is anyone on the network) to log in and change the configuration options.

Configuring Access with User Access Control

To use the Modeler Administration Console to make updates to a SPSS Modeler Server configuration installed on a Windows machine that has User Access Control (UAC) enabled, you must have read, write, and execute permissions defined on the *config* directory and on the *options.cfg* file. These (NTFS) permissions must be defined at the specific user level and not at group level, this is due to the way that UAC and NTFS permissions interact.

The Modeler Administration Console is included in IBM SPSS Deployment Manager.

SPSS Modeler Server connections

You must specify a connection to each SPSS Modeler Server on your network that you want to administer. You must then log in to each server. Although the server connection is remembered across Modeler Administration Console sessions in IBM SPSS Deployment Manager, the login credentials are not. You must log in every time you start IBM SPSS Deployment Manager.

To set up a server connection

1. Ensure that the IBM SPSS Modeler Server service is started.
2. From the File menu, choose **New** and then **Administered Server Connection**.
3. On the first page of the wizard, enter a name for your server connection. The name is for your own use and should be something descriptive; for example, *Production Server*. Ensure that Type is set to **Administered IBM SPSS Modeler Server**, then click **Next**.
4. On the second page, enter the hostname or IP address of the server. If you have changed the port from the default, enter the port number. Click **Finish**. The new server connection is shown in the Server Administrator pane.

To perform administration tasks, you must now log in.

To log in to the server

1. In the Server Administrator pane, double-click to select the server to which you want to log in.
2. In the Login dialog box, enter your credentials. (Use your user account for the server host.) Click **OK**.

If the login fails with the message **Unable to obtain administrator rights on server**, the most likely cause is that administrator access has not been configured correctly. See the topic “Configuring Access with Modeler Administration Console ” on page 195 for more information.

If the login fails with the message **Failed to connect to server '<server>'**, make sure that the user ID and password are correct, then make sure that the IBM SPSS Modeler Server service is running. For example, under Windows, go to Control Panel > Administrative Tools > Services and check the entry for IBM SPSS Modeler Server. If the Status column does not show **Started**, select this line on the screen and click **Start**, then retry the login.

Once you log in to your IBM SPSS Modeler Server, two options are shown below the server name, Configuration and Monitoring. Double-click one of these options.

SPSS Modeler Server Configuration

The Configuration pane shows configuration options for SPSS Modeler Server. Use this pane to change the options as desired. Click **Save** on the toolbar to save the changes. Note that changing any option marked with an asterisk (*) requires a server restart to take effect.

The options are described in the following sections, with the corresponding line in `options.cfg` given in parentheses for each option. Options that are visible only in `options.cfg` are described at the end of this section.

Note: If a non-root user wants to change these options, write permission is required for the SPSS Modeler Server **`config`** directory.

Connections/Sessions

Modeler port number. (`port_number`) The port number for SPSS Modeler Server to listen on. Change if another application already uses the default. End users must know the port number in order to use SPSS Modeler Server.

Embedded DataView service port number. (`data_view_port_number`) The port number for the DataView service embedded in SPSS Modeler Server to listen on. Change if another application already uses the default.

Maximum number of connections. (`max_sessions`) Maximum number of server sessions at one time. A value of –1 indicates no limit.

Analytic Server connection

Enable Analytic Server SSL (`as_ssl_enabled`). Specify Y to encrypt communications between Analytic Server, and SPSS Modeler otherwise, N.

Host (`as_host`). The IP address of the Analytic Server.

Port Number (`as_port`). The Analytic Server port number.

Context Root (`as_context_root`). The context root of the Analytic Server.

Tenant (`as_tenant`). The tenant that the SPSS Modeler Server installation is a member of.

Realm (`as_realm`). The realm used for this Analytic Server.

Prompt for Password (`as_prompt_for_password`). Specify N if the SPSS Modeler Server is configured with the same authentication system for users and passwords as the system that is used on Analytic Server; for example, when you use Kerberos authentication, otherwise, Y.

Note: If you intend to use Kerberos SSO, you must set extra options in the `options.cfg` file. For more information, see the topic "Options visible in options.cfg " later in this chapter.

Note: To connect SSL enabled Analytic Server, extra steps are required as follows:

1. Use the following command to extract the certificate file `trust.cer` from the JKS file (i.e. `trust.jks`):

```
/bin/keytool -export -alias server-alias -storepass pass4jks -file /home/sslkeys/trust.cer
-keystore /home/sslkeys/trust.jks
```

2. Import the `trust.cer` file into `cacerts` in the JRE used by your application server.
3. Import the `trust.cer` file into `caserts` in the JRE used by the SPSS Modeler Server.
4. Restart the SPSS Modeler Server and the IBM SPSS Collaboration and Deployment Services Repository Server.

Data file access

Python executable path for Extension nodes and Custom Dialog Builder..
(eas_pyspark_python_path) Full path to the Python executable file, including the name of the file. [program_files_restricted] might need to be set to No depending on your Python installation location.

Restrict access to only data file path. (data_files_restricted) When set to **yes**, this option restricts data files to the standard data directory and any files that are listed in the **Data File Path**. If you want to use the **View Data** feature while this restriction is enabled, you must set matching temporary file paths in both the temp_directory and data_file_path parameters.

Data file path. (data_file_path) A list of directories to which clients are allowed to read and write data files. This option is ignored unless the **Restrict Access to Only Data File Path** option is turned on. Use forward slashes in all path names. On Windows, specify multiple directories by using semicolons (for example, [server install path]/data;c:/data;c:/temp). On Linux and UNIX, use colons (:) instead of semicolons. The data file path must include any path that is specified in the temp_directory parameter.

Restrict access to only program files path. (program_files_restricted) When set to **yes**, this option restricts program file access to the standard bin directory and any files that are listed in the **Program files path**. As of release 17, the only program file to which access is restricted is the Python executable (see the **Python executable path**).

Program files path. (program_file_path) A list of directories from which clients are allowed to execute programs. This option is ignored unless the **Restrict Access to Only Program Files Path** option is turned on. Use forward slashes in all path names. Specify multiple directories by using semicolons.

Maximum file size. (max_file_size) Maximum size (in bytes) of temporary and exported data files that are created during stream execution (does not apply to SAS and SPSS Statistics data files). A value of –1 indicates no limit.

Temporary directory. (temp_directory) The directory used to store temporary data files (cache files). Ideally, this directory is on a separate high-speed drive or controller because speed of access to this directory can have a significant impact on performance. You can specify multiple temporary directories, separating each with a comma (for example: temp_directory, "D:/Modeler_temp, C:/Program Files/IBM/SPSS/ModelerServer/<version>/Tmp"). These directories should be on different disks. The first directory is used most often, and the other directories are used to store temporary work files when certain data preparation operations (such as sort) use parallelism during execution. Allowing each execution thread to use separate disks for temporary storage can improve performance. Use forward slashes in all path specifications.

Note:

- Temporary files are generated in this directory during startup of SPSS Modeler Server. Ensure that you have the necessary access rights to this directory (for example, if the temporary directory is a shared network folder), otherwise SPSS Modeler Server startup fails.
- The temp_directory setting does not apply when you run Evaluation streams through IBM SPSS Collaboration and Deployment Services jobs. When you run such a job, a temporary file is created. By default, the file is saved to the IBM SPSS Modeler Server installation directory. You can change the default data folder that the temp files are saved to when you create the IBM SPSS Modeler Server connection in IBM SPSS Modeler.

Python executable path for bulk loading. (python_exe_path) Full path to the Python executable including the executable name. If access to program files is restricted, then you must add the directory that contains the Python executable to the program_file_path parameter (see **Restrict access to only program files path**).

Path to OPL library of full version of CPLEX. (cplex_opl_lib_path) Path to the OPL library for the full version of CPLEX.

Performance/Optimization

Stream rewriting. (`stream_rewriting_enabled`) Allows the server to optimize streams by rewriting them. For example, the server might push data reduction operations closer to the source node to minimize the size of the dataset as early as possible. Disabling this option is normally recommended only if the optimization causes an error or other unexpected results. This setting overrides the corresponding client optimization setting. If this setting is disabled in the server, then the client cannot enable it. But if it is enabled in the server, the client can choose to disable it.

Parallelism. (`max_parallelism`) Describes the number of parallel worker threads that SPSS Modeler is allowed to use when running a stream. Setting this to 0 or any negative number causes IBM SPSS Modeler to match the number of threads to the number of available processors on the computer; the default value for this option is –1. To turn off parallel processing (for machines with multiple processors), set this option to 1. To allow limited parallel processing, set it to a number smaller than the number of processors on your machine. Note that a hyperthreaded or dual-core processor is treated as two processors.

Buffer size (bytes). (`io_buffer_size`) Data files transferred from the server to the client are passed through a buffer of this number of bytes.

Cache compression. (`cache_compression`) An integer value in the range 0 to 9 that controls the compression of cache and other files in the server's temporary directory. Compression reduces the amount of disk space used, which can be important when space is limited. Compression increases processor time, but this is almost always made up by the reduction in disk access time. Note that only certain caches, those accessed sequentially, can be compressed. This option does not apply to random-access caches, such as those used by the network training algorithms. A value of 0 disables compression entirely. Values from 1 upward provide increasing degrees of compression but with a corresponding cost in access time. The default value is 1; higher values may be needed where disk space is at a premium.

Memory usage multiplier. (`memory_usage`) Controls the proportion of physical memory allocated for sorting and other in-memory caches. The default is 100, which corresponds to approximately 10% of physical memory. Increase this value to improve sort performance where free memory is available, but be careful of increasing it so high as to cause excessive paging.

Modeling memory limit percentage. (`modelling_memory_limit_percentage`) Controls the proportion of physical memory allocated for training Kohonen and *k*-means models. The default is 25%. Increase this value to improve training performance where free memory is available, but be careful of increasing it so high as to cause excessive paging when data spills onto the disk.

Allow modeling memory override. (`allow_modelling_memory_override`) Enables or disables the **Optimize for Speed** option in certain modeling nodes. The default is enabled. This option allows the modeling algorithm to claim all available memory, bypassing the percentage limit option. You may want to disable this if you need to share memory resources on the server machine.

Maximum and minimum server port. (`max_server_port` and `min_server_port`) Specifies the range of port numbers that can be used for the additional socket connections between client and server that are required for interactive models and stream execution. These require the server to listen on another port; not restricting the range could cause problems for users on systems with firewalls. Default value for both is –1, meaning "no restriction." Thus, for example, to set the server to listen on port 8000 or above, you would set `min_server_port` to 8000 and `max_server_port` to –1.

Note that you must open additional ports over the main server port to open or execute a stream, and correspondingly more ports if you want to open or execute concurrent streams. This is required in order to capture feedback from the stream execution.

By default, IBM SPSS Modeler will use any open port that is available; if it does not find one (for example, if they are all closed by a firewall), an error is displayed when you execute the stream. To configure the range of ports, IBM SPSS Modeler will need two open ports (in addition to the main server port) available per concurrent stream, plus 3 additional ports for each ODBC connection from within any connected client (2 ports for the ODBC connection for the duration of that ODBC connection, and an additional temporary port for authentication).

Note: An ODBC connection is an entry in the database connections list, and can be shared between multiple database nodes specified with the same database connection.

Note: It is possible that the authentication ports can be shared if the connections are made at different times.

Note: Best practice dictates that the same ports should be used to communicate with both IBM SPSS Collaboration and Deployment Services and SPSS Modeler Client. These can be set as `max_server_port` and `min_server_port`.

Note: If you change these parameters, you need to restart SPSS Modeler Server for the change to take effect.

Array fetch optimization. (`sql_row_array_size`) Controls the way that SPSS Modeler Server fetches data from the ODBC datasource. The default value is 1, which fetches a single row at a time. Increasing this value causes the server to read the information in larger chunks, fetching the specified number of rows into an array. With some operating system/database combinations, this can result in improvements to the performance of SELECT statements.

SQL

Maximum SQL string length. (`max_sql_string_length`) For a string imported from the database with SQL, the maximum number of characters that are guaranteed to be passed successfully. Depending on the operating system, string values longer than this may be truncated on the right without warning. The valid range is between 1 and 65,535 characters. This property also applies to the Database export node.

Note: The default value for this parameter is 2048. If the text you are analyzing is longer than 2048 characters (for example, this may occur if using the SPSS Modeler Text Analytics Web Feed node) we recommend increasing this value if working in native mode otherwise your results may be truncated. If you are using a database and user-defined functions (UDF), this restriction does not occur; this can account for differences in results between native and UDF modes.

Automatic SQL generation. (`sql_generation_enabled`) Allows automatic SQL generation for streams, which may substantially improve performance. The default is enabled. Disabling this option is recommended only if the database is not able to support queries submitted by SPSS Modeler Server. Note that this setting overrides the corresponding client optimization setting; also note that for purposes of scoring, SQL generation must be enabled separately for each modeling node regardless of this setting. If this setting is disabled in the server, then the client cannot enable it. But if it is enabled in the server, the client can choose to disable it.

Default SQL string length. (`default_sql_string_length`). Specifies the default width of string columns that will be created within database cache tables. String fields in database cache tables will be created with a default width of 255 if there is no upstream type information. If you have wider values than this in your data, either instantiate an upstream Type node with those values, or set this parameter to a value that is large enough to accommodate those string values.

Enable Database UDF. (`db_udf_enabled`). If set to Y (default), causes the SQL generation option to generate user-defined function (UDF) SQL instead of pure SPSS Modeler SQL. UDF SQL usually outperforms pure SQL.

SSL

Enable SSL. (`ssl_enabled`) Enables SSL encryption for connections between SPSS Modeler and SPSS Modeler Server.

SSL keystore file. (`ssl_keystore`) The SSL key database file to be loaded when the server starts (either a full path or a relative path to the SPSS Modeler installation directory).

SSL keystore stash file. (`ssl_keystore_stash_file`) The name of the key database password stash file to be loaded when the server starts up (either a full path or a relative path to the SPSS Modeler installation directory). On Linux, if you want to leave this setting blank and be prompted for the password when starting the SPSS Modeler Server, see the following instructions:

- On Linux/UNIX:
  1. Make sure the `ssl_keystore_stash_file` setting in `options.cfg` does not have a value.

2. Locate the following line in the `modelersrv.sh file`:

   ```
   if "$INSTALLEDPATH/$SCLEMDNAME" -server $ARGS; then
   ```

3. Add the `-request_ssl_password` switch as follows:

   ```
   if "$INSTALLEDPATH/$SCLEMDNAME" -request_ssl_password -server $ARGS; then
   ```

4. Restart the SPSS Modeler Server. You will be prompted for a password. Enter the correct password, click **OK**, and the server will start.

Keystore certificate label. (`ssl_keystore_label`) Label for the specified certificate.

Note: To use the Administration Console with a server setup for SSL, you must import any certificates required by SPSS Modeler Server into the Deployment Manager trust store (under `../jre/lib/security`).

Note: If you change these parameters, you need to restart SPSS Modeler Server for the change to take effect.

Coordinator of Processes configuration

Host. (`cop_host`) The hostname or IP address of the Coordinator of Processes service. The default "spsscop" is a vanity name which administrators can choose to add as an alias for the IBM SPSS Collaboration and Deployment Services host in DNS.

Port number. (`cop_port_number`) The port number of the Coordinator of Processes service. The default, 8080, is the IBM SPSS Collaboration and Deployment Services default.

Context root. (`cop_context_root`) The URL of the Coordinator of Processes service.

Login name. (`cop_user_name`) The user name for authentication to the Coordinator of Processes service. This is an IBM SPSS Collaboration and Deployment Services login name so may include a security-provider prefix (for example: ad/jsmith).

Password. (`cop_password`) The password for authentication to the Coordinator of Processes service.

Note: If you update the `options.cfg` file manually instead of using the Modeler Administration Console in IBM SPSS Deployment Manager, you must manually encode the `cop_password` value that you specify in the file. Plain text passwords are invalid and cause registration with the Coordinator of Processes to fail.

Follow these steps to manually encode the password:

1. Open a Command Prompt, navigate to the SPSS Modeler `./bin` directory, and run the command `pwutil.bat/sh`.
2. When requested, type the user name (the `cop_user_name` you are specifying in `options.cfg`) and press Enter.
3. When requested, type the password for that user.

   The encoded password is displayed between double quotes on the command line as part of the returned string. For example:

   ```
   C:\Program Files\IBM\SPSS\Modeler\18\bin>pwutil
   User name: copuser
   Password: Pass1234
   copuser, "0Tqb4n.ob0wrs"
   ```

4. Copy the encoded password, without the double quotes, and paste it between the double quotes that already exist for the `cop_password` value in the `options.cfg` file.

Enabled. (`cop_enabled`) Determines whether the server should attempt to register with the Coordinator of Processes. The default is *not* to register because the administrator should choose which services are advertised through the Coordinator of Processes.

SSL Enabled. (`cop_ssl_enabled`) Determines whether SSL is used to connect to the Coordinator or Processes server. If this option is used, you must import the SSL certificate file into the SPSS Modeler

Server JRE. To do this, you must obtain the SSL certificate file and its alias name and password. Then run the following command on the SPSS Modeler Server:

```
$JAVA_HOME/bin/keytool -import -trustcacerts -alias $ALIAS_NAME -file
$CERTIFICATE_FILE_PATH -keystore $ModelerServer_Install_Path/jre/lib/security/
cacerts
```

Server name. (`cop_service_name`) The name of this SPSS Modeler Server instance; the default is the host name.

Description. (`cop_service_description`) A description of this instance.

Update interval (min). (`cop_update_interval`) The number of minutes between keep-alive messages; the default is 2.

Weight. (`cop_service_weight`) The weight of this instance, specified as an integer between 1 and 10. A higher weight attracts more connections. The default is 1.

Service host. (`cop_service_host`) The fully-qualified host name of the IBM SPSS Modeler Server host. The default of the host name is derived automatically; the administrator can override the default for multi-homed hosts.

Default data path. (`cop_service_default_data_path`) The default data path for a Coordinator of Processes registered IBM SPSS Modeler Server installation.

Options visible in options.cfg

Most configuration options can be changed using the IBM SPSS Modeler Administration Console included with IBM SPSS Deployment Manager. But there are some exceptions, such as those described in this section. The options in this section must be changed by editing the `options.cfg` file. See "IBM SPSS Modeler Server Administration" on page 195 and "Using the options.cfg file" on page 204 for more information. Note that there may be additional settings in `options.cfg` that are not listed here.

Note: This information only applies to a remote server (IBM SPSS Modeler Server, for example).

`administrators`. Specify the user names of those users to whom you want to grant administrator access. See the topic "Configuring Access with Modeler Administration Console " on page 195 for more information.

`allow_config_custom_overrides`. Do not modify unless instructed to do so by a technical-support representative.

`data_view_port_number`. You can right-click a data node and select **View Data** to examine and refine your data in interesting ways with advanced data visualizations. This feature uses the port number `28900` by default. Modify the value for this `data_view_port_number` configuration option if you need to use a different port number. We recommend using the default, if possible.

`fips_encryption`. Enables FIPS compliant encryption. The default is `N`.

`group_configuration`. When enabled, IBM SPSS Modeler Server checks the `groups.cfg` file which controls who can log on to the server.

`max_transfer_size`. For internal system use only. **Do not modify**.

`shell`. (UNIX servers only) Overrides the default setting for the UNIX shell, for example `shell, "/usr/bin/ksh"`. By default, IBM SPSS Modeler uses the shell defined in the user profile of the user who is connecting to IBM SPSS Modeler Server.

`start_process_as_login_user`. Set this to Y if you are running SPSS Modeler Server with a private password database, starting the server service from a non-root account.

`use_bigint_for_count`. When the number of the records to be counted is larger than what a normal integer (2^31-1) can hold and is within the range of -2^63 to 2^63-1, set this option to Y. When this option is set to Y, and a stream is connected to Db2; SQL Server; or a Teradata, Oracle, or Netezza database, a function is used where a record count is needed (for example, the **Record_Count** field generated by the Aggregate node).

When this option is enabled, and if working with either Db2 or SQL Server, SPSS Modeler uses COUNT_BIG() for record counting. If working with Teradata, Oracle, or Netezza, SPSS Modeler will use COUNT(). For all other databases, there is no SQL pushback for the function. The difference is that when `use_bigint_for_count` is enabled, all record counts are saved as BIG INT (or LONG) type, as compared to normal integer when the options is disabled.

From the point of view of database numeric, see the following table for the results of setting `use_bigint_for_count` to Y and N for the same data.

Table 17. Database numeric interpretations when use_bigint_for_count is set to Y and N

Database Numeric	**Interpretation for data when `use_bigint_for_count` is set to Y**	**Interpretation for data when `use_bigint_for_count` is set to N**
Numeric(18,0)	int	int
Numeric(19,0)	int	string
Numeric(38,0)	int	string

`cop_ssl_enabled`. Set this option to Y if you are using SSL to connect to the Coordinator of Processes Service. If this option is used, you must import the SSL certificate file into the SPSS Modeler Server JRE. To do this, you must obtain the SSL certificate file and its alias name and password. Then run the following command on the SPSS Modeler Server:

```
$JAVA_HOME/bin/keytool -import -trustcacerts -alias $ALIAS_NAME -file
$CERTIFICATE_FILE_PATH -keystore $ModelerServer_Install_Path/jre/lib/security/
cacerts
```

`cop_service_default_data_path`. You can use this option to set the default data path for a Coordinator of Processes registered IBM SPSS Modeler Server installation.

Users can create their own Analytic Server connections in SPSS Modeler via **Tools** > **Analytic Server Connections**. Administrators can also define a default Analytic Server connection using the following properties:

`as_ssl_enabled`. Y or N.

`as_host`. Specify the Analytic Server host name or IP address.

`as_port`. Specify the Analytic Server port number.

`as_context_root`. Specify the Analytic Server context root.

`as_tenant`. Specify the name of the tenant that the IBM SPSS Modeler Server is a member of

`as_prompt_for_password`. Y or N.

By default, Analytic Server authentication using the Kerberos method is not enabled. To enable Kerberos authentication, use the three following properties:

`as_kerberos_auth_mode`. To enable Kerberos authentication, set this option to Y.

`as_kerberos_krb5_conf`. Specify the path to the Kerberos configuration file that Analytic Server should use; for example, `c:\windows\krb5.conf`.

`as_kerberos_krb5_spn`. Specify the Analytic Server Kerberos SPN; for example, `HTTP/ashost.mydomain.com@MYDOMAIN.COM`.

SPSS Modeler Server Monitoring

The monitoring pane of Modeler Administration Console in IBM SPSS Deployment Manager shows a snapshot of all processes running on the SPSS Modeler Server computer, similar to the Windows Task Manager. To activate the monitoring pane, double-click the Monitoring node beneath the desired server in the Server Administrator pane. This populates the pane with a current snapshot of data from the server.

The data refreshes at the rate shown (one minute by default). To refresh the data manually, click the **Refresh** button. To show only SPSS Modeler Server processes in this list, click the **Filter out non-SPSS Modeler processes** button.

Using the options.cfg file

The `options.cfg` file is located in the `[server install path]/config` directory. Each setting is represented by a comma-separated name-value pair, where the `name` is the name of the option and the `value` is the value for the option. Pound (hash) signs (#) indicate comments.

Note: Most configuration options can be changed using IBM SPSS Modeler Administration Console in IBM SPSS Deployment Manager, rather than this configuration file, but there are a few exceptions. See the topic "Options visible in options.cfg" on page 202 for more information.

By using IBM SPSS Modeler Administration Console, you can avoid server restarts for all options except the server port. See the topic "IBM SPSS Modeler Server Administration" on page 195 for more information.

Note: This information only applies to a remote server (IBM SPSS Modeler Server, for example).

Configuration options that can be added to the default file

By default, in-database caching is enabled with IBM SPSS Modeler Server. You can disable this feature by adding the following line to the `options.cfg` file:

```
enable_database_caching, N
```

Doing so causes temporary files to be created on the server and not in the database.

To view or change the IBM SPSS Modeler Server configuration options:

1. Open the `options.cfg` file with a text editor.
2. Locate the options of interest. For a full list of options, see " SPSS Modeler Server Configuration" on page 197.
3. Edit the values, as appropriate. Note that all pathname values must use a forward slash (/) rather than a backslash as the pathname separator.
4. Save the file.
5. Stop and restart IBM SPSS Modeler Server so that the changes will take effect.

Closing unused database connections

By default, IBM SPSS Modeler caches at least one connection to a database once that connection has been accessed. The database session is held open even when streams requiring database access are not being executed.

Caching database connections can improve execution times by removing the need for IBM SPSS Modeler to reconnect to the database each time a stream is executed. However, in some environments, it is important for applications to release database resources as quickly as possible. If too many IBM SPSS Modeler sessions maintain connections to the database that are no longer used, database resources may become exhausted.

You can avoid this possibility by turning off the IBM SPSS Modeler option `cache_connection` in a custom database configuration file. Doing so can also make IBM SPSS Modeler more resilient to faults in the database connection (such as timeouts) that can occur when connections are used over a long period of time by an IBM SPSS Modeler session.

To cause unused database connections to be closed:

1. Locate the *[server install path]/config* directory.
2. Add the following file (or open it, if it already exists):

odbc-custom-properties.cfg

3. Add the following line to the file:

```
cache_connection, N
```

4. Save and close the file.
5. Restart IBM SPSS Modeler Server.

Note:

In-database caches are saved in database as either a regular table or a temporary table, depending on each database's implementation. For example, temporary tables are used for Db2, Oracle, Amazon Redshift, Sybase, and Teradata. For these databases, setting `cache_connection` to `N` does not work as expected because the temporary table is only valid within a session (it will be cleaned automatically by the database when the database connection is closed).

So when running an SPSS Modeler stream against one of these databases with `cache_connection` set to `N`, an error such as **Failed to create table for in-database caching. Using file cache instead.** may result. This indicates that SPSS Modeler failed to created the in-database cache. Also, in some cases for an SPSS Modeler-generated SQL query, a temporary table is used but the table is empty.

To work around this issue, you can choose to use a regular database table for in-database caches. To do this, create a custom database property configuration file that contains the following line:

```
table_create_temp_sql, 'CREATE TABLE <table-name> <(table-columns)>'
```

This forces a regular database table to be used for the in-database cache, and the table will be dropped when all connections to the database are closed or when the working stream is closed.

Chapter 22. Accessibility features

Accessibility features help a user who has a physical disability, such as a visual impairment or limited mobility, to use software products successfully.

This section provides an overview of alternative methods for accessing the functionality of the product. More specifically, the following topics are covered:

- Keyboard navigation of the user interface
- Special issues for visually impaired users
- Special issues for blind users

Keyboard navigation

Keyboard shortcuts allow you to navigate the user interface without using a mouse. At the most basic level, you can press the Alt key plus the appropriate key to activate window menus or press the Tab key to scroll through dialog box controls.

The following table describes general shortcuts that may be used throughout IBM SPSS Deployment Manager.

Table 18. General keyboard shortcuts

Shortcut	**Action**
Shift + F10	Brings up a menu, similar to a right-click.
Ctrl + F7	Brings up a list of active views and allows you to select one.
Tab	Iterates through the different controls on the view, such as tables or toolbars
Shift + Tab	Iterates through the different controls on the view in reverse.
Ctrl + Tab	Advances to the next control on the view when focus is on a control that accepts a Tab as input. For example, the Message text box of the notification dialog boxes allows tab entries. To move to the next control instead of inserting a tab character, use Ctrl+Tab.
Ctrl + M	Maximizes the active view or returns a maximized view to its original size
Alt	Highlights mnemonics. The menu titles in the menu bar have their mnemonic letters underlined.
Alt + mnemonic	Selects the control that is mapped to the mnemonic, such as a menu item, button, check box, or text field. for example, Alt + F opens the *File* menu. Pressing Alt + N in the Roles editor brings up the Create New Role dialog box.

Navigating the content explorer

The following table describes keyboard shortcuts that may be used within the Content Explorer.

Table 19. Content Explorer shortcuts

Shortcut	**Action**
Up/down Arrows	Highlights items in the tree.
Enter	Opens the selected item.

Navigating tables

The following table describes shortcuts that may be used within tables appearing in views and dialog boxes. Use the Tab key to select the table for subsequent navigation.

Table 20. Table shortcuts

Shortcut	Action
Up/down Arrows	Highlights items in the tables.
Right arrow	Expands a selected item, if the item can be expanded.
Left arrow	Collapses an expanded item.
Ctrl+<column number>	For a highlighted row, moves focus to the cell in the specified column if the cell value can be modified. Both keys must be released simultaneously. If the value cannot be modified, the focus is unchanged.

Navigating the job history and job schedule views

The following table describes keyboard shortcuts that may be used within the job history and job schedule views.

Table 21. Job History and Job Schedule shortcuts

Shortcut	Action
Up/down Arrows	Highlights items in the tables.
Right arrow	Highlights items in tables. Expands a selected item, if the item can be expanded.
Left arrow	Collapses an expanded item.
Enter	Opens a selected item in the editor. (*Note*: For jobs, use Shift + F10 to open a menu for a selected Job and select **Open In Job Editor**.)

Navigating the job editor

The following table describes keyboard shortcuts that may be used within the job editor.

Table 22. Job Editor shortcuts

Shortcut	Action
Up/down Arrows	Highlights options in the job palette. Moves the cursor around in the graphical editor. If the Move anchor is selected, repositions a job step in the editor.
Left-arrow/right-arrow	Moves the cursor around in the graphical editor. Cycles the selection through the different job steps and the event flow items as well. If the Move anchor is selected, repositions a job step in the editor.
Period (**.**)	Cycles through the move options for a selected job step.
Plus sign (**+**)	Adds a new step of the type that is selected in the palette to the job.

Navigating the help system

The following table describes keyboard shortcuts that may be used within the help system.

Table 23. Help system shortcuts	
Shortcut	**Action**
Up/down Arrows	On the Contents tab, moves up and down through the topics.
Right Arrow	On the Contents tab, expands any subtopics of the current topic.
Left Arrow	On the Contents tab, moves to the parent of the current topic. If the current topic has expanded subtopics, the subtopics are closed.
Shift+F10	On the Contents tab, displays the menu from which the selected topic can be opened.

Accessibility for visually impaired users

IBM SPSS Deployment Manager honors the contrast settings specified for your operating system. To adjust these settings, consult your operating system documentation.

In addition, you have control over the magnification in PDF files viewed in Acrobat Reader. To set the magnification in Acrobat Reader:

1. From the View menu, choose the **Zoom To** option on the Zoom submenu.
2. Specify the magnification level.

Accessibility issues for blind users

Support for blind users is predominately dependent on the use of a screen reader.

IBM SPSS Collaboration and Deployment Services has been fully tested with a screen reader. For information on installing and configuring your screen reader, consult the vendor documentation.

Notices

This information was developed for products and services offered in the US. This material might be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you provide in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBMproducts. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to actual people or business enterprises is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

Privacy policy considerations

IBM Software products, including software as a service solutions, (“Software Offerings”) may use cookies or other technologies to collect product usage information, to help improve the end user experience, to tailor interactions with the end user or for other purposes. In many cases no personally identifiable information is collected by the Software Offerings. Some of our Software Offerings can help enable you to collect personally identifiable information. If this Software Offering uses cookies to collect personally identifiable information, specific information about this offering’s use of cookies is set forth below.

This Software Offering does not use cookies or other technologies to collect personally identifiable information.

If the configurations deployed for this Software Offering provide you as customer the ability to collect personally identifiable information from end users via cookies and other technologies, you should seek your own legal advice about any laws applicable to such data collection, including any requirements for notice and consent.

For more information about the use of various technologies, including cookies, for these purposes, See IBM’s Privacy Policy at http://www.ibm.com/privacy and IBM’s Online Privacy Statement at http://www.ibm.com/privacy/details the section entitled “Cookies, Web Beacons and Other Technologies” and the “IBM Software Products and Software-as-a-Service Privacy Statement” at http://www.ibm.com/software/info/product-privacy.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be

trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at "Copyright and trademark information" at www.ibm.com/legal/copytrade.shtml.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

Other product and service names might be trademarks of IBM or other companies.

Glossary

This glossary includes terms and definitions for IBM SPSS Collaboration and Deployment Services.

The following cross-references are used in this glossary:

- See refers you from a term to a preferred synonym, or from an acronym or abbreviation to the defined full form.
- See also refers you to a related or contrasting term.

To view glossaries for other IBM products, go to www.ibm.com/software/globalization/terminology (opens in new window).

A

access control list (ACL)
In computer security, a list associated with an object that identifies all the subjects that can access the object and their access rights.

ACL
See access control list.

action
A permission for an aspect of system functionality. For example, the ability to set up notifications is defined as an action. Actions are grouped and assigned to users through roles. See also role.

Active Directory (AD)
A hierarchical directory service that enables centralized, secure management of an entire network, which is a central component of the Microsoft Windows platform.

AD
See Active Directory.

allowed user
A subset of the users defined in a remote directory, such as SiteMinder or Windows Active Directory, that are allowed access to SPSS Predictive Enterprise Services. Allowed users are defined when only a few users in a remote directory need access to the application.

API
See application programming interface.

appender
A component that receives logging requests from a logger and writes log statements to a specified file or console. See also logger.

application programming interface (API)
An interface that allows an application program that is written in a high-level language to use specific data or functions of the operating system or another program.

B

batch file
A file that contains instructions that are processed sequentially, as a unit.

binary large object (BLOB)
A data type whose value is a sequence of bytes that can range in size from 0 bytes to 2 gigabytes less 1 byte. This sequence does not have an associated code page and character set. BLOBs can contain, for example, image, audio, or video data.

BLOB
See binary large object.

break group
A set of rows of returned data that are grouped according to a common column value. For example, in a column of states, the rows of data for each state are grouped together.

burst report
A report that generates multiple output files during a single run by using multiple input parameters taken from break groups in the report.

C

cascading permission
A permission of a parent folder in the content repository that has been propagated to its child objects.

character large object (CLOB)
A data type whose value is a sequence of characters (single byte, multibyte, or both) that can range in size from 0 bytes to 2 gigabytes less 1 byte. In general, the CLOB data type is used whenever a character string might exceed the limits of the VARCHAR data type.

CLOB
See character large object.

common warehouse metamodel (CWM)
A metamodel written to be a common standard by the Object Management Group (OMG).

content repository
A centralized location for storing analytical assets, such as models and data. Content repository includes facilities for security and access control, content management, and process automation.

context data
Input data that is passed with a scoring request in real time. For example, when a score is requested for a customer based on credit rating and geocode, the credit score and geocode will be the context data for the request.

credential
Information acquired during authentication that describes a user, group associations, or other security-related identity attributes, and that is used to perform services such as authorization, auditing, or delegation. For example, a user ID and password are credentials that allow access to network and system resources.

CWM
See common warehouse metamodel.

D

data warehouse
A subject-oriented collection of data that is used to support strategic decision making. The warehouse is the central point of data integration for business intelligence. It is the source of data for data marts within an enterprise and delivers a common view of enterprise data.

distinguished name (DN)
The name that uniquely identifies an entry in a directory. A distinguished name is made up of attribute:value pairs, separated by commas. For example, CN=person name and C=country or region.

DN
See distinguished name.

Document Object Model (DOM)
A system in which a structured document, for example an XML file, is viewed as a tree of objects that can be programmatically accessed and updated. See also Simple API for XML.

document type definition (DTD)
The rules that specify the structure for a particular class of SGML or XML documents. The DTD defines the structure with elements, attributes, and notations, and it establishes constraints for how each element, attribute, and notation can be used within the particular class of documents.

DOM
See Document Object Model.

dormant schedule
A schedule associated with a deleted or unlabeled version of a job. A dormant schedule cannot be used until it is associated with a valid labeled job version.

DTD
See document type definition.

E

EAR
See enterprise archive.

enterprise archive (EAR)
A specialized type of JAR file, defined by the Java EE standard, used to deploy Java EE applications to Java EE application servers. An EAR file contains EJB components, a deployment descriptor, and web archive (WAR) files for individual web applications. See also Java archive, web archive.

execution server
A server that enables analytical processing of resources stored in the repository. For example, to execute an IBM SPSS Statistics syntax in an IBM SPSS Collaboration and Deployment Services job, an IBM SPSS Statistics execution server must be designated.

export
The process of storing objects and metadata from the content repository to an external file.

extended group
A locally-defined group of remote users. Extended groups are defined when groups in the remote directory are not fine-grained enough.

Extensible Markup Language (XML)
A standard metalanguage for defining markup languages that is based on Standard Generalized Markup Language (SGML).

Extensible Stylesheet Language (XSL)
A language for specifying style sheets for XML documents. Extensible Stylesheet Language Transformation (XSLT) is used with XSL to describe how an XML document is transformed into another document.

F

field content assist
A feature that provides predefined system and variable values for entry fields.

G

general job step
A method for running native operating system commands and executable programs on a host or a remote process server. General jobs have access to files stored within the repository and on the file system and can be used to control the input/output of analytical processing.

I

import
The process of adding objects and metadata defined in an external file generated by export, to the content repository.

iterative consumer reporting job step
A job step that is passed a set of input values generated by a preceding iterative producer reporting job step. The report in iterative consumer job step is executed for each tuple in the received data set.

iterative producer reporting job step
A job step that generates a set of values passed as input parameters to a following iterative consumer job step.

J

JAAS
See Java Authentication and Authorization Service.

JAR
See Java archive.

Java archive (JAR)
A compressed file format for storing all of the resources that are required to install and run a Java program in a single file. See also enterprise archive, web archive.

Java Authentication and Authorization Service (JAAS)
In Java EE technology, a standard API for performing security-based operations. Through JAAS, services can authenticate and authorize users while enabling the applications to remain independent from underlying technologies.

Java Generic Security Services (JGSS)
A specification that provides Java programs access to the services that include the signing and sealing of messages and a generic authentication mechanism.

Java Naming and Directory Interface (JNDI)
An extension to the Java platform that provides a standard interface for heterogeneous naming and directory services.

JGSS
See Java Generic Security Services.

JNDI
See Java Naming and Directory Interface.

job
A mechanism for automating analytical processing. A job consists of job steps, executed sequentially or conditionally. Input parameters can be defined for a job. A job can be run on demand or triggered by time-based or message-based schedules, with records of job execution stored as job history.

job step
A discrete unit of processing in a job. Depending on the type, job steps can be run on the content repository host or specially defined execution or remote process servers. Objects stored in the repository or the file system can provide input for a job step, and job step output can be stored in the repository or written to the file system.

K

KDC
See key distribution center.

Kerberos
A network authentication protocol that is based on symmetric key cryptography. Kerberos assigns a unique key, called a ticket, to each user who logs on to the network. The ticket is embedded in messages that are sent over the network. The receiver of a message uses the ticket to authenticate the sender.

key distribution center (KDC)
A network service that provides tickets and temporary session keys. The KDC maintains a database of principals (users and services) and their associated secret keys. It is composed of the authentication server and the ticket granting ticket server.

keystore
In security, a file or a hardware cryptographic card where identities and private keys are stored, for authentication and encryption purposes. Some keystores also contain trusted or public keys.

L

LDAP
See Lightweight Directory Access Protocol.

Lightweight Directory Access Protocol (LDAP)
An open protocol that uses TCP/IP to provide access to directories that support an X.500 model and that does not incur the resource requirements of the more complex X.500 Directory Access Protocol (DAP). For example, LDAP can be used to locate people, organizations, and other resources in an Internet or intranet directory.

lock
The process by which integrity of data is ensured by preventing more than one user from accessing or changing the same data or object at the same time.

logger
A component that prepares log statements to be written to console or log file. See also appender.

M

message-based schedule
A schedule that is used to trigger job execution by an event signalled by a Java Messaging Service (JMS) message. For example, when a job relies on the input from a third-party application, the application must send a JMS message when the input file is ready for processing.

metamodel
A model that defines the language for expressing a model.

meta-object
An instance of an XMI class as defined in the metamodel.

meta-object facility (MOF)
A generalized facility and repository for storing abstract information about concrete object systems; dealing mostly with construction, standardized by the Object Management Group (OMG).

MIME
See Multipurpose Internet Mail Extensions.

MOF
See meta-object facility.

Multipurpose Internet Mail Extensions (MIME)
An Internet standard that allows different forms of data, including video, audio, or binary data, to be attached to email without requiring translation into ASCII text.

N

notification
A mechanism that is used to generate email messages informing users of specific types of system events, such as changes to content repository objects and processing success and failure. Unlike subscriptions, notifications can be set up to send email to multiple users.

O

Object Management Group (OMG)
A non-profit consortium whose purpose is to promote object-oriented technology and the standardization of that technology. The Object Management Group was formed to help reduce the complexity, lower the costs, and hasten the introduction of new software applications.

ODS
See Output Delivery System.

OMG
See Object Management Group.

Output Delivery System (ODS)
A method of controlling the destination for output within SAS. ODS can route SAS output to a SAS data file, a text listing file, HTML files, and files optimized for high-resolution printing.

P

package
An installable unit of a software product. Software product packages are separately installable units that can operate independently from other packages of that software product.

principal
An entity that can communicate securely with another entity. A principal is identified by its associated security context, which defines its access rights.

R

remote process server
A remote system that is designated for running native operating system commands and executable programs.

repository content adapter
An optional software package that enables storing and processing content from other IBM SPSS applications, such as Statistics, Modeler, and Data Collection, as well as third parties.

repository database
A relational database that is used for storing content repository objects and metadata.

resource
A content repository object.

resource definition
A subset of content repository resources used to enable analytical processing, such as definitions of data sources, credentials, execution servers, and JMS message domains.

role
A set of permissions or access rights. See also action.

S

SAX
See Simple API for XML.

schedule
A content repository object that triggers job execution.

scoring configuration
A configuration that defines model-specific settings for generating real-time scores, such as input data, processing rules, outputs, logging, etc.

security provider
A system that performs user authentication. Users and groups can be defined locally (in which case, IBM SPSS Collaboration and Deployment Services itself is the security provider) or derived from a remote directory, such as Windows Active Directory or OpenLDAP.

service provider interface (SPI)
An API that supports replaceable components and can be implemented or extended by a third party.

SGML
See Standard Generalized Markup Language.

shell script
A program, or script, that is interpreted by the shell of an operating system.

Simple API for XML (SAX)
An event-driven, serial-access protocol for accessing XML documents, used. A Java-only API, SAX is used by most servlets and network programs to transmit and receive XML documents. See also Document Object Model.

single sign-on (SSO)
An authentication process in which a user can access more than one system or application by entering a single user ID and password.

SOAP
A lightweight, XML-based protocol for exchanging information in a decentralized, distributed environment. SOAP can be used to query and return information and invoke services across the Internet.

SPI
See service provider interface.

SSO
See single sign-on.

Standard Generalized Markup Language (SGML)
A standard metalanguage for defining markup languages that is based on the ISO 8879 standard. SGML focuses on structuring information rather than presenting information; it separates the structure and content from the presentation. It also facilitates the interchange of documents across an electronic medium.

stop word
A word that is commonly used, such as "the," "an," or "and," that is ignored by a search application.

subscription
Email notices and Really Simple Syndication (RSS) feeds that repository users create to receive when the state of an asset changes.

T

TGT
See ticket-granting ticket.

ticket-granting ticket (TGT)
A ticket that allows access to the ticket granting service on the key distribution center (KDC). Ticket granting tickets are passed to the principal by the KDC after the principal has completed a successful request. In a Windows 2000 environment, a user logs on to the network and the KDC will verify the principal's name and encrypted password and then send a ticket granting ticket to the user.

time-based schedule
A schedule that triggers job execution at a specified time or date. For example, a time-based schedule may run a job at 5:00 pm every Thursday.

U

Universally Unique Identifier (UUID)
The 128-bit numeric identifier that is used to ensure that two components do not have the same identifier.

UUID
See Universally Unique Identifier.

V

Velocity
A Java-based template engine that provides a simple and powerful template language to reference objects defined in Java code. Velocity is an open source package directed by the Apache Project.

W

W3C
See World Wide Web Consortium.

WAR
See web archive.

web archive (WAR)
A compressed file format, defined by the Java EE standard, for storing all the resources required to install and run a web application in a single file. See also enterprise archive, Java archive.

Web Services Description Language (WSDL)
An XML-based specification for describing networked services as a set of endpoints operating on messages containing either document-oriented or procedure-oriented information.

World Wide Web Consortium (W3C)
An international industry consortium set up to develop common protocols to promote evolution and interoperability of the World Wide Web.

WSDL
See Web Services Description Language.

X

XMI
See XML Metadata Interchange.

XML
See Extensible Markup Language.

XML Metadata Interchange (XMI)
A model-driven XML integration framework for defining, interchanging, manipulating, and integrating XML data and objects. XMI-based standards are in use for integrating tools, repositories, applications, and data warehouses.

XSL
See Extensible Stylesheet Language.

Index

A

A/B split testing
  for scoring 108, 109
accessibility
  blind users 209
  keyboard navigation 207
  visually impaired users 209
accessing
  expired files 30
actions
  promotion 67
active log file 190
adding
  administered servers 183
  job variables 117
  principals for label 32
  promotion policies 53
adding steps 118
administered servers
  adding 183
  deleting 185
  logging in 185
  logging out 185
  properties 185
  server information 184
  types 183
administration 185
administration application
  connections 186
administrator access
  for IBM SPSS Modeler Server 195
  with User Access Control (UAC) 196
administrator group 188
advanced mode
  for analytic data views 81
advanced search 17–21
aliases
  creating 102
  deleting 103
  editing 103
  for scoring configurations 101
  in A/B split testing 108, 109
  viewing 110
allow_modelling_memory_override
  options.cfg file 199
analytic data view
  champion challenger 178
analytic data views
  adding attributes 82
  adding tables 81
  advanced mode 81
  business object models 88, 89
  creating 70
  data access plans 71, 72, 76, 77, 79
  data mapping 74
  data models 73, 79, 86, 87
analytic data views *(continued)*
  deleting attributes 87
  deleting tables 87
  derived attributes 84
  execution object models 90, 91
  table relationships 83
application
  exiting 11
  navigating 9
  starting 9
application server data sources
  for data access plans 77
applying
  locks 21, 22
association sets
  creating 107
  deleting 108
  editing 107
  for scoring configurations 104
  in A/B split testing 108, 109
  viewing 110
ATOM feeds 34
attachment 143, 146
auditing
  for scoring 99

B

batch
  data access plans 72
batch files 157
batch scoring 101
blank cells
  in job history 133
BOM, *See* business object models
broadcasting a message to users 194
bulk property updates
  general 43
  permission 44
  version 43
business object models
  exporting 81, 89
  importing 81, 89

C

cache compression 188, 192, 199
cache sizes
  for scoring 101
cache_compression
  options.cfg file 199
cache_connection option 204
caching, in-database 204
canceling
  jobs 133
cardinality
  data model relationships 79

cells
blank 133
challenger models
data sources 176
champion challenger 134, 135, 173–176, 178, 179
champion models 173, 176
changing
passwords 15
clean up 151
client
starting 9
client data access 188
clusters 58
Cognos
champion challenger 179
collaboration 1
columns 50, 51
comparing
models 173
compiled transformations 188
computation time
for scoring configurations 99
computation wait time
for scoring configurations 99
conditional connector 120
configuration file
location 187
configuration options
automatic SQL generation 200
connections 186
connections and sessions 197
coordinator of processes 201
COP 189, 201
data file access 187, 198
logging 190
login attempts 197
memory management 199
overview 186, 197
parallel processing 199
performance and optimization 199
port number 197
SQL string length 200
SSL data encryption 200
stream rewriting 199
temp directory 187, 198
users 188
configurations
for scoring models 96
conflict resolution
duplicates 65, 66
individual 65, 66
invalid versions 65, 66
conflicts
duplicates 65, 66
import 64–66
invalid versions 65, 66
connection
existing 15
new 14
server 13–15
terminating 15
connection timeout 186
connections
server 31, 34
connectors
conditional 120
deleting 121
fail 120
pass 120
relationship 119–121
sequential 120
constant filters
for data access plans 77
content assist 10
content explorer
deleting files 17
files 16, 17
jobs 115, 116
keyboard navigation 207
organization 13
overview 13
permissions 16
properties 23
search 17–21
searching 20
server connections 13–15
versions 28, 29, 35
content object
properties 23
content objects
definition 13
permissions 16
storage 13, 16, 17
versions 27, 29
content repository
deleting files 17
Content Repository
execution server 56
server definition 56
context data sources
for data access plans 77
conventions
naming 10
Coordinator of Processes 189
coordinator of processes configuration
for IBM SPSS Modeler Server 201
COP configuration
for IBM SPSS Modeler Server 201
cop_enabled
options.cfg file 201
cop_host
options.cfg file 201
cop_password
options.cfg file 201
cop_port_number
options.cfg file 201
cop_service_description
options.cfg file 201
cop_service_host
options.cfg file 201
cop_service_name
options.cfg file 201
cop_service_weight
options.cfg file 201
cop_update_interval
options.cfg file 201
cop_user_name
options.cfg file 201

copying
topics 41
CPU priority 192
creating
custom properties 36
schedules 126
scoring configuration aliases 102
scoring configuration associations 107
server clusters 58
topics 40, 41
credential definitions 45, 46
credential destination 45
credentials
for schedules 126
job 114
promoting 67
custom properties
accessing 36
creating 25, 36, 38
defining 25
deleting 39
editing 25, 38
existing 25, 38
label 36
modifying 25, 38
new 36
overview 36
property type 36
property types 25
searching 38
selection values 25, 36, 38
setting values 36
customizing
notification messages 146

D

daily schedules 126
data
filtering model management views 139
data access 188
data access plans
creating 72
deleting 79
for analytic data views 72, 76, 77, 79
for scoring 97
overriding data sources 76, 77
previewing data 79
real-time 77
data access time
for scoring configurations 99
Data Editor
hiding data in the client 192
data files
champion challenger 178
data initialization time
for scoring configurations 99
data models
adding attributes 82
adding derived attributes 84
adding tables 73, 81
data mapping 74
defining relationships 83
deleting attributes 87
data models *(continued)*
deleting tables 87
for analytic data views 79
modifying properties 86
relationships 79
stream fields 75
table attributes 75
data service data sources
for data access plans 77
keys 51
tables 51
data source definitions 45
data sources
application server data sources 46, 50
data service data sources 46, 50, 51
JDBC data sources 46, 47, 49
modifying 51
ODBC data sources 46, 47
promoting 67
data types 50, 51
data_file_path
options.cfg file 198
data_files_restricted
options.cfg file 198
database access
limiting 188
database caching
controlling from options.cfg 204
database connections
closing 204
databases 114
date range 18, 19
dates
expiration 29–31
defining tables 50
delayed promotion 53, 55
delete search terms 20
deleting
administered servers 185
custom properties 39
files 17
groups 27
job variables 118
principals for label 32
schedules 129
scoring configuration aliases 103
scoring configuration associations 108
scoring configurations 111
search terms 20
topics 41
users 27
versions 29
delivery failure 148
deployment 2
description 35
destination
credentials 45
server 56
destination name 52, 53
DEVICE 154, 155
disconnecting users 194
distribution channels 34
drag and drop 9
DSNs

DSNs *(continued)*
restricting access 188
duplicates 65, 66
durable subscriptions 127

E

edit search terms 20
editing
custom properties 25, 38
job variables 117
permissions for label 32
properties 24
schedules 128
scoring configuration aliases 103
scoring configuration associations 107
scoring configurations 111
server clusters 59
topics 25
email delivery failure 148
enabling
filtering 137, 138
encryption
FIPS 202
end date
for schedules 126
enter key 9
error messages
for schedules 132
error on stream execution 199
evaluation
results 134–136
evaluation type
filtering model management views 139
exclusions
search terms 17, 21
executables 157
executing jobs 119
execution object models
exporting 91
execution servers
remote process 2, 4
remote process servers 153, 157
SAS 2, 4, 153
existing jobs 116
exiting
application 11
client 11
system 11
expiration
submitted jobs 181
expiration date 35
expiration dates
exporting 30
importing 30
modifying 30
overview 29
reactivating 31
restrictions 30
searching 30
setting 30
viewing 30
export
external references 61
export *(continued)*
folders 61, 62, 64–66
jobs 61
exporting
business object models 81, 89
conflicts 64, 65
execution object models 91
expiration dates 30
resource definitions 59
restrictions 62
external files
in SAS steps 155
external references
exporting 61
importing 61

F

F1 help 9
fail connector 120
field content assist 10
file
filtering model management views 139
permissions 123
subscriptions 146
versions 27–29, 35
FILENAME 155
fileref 155
files
.pes 61, 62
adding 16
adding to repository 61
copying 16
deleting 17
downloading 17, 61
expired 30
expiring 29–31
locking 21, 22
moving 16
opening 16
permissions 16
reactivating 31
unlocking 22
filtering
enabling 137, 138
job history 138
job schedule 137
scoring view 110
filters
data 139
evaluation type 139
file 139
index 139
job 137
label 137
model execution date 139
trend 139
FIPS encryption 202
firewall settings
options.cfg file 199
folder
resource definitions 45, 46, 55–57
search 17–21
searching 20

folder options 143
folders
  child 26
  content repository 181
  expiration of 181
  exporting 61, 62, 64–66
  importing 61–66
  parent 26
  resource definitions 13
  restrictions 181
  submitted jobs 181

G

General job steps
  adding to jobs 157
  defining properties 157
  examples 159, 160
  input files 158
  naming 157
  output files 158, 159
  working directory 157–159
general properties 24
global conflict resolution 64
glossary 215
group authorization
  URL 186
group search terms 19
group_configuration 202
groups
  creating 191
  deleting 27
  editing 191
  existing 26
  new 26
  overview 191
  permissions 123
  settings 192
GSFNAME 155
guidelines
  naming 10

H

help
  accessing 9
  F1 9
  keyboard navigation 208
history
  job 133, 138
  job step 134
  server status 140
host 186
hourly schedules 126
HTML output
  for IBM SPSS Statistics steps 155

I

IBM Operational Decision Manager
  business object models 88, 89
  execution object models 90, 91
IBM SPSS Analytic Server
IBM SPSS Analytic Server *(continued)*
  configuration options 197
IBM SPSS Collaboration and Deployment Services Deployment Manager 2, 3
IBM SPSS Collaboration and Deployment Services Deployment Portal 2, 4
IBM SPSS Collaboration and Deployment Services Repository 2, 3
IBM SPSS Modeler Administration Console
  administrator access 195
  User Access Control access 196
IBM SPSS Modeler Server
  administration of 195
  administrator access 195
  configuration options 197
  coordinator of processes configuration 201
  COP configuration 201
  monitoring usage 203
  port number 197
  server processes 203
  temp directory 198
  User Access Control access 196
IBM SPSS Statistics Administration Console
  connections 186
immediate promotion 53, 55
import
  external references 61
  folders 61–66
  jobs 61
  security permissions 62
import order 62
importing
  business object models 81, 89
  conflicts 64, 65
  expiration dates 30
  resource definitions 59
  restrictions 62
in-database caching 204
index
  filtering model management views 139
individual conflict resolution 65, 66
inheritance
  permissions 26
input files
  General job steps 158
invalid versions 65, 66
io_buffer_size
  options.cfg file 199
iterative principals 123

J

Java memory 188, 192
JDBC data sources
  for data access plans 77
  third-party drivers 49
JMS
  mapping variables 128
JMS message domain 129
job
  notifications 141
  subscriptions 146
Job Editor
  keyboard navigation 208

job history
filtering 138
keyboard navigation 208
limiting 138
reordering 138
viewing 133, 138
job labels
for schedules 126
job output 134
job schedule
filtering 137
keyboard navigation 208
limiting 137
reordering 137
viewing 137
job step
notifications 142
permissions 123
job step history
viewing 133, 134
job step name
champion challenger 174
job step output 122
job steps
history 134
results 134
job variables
in schedules 128
logging values 134
jobs
adding steps 116, 118
canceling 133
components 113
content explorer 115, 116
creating 115
definition 113
dependencies 114
editing 116, 118
existing 116
exporting 61
filtering 137, 138
history 133, 134, 138
importing 61
job editor 116, 118–121
logs 134
new 115
notifications 125
opening 115, 116
overview 113
prerequisites 114
process 114
promoting 67
properties 116
relationships 119–121
running 119, 125–127, 132
saving 122
schedule 137
scheduling 125–127, 132
searching 137, 138
status 133, 134
submitted 181
variables 117, 125
JVM memory 188, 192

K

Kerberos 202
keyboard navigation
in content explorer 207
in help system 208
in job editor 208
in job history view 208
in tables 208
keywords 35

L

label
notifications 143
labels
applying 28
filtering 137, 138
for versions 27–29
LATEST 27
permissions 32
recommendations 28
removing 28
security 32
server version 34
LATEST label 27
locations
configuration file 187
profile file 187
temporary files directory 187
locking objects 21, 22
locks
applying 21, 22
files 21, 22
objects 21, 22
removing 22
resource definitions 21, 22
log in
server 15
log off
server 15
logging
for scoring 99
logs
configuration options 190
job 134
output 134
viewing 190
viewing performance information 190
Lotus 144

M

manage label permission
for labels 32
max_file_size
options.cfg file 198
max_login_attempts
options.cfg file 197
max_parallelism
options.cfg file 199
max_sessions
options.cfg file 197

max_sql_string_length
options.cfg file 200
maximum JVM memory 188, 192
maximum threads 188, 192
memory 188, 192
memory management
administration options 199
memory_usage
options.cfg file 199
message
broadcast to users 194
message customization
body text 146
errors 146
event property variables 146
HTML formatting 146
message customization 146
preview 146
subject 146
template 146
message domain 52, 53, 167
message domain attributes 52, 53
message domain name 52
message domain properties 52
message domains
for schedules 127
promoting 67
message filters
for schedules 127
message selector 167
message template 170
message text 167
message-based job 52
message-based processing example 129
message-based step 52
message-driven step
defining properties 167
metadata
user-defined 25, 36, 38–42
version properties 35
metrics
for scoring performance 99
MIME type filtering 54
MIME types
promotion policies 54
model evaluation
notifications 143, 169
return codes 143, 169
model execution date
filtering model management views 139
model management 134, 135
modeling
memory management 199
modelling_memory_limit_percentage
options.cfg file 199
models
champion challenger 173
comparing 173
evaluating 173
IBM SPSS Modeler 173
monitoring 173
modifying
permissions 26
server clusters 59
monitoring users 193
monthly schedules 126
mouse 9
multiple stream execution 199

N

naming conventions 10
naming factory 52, 53
naming service 52, 53
narrow search 18, 19
narrowing 18, 19
navigation
enter key 9
mouse 9
Netezza 49
new
custom properties 36
topics 40, 41
new jobs 115
NOSPLASH 154
NOSTATUSWIN 154
notification job step
adding to a job 169
body 170
connecting to other job steps 169
customizing 170
email attachments 169
from address 170
general information 170
iterative consumer 170
message template 170
Notification tab 170
outcome-based 169
recipients 170
subject 170
use in model evaluation 169
notifications
action 141
attachment 143
body text 146
content 141
content notifications 142
defining 142
disabling 125
entering recipient email 144
errors 146
event property variables 146
folder events 142
folder notifications 142
folder options 143
folder structure changes 142
for exported and imported objects 141
HTML formatting 146
job 141
job step 142
label 141, 143
message customization 143, 146
message subject 143
model evaluation 143
modifying 142
preview 143, 146
previewing message 146
recipients 143

notifications *(continued)*
removing 142, 143
reverting to the default template 143
selecting recipients from Lotus Notes 144
selecting recipients from security subscribers 145
sender 143
subject 146
template 143, 146
notifications delivery failure 148

O

object permissions
cascading 26
deleting 27
modifying 26
object properties
editing 24
viewing 23, 24, 31, 34
objects
content 13, 16, 17
deleting 17
files 16, 17
locking 21, 22
permissions 16
search 17–21
unlocking 22
versions 35
ODBC 114
ODBC data sources
champion challenger 178
ODS 155
options.cfg 202
options.cfg file 204
output
log files 134
output file permissions 123
output files
General job steps 158, 159
output permissions 123
overriding data sources
for data access plans 76, 77
overview
content explorer 13
jobs 113, 114

P

parallel processing
controlling 199
parameter filters
for data access plans 77
pass connector 120
passwords
changing 15
modifying 15
new 15
pause 186
PDF settings
for the visually impaired 209
performance log
interval 190
viewing 190
permission inheritance 26
permissions
applying 26
assigning 26
cascading 26
creating 26
deleting 27
editing for label 32
for labels 32
for SAS results 154
for visualization results 150
inheritance 16
job step 123
modifying 26, 27
object 26, 27
output file 123
removing 27
specifying as iterative variables 123
updating in bulk 44
when importing 62
PMML 93
PMML files 159, 160
policy 53–55, 67
port number
IBM SPSS Modeler Server 197
port settings
options.cfg file 199
port_number
options.cfg file 197
predictor effectiveness 136
preferences
distribution channels 34
prerequisites
jobs 114
preview 143
previewing notification message 146
principals
adding for label 32
deleting for label 32
modifying label permissions 32
privileges
access 36, 40, 41
administrative 36, 40, 41
processors
multiple 199
profile file 187, 191
profiles
creating 191
editing 191
overview 191
settings 192
program_file_path
options.cfg file 198
program_files_restricted
options.cfg file 198
promotion
actions for roles 67
delayed 53, 55, 67
dependent object 54
immediate 53, 55
MIME types 54
policy 67
promotion policies
adding 53

promotion policies *(continued)*
deleting 55
resource definitions 54
timing 53
promotion policy
modifying 55
properties
bulk updates 42–44
content object 23
custom 25, 36, 38, 39, 41
definition 23
deleting 20
editing 20, 24
general 24
object 23, 24
overview 23
server 25, 31, 34, 36, 38–42
topics 40–42
viewing 23, 24, 31, 34
property 18, 19
property types
custom properties 25
property variables 10

R

range of recurrence
for schedules 126
reactivating
expired files 31
real-time
data access plans 72, 77
recipient
of notification 143
recipients 143, 146
recurrence pattern
for schedules 126
reference filters
for data access plans 77
references
external 61
refine search 18–20
relationships
between steps 119–121
deleting 121
in data models 79
remote process
execution servers 2, 4
remote process servers
assigning to job steps 153, 157
server definition 57
remotely-deployed scoring servers 5
removing
custom properties 39
locks 22
topics 41
renaming
topics 41
reorder search terms 20
reordering
search terms 20
reporting steps 149
reports
external 181
reports *(continued)*
submitted 181
repository
connecting to 13–15
content 13, 16, 17
deleting 17
disconnecting from 15
files 16, 17
locking objects 21, 22
objects 21, 22
search 17–21
unlocking objects 22
resolving conflicts
globally
individually 64
individually 65, 66
resource definitions
credential definitions 45, 46
data source definitions 45
data sources 46
exporting 59
importing 59
in promotion policies 54
locking 21, 22
server definitions 45, 55–57
resource definitions folder 13
restart 186
restarting the web service 195
restrictions
expiration dates 30
exporting 62
importing 62
submitted jobs 181
results
analysis 134–136
champion challenger 134, 135
job 133, 134
job step 134
model evaluation 134
predictor effectiveness 136
search 20
streams 134, 135
resuming
scoring configurations 111
return codes
model evaluation 143
notifications 143
reverting to the default template 143
rows
filtering 137, 138
RSS feeds 34
running jobs 119, 125, 132

S

SAS
execution server 2, 4, 57
server definition 57
SAS execution servers
assigning to job steps 153
SAS steps
adding to jobs 153
controlling processing 153–155
creating graphs 154, 155

SAS steps *(continued)*
defining properties 153, 154
example 155
external files 155
HTML output 155
naming 153
naming results 154
result locations 154
result permissions 154
text output 155
versions 153
saving jobs 122
schedule
job 137
schedules
creating 126
credentials 126
deleting 129
dormant 129
editing 128
job labels 126
job variables 128
message-based 127
time-based 126
scheduling jobs 125, 132
scoring
A/B split testing 108, 109
configurations 96
models 93
performance 111
using PMML files 160
scoring configurations
aliases 101
association sets 104, 107, 108
auditing 99
batch scoring 101
cache sizes 101
configuration names 97
creating 96
data 97
data providers 97
deleting 111
editing 111
input 98
labels 97
logging 99
models 97
output 98
resuming 111
suspending 111
viewing 110
scoring graph view 111
scoring servers 5
scoring view
deleting configurations 111
editing configurations 111
filtering 110
resuming configurations 111
suspending configurations 111
screen readers 209
scripting
champion challenger 173
IBM SPSS Modeler 173
search
search *(continued)*
accessing 17, 18
advanced 17–21
AND 18, 19
content objects 17–21
date range 18, 19
deleting 20
dialog 17, 18
expiration dates 30
files 17–21
group 19
OR 18, 19
property 18, 19
refining 18–20
reordering 20
simple 17–19, 21
time 18, 19
ungroup 19
search terms
deleting 20
editing 20
group 19
properties 20
reordering 20
stopwords 17, 21
ungroup 19
searching
content objects 20
custom properties 38
files 20
for jobs 137, 138
submitted jobs 181
topics 42
Secure Sockets Layer (SSL) 186
security
for labels 32
security providers 123
security subscribers 145
selecting
versions 29
selection
drag and drop 9
selection values
deleting 25, 38
editing 25, 38
modifying 25, 38
removing 25, 38
sender
of notification 143
sequential connector 120
server
change password 15
connection 31, 34
connections 31
custom properties 25, 36, 38, 39
log in 13–15
log off 13, 15
properties 31, 34
topics 25, 40, 41
URL 31
version label 34
server clusters
adding servers 59
creating 58

server clusters *(continued)*
modifying 59
name 59
promoting 67
removing servers 59
settings 59
weights 59
server connection
changing password 15
ending 15
existing 15
new 14
terminating 15
server connections
jobs 114
server definitions
adding 55
content repository 56
location 56
modifying 57
names 56
remote process 57
SAS 57
types 56
server destination 56
server port settings
options.cfg file 199
server processes 193
server software
pause 186
restart 186
shutdown 186
server status 140
servers
promoting 67
remote process servers 153, 157
SAS 153
shortcuts
keyboard 207
shutdown 186
simple search 17–19, 21
sorting 188
split testing
for scoring 108, 109
SQL generation
enabling for IBM SPSS Modeler Server 200
sql_generation_enabled
options.cfg file 200
SSL 31, 186
SSL data encryption
enabling for IBM SPSS Modeler Server 200
ssl_certificate_file
options.cfg file 200
ssl_enabled
options.cfg file 200
ssl_private_key_file
options.cfg file 200
ssl_private_key_password
options.cfg file 200
standard error 157
standard output 157
start date
for schedules 126
start time
start time *(continued)*
for schedules 126
starting
application 9
client 9
system 9
status
filtering 137, 138
job 133
job step 134
server 140
stderr 157
stdout 157
steps
adding 118
executing 119
job 113–116, 118, 119
relationships 119–121
stream_rewriting_enabled
options.cfg file 199
streams
results 134, 135
status 134, 135
subject 143
submitted jobs
expiration dates 181
restrictions 181
searching 181
subscribing to files 147
subscription delivery failure 148
subscriptions
email address for 147
files 146
for exported and imported objects 141
jobs 146
managing 147
modifying 147
removing 147
subscribing to files 147
unsubscribing 147
user preferences 147
suspending
scoring configurations 111
syndication feeds 34
system
content explorer 13
exiting 11
help 9
jobs 113
launching 9
naming files 10
navigating 9
starting 9, 13

T

table
search results 20
table properties 50, 51
tables
champion challenger 134, 135
filtering 137, 138
job history 133, 138
job schedule 137

tables *(continued)*
job step history 133, 134
keyboard navigation 208
model evaluation 134
predictor effectiveness 136
server status 140
tables keys 51
temp directory
for IBM SPSS Modeler Server 198
temp_directory
options.cfg file 198
template 143
temporary files directory 192
Teradata 49
terms
deleting 20
editing 20
group 19
reordering 20
search 18, 19
search exclusions 17, 21
ungroup 19
text output
for SAS steps 155
third-party sorting 188
time 18, 19
timeout
connection 186
topics
accessing 40
copying 41
creating 25, 40, 41
deleting 41
editing 25
existing 25
label 41
modifying 25
new 40
property type 41
renaming 41
searching 42
selection values 25, 41
transformations
compiling 188
trend
filtering model management views 139

U

umask 188
umask setting 192
UNC file references 114
ungroup search terms 19
UNIX
restarting the web service 195
UNIX shell 202
unlocking 22
unlocking objects 22
use versions permission
for labels 32
user groups
creating 191
editing 191
overview 191
user groups *(continued)*
settings 192
user preferences
distribution channels 34
email 147
user profiles
creating 191
editing 191
overview 191
settings 192
users
broadcasting a message to 194
deleting 27
disconnecting 194
existing 26
monitoring 193
new 26
permissions 123

V

variables
adding to jobs 117
editing 117
in jobs 117, 125
removing from jobs 118
version
description 35
expiration date 35
keywords 35
metadata 35
properties 35
values 35
version label
filtering 137, 138
version properties
updating in bulk 43
versions
deleting 29
invalid 65, 66
labeling 34
labels 27–29
promoting 67
selecting 29
view data 192
viewing
expired files 30
job history 133, 138
job output 134
job schedule 137
job step history 133, 134
logs 134
properties 23, 24, 31, 34
search results 20
server properties 185
server status 140
views
champion challenger 134, 135
model evaluation 134
model management 134, 135
predictor effectiveness 136
server status 140
visualization reports
adding to jobs 149

visualization reports *(continued)*
clean-up 151
data source 149
defining general properties 149
dimensions 150
metadata 150
naming 149
notifications 151
output file location 150
parameters 150
prompts 150
result permissions 150
results 150
single 150
type 150
variables 150
versions 149
visualization reports steps
output file format 150

W

web service - restarting 195
weely schedules 126
weights
in server clusters 59
Windows
restarting the web service 195
working directory
for General job steps 157–159

X

XOM, *See* execution object models