

SAS® Visual Analytics 5.2

User's Guide

The correct bibliographic citation for this manual is as follows: SAS Institute Inc. 2012. *SAS® Visual Analytics 5.2: User's Guide*. Cary, NC: SAS Institute Inc.

SAS® Visual Analytics 5.2: User's Guide

Copyright © 2012, SAS Institute Inc., Cary, NC, USA

All rights reserved. Produced in the United States of America.

For a hard-copy book: No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher, SAS Institute Inc.

For a Web download or e-book: Your use of this publication shall be governed by the terms established by the vendor at the time you acquire this publication.

The scanning, uploading, and distribution of this book via the Internet or any other means without the permission of the publisher is illegal and punishable by law. Please purchase only authorized electronic editions and do not participate in or encourage electronic piracy of copyrighted materials. Your support of others' rights is appreciated.

U.S. Government Restricted Rights Notice: Use, duplication, or disclosure of this software and related documentation by the U.S. government is subject to the Agreement with SAS Institute and the restrictions set forth in FAR 52.227-19, Commercial Computer Software-Restricted Rights (June 1987).

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513.

June 2012

SAS® Publishing provides a complete selection of books and electronic products to help customers use SAS software to its fullest potential. For more information about our e-books, e-learning products, CDs, and hard-copy books, visit the SAS Publishing Web site at support.sas.com/publishing or call 1-800-727-3228.

SAS® and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.

Contents

<i>Using This Book</i>	<i>xi</i>
<i>What's New in SAS Visual Analytics</i>	<i>xiii</i>
<i>What's New in SAS Visual Analytics</i>	<i>xvii</i>

PART 1 Introduction to SAS Visual Analytics 1

<i>Chapter 1 / About SAS Visual Analytics</i>	3
What Is SAS Visual Analytics?	3
Benefits of Using SAS Visual Analytics	4
How Does SAS Visual Analytics Work?	4
<i>Chapter 2 / Manage Access to SAS Visual Analytics</i>	7
Access to SAS Visual Analytics Capabilities	7
Log On to SAS Visual Analytics	8
Log Off from SAS Visual Analytics	9
<i>Chapter 3 / SAS Visual Analytics Interface</i>	11
Your First Look at the SAS Visual Analytics Home Page	11
About the Availability of Menus and Menu	
Selections in SAS Visual Analytics	13
Managing Your Content on the Home Page	14
Specifying Your Preferences	15
Discovering Details about Objects on the Home Page	20
Adding Comments to Reports and Explorations on the Home Page	22
Searching from the Home Page	25

PART 2 Preparing Data 27

Chapter 4 / Overview of the Data Preparation Interface	29
About the SAS Visual Analytics Data Preparation Interface	30
Overview of the Data Preparation Interface	31
Specifying Your Preferences for the Data Preparation Interface	32
About Access to DBMS Data	33
About Exporting Code	34
About Jobs	35
About Table Metadata	36
Working with User-Defined Formats	37
Chapter 5 / Creating Prepared Data	39
Features in Data Preparation	39
Concepts in Data Preparation	40
Prepare Data	41
Joining Tables	45
Adding Calculated Columns	46
About Previewing Data	47
Distributing Data	47
Chapter 6 / Loading Data	49
Concepts in Loading Data	49
Load from HDFS	50
Load from Table	51
Unload a Table	53
Chapter 7 / Server Management	55
Features in Server Management	55
SAS LASR Analytic Server Concepts	56
Best Practices for Starting Servers and Loading Tables	57
What Does the In-Memory Data Tab Show?	58
Managing Server Description Files	60
Start a Server	61
Stop a Server	63

PART 3 Exploring Data 65

Chapter 8 / Overview of SAS Visual Analytics Explorer	67
What Is the SAS Visual Analytics Explorer Interface?	67
Accessing the SAS Visual Analytics Explorer Interface	68
The Welcome Window	68
Overview of SAS Visual Analytics Explorer Interface	69
Chapter 9 / Opening a Data Source	73
Overview of Opening a Data Source	73
Select a Data Source	73
Select Columns	74
Reduce Your Data	74
Chapter 10 / Managing Data	79
Refresh Your Data Source	79
Change the Data Source for Your Visual Exploration	80
Managing Data Properties	80
Define a Geography Data Item	84
Chapter 11 / Working with Visualizations	87
Overview of Visualizations	89
Working with Visualizations	95
Managing Visualization Data Roles	99
Working with Filters	101
Working with Automatic Charts	102
Working with Bar Charts	103
Working with Line Charts	105
Working with Tables	108
Working with Crosstabs	109
Working with Scatter Plots	110
Working with Bubble Plots	112
Working with Histograms	115
Working with Box Plots	117
Working with Heat Maps	120

Working with Geo Maps	122
Working with Correlation Matrices	124
Chapter 12 / Working with Visualization Filters	127
About Filters	127
Create a Filter	128
Edit a Filter By Using Visual Filter Controls	129
Edit a Filter By Using the Edit Filter Window	131
Set the Scope of a Filter (Global or Local)	133
Remove a Filter	134
Reset a Filter	134
Chapter 13 / Exporting Data	135
Overview of Exporting Data	135
Export a Visual Exploration as a Report	136
Export a Visual Exploration as a PDF	136
Save a Visualization as an Image File	137
E-mail a Visual Exploration as a Link	138
Chapter 14 / Managing Hierarchies	139
Overview of Hierarchies in the SAS Visual Analytics Explorer Interface	139
Create a New Hierarchy	140
Edit a Hierarchy	141
Delete a Hierarchy	141
Chapter 15 / Managing Visual Explorations	143
What Is a Visual Exploration?	143
Create a New Visual Exploration	143
Saving Your Visual Exploration	144
Deleting Visual Explorations	144
Chapter 16 / Performing Data Analysis	145
Overview of Analyzing Your Data	145
Add Analysis to an Existing Visualization	147
Correlate Multiple Measures as a New Visualization	148

Chapter 17 / Specifying Your Preferences	149
Specifying Global Preferences	149
Specify Preferences That Are Specific to the Explorer Interface	149

PART 4 Designing Reports 151

Chapter 18 / Overview of the Designer Interface	153
About the SAS Visual Analytics Designer Interface	153
Your First Look at the Designer Interface	154
Open Window	157
Search Window	159
Chapter 19 / Opening Reports	161
Opening Reports Created in the Designer Interface	161
An Example Report	162
Chapter 20 / Creating and Saving Reports	163
Creating a New Report	163
Basing a New Report on Report Objects Imported from One or More Existing Reports	164
Saving a Report	168
Chapter 21 / Selecting Data for Reports	169
Overview of Queries Based on Data Items	169
Adding a Data Source or Data Sources	170
Changing a Data Source	172
Removing a Data Source	173
Using Data Items in a Query	174
Modifying Data Role Assignments in Report Objects	182
Removing Data Assignments from Report Objects	183
Chapter 22 / Displaying Results in a Table, Graph, or Gauge	185
Using Tables, Graphs, and Gauges	186

Using Tables to Display Results	187
Using Graphs to Display Results	191
Using Gauges to Display Results	204
Adding Text and Images	212
Adding Comments to Report Objects	216
Chapter 23 / Sorting Data Items in Reports	219
How Sorting Can Help with Analysis	219
Sorting Data in a List Table	219
Sorting Data in a Crosstab	220
Sorting Data in a Graph	221
Chapter 24 / Working with Interactions	223
Overview of Interactions	223
Creating Interactions	224
Deleting Interactions	229
Chapter 25 / Maintaining Multi-Section Reports	233
Overview of Report Sections	233
Adding a Section to a Report	233
Renaming a Report Section	234
Deleting a Section from a Report	235
Moving or Dragging a Report Object to Another Section	235
Chapter 26 / Exporting Data from Report Objects	237
Overview of Exporting from the Designer Interface	237
Exporting Data from Report Objects	238
PART 5 Viewing Reports 247	
Chapter 27 / Viewing Reports on a Mobile Device	249
Opening the App for the First Time	249
Connecting to Your Library Server for the First Time	250

Chapter 28 / Viewing Reports in the Web Viewer	251
Overview of Viewing Reports in the Web Viewer	251
Opening a Report in the Web Viewer	251
Adding Comments to a Report in the Web Viewer	253

PART 6 Managing the Environment 255

Chapter 29 / Working with HDFS	257
What Is the HDFS Content Explorer?	257
What Can I Do with the HDFS Content Explorer?	258
Navigating HDFS Folders	259
Adding a Table to HDFS	259
Deleting a Table from HDFS	260
Viewing HDFS System Properties	260
Viewing Basic File Information	262
Viewing Table Information	264
Viewing Block Information	266
Chapter 30 / Monitoring Resources	271
Utilization History	271
Real-Time View	272
Chapter 31 / Monitoring Processes	275
Viewing Memory and CPU Utilization	275
CPU and Memory Utilization by Instance	276
Chapter 32 / Managing Mobile Devices	277
Features in Mobile Device Administration	277
Managing Mobile Devices	278

**PART 7 SAS High-Performance Computing
Management Console 281**

Chapter 33 / Overview of the Console	283
About the Console	283
Accessing the Console	284
Specifying Console Preferences	284
Configuring Console Network Settings	286
Configuring Console Logging	287
Chapter 34 / Managing Console Users and Groups	291
Managing Console Users and Groups	291
Force Off Console Sessions	300
Chapter 35 / Viewing Console Logs	303
View Logs for an Active Session	303
Search Console Logs	304
Understanding Log Details	305
Chapter 36 / Managing Users and Groups	309
About Managing Users and Groups	309
Managing Users and Groups	311
Managing User Sessions	319
Perform an SSH Lockout	321
Chapter 37 / Managing Database Resources	325
Managing Active Queries and Connections	325
Managing Temporary Tables	329
Chapter 38 / Gridhosts File Management	331
What Is the Gridhosts File?	331
Managing the Gridhosts File	332
Glossary	335
Index	339

Using This Book

Audience

The features of SAS Visual Analytics are designed for the following users:

- Persons responsible for managing SAS servers and managing the SAS Visual Analytics environment.
- Persons needing to explore data in support of ad hoc business questions.
- Persons responsible for designing and creating reports for their enterprise.
- Persons responsible for analyzing report data and making decisions based on that data.

Prerequisites

Here are the prerequisites for using SAS Visual Analytics:

- A user ID and password for logging on to SAS Visual Analytics.
- A supported browser installed on your desktop client.
- Access to data sources that can be used to obtain data for exploration or reports.

If you have questions about whether you are ready to use SAS Visual Analytics, contact your system administrator.

What's New

What's New in SAS Visual Analytics

Enhancements to the Data Preparation Interface

Support for Jobs

The data preparation interface adds support for saving the data preparation activities as a job. The job is stored as an object in metadata. Metadata security applies to the object so that the jobs are visible to users that are authorized to use them only. Scheduling can be used to perform routine tasks such as restaging data, starting and stopping SAS LASR Analytic Server instances, and loading and unloading data.

For deployments that include SAS Data Integration Studio, jobs can be deployed for scheduling. Alternatively, you can export the code for the job to a directory and use it with a SAS DATA Step Batch Server. The batch server and the Schedule Manager support scheduling jobs. SAS Visual Analytics also includes support for using operating system scheduling and Platform Suite for SAS.

Enhancements to Preparing Data

The following enhancements are available when preparing data:

- Tables can be joined with a variety of join conditions such as inner join, left outer join, right outer join, and full outer join.
- Output data can be stored in Teradata and EMC Greenplum through SAS/ACCESS interfaces. The SAS LASR Analytic Server supports these two third-party vendor databases as a co-located data provider.

Support for New Co-Located Data Providers

SAS Visual Analytics Hadoop continues to be supported as the distributed data provider that is available from SAS. This release adds support for using Teradata and EMC Greenplum as co-located data providers. Data access is provided by a SAS/ACCESS interface that is used by the data preparation interface for reading and writing tables. Once the data is distributed to one of the third-party vendor appliances, the data preparation interface can be used to load the distributed data into a SAS LASR Analytic Server instance.

Enhancements to SAS High-Performance Computing Management Console

Enhancements to the console include:

- support for HTTPS
- support for Teradata and EMC Greenplum co-located data providers.
- support for generating Secure Shell (SSH) keys for existing operating system user accounts.

- support for distributing the shared mid-tier key to existing operating system user accounts. The key can be distributed independently from generating SSH keys.
- support for monitoring long-running queries and temporary tables in EMC Greenplum.

Enhancements to the Home Page

New and enhanced features to the home page include:

- Thumbnail images are available for explorations.
- The object inspector has been enhanced. For example, it displays thumbnail images for explorations.
- The search has been improved.

Enhancements to the Explorer Interface

The explorer interface has a new crosstab visualization type.

Enhancements to the Designer Interface

New and enhanced features in the designer interface include:

- You can add text, images, and titles to reports.
- You can add and update interactions for report objects using a new **Interactions** tab in the right pane.

- You can add and reply to comments on reports using a new **Comments** tab in the right pane.
- You can add and update data items for report objects using a new **Roles** tab in the right pane.
- You can configure the tabs in the right pane. For example, you can choose to display icons instead of the tab names.
- Using the updated data pane, you can get details about the measures in the report.
- You can replace the data source for a report.
- You can add duplicate data items.
- The interface is more accessible via a keyboard.
- The count aggregation is available for data items.
- There is a frequency assignment for data items.
- You can export data to Microsoft Excel, tab-separated values (TSV) files, and comma-separated values (CSV) files.

Accessibility

Accessibility Features of SAS Visual Analytics

Overview

SAS Visual Analytics has been tested with assistive technology tools. It includes accessibility and compatibility features that improve the usability of the product for users with disabilities, with exceptions noted below. These features are related to accessibility standards for electronic information technology that were adopted by the U.S. Government under Section 508 of the U.S. Rehabilitation Act of 1973 (2008 draft proposal initiative update). Applications are also tested against Web Content Accessibility Guidelines (WCAG) 2.0, part of the Web Accessibility Initiative (WAI) of the Worldwide Web Consortium (W3C). For detailed information about the accessibility of this product, send e-mail to accessibility@sas.com or call SAS Technical Support.

Documentation Format

Please contact accessibility@sas.com if you need this document in an alternative digital format.

User Interface Layout

SAS Visual Analytics contains the following interfaces:

- The SAS Visual Analytics *home page* contains three main sections:
 - The top of the home page contains the menu bar, which contains menus, a **Search** field, and a **Log Off** button.
 - The center of the home page contains the workspace, where you can create content, view your content, or view content created by other users.
 - The right side of the home page contains a **Common Actions** pane, which provides an alternate way to create reports, explore data, and manage your environment or favorites. The **Videos** pane links you to videos for SAS Visual Analytics. The **SAS Resources** pane links you to resources on the SAS Web site and to social media.
- The *data preparation interface* in SAS Visual Analytics enables you to perform data preparation operations and to manage the environment. The application window contains four main sections:
 - The top of the window contains the application name and an application bar that includes a menu bar and a **Log Off** button.
 - The left side of the window contains a collapsible navigation pane. This pane contains a view of the SAS Folders tree. You can navigate the tree to select tables (one at a time) for use in data preparation or loading into a SAS LASR Analytic Server instance.
 - The center of the window (the work area) contains tabs that you use for preparing data and managing the environment. Each data preparation activity is performed on a tab. A tab is also opened for working with HDFS, for monitoring resources and processes, and for managing mobile devices. The HDFS tab is available for deployments that use SAS Visual Analytics Hadoop only. The work area can be split (either vertically or horizontally) into two groups of tabs.
 - The bottom of the window contains a status bar that displays information about your connection to the metadata server.

- The *explorer interface* in SAS Visual Analytics enables you to explore data. The application window contains four main sections:
 - The top of the window contains the Home bar, the menu bar, and the toolbar. The Home bar contains a **Home** button that enables you to return to the SAS Visual Analytics home page. The menu bar contains menus and a **Log Off** button.
 - The left side of the window contains the data pane. The data pane contains all of the data items in your visual exploration. A table at the bottom of the data pane enables you to modify the properties of the data item that is currently selected.
 - The center of the window contains the workspace. The workspace contains the visualizations (charts, tables, histograms, and maps) for your visual exploration. A dock bar at the bottom of the workspace enables you to manage your visualizations.
 - The right side of the window is a pane that contains the **Roles**, **Filters**, and **Properties** tabs. The **Roles** tab enables you to manage data roles for the current visualization. The **Filters** tab enables you to create filters for the current visualization. The **Properties** tab enables you to manage properties for the current visualization.
- The *designer interface* in SAS Visual Analytics provides a framework for working with data and reports. The application window contains four main sections:
 - The top of the window contains the Home bar and the menu bar. The Home bar contains a **Home** button that enables you to return to the SAS Visual Analytics home page. The menu bar contains menus and a **Log Off** button.
 - The left side of the window contains two panes. The upper pane contains the **Objects** and **Imports** tabs. These tabs enable you to select report objects or import other reports or report objects. The lower pane contains the **Data** and **Display Rules** tabs. These tabs enable you to select a data source (or data sources) and to create a new display rule. A table at the bottom of the **Data** tab enables you to modify the properties of the data item that is currently selected.
 - The center of the window is the canvas (or workspace) where you design reports.

- The right side of the window contains a pane for the **Properties** tab, which lists the properties for the currently selected report object. The **Styles** tab enables you to specify styles for the data, frame, and text, as well as colors for the data. The **Display Rules** tab enables you to populate or add intervals for the report object, like a gauge, which is currently selected in the canvas. The **Comments** tab enables you to add comments to report objects after you save the report. The **Interactions** tab enables you to add or update filter or brush interactions to the report objects in a section. The **Roles** tab enables you to add or update role assignments when you have a report object and a data source selected.

To customize the application window and its features, select **File** ▶ **Preferences**. On the SAS Visual Analytics home page, click either  or **Edit Preferences** to open the Preferences window. For more information about the layout and features of SAS Visual Analytics, see “[Your First Look at the SAS Visual Analytics Home Page](#)” on page 11.

Themes

An application’s theme is the collection of colors, graphics, and fonts that appear in the application. The following themes are provided with this application: SAS Corporate, SAS Blue Steel, SAS Light, and SAS Dark. To change the theme for the application, select **File** ▶ **Preferences** and go to the **Global Preferences** page. On the SAS Visual Analytics home page, click either  or **Edit Preferences** to open the Preferences window. For more information about the layout and features of SAS Visual Analytics, see “[Your First Look at the SAS Visual Analytics Home Page](#)” on page 11.

You can also use keyboard shortcuts to magnify the contents of the browser window or to invert the application colors. For more information, see “[Keyboard Shortcuts](#)” on page xx.

Note: If you have special requirements for your themes, then contact your system administrator or visual designer about using the SAS Theme Designer for Flex application to build custom themes. SAS Theme Designer for Flex is installed with SAS themes. For more information about this tool, see *SAS Theme Designer for Flex: User’s Guide*.

Keyboard Shortcuts

The following table contains the keyboard shortcuts for the application. In the user interface, the shortcuts are displayed within parentheses in tooltips and menu labels.

Note: Some application-level keyboard shortcuts do not work when you first open an application. When that happens, press Tab to place the focus on the application, and then try the keyboard shortcut again.

Note: When you use a keyboard shortcut to activate a button, first give the focus to the field or section that the button is associated with before you use the keyboard shortcut.

For example, if a table has an associated  button, you must first move the focus to the table before you press Ctrl+?.

Task	Keyboard Shortcut
Open a Help pop-up window from the  button.	Ctrl+? Note: This shortcut does not work on some keyboards (for example, the Italian keyboard).
Zoom in.	Ctrl++
Zoom out.	Ctrl+-
Reset the zoom state.	Ctrl+0
Hide or show the application bar (includes the menu bar and the workspace bar), category pane, and tile pane.	Ctrl+Alt+Shift+M
Open a pop-up menu.	Shift+F9 (if a menu is available in that context) Note: If you use Shift+F9 to display the pop-up menu, then it is always displayed in the top left corner of the user interface control that you are using.

Task	Keyboard Shortcut
Open the Landmarks window and move focus to the list of landmarks. Each landmark represents an area of the application window. You can quickly navigate to that area of the application by selecting a landmark.	Ctrl+F6
Temporarily invert or revert application colors (for the current session only). Note: You can set the Invert application colors preference in the Preferences window if you want the color change to persist across sessions.	Ctrl+~
Rename the selected tab.	Make sure that focus is on the tab. Press F2 and specify the new name. To commit your changes, press Enter. To cancel your changes, press Esc.
Close the selected tab.	Make sure that focus is on the tab, and then press Delete. Note: Some tabs cannot be closed.
Switch in and out of Edit mode for a table cell.	To enter Edit mode, select a cell and press F2. To exit Edit mode, press Esc.
Navigate to the column headings in a table.	For a two-dimensional table, first make sure that focus is on the table and that you are not in Edit mode. Press Ctrl+F8 to switch focus between column headings and table cells. Use the arrow keys to navigate from heading to heading. For a multidimensional table, first make sure that focus is on a table cell and that you are not in Edit mode. Press Ctrl+F8 to switch focus between column headings, row headings, and table cells. Use the arrow keys to navigate from heading to heading.

Task	Keyboard Shortcut
Navigate the content rows of a table	When table cells are in Edit mode: ■ Press Tab and Shift+Tab to move from cell to cell horizontally across columns. ■ Press Enter and Shift+Enter to move from cell to cell vertically across rows. When table cells are not in Edit mode, use the arrow keys to move from cell to cell.
Sort columns in a table.	To sort a single column, navigate to the column heading of the column that you want to sort. Press spacebar to sort the column. To sort additional columns, navigate to the column heading of each additional column that you want to sort. Press Ctrl+spacebar.

Exceptions to Accessibility Standards

Exceptions to accessibility standards are documented in the following table.

Note: The JAWS issues occur when JAWS is used with Internet Explorer. Other browsers were not tested with JAWS, unless noted.

Accessibility Issue	Workaround
Using JAWS with this application can sometimes cause Internet Explorer to stop responding.	Using JAWS 13 or later might resolve the issue.
JAWS cannot read some of the controls in the application, such as images, icons, and buttons.	
JAWS cannot read the tooltips of items in trees and lists.	

Accessibility Issue	Workaround
JAWS refers to table controls as list boxes.	When JAWS reports that a control is a list box, keep in mind that it might actually be a table.
The keyboard shortcuts that are used to interact with editable tables can conflict with keyboard shortcuts for JAWS' forms mode.	As a best practice, disable the JAWS virtual PC cursor when you work with tables. Tab to the table and press Insert+Z to disable the JAWS virtual PC cursor. When you finish interacting with the table, press Insert+Z to re-enable the JAWS virtual PC cursor.
JAWS cannot read two-column property tables.	
JAWS does not correctly read the information in a table: <ul data-bbox="102 783 692 1200" style="list-style-type: none"><li data-bbox="102 783 692 845">■ JAWS cannot read the column headings of a table.<li data-bbox="102 854 692 960">■ When table cells are not editable and focus is on the body of the table, JAWS reads an entire row at a time instead of cell by cell.<li data-bbox="102 969 692 1200">■ When table cells are editable, and focus is on the body of the table, JAWS reads only the first row of the table. If you use the arrow keys to select a cell or row, then JAWS does not read anything. If you press Enter to edit a cell, then JAWS reads the row that contains the edited cell.	
When you are in a table cell, if you press Home, End, Page Up, or Page Down, the selected cell will change to be one in the first column of the currently displayed columns for the table.	Use the arrow keys to navigate through the cells of the table.
You cannot use the keyboard to activate the links within how-to topics and Help pop-up windows.	Use the Help menu to access the linked documents.
You cannot use Shift+F10 to open a pop-up menu.	Use Shift+F9 to open pop-up menus that are created for the SAS application. The generic menu that is provided by the Flash player cannot be opened by Shift+F9.

Accessibility Issue	Workaround
You cannot use the keyboard to access the close (x) button that is in the top right corner of a tab.	Make sure that focus is on the tab, and then press Delete to close the tab.
Visual focus for the menu bar is indicated with an outline around the entire menu bar. To select individual menus, use the left or right arrow key.	
Sometimes, after the focus leaves an area of the Web browser that is controlled by the Flash player, you will not be able to use the keyboard to return the focus to that area.	Use your pointer to bring the focus back into the area.
Sometimes, you cannot use F2 to put a table cell into Edit mode.	When you are in the cell that you want to edit, press Ctrl+F8 twice (for two-dimensional tables) and three times (for multidimensional tables).
When you create a row filter with the data preparation interface, you cannot use the keyboard to select the column to use with the filter.	
You cannot use the keyboard to access the process details window for an application instance on the Process Monitor in the data preparation interface.	
You cannot use the Tab key to move the focus outside of a code or expression editor. Pressing Tab within the editor only inserts tabs.	For Internet Explorer, press Shift+F10 and then press Esc to move the focus outside of the editor. For Firefox, press Alt+Tab to switch to another application. When you switch back, the focus will be outside of the editor.
The designer and data preparation interfaces do not have usable landmarks.	

Part 1

Introduction to SAS Visual Analytics

Chapter 1

<i>About SAS Visual Analytics</i>	3
---	---

Chapter 2

<i>Manage Access to SAS Visual Analytics</i>	7
--	---

Chapter 3

<i>SAS Visual Analytics Interface</i>	11
---	----

1

About SAS Visual Analytics

<i>What Is SAS Visual Analytics?</i>	3
<i>Benefits of Using SAS Visual Analytics</i>	4
<i>How Does SAS Visual Analytics Work?</i>	4

What Is SAS Visual Analytics?

SAS Visual Analytics is an easy to use, Web-based product that leverages SAS high performance analytic technologies and empowers organizations to explore huge volumes of data very quickly in order to see patterns and trends, and to identify opportunities for further analysis. The highly visual, drag and drop data exploration interface combined with the speed of the SAS LASR Analytic Server accelerates analytic computations and enables organizations to derive value from massive amounts of data. This creates an unprecedented ability to solve difficult problems, improve business performance, and mitigate risk rapidly and confidently.

SAS Visual Analytics empowers business users, business analysts, and IT administrators to accomplish tasks from an integrated suite of applications that are accessed from a home page. The central entry point for SAS Visual Analytics enables users to perform a wide variety of tasks such as preparing data sources, exploring data, designing reports, as well as analyzing and interpreting data. Most importantly, reports can be displayed on a mobile device or in a Web viewer.

Benefits of Using SAS Visual Analytics

Using SAS Visual Analytics, users can explore new data sources, investigate them, and create visualizations to uncover relevant patterns, and then easily share those visualizations in reports. Traditional reporting is prescriptive. That is, you know what you are looking at and what you need to convey. However, data discovery invites you to plumb the data, its characteristics, and relationships. Then, when useful visualizations are achieved, you can incorporate those visualizations into reports that are available on a mobile device or in a Web viewer.

SAS Visual Analytics provides users with the following benefits:

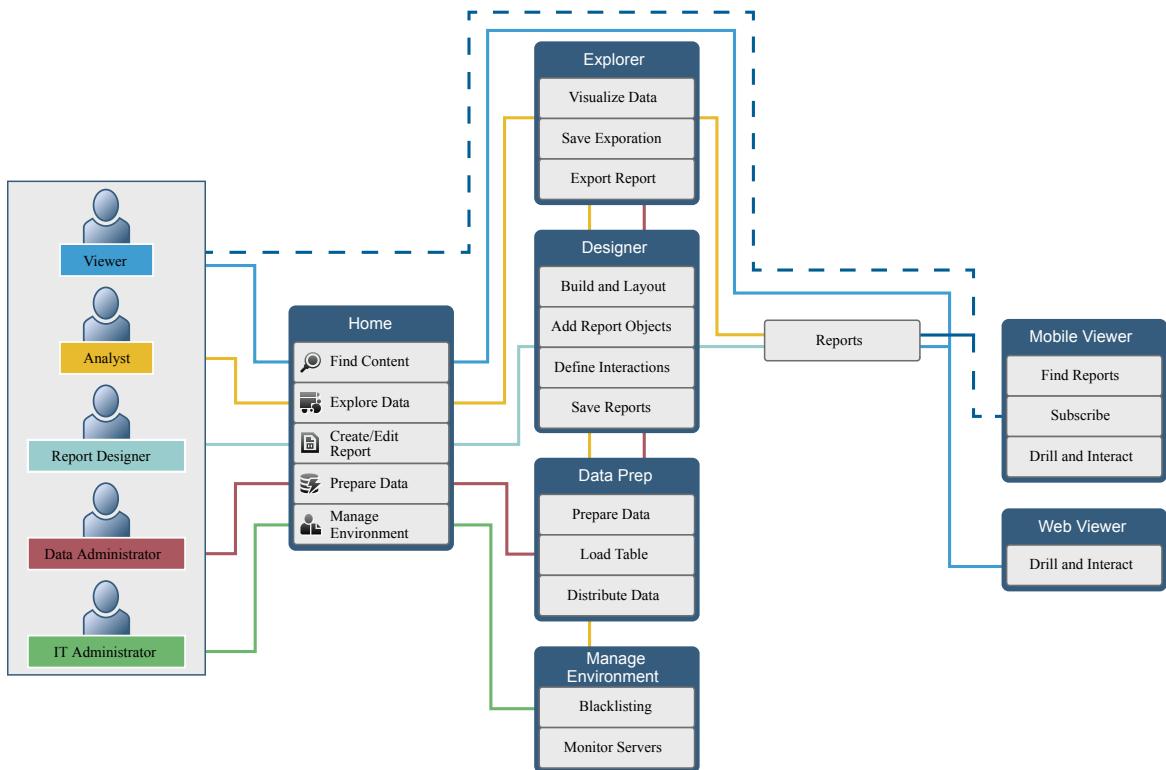
- enables users to apply the power of SAS analytics to massive amounts of data
- empowers users to visually explore data, based on any variety of measures, at amazingly fast speeds
- enables users to share insights with anyone, anywhere, via the Web or a mobile device

How Does SAS Visual Analytics Work?

You can use SAS Visual Analytics to explore and view data, interact with and create reports, and display reports on a mobile device or in a Web viewer. You can explore your data by using interactive visualizations such as charts, histograms, and tables. Report authors can easily point and click to query central sources of data. You can add filters and design the layout using tables, graphs, and gauges using drag and drop to create a well-formatted report.

The following figure illustrates how the different pieces of SAS Visual Analytics work together. It also shows how users interact with the different interfaces.

Figure 1.1 Overview of SAS Visual Analytics



2

Manage Access to SAS Visual Analytics

<i>Access to SAS Visual Analytics Capabilities</i>	7
<i>Log On to SAS Visual Analytics</i>	8
<i>Log Off from SAS Visual Analytics</i>	9

Access to SAS Visual Analytics Capabilities

Users might have access to different functionality, depending on their assigned roles. Roles are mapped to capabilities. A *capability*, also known as an *application action*, defines the operations a user can perform.

SAS Visual Analytics is shipped with three predefined roles—Report Viewing, Analysis, and Administration. A predefined set of capabilities is available for each role. A system administrator can modify these roles and specify the capabilities that meet the guidelines for your company. They can also define new roles. If you have questions about your assigned role, contact your system administrator.

For more information about the roles and capabilities that are available, see the *SAS Visual Analytics: Administration Guide*.

Note: This user's guide discusses tasks that you might be able to perform, depending on your role.

Log On to SAS Visual Analytics

SAS Web Visual Analytics uses the standard logon window for SAS applications.

To log on to SAS Visual Analytics:

- 1 To display the SAS Visual Analytics logon window, click on the URL that is supplied by your system administrator or paste it into the address field of your browser. For example, you might enter `http://server01.abc.com:8080/SASVisualAnalyticsHub/`

Display 2.1 Log On Window for SAS Visual Analytics



- 2 To log on to SAS Visual Analytics:

- In the **User ID** field, enter your user ID.
- In the **Password** field, enter the password for the user ID that you just entered.
- Click **Log On**.

The SAS Visual Analytics home page appears.

Note: Your password is case-sensitive. Your user ID might be case-sensitive, depending on the operating system that is used to host the Web application server. If you need assistance, contact your system administrator.

Log Off from SAS Visual Analytics

To log off from SAS Visual Analytics, click **Log Off** in the upper right corner of the user interface.

Note: When you select **Log Off**, you are logged off from all environments.

If you are prompted about unsaved changes, click **Log Off** to exit without saving or click **Continue** to return to SAS Visual Analytics and keep working. Here is an example of the time-out warning message:

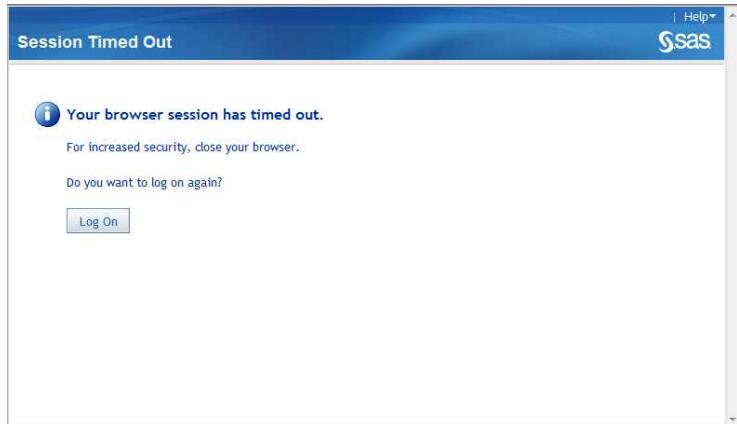
Display 2.2 Session Time-out Warning Message



If you lose your connection to SAS Visual Analytics (for example, your session times out), then you must begin again at the point where you last saved your work. By default, if there is no activity for 30 minutes, then SAS Visual Analytics automatically logs you off and displays the logon window. Your system administrator can change the inactivity period and whether the **Log On** button is available. As a best practice, save your work frequently.

The following example shows the message that appears after your session has timed out:

Display 2.3 Session Timed Out Message for SAS Visual Analytics



3

SAS Visual Analytics Interface

<i>Your First Look at the SAS Visual Analytics Home Page</i>	11
<i>About the Availability of Menus and Menu Selections in SAS Visual Analytics</i>	13
<i>Managing Your Content on the Home Page</i>	14
<i>Specifying Your Preferences</i>	15
Specifying Global Preferences	15
Specifying the SAS Visual Analytics General Preferences	17
Specifying the SAS Visual Analytics Initial Page Preference	18
Specifying Preferences for the Data Preparation Interface	19
Specifying Preferences for the Explorer Interface	19
<i>Discovering Details about Objects on the Home Page</i>	20
<i>Adding Comments to Reports and Explorations on the Home Page</i>	22
<i>Searching from the Home Page</i>	25

Your First Look at the SAS Visual Analytics Home Page

After you log on to SAS Visual Analytics using the standard window for SAS applications, you will see the home page. The home page enables you to create new

content in SAS Visual Analytics and enables you to access content that you and others have created.

Here are the features of the SAS Visual Analytics home page:

Figure 3.1 SAS Visual Analytics Home Page



- 1 The menu bar enables you to access task options, such as creating reports, exploring data, managing your environment or favorites, and getting help on using SAS Visual Analytics. You can also search all SAS content from the menu bar, and you can log off from SAS Visual Analytics.
- 2 The **Create Content** area provides icons to let you quickly create a new report or to explore data.
- 3 The **My Content** area lists any explorations or reports that you have opened or created recently. It also lists any content that you have marked as a favorite. Click **Browse** to explore folders to find a report or an exploration.
- 4 The **Other Content** area displays lists of explorations or reports that your administrator has added for you. If your administrator has not created any lists, then you will not see this content area. If you have an Administration role, then you will also see a **Manage** link that enables you to manage lists.

- 5 The **Common Actions** pane provides an alternate way for you to create reports, explore data, and manage your environment or favorites. If you have an Administration role, then you will also see a link for preparing data.
- 6 The **Videos** pane links you to videos for SAS Visual Analytics.
- 7 The **SAS Resources** pane links you to resources on the SAS Web site and to social media.

Once you start using the home page, thumbnails enable you to open explorations or reports that you have created. You can mark content as favorites, and you can search for content on the home page.

For more information about other parts of the SAS Visual Analytics interface, see the following topics:

- “[Overview of SAS Visual Analytics Explorer Interface](#)” on page 69.
- “[Your First Look at the Designer Interface](#)” on page 154.

About the Availability of Menus and Menu Selections in SAS Visual Analytics

All of the following conditions influence whether a SAS Visual Analytics menu or menu selection is available to use:

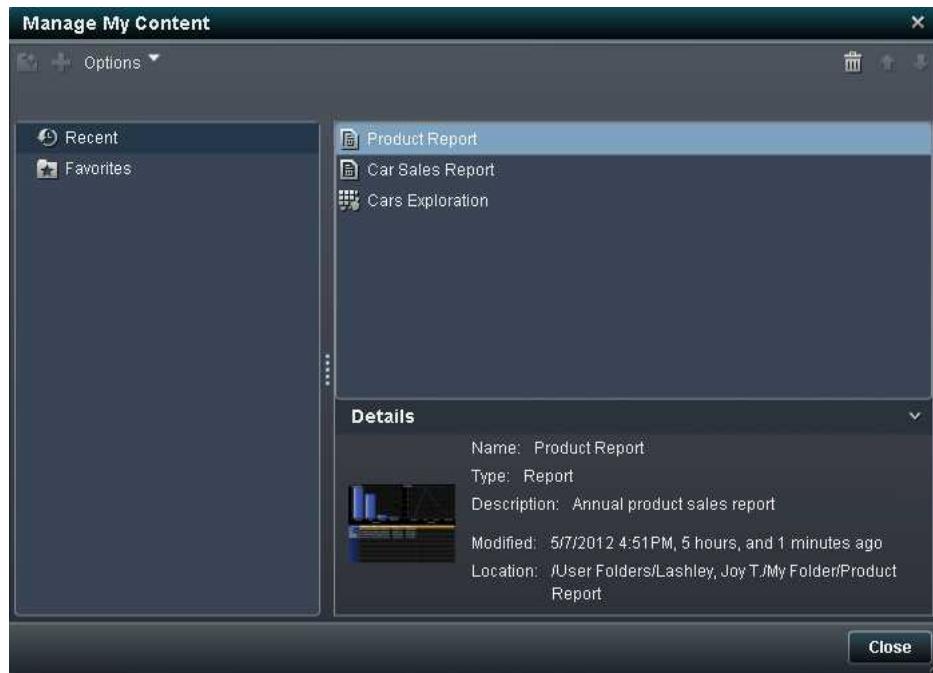
- your role and the associated capabilities. For example, you must have an Administration role to prepare data.
- your location in SAS Visual Analytics. For example, some tasks are available only if you are designing a report.
- the currently selected report object. For example, ranges are not available for list tables.
- whether the data for a report has been defined. For example, if the data has not been selected, then you cannot create a filter.

For more information about roles and capabilities, see the *SAS Visual Analytics: Administration Guide*.

Managing Your Content on the Home Page

The home page displays recent reports and explorations, as well as favorites, under the **My Content** heading on the home page. To manage your content, click **Manage** to the right of the **My Content** heading. This displays the Manage My Content window.

Display 3.1 Manage My Content Window



The icons in the Manage My Content window let you create a new favorites group, add a favorite, or select options.

Icon	Description
	Creates a new folder for your reports and explorations.

Icon	Description
	Adds a report or exploration to your list of favorites.
Options	Enables you to select either Clear Recent History or Clear Favorites .
	Enables you to delete a favorites folder, a report, an exploration, or recent content one at a time. You can delete multiple favorite items by pressing the Ctrl key.
	Enables you to move favorites folders, reports, and explorations up in the list to change what is displayed under the My Content heading on the home page.
	Enables you to move favorites folders, reports, and explorations down in the list to change what is displayed under the My Content heading on the home page.

When you select a report or exploration in the Manage My Content window, the details such as the name and location of that report are displayed. If you select a favorite group, the details include information about how many favorites the folder contains and its location.

Specifying Your Preferences

Specifying Global Preferences

You can specify global preferences that are applied to all SAS Web applications that are displayed with the Adobe Flash player. These preferences are set by each user.

To specify global preferences:

- 1 Click either or **Edit Preferences** to open the Preferences window.

Note: If you are in the data preparation, exploration, or designing environments, then select **File** ▶ **Preferences** to open the Preferences window.

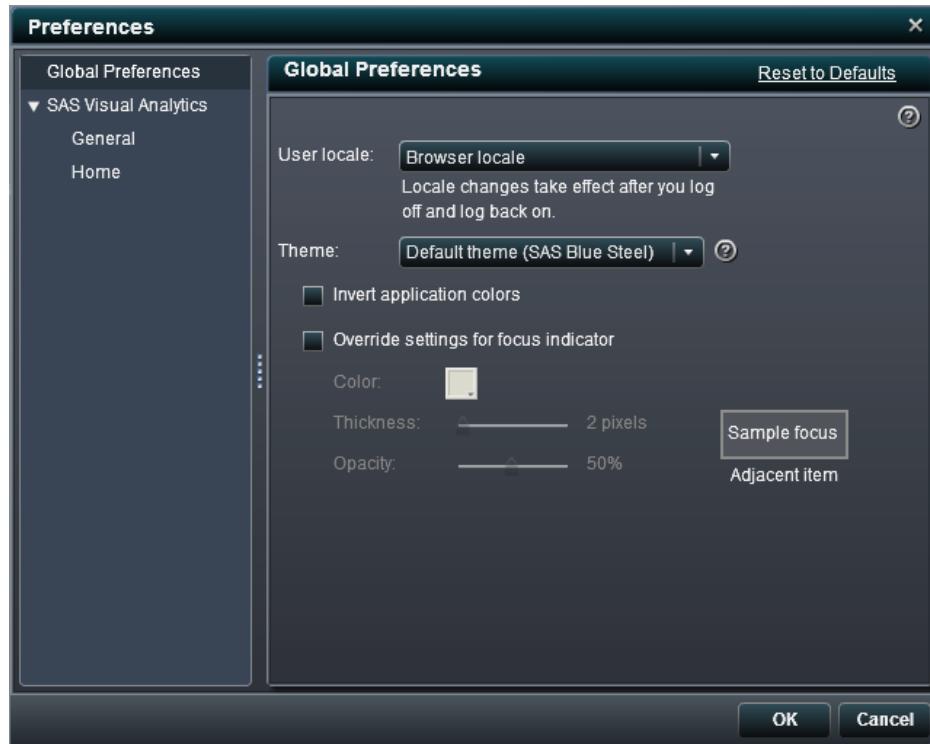
- 2 Make sure that **Global Preferences** is selected in the left pane.
- 3 Select a **User locale** to specify your language and geographic region.

Select a **Theme** to change the color scheme and other visual settings for all of your SAS Web applications.

Select **Invert application colors** to invert all of the colors in your SAS Web applications.

Select **Override setting for focus indicator** to change the color, thickness, and opacity of the focus in your SAS Web applications.

Display 3.2 Global Preferences in the Preferences Window



4 Click **OK** to apply your changes.

Note: To restore all of the Global Preferences options, click **Reset to Defaults**.

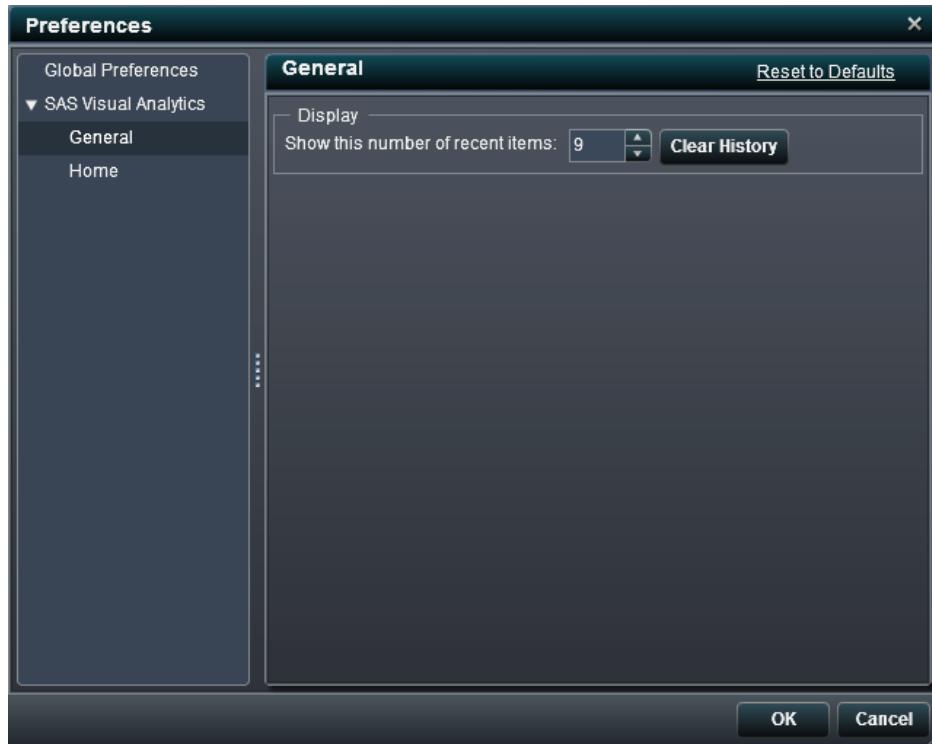
Specifying the SAS Visual Analytics General Preferences

Using the SAS Visual Analytics home page, you can specify your general preferences for SAS Visual Analytics.

To specify your general preferences:

- 1 Click either  or **Edit Preferences** to open the Preferences window.
- 2 Click **General** in the left pane.
- 3 Specify your preference for **Show this number of recent items**. The default setting is 9 items.
Click **Clear History** to reset your history.

Display 3.3 Preferences Window for SAS Visual Analytics General Preferences



4 Click **OK** to apply your changes.

Note: To restore the SAS Visual Analytics general preference, click **Reset to Defaults**.

Specifying the SAS Visual Analytics Initial Page Preference

Using the SAS Visual Analytics home page, you can specify your preference for the initial screen that you want to see after you log on to SAS Visual Analytics.

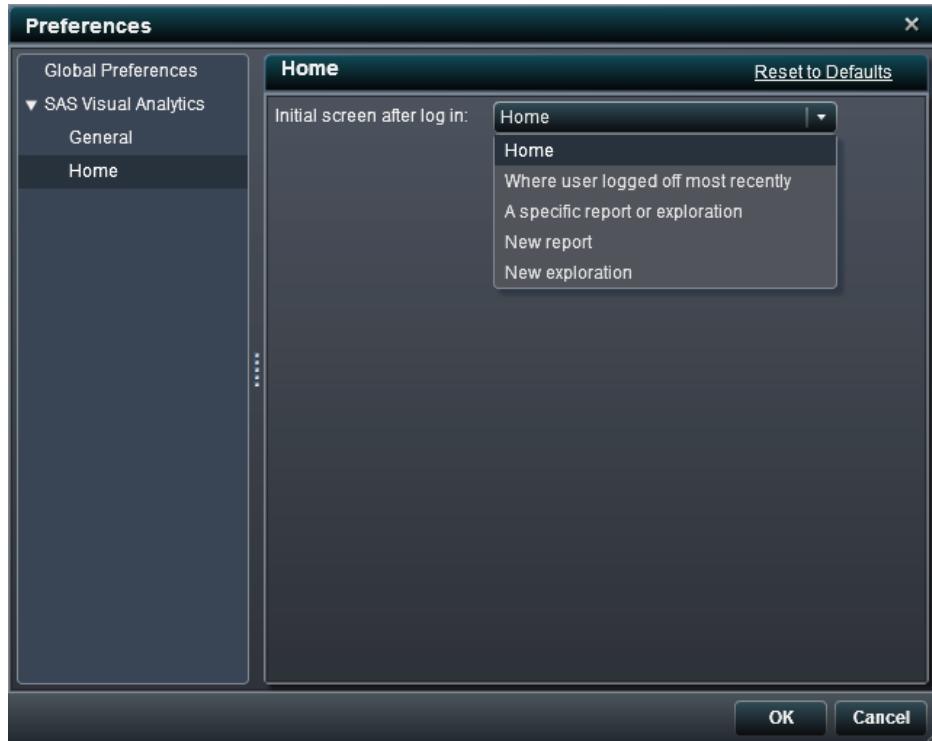
To specify your initial page preference:

- 1 Click either  or **Edit Preferences** to open the Preferences window.
- 2 Click **Home** in the left pane.
- 3 Select one of the options from the **Initial screen after log in** drop-down list.

The following options are available:

- **Home**
- **Where user logged off most recently**
- **A specific report or exploration**
- **New report**
- **New exploration**

Display 3.4 Preferences Window for SAS Visual Analytics Initial Page



4 Click **OK** to apply your changes.

Note: To restore the SAS Visual Analytics option, click **Reset to Defaults**.

Specifying Preferences for the Data Preparation Interface

See “Specifying Your Preferences for the Data Preparation Interface” on page 32.

Specifying Preferences for the Explorer Interface

See “Specify Preferences That Are Specific to the Explorer Interface” on page 149.

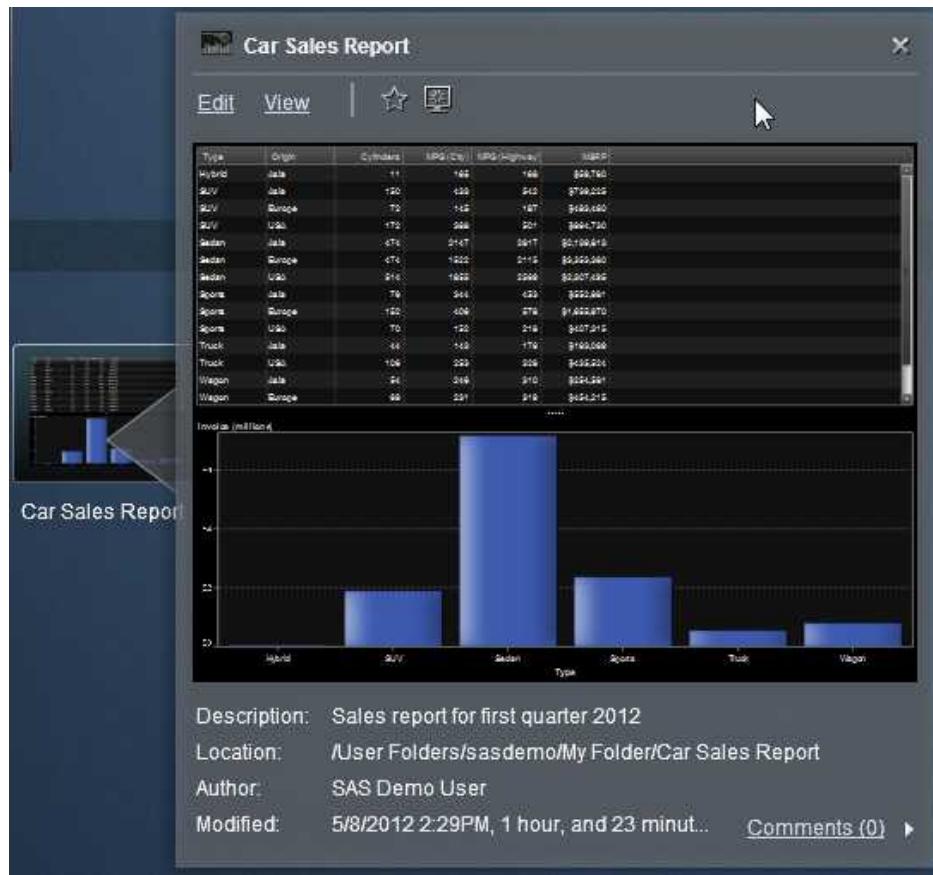
Discovering Details about Objects on the Home Page

The home page provides a feature that enables you to quickly find more information about an object when you click it. You can see a description of the report or exploration, where the report or exploration is saved, who the author is, and when the report or exploration was modified. A small toolbar lets you add the report or exploration to your favorites or set it as your preferred starting page. Note that when you set a favorite, the icon turns yellow (★).

Your role determines the availability of other actions on the toolbar (such as viewing, editing, or opening) in the object inspector.

Here is an example of what you might see for a report:

Display 3.5 The Object Inspector on the Home Page



If you have the Report Viewing role, then the toolbar enables you to open a report in the Web viewer by clicking **Open**.

If you have the Analysis or Administration roles, then the toolbar in the object inspector enables you to do the following:

- **Edit** a report, which opens the report in the designer interface so that you can edit or change the objects in the report.
- **View** a report, which opens the report in the Web viewer.
- **Open** for an exploration, which displays the explorer interface.

- Click  to add the report or exploration to your list of favorites. When the icon is yellow () click to remove the object from your list of favorites.
- Click  to set the report or exploration as the initial screen when you start SAS Visual Analytics.

All users can add **Comments** to a report or view existing comments about the report using the link in the lower right of the object inspector.

Adding Comments to Reports and Explorations on the Home Page

You can use the object inspector on the home page to add comments to a report or to respond to existing comments for a report or an exploration.

To add a comment:

- 1 Click a report or an exploration to open the object inspector.
- 2 Click **Comments** in the lower right corner to expand the window.

Display 3.6 Comments on the Home Page

The screenshot shows the SAS Home Page with a report titled "Car Sales Report". The report includes a table of car data and a bar chart. To the right, a "Comments for: Car Sales Report" window is open, showing a search bar and a blank comment area. Below the report, there is a description and metadata.

Description: Sales report for first quarter 2012

Location: /User Folders/sasdemo/My Folder/Car Sales Report

Author: SAS Demo User

Modified: 5/8/2012 2:29PM, 1 hour, and 23 minut... [Comments \(0\)](#)

New Topic

3 Click **New Topic** to open the New Topic window.

4 Enter a topic name and a comment.

Here is an example:

Display 3.7 New Topic Window for Adding Comments

The "New Topic" window is open, showing a topic name and a comment area. At the bottom, there are buttons for "Normal" (with a dropdown arrow), "Attachment", and "Save".

Topic: 2012 Data

Comment: Please add data for the second quarter of 2012.

Normal | **Attachment** **Save**

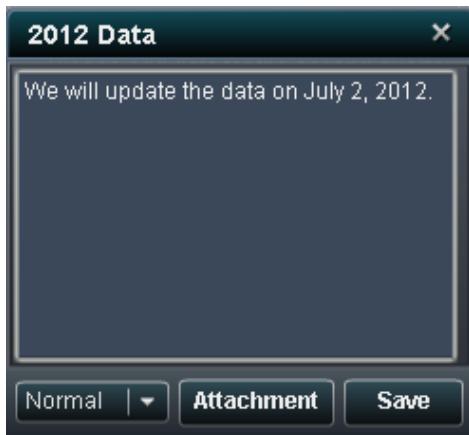
- 5 (Optional) Set a priority for the comment using the drop-down list. You can select **Normal**, **High**, or **Low**. If you set the priority to **High**, then a visual flag will be added to the comment.
- 6 (Optional) Click **Attachment** to add a file or an image to your comment.
- 7 Click **Save**. The comment link in the object inspector will update to show that there is a comment.

To respond to an existing comment:

- 1 Click a report to open the object inspector.
- 2 Click **Comments** in the lower right corner to expand the window.
- 3 Click **Reply** to open the comment window.
- 4 Enter a comment.

Here is an example:

Display 3.8 Window for Replying to a Comment



- 5 (Optional) Set a priority for the reply using the drop-down list. You can select **Normal**, **High**, or **Low**.
- 6 (Optional) Click **Attachment** to add a file or image to your reply.

- 7 Click **Save**. The comment link in the object inspector will update to show the number of comments.

Searching from the Home Page

You can search all SAS content from the menu bar on the home page. The following types of searches are supported:

- single-word
- multiple-word
- in-word wildcards
- spaces in quoted strings
- plus (+) and minus (-) syntax

The plus (+) and minus (-) are prefix operators. This means that the operator precedes the search term that is required or excluded rather than following it. For example:

- Using "2009 2010 2011" matches documents that contain any of the three terms. The more terms that match, then the higher the ranking.
- Using "+2009 +2010 +2011" matches only documents that contain all three terms.
- Using "2009 2010 -2011" matches documents that contain 2009 or 2010 but only if they do not contain 2011.

The search feature is not case sensitive.

Only 1000 items can be returned in a single search. If your search exceeds 1000 items, then a message is displayed at the top of the window. Narrowing your search enables the relevant search results to display.

Part 2

Preparing Data

<i>Chapter 4</i>	
<i>Overview of the Data Preparation Interface</i>	29
<i>Chapter 5</i>	
<i>Creating Prepared Data</i>	39
<i>Chapter 6</i>	
<i>Loading Data</i>	49
<i>Chapter 7</i>	
<i>Server Management</i>	55

4

Overview of the Data Preparation Interface

<i>About the SAS Visual Analytics Data Preparation Interface</i>	30
<i>Overview of the Data Preparation Interface</i>	31
<i>Specifying Your Preferences for the Data Preparation Interface</i>	32
<i>About Access to DBMS Data</i>	33
<i>About Exporting Code</i>	34
<i>About Jobs</i>	35
Creating Jobs	35
Deploying Jobs	35
Editing Jobs	35
<i>About Table Metadata</i>	36
<i>Working with User-Defined Formats</i>	37
Make User-Defined Formats Available	37
Limitations	37

About the SAS Visual Analytics Data Preparation Interface

The data preparation interface enables an administrator to prepare data by performing joins, adding calculated columns, subsetting data, and sorting. Once the data preparation operations are entered, administrators can load tables to the SAS LASR Analytic Server. In addition, for large data sets, the prepared data can be distributed to a co-located data provider such as SAS Visual Analytics Hadoop. The SAS LASR Analytic Server is able to read large data sets in parallel to memory from data that is distributed in a co-located data provider.

The data preparation interface is also used to start and stop server instances, as well as to load and unload tables from the server instances. Once a server is started, administrators can view information about the server instance. This includes the user account that was used to start the server and the tables that are resident in memory. The information about tables includes the number of rows and columns for an at-a-glance evaluation of the number of server instances and tables available for use.

For deployments that use SAS Visual Analytics Hadoop, navigating files in HDFS is also available from the data preparation interface. The HDFS content explorer enables administrators to browse the tables that were prepared and added to HDFS. Tables that are already prepared for analytics can be added to HDFS directly. Once stored in HDFS, tables can be loaded into memory in a server instance from the HDFS content explorer. Because SAS uses a special file format for the data that is stored in HDFS, the HDFS content explorer also provides information about the columns and row count for the prepared data.

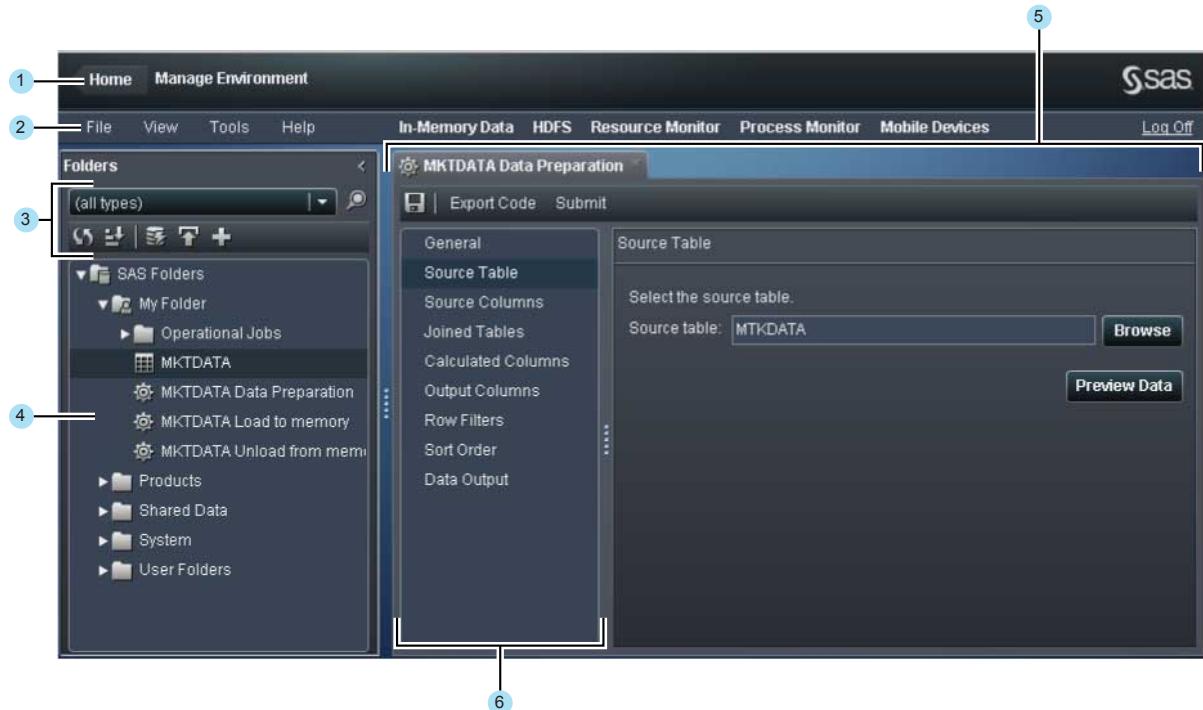
Resource and process monitors are provided so that administrators can view real-time statistics and visualizations of resource use like central processing unit (CPU) utilization, memory utilization, and input/output (I/O) rates. A line plot of resource utilization against time is provided. In addition, a real-time view creates a visualization for each blade that is used in the cluster and each CPU on a blade.

Controlling access by mobile devices to data is enabled with the mobile device administration user interface. This enables administrators to easily manage access to

data through SAS Visual Analytics applications by mobile devices. Information about the mobile devices in your organization is available to administrators. The information includes user ID, device information, and timestamp of last access.

Overview of the Data Preparation Interface

The data preparation interface contains the following elements:



- 1 The application bar enables you to return to the SAS Visual Analytics Home page.
- 2 The menu bar contains menus and buttons that enable you to perform tasks.
- 3 The navigation pane controls enable you to filter, sort, and search for objects. When you select a table, buttons appear that enable you to perform actions on the table.
- 4 The navigation pane displays a tree of tables and jobs.

- 5 The workspace enables you to specify job settings, and monitor and view management applications.
- 6 The job navigation pane lists the pages that enable you to define the settings for a job.

Specifying Your Preferences for the Data Preparation Interface

The data preparation interface stores the parameters for performance monitoring in user preferences. To set the performance monitoring parameters:

- 1 Select **File** ▶ **Preferences** to open the Preferences dialog box.
- 2 Select **SAS Visual Analytics** ▶ **Application Server**.

Select a SAS Application Server to use from the menu. If you have added SAS Application Server instances to your deployment, make sure that Job Execution Services has been configured to use the application server. For more information, see *SAS Intelligence Platform: Middle-Tier Administration Guide*.

- 3 Select **SAS Visual Analytics** ▶ **Manage Environment**. Enter values for the following preferences:

Parameter	Sample Value	Description
Resource monitor sample rate (ms)	1000	Specify the sampling rate that the resource monitor uses for polling the machines in the cluster.
Process monitor sample rate (ms)	6000	Specify the sampling rate that the performance monitor uses for polling application instances.

Parameter	Sample Value	Description
Show the processes that measure performance	Selected	By default, the processes that measure performance are filtered from the graphs in the Process monitor. However, if several instances of the processes that measure performance are running, they can reduce performance. By selecting the check box, you can view and monitor the processes that measure performance.

- 4 Click **OK** to apply your changes.

About Access to DBMS Data

The data preparation interface can be used to read source tables from third-party vendor databases and to write tables to them as well. The SAS/ACCESS Interface products provide this feature. Be aware that the SAS/ACCESS Interface products do not support creating views. On the Data Output page of a prepare data job, you can select **View** when the output library is a DBMS library, but it results in an error.

When writing to tables on the Data Output page of a prepare data job, the data preparation software writes the table to the database and then updates the metadata for the table. The update is performed with the METALIB procedure. In order for the metadata update to succeed, the user account that is performing the action must have a stored login in metadata.

For more information, see "How to Store Passwords for a Third-Party Server" in *SAS Intelligence Platform: Security Administration Guide*.

About Exporting Code

There are many opportunities to export code to a file when you use the data preparation interface. This feature enables administrators to review the SAS statements for a task. Some of these tasks are starting and stopping a SAS LASR Analytic Server instance, and loading and unloading tables. The statements in the file can be modified and scheduled. For more information about the LASR, OLIPHANT, and HPDS2 procedures, see the *SAS LASR Analytic Server: Administration Guide*.

Operations that enable saving the SAS statements have an **Export Code** button. Click the **Export Code** button to view the Export Code window. On the window, the **Export** button saves the SAS statements to the file that is specified in the **Location** field but does not submit the SAS statements for execution.

Note: The exported code does not include information about the metadata server connection. The connection information is often supplied in options such as METASERVER=, METAPORT=, METAUSER=, or METAPASS=. These options are needed to run the PROC METALIB statement that is included in jobs that register tables. The affected jobs are the prepare data and add to data server jobs. For information about the metadata server connection options, see *SAS Language Interfaces to Metadata*.

Note: The SAS statements are saved to a file with the credentials of the user account that is used to run the JBoss Application Server. This user account must have Write access to the location, and the location must be accessible from the machine that is running JBoss.

It is also possible to configure logging so that SAS statements are stored to a log file. Only the SAS statements that are generated by the data preparation interface are available to be logged. For information about enabling this feature, see *SAS Visual Analytics: Administration Guide*.

About Jobs

Creating Jobs

Most tasks that can be performed with the data preparation interface can be saved as a job. For example, when you start a SAS LASR Analytic Server instance, the data preparation interface opens a tab for the job. By default, the tab is named according to the action to perform (for example, `Start server job`) followed by a random number.

The General page on each tab is used to specify a name for the job, a metadata folder location, and a description. Jobs are saved by clicking . The job name field is limited to 60 characters.

Deploying Jobs

Once a job is saved, you can edit the job and modify the task that the job performs. If your deployment includes SAS Data Integration Studio, you can deploy the jobs for scheduling.

For deployments that do not have SAS Data Integration Studio, it is still possible to export the code for the job. You can then use the Schedule Manager plug-in to SAS Management Console to deploy the code as a SAS DATA Step Program. For more information about deploying jobs, see *Scheduling in SAS*.

Editing Jobs

Once a job is stored in metadata, you can open the job to edit it. This might be necessary if you need to add or remove columns, change a row filter, or change the column expression in a calculated column. If a job is saved with another application that can edit jobs, such as SAS Data Integration Studio, then the data preparation interface can no longer edit the job. You can use other SAS applications to modify properties such as the description, but saving the job prevents it from opening in the data preparation interface.

When you save the job, if you change the job name or folder location, the existing job object is updated with the new values. However, you cannot delete job objects from metadata with the data preparation interface. If you need to delete objects from metadata (such as jobs, libraries, tables, or folders) use SAS Management Console.

When a job is saved in metadata, information about the deployment is saved with the job. For example, the host name of the SAS LASR Analytic Server, the path to the server descriptions files, and so on, are saved with the job. If the SAS Deployment Manager is used to reconfigure the deployment, the existing values for the related properties in saved jobs are not modified. If you edit a job and the properties have changed since the job was saved, then the new values for the properties replace the existing values when the job is saved. For a list of the related properties, see "SAS LASR Analytic Server Administration" in the *SAS Visual Analytics: Administration Guide*.

About Table Metadata

The data preparation interface uses table metadata for information about table properties such as the column names and data types. Source tables are registered in metadata with SAS Management Console before you use the data preparation interface. Afterward, you use the data preparation interface to select a source table to use for creating a job.

When you create a prepare data job, you must specify an output table on the Data Output page of the prepare data job. For deployments that use a third-party vendor database as a co-located data provider, the prepare data job can also add the table to the data server. These tables must be registered in metadata so that you can select the table in the data preparation interface in order to load the table to memory on the SAS LASR Analytic Server.

Output tables on prepare data jobs and add to data server jobs are registered whenever you perform any of the following actions:

- save the job
- submit the job
- export the code for the job

Working with User-Defined Formats

Make User-Defined Formats Available

The preferred method for making user-defined formats available to a SAS Application Server is to name the format catalog formats.sas7bcat and to place it in *SAS-config-dir/Lev1/SASApp/SASEnvironment/SASFormats*. For more information about using user-defined formats, see *SAS Intelligence Platform: Data Administration Guide*.

Limitations

The following list identifies the limitations for using user-defined formats:

- If the output library for a prepare data job is WORK, then format information for the table is lost. This affects all co-located data providers, regardless of the vendor.
- If the output library for a prepare data job is a Base SAS library, then the format information is preserved with the table.
- For deployments that use SAS Visual Analytics Hadoop as the co-located data provider, format information is stored with the table when it is added to HDFS. This is true only when the output library is a Base SAS library.
- For deployments that use a third-party vendor database as the co-located data provider, format information is lost when the table is added to the data server.
- The Preview Data feature that is available when you create a prepare data job does not show user-defined formats.

You can add the format information back. You can export the code for a load table job and modify the LASR procedure code to include the format statements.

5

Creating Prepared Data

<i>Features in Data Preparation</i>	39
<i>Concepts in Data Preparation</i>	40
Source Data	40
Data Output	40
Output to SAS Visual Analytics Hadoop	41
<i>Prepare Data</i>	41
<i>Joining Tables</i>	45
<i>Adding Calculated Columns</i>	46
<i>About Previewing Data</i>	47
<i>Distributing Data</i>	47

Features in Data Preparation

The chapter describes the data preparation activities that can be performed with the data preparation interface. The following list identifies the available actions:

- select a source table and join tables based on join conditions
- add calculated columns and drop columns from source and joined tables
- control the sort order for the output data and apply row filters to subset the data
- preview the output data

- add the prepared data to a co-located data provider

With this feature set, the data preparation interface enables administrators to access data from the variety of data sources that SAS and the SAS/ACCESS interfaces can read and then operate on the data. Administrators can control the variables needed for analytics and make the data available to the SAS LASR Analytic Server for in-memory analytic processing.

Concepts in Data Preparation

Source Data

The data preparation interface enables administrators to access a wide variety of data sources for input to the data preparation process. The SAS/ACCESS interfaces can be used to interact with database management systems (DBMS). SAS data sets and SAS Scalable Performance Data Server can be used to construct a data warehouse. These data sources are available as source data for the analytic processing that can then be performed by the SAS LASR Analytic Server. The source data must be registered in SAS metadata as libraries and tables. For information about registering libraries and tables, see *SAS Intelligence Platform: Data Administration Guide*.

Data Output

The unique value of the data preparation interface is that it enables administrators to add data to a data provider that is co-located with the SAS LASR Analytic Server. SAS Visual Analytics Hadoop is a co-located data provider that is available from SAS. In this case, the Hadoop Distributed File System (HDFS) is used for data output. The purpose of adding the data to a co-located data provider is that the server can read data in parallel at very impressive rates from a co-located data provider. In addition to SAS Visual Analytics Hadoop, Teradata Enterprise Data Warehouse and EMC Greenplum Data Computing Appliance are supported co-located data providers. When a third-party vendor appliance is used, the SAS/ACCESS interface for the database must be licensed and configured.

Output to SAS Visual Analytics Hadoop

In addition to the performance advantage of using co-located data, when SAS Visual Analytics Hadoop is used, it provides data redundancy. By default, two copies of the data are stored in HDFS. If a machine in the cluster becomes unavailable, another machine in the cluster retrieves the data from a redundant block and loads the data into memory.

The data preparation interface distributes blocks evenly to the machines in the cluster so that all the machines acting as the server have an even workload. The block size is also optimized based on the number of machines in the cluster and the size of the data that is being stored. Before the data is transferred to HDFS, SAS software determines the number of machines in the cluster, row length, and the number of rows in the data. Using this information, SAS software calculates an optimal block size to provide an even distribution of data. However, the block size is bounded by a minimum size of 1 KB and a maximum size of 64 MB.

For very small data sets, the data is not distributed evenly. It is transferred to the root node of the cluster and then inserted into SAS Visual Analytics Hadoop. SAS Visual Analytics Hadoop then distributes the data in blocks according to the default block distribution algorithm.

Prepare Data

To prepare data:

- 1 Select a source table from the navigation pane, right-click, and select **Prepare Data**.
To preview the data, click **Preview Data** on the Source Table page. The preview enables you to confirm that you have Read access to the data before you begin the data preparation.
- 2 On the **General** page, specify values for the **Job name**, **Location**, and **Description** if you intend to save this action as a job.

- 3 On the Source Columns page, deselect the check boxes for any columns that you do not want to include in the prepared data.
- 4 On the Joined Tables page, click  to select a table to join. For more information about how to use this page, see “[Joining Tables](#)” on page 45.
- 5 On the Calculated Columns page, click  to add new columns to the prepared data. For more information about how to use this page, see “[Adding Calculated Columns](#)” on page 46.
- 6 On the Output Columns page, remove, reorder, or edit the output column information. You can specify the output column name, description, format, and length.

TIP The default length for numeric variables is 8. The length can be set to a lower number to reduce storage size. However, length for numeric variables is related to the precision, and reducing their lengths arbitrarily can cause precision loss. For more information, see "Numeric Variable Length and Precision in UNIX Environments" in *SAS Companion for UNIX Environments*.

- 7 On the Row Filters page, add filters to subset the input data. Click  to add a new filter. Select the column name to filter on, select the filter criteria, and enter the filter value. If more than one filter is added, the filters are applied using AND logic. You cannot create a row filter on a calculated column.
- 8 On the Sort Order page, select the column that you want to sort by. The default sort order is ascending. Use the menu to beside the selected column name to set the sort order to descending.
- 9 On the Data Output page, set the following parameters:

Parameter	Sample Value	Description
Output table	<i>tablename</i>	The output table name is populated automatically with a default value. Click Browse to select a different table, or enter the table name that you want to use.

Parameter	Sample Value	Description
Library	WORK	Click Browse to specify a library for the output table.
Location	/Shared Data	Click Browse to specify a folder for the output table. The button becomes active when the Library value is changed from WORK to another library.
Type of output	Table	<p>For deployments that use SAS Visual Analytics Hadoop, when you select Table, the size of the data is calculated and distributed evenly in HDFS with an optimal block size. If you select View, then the data is distributed in 32-megabyte blocks. However, for large data sets, any reduction in block utilization might be considered worthwhile in order to avoid moving a large data set.</p> <p>For deployments that use a third-party vendor database, do not select View when the output library is for a DBMS. Views are not supported by the SAS/ACCESS Interfaces.</p>

When the co-located data provider is SAS Visual Analytics Hadoop, the following options are available:

Table 5.1 SAS Visual Analytics Hadoop Options

Parameter	Sample Value	Description
Add to HDFS	Selected	Select this check box to distribute the data to HDFS.
HDFS output path	/user/	Enter the fully qualified path in HDFS to use for storing the prepared data. The path is case-sensitive.

Parameter	Sample Value	Description
HDFS filename	<i>tablename</i>	The filename for the table is automatically populated. It must match the name of the output table.
Table description	<i>Prepared data for tablename</i>	Specify a description to associate with the prepared data. The description is displayed beside the table name in the explorer interface.

When the co-located data provider is a third-party vendor database, the following options are available:

Table 5.2 Data Server Options

Parameter	Sample Value	Description
Add to data server	Selected	Select this check box to distribute the data to the data server.
Output table	<i>tablename</i>	The output table name is populated automatically with a default value. Click Browse to select a different table, or enter the table name that you want to use.
Library	SAS Visual Analytics Distributed Data	Make sure that you select a library that associated with the co-located data provider and that the library is configured to distribute data.
Location	/Shared Data	Click Browse to specify a folder for the output table.

10 Click Submit.

Joining Tables

On the Join Tables page, you can add multiple tables to join with the source table that you have selected. The source table can also be joined to itself. When a table is added to the **Join tables** list, the application scans the column names in the join table and attempts to find a matching column name and type in the source table. If a match is found, then the **Source table column** and **Join table column** menus are set to the column name. If no match is found or you prefer a different join condition, set the menus appropriately. The default join type is an inner join. You can specify one join condition for each table that is joined with the source table.

TIP In the **Available items** and **Selected items** lists, make sure that you select the columns that you want to include in the prepared data and move them to the **Selected items** list.

- 1 On the Joined Tables page, click  and then browse for the table to join with the source table.
- 2 The default join type is an inner join. To use a different join type, select **Inner join** in the table to open the **Join Type** menu.
- 3 Use the **Source table column** and **Join table column** menus to set the join condition.
- 4 Be sure to select the columns from the **Available items** list and move them to the **Selected items** list.

Adding Calculated Columns

Variables can be added to the prepared data. The values can be based on variables from the source table or a joined table. The added variable can even be based on a variable that is dropped from a joined table or the source table.

The following fields are available on the Calculated Columns Drill-In page after  is clicked:

Table 5.3 Calculated Columns Drill-In Fields

Field Name	Description
Column name	Specify a name for the calculated column to add to the prepared data. A default name of <code>DataMember</code> is provided. This field is limited to 32 characters.
Description	Specify a description for the column. This value is used as a column label.
Type	Select a value from the menu.
Functions tab	Select a value from the Category menu to reduce the list of functions to those associated with the category. When you select a function from the list, the function description is provided in the area below the function list. Click Insert or double-click the function name to use the function in the column expression.
Columns tab	Click the Columns tab to select a column from the source or joined tables to use. Click Insert or double-click the column name to use it in the column expression.
Arithmetic and logic operators	Click the button for the arithmetic or logic operator to insert. Buttons for inserting a concatenation operator and enclosing expressions in quotation marks and parenthesis are also available.

TIP Once the column expression is complete, click the job name from the breadcrumb trail to return to the Calculated Columns page.

On the Calculated Columns page, you can specify a format as well as change any of the fields for the calculated column.

About Previewing Data

When you create a data preparation job, you can preview the data from the Source Table page and the Data Output page. If you click the **Preview Data** button, a window appears that shows the first 25 rows of data. When the button is clicked from the Data Output page, all of the data preparation steps are performed (for example, joining tables, filtering rows, and so on) before the data is displayed.

After the window appears, you can increase the preview row limit, up to 1000, and then click **Preview** to see more rows. You can also click a column heading to sort the previewed data in ascending or descending order. Be aware that this sorts the data that has already been fetched and is already visible. Sorting in the Preview Data window does not fetch the data again in the sorted order.

The Preview Data window also does not apply user-defined formats when displaying data.

Distributing Data

An important feature of the data preparation interface is the ability to distribute data to a co-located data provider. Some of the advantages are described in “[Concepts in Data Preparation](#)”. The data preparation interface provides two ways to distribute data to a co-located data provider.

When you create a prepare data job, the Data Output page enables you to distribute data. For deployments that use SAS Visual Analytics Hadoop, there is an **Add to HDFS** check box and a series of fields for specifying HDFS parameters. For deployments that

use a third-party vendor database, the check box is labeled **Add to data server** and fields are provided for specifying the DBMS library. For more information, see “[Prepare Data](#)”.

If you have tables or views that are already prepared for analysis, then you can distribute the data without creating a prepare data job. Make sure the table or view is registered in metadata so that it appears in the SAS Folders tree. You can select the table object and select either **Add to HDFS** or **Add to Data Server**. The menu choice depends on the co-located data provider.

For deployments that use a third-party vendor database as a co-located data provider, a default library for the database is registered in metadata during the installation and configuration time frame. This library is located in SAS metadata at **/Products/SAS Visual Analytics Data Preparation/SAS Visual Analytics Distributed Data**. This library is configured with options that are specific to operating with a distributed database. For information about the options, see [SAS Visual Analytics: Administration Guide](#). When you create an add to data server job, remember to save the job so that the output table is registered in metadata with the distributed data library.

6

Loading Data

<i>Concepts in Loading Data</i>	49
<i>Load from HDFS</i>	50
<i>Load from Table</i>	51
<i>Unload a Table</i>	53

Concepts in Loading Data

SAS Visual Analytics can load data from registered tables or SASHDAT files stored in HDFS to the SAS LASR Analytic Server. The following table identifies some of the advantages and disadvantages of each method:

Table 6.1 Comparison of Registered Tables and HDFS Data Sources

Data Source	Advantages	Disadvantages
Registered table (from a library that does not use a co-located data provider)	Provides a rapid method for loading tables.	Appropriate for smaller data sets because the data must be transferred over the network. If the table is unloaded or the server stops, the data must be transferred over the network again.

Data Source	Advantages	Disadvantages
SASHDAT file in HDFS or a registered table from a library that uses a co-located data provider	<p>Impressive performance for loading very large data sets in parallel.</p> <p>If a SAS LASR Analytic Server is stopped or the table is unloaded from the server, loading it again is also very fast.</p>	Requires a separate step to add the data to HDFS or the co-located third-party vendor database and then load it before the data is available to SAS clients.

Load from HDFS

SAS LASR Analytic Server can load data from the Hadoop Distributed File System (HDFS) that is co-located on the machines in the cluster. Before loading the data, you must add the data to HDFS. To load from HDFS:

- 1 Navigate to the SASHDAT file to use in the HDFS content explorer. Select the file and click .
- 2 On the General page, specify values for the **Job name**, **Location**, and **Description** if you intend to save this action as a job.
- 3 Select the Settings page and specify the SAS LASR Analytic Server and data parameters:

Field	Description
Configuration directory	This field specifies the location of the server description files.
Server description file	This field specifies the name of the server description file.
HDFS file	This path to the SASHDAT file is populated automatically.

Field	Description
Table description	Specify a description for the data. This description overrides the description that was associated with the data set before it was added to HDFS. The description is displayed beside the table name in the explorer interface.

4 Click Submit.

The following items provide additional information about loading from HDFS:

- Before the data can be loaded from HDFS, it must be added to HDFS. For more information, see [“Prepare Data” on page 41](#).
- The data can also be added to HDFS with the OLIPHANT procedure from a SAS session.
- SAS LASR Analytic Server reads SASHDAT files from HDFS only. It does not read any other data that is stored in HDFS such as the output of MapReduce jobs.

Load from Table

SAS Visual Analytics can load tables into memory on a SAS LASR Analytic Server from a table that is registered in SAS metadata. If the table is from a DBMS library that is co-located with the SAS LASR Analytic Server instance, then the server instance recognizes that the table is from a co-located data provider and reads the data into memory in parallel. If the table is not from a co-located data provider (for example, a SAS data set), then the table is read from the data source, transferred to the root node on the cluster, distributed to the machines on the cluster, and then loaded into memory.

To load from a table:

- 1 Locate the table using the SAS Folders view or by searching. Right-click the table icon and select **Load from Table**.

- 2 On the General page, specify values for the **Job name**, **Location**, and **Description** if you intend to save this action as a job.
- 3 Select the Settings page and specify the SAS LASR Analytic Server and data parameters:

Field	Description
Configuration directory	This field specifies the location of the server description files.
Server description file	This field specifies the name of the server description file.
Table	The table name is populated automatically.
Table description	Specify a description for the data. This description overrides the description that is stored with the table. The description is displayed beside the table name in the explorer interface.

- 4 Click **Submit**.

The following items provide additional information about loading from a table:

- The table to use as the data source must be registered in metadata before you can perform this task.
- For deployments that use a co-located data provider, the SAS LASR Analytic Server performs a comparison of IP address to determine whether a parallel read is possible. The IP address that is in the SERVER= option of the generated LIBNAME statement must match the IP address of the SAS LASR Analytic Server root node.
- When the data is loaded from a data source that is not a co-located data provider, it is transferred to the root node of the server, distributed evenly among the machines in the cluster, and then loaded into memory.

Unload a Table

To unload a table:

- 1 Navigate on the In-Memory Data page to the table to unload.
- 2 Select the table and click .
- 3 On the General page, specify values for the **Job name**, **Location**, and **Description** if you intend to save this action as a job.
- 4 Select the Settings page and specify the following parameters:

Field	Description
Configuration directory	This field specifies the location of the server description files.
Server description file	This field specifies the name of the server description file.
Table	The table information is populated automatically.

- 5 Click **Submit**.

7

Server Management

<i>Features in Server Management</i>	55
<i>SAS LASR Analytic Server Concepts</i>	56
Server Process Run Time	56
Server Description Files	56
<i>Best Practices for Starting Servers and Loading Tables</i>	57
<i>What Does the In-Memory Data Tab Show?</i>	58
<i>Managing Server Description Files</i>	60
Expected Behavior	60
Orphan Server Description Files	60
Servers without a Server Description File	61
<i>Start a Server</i>	61
<i>Stop a Server</i>	63

Features in Server Management

This chapter addresses the use of the SAS Visual Analytics data preparation interface to manage the SAS LASR Analytic Server. The server management functions include the following:

- starting and stopping SAS LASR Analytic Server instances
- listing the tables that are loaded into memory on a server instance

- managing server description files for server instances that stopped due to inactivity

SAS LASR Analytic Server Concepts

Server Process Run Time

When a SAS LASR Analytic Server is started with the data preparation interface, the server is subject to a maximum run time that was set when the SAS software was installed on the machines in the cluster. A typical value is 7200 seconds (2 hours). The maximum run-time setting acts as an inactivity time out. Whenever a SAS client initiates an action with the server, the run time is reset.

Server Description Files

When a server is started, a server description file is written to disk. The file describes the SAS LASR Analytic Server instance that was started. This file is not intended to be read by any application other than SAS clients. The SAS client applications read the file to determine how to connect to the server instance. These files are located in a specific directory so that the SAS clients can coordinate management of server instances being started and stopped.

When a server is stopped with the data preparation interface, the server description file is removed as the server instance is stopped. If a server instance stops because it reached its maximum run time, the server description file is not removed. When a server description file is not removed, it is identified on the **In-Memory Data** tab of the data preparation interface as an orphaned file. Manage these orphaned files by either deleting them from the **In-Memory Data** tab, or using them to restart server instances.

Best Practices for Starting Servers and Loading Tables

The following items summarize some best practices for using the data preparation interface and SAS LASR Analytic Server in your environment:

- Start as few server instances as possible. You can load many tables into a server instance.
- Start additional servers when you need to limit access to data that is loaded into memory. Start an additional server instance and load the data for the group of users that are authorized to view the data. Information about implementing group-level security is provided in the *SAS Visual Analytics: Administration Guide*.
- Because a server description file represents a server instance, use descriptive filenames for them. For example, if data analysis is being performed by a team in a marketing organization and the group wants to limit access to the data, then start a separate server instance for the team and use a server description filename such as marketingTeam1.
- Whenever a group, such as the marketing team, needs additional data for analysis, use the same server description file so that you load it into the same server instance. If the server instance must be restarted for the group, use the same server description file.
- Once explorations or reports are created using a particular data source, do not change the name of the server description file or the data source. If either of these are changed, then the exploration or report will not work.

What Does the In-Memory Data Tab Show?

The **In-Memory Data** tab is the part of the data preparation interface that is used for starting and stopping SAS LASR Analytic Server instances. It is also used for loading and unloading tables to those server instances.

The following display shows an example of the **In-Memory Data** tab.

Display 7.1 Data Preparation In-Memory Data Tab

Note: If the **In-Memory Data** tab does not show any server instances and reports an error, then contact your system administrator to make sure the SAS LASR Analytic Server Monitor is running.

You can click to have the table display the latest information that is available from the SAS LASR Analytic Server Monitor. You can select the **Show tables only** check box to reorganize the table by table name rather than user ID and server instance.

The following list identifies the fields shown in the table:

Tables by Owner

This column lists the unique user IDs that have started server instances. You can expand the node to view the individual server instances started by the user ID and the tables that are loaded into memory on the server instance.

Status

The status column is used to provide a graphical representation of the status for a server instance.

The following icons are used in the status column:

- ● indicates that the server instance is running and ready for use.
- ■ indicates that the server instance is stopped. These servers were not stopped manually, but ran until they reach the maximum run time. These items are grouped under the **orphan** user ID to represent that the server instance is not running. For more information, see [“Orphan Server Description Files” on page 60](#).
- ♦ indicates that the server instance did not respond to network communication for 20 seconds.

Server Description File

For servers that are in a ready state, this field shows the server description file name. For servers that are listed under the **orphan** user ID, this field lists the fully qualified path to the server description file.

Host

This field specifies the host name for the machine that is used as the SAS LASR Analytic Server root node.

Port

This field specifies the network port that is used for communication with the server instance.

Description

This field specifies the table description. This description is visible in the explorer interface when users select a table to explore. Some tables do not have a description because either a description was not stored with the data set, or the data source does not support associating a description with a table.

Rows

This field specifies the number of rows in the table.

Columns

This field specifies the number of columns in the table.

Start Time

For server instances, this field specifies the date and time that the server instance was started. For tables, this field specifies the date and time that the table was loaded into memory.

Last Access

This field specifies the date and time that the server instance was accessed.

Managing Server Description Files

Expected Behavior

In most cases, there is no action to take for managing server description files. When you start a server instance with the data preparation interface, you specify the server description filename. The file is always written to a single configuration directory that is used for server description files.

When you stop a server instance, the software determines the server description file in the configuration directory that is associated with the server instance. The software deletes the server description file as the server is stopped.

This pattern of starting and stopping server instances with the data preparation interface results in server description files being written and deleted from the configuration directory in a controlled manner.

Orphan Server Description Files

If a server instance is not stopped from the data preparation interface or with a scheduled job, the server can reach its maximum run time. When this happens, the server instance exits but does not delete the server description file as it stops. This

results in server description files being left in the configuration directory. If any of these server description files are found, **In-Memory Data** tab shows them under the **orphan** user ID.

An orphaned server description files slows down SAS Visual Analytics clients because the file is read every time a client attempts view the in-memory tables. Each client attempts to connect to the server instance that is described in the file. A best practice is to delete the orphan server description files.

You can select an orphan server description file from the table and then click . You must have UNIX file system permission to remove the file.

Servers without a Server Description File

For an example of a server without a server description file, see the second server instance (the fifth row) in [Display 7.1](#).

Server instances that do not show a server description file in the **In-Memory Data** tab are unreachable from the data preparation interface. Do not attempt to stop the server, load tables, or unload tables with the server instance. These operations fail when the server description file is not available.

A server without a server description file is typically started with a different SAS application or a different version of the data preparation software. Determine how the server instance was started and manage the server with that SAS client.

Start a Server

To start a server:

- 1 Click the **In-Memory Data** button and then click .
- 2 On the **General** page, specify values for the **Job name**, **Location**, and **Description** if you intend to save this action as a job.
- 3 Click **Settings** and specify the SAS LASR Analytic Server parameters:

Field	Description
Configuration directory	This field specifies the path to the server description files.
Server description file	Use the menu to select a server instance to start again or to specify the name of the server description file. Each server description file represents a server instance. The server description file is created in the directory identified in the Configuration directory field.
Override if server description file exists	Select this check box to force starting a server instance. For more information about this option, see “Server Description Files” on page 56 .
Security root directory	This field specifies the root of the security path on the machines in the cluster that the server uses for signature files.
Group security directory	For deployments that implement group-level security by creating directories that are relative to the Security root directory value, specify the relative path to the group security directory.
Number of machines to use	Use the default value ALL unless directed to change the value by SAS technical support.

4 Click Submit.

The following items provide additional information about starting a server instance:

- If a server instance is running and you attempt to start a second server instance using the same server description file, the second server instance is not started. An error message is provided to let you know that a server instance might be running. If you select the **Override if server description file exists** check box and retry, an attempt is made to stop the server instance that is described in the file, and then the second server instance is started.

- If a server instance stops because it reached its maximum run time, then the server description file for the server is not removed. In this case, attempting to start the server again, but not selecting the **Override if server description file exists** check box, results in an error.

Stop a Server

To stop a server:

- 1 Click the **In-Memory Data** button, select the server instance to stop, and click ■.
- 2 On the **General** page, specify values for the **Job name**, **Location**, and **Description** if you intend to save this action as a job.
- 3 Click **Settings** and review the value in the **Server description file** field. If the data preparation software can determine which server description file matches the server instance, the field is static. If the server description file cannot be automatically determined, a menu is provided.
- 4 Click **Submit**.

The following items provide additional information about stopping a server instance:

- When a server is stopped, the server description file is removed. The server description filename remains listed in the menu, so it can be started again without retyping the filename.
- The tables that were loaded in memory on the server are no longer available for selection in the SAS Visual Analytics explorer interface.

Part 3

Exploring Data

<i>Chapter 8</i>	
<i>Overview of SAS Visual Analytics Explorer</i>	67
<i>Chapter 9</i>	
<i>Opening a Data Source</i>	73
<i>Chapter 10</i>	
<i>Managing Data</i>	79
<i>Chapter 11</i>	
<i>Working with Visualizations</i>	87
<i>Chapter 12</i>	
<i>Working with Visualization Filters</i>	127

<i>Chapter 13</i>	
<i>Exporting Data</i>	135
<i>Chapter 14</i>	
<i>Managing Hierarchies</i>	139
<i>Chapter 15</i>	
<i>Managing Visual Explorations</i>	143
<i>Chapter 16</i>	
<i>Performing Data Analysis</i>	145
<i>Chapter 17</i>	
<i>Specifying Your Preferences</i>	149

8

Overview of SAS Visual Analytics Explorer

<i>What Is the SAS Visual Analytics Explorer Interface?</i>	67
<i>Accessing the SAS Visual Analytics Explorer Interface</i>	68
<i>The Welcome Window</i>	68
<i>Overview of SAS Visual Analytics Explorer Interface</i>	69

What Is the SAS Visual Analytics Explorer Interface?

The SAS Visual Analytics explorer interface is a component of SAS Visual Analytics that enables you to explore your data sources. You can explore your data by using interactive visualizations such as charts, histograms, and tables.

You can use the explorer interface to open saved *visual exploration* objects or it can explore in-memory tables from the SAS LASR Analytic Server.

You can export your exploration results as reports, which can be refined in the designer interface or viewed directly. Reports can be viewed on a mobile device or in the SAS Visual Analytics Web viewer.

You can also share your saved visual explorations via e-mail and export your visualizations as image files.

Accessing the SAS Visual Analytics Explorer Interface

You can access the SAS Visual Analytics explorer interface by using any of the following methods:

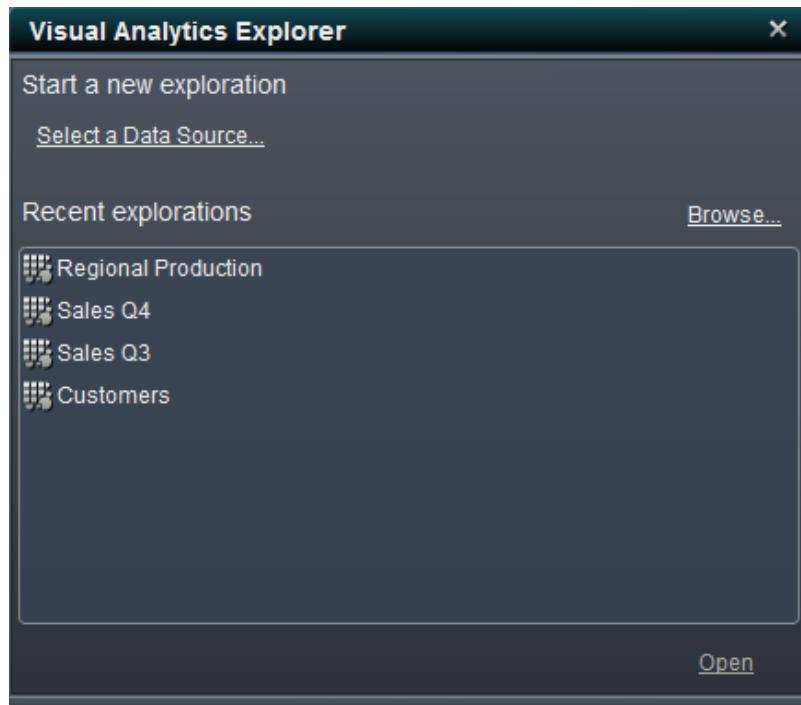
- On the Visual Analytics Home page, click **Explore Data** to open the explorer interface.
You can select a data source by using the [Welcome window](#).
- On the Visual Analytics Home page, double-click a visual exploration to open that visual exploration in the explorer interface.
- Follow a Web link or enter the URL to access the explorer interface directly. For example, you might enter
`http://server.abc.com/SASVisualAnalyticsExplorer`.

After you have logged in to SAS Visual Analytics, you can choose a data source from the [Welcome window](#).

- From an e-mail, follow a link directly to a specific visual exploration. After you have logged in to SAS Visual Analytics, the visual exploration opens in the explorer interface automatically.

The Welcome Window

If you enter the Visual Analytics explorer interface without a reference to a specific visual exploration, then the Welcome window appears.



The Welcome window enables you perform the following tasks:

- Create a new visual exploration by clicking **Select a data source**. See “[Opening a Data Source](#)” on page [73](#) for more information about creating a new visual exploration.
- Open a recently used visual exploration by double-clicking the visual exploration name in the **Recent explorations** list.
- Select an existing visual exploration by clicking **Browse**.

Overview of SAS Visual Analytics Explorer Interface

The SAS Visual Analytics explorer interface contains the following elements:



- 1 The home bar enables you to return to the SAS Visual Analytics Home page and to select from your recent visual explorations.
- 2 The menu bar contains menus that enable you to perform tasks.
- 3 The toolbar contains icons that enable you to manage your visual explorations and visualizations.
- 4 The data pane enables you to manage the data that is used in your visualizations.
- 5 The data properties table enables you to manage the properties of the data item that is selected in the data pane
- 6 The workspace displays your visualizations.

- 7** The right pane contains tabs that enable you to change the properties of your visualizations, subset the data that is currently being used, and set the data roles for your visualizations.
- 8** The dock pane contains any visualizations that you have minimized and enables you to manage your visualizations.

9

Opening a Data Source

<i>Overview of Opening a Data Source</i>	73
<i>Select a Data Source</i>	73
<i>Select Columns</i>	74
<i>Reduce Your Data</i>	74

Overview of Opening a Data Source

When you open a data source, you can perform the following tasks:

- “Select a Data Source”
- “Select Columns”
- “Reduce Your Data”

After you have opened a data source, you can specify properties for the data items such as data formats and default aggregation types. See “[Managing Data Properties](#)” on [page 80](#).

Select a Data Source

In step 1 of the Select Data wizard, select the table that you want to import.

Click **Next** to proceed to the next step of the wizard, or select **Finish** to accept the default import settings for the remaining steps.

Select Columns

In step 2 of the Select Data wizard, select the columns that you want to import from the source data.

To select or deselect all of the columns, select the **All** check box.

To select or deselect individual columns, select the check box beside the column name.

A summary at the bottom of the window displays the number of columns that will be imported.

Click **Next** to proceed to the next page of the wizard, or click **Finish** to accept the default import settings for the remaining steps.

Reduce Your Data

In step 3 of the Select Data wizard, create filters to reduce the data that you will import.

To create a filter:

- 1 Select the column that you want to filter from the **Choose column** drop-down list. If you want to filter on a column that you excluded in step 2, then select **Include hidden columns**.
- 2 Select the operator for the filter from the **Choose operator** drop-down list.

Table 9.1 Operators for Datetime Data

Is	Specifies that a matching value must match the filter value exactly.
Is not	Specifies that a matching value must not match the filter value exactly.

Ends before	Specifies that a matching value must be before the filter value.
Ends on or before	Specifies that a matching value must be before or equal to the filter value.
Starts after	Specifies that a matching value must be after the filter value.
Starts on or after	Specifies that a matching value must be after or equal to the filter value.

Table 9.2 Operators for Character Data

Is	Specifies that a matching value must match the filter value exactly.
Is not	Specifies that a matching value must not match the filter value exactly.
Begins with	Specifies that a matching value must contain the filter value at the start of the value.
Ends with	Specifies that a matching value must contain the filter value at the end of the value.
Contains	Specifies that a matching value must contain the filter value.
Does not contain	Specifies that a matching value must not contain the filter value.
Is missing	Specifies that a missing value matches the filter.
Is not missing	Specifies that any nonmissing value matches the filter.

Table 9.3 Operators for Numeric Data

Is	Specifies that a matching value must match the filter value exactly.
Is not	Specifies that a matching value must not match the filter value exactly.
Between	Specifies that a matching value must be greater than or equal to the first filter value and less than or equal to the second filter value.

Less than	Specifies that a matching value must be less than the filter value.
Less than or equal	Specifies that a matching value must be less than or equal to the filter value.
Greater than	Specifies that a matching value must be greater than the filter value.
Greater than or equal	Specifies that a matching value must be greater than or equal to the filter value.
Is missing	Specifies that a missing value matches the filter.
Is not missing	Specifies that any nonmissing value matches the filter.

3 Depending on the operator that you specified, select or specify a filter value:

Is, Is not

For character data, select the filter value from a drop-down list. For numeric data, enter the filter value in the text field.

Is missing, Is not missing

No filter value is needed.

(All date operators)

For all date filters, click  and then select a date from the calendar.

(All datetime operators)

For all datetime filters, click  and then select a date and time from the calendar.

(All time operators)

For all time filters, click  and then select a time.

Between

Enter the lower value in the first text field and the higher value in the second text field.

For all other operators, enter the filter value in the text field.

4 Click **Apply** to create the filter.

- 5 (Optional) To delete a filter, click the  icon next to the filter that you want to delete.
- 6 (Optional) To preview your filter results, click **Preview Data**. To close the preview, click **Hide Data**.
- 7 When you are finished managing your filters, click **Finish** to import your data.

10

Managing Data

<i>Refresh Your Data Source</i>	79
<i>Change the Data Source for Your Visual Exploration</i>	80
<i>Managing Data Properties</i>	80
Overview of Managing Data Properties	80
Manage Data Properties by Using the Data Properties Window	80
Manage Data Properties By Using the Data Pane	82
Select a Numeric Format	83
Select a Datetime Format	83
<i>Define a Geography Data Item</i>	84
About Geography Data Items	84
Define a Geography Data Item by Using a Predefined Role	84
Define a Geography with a Custom Role	85

Refresh Your Data Source

At any time, you can refresh your data source to read the latest changes.

To refresh your data, select **Data** ▶ **Refresh Data**.

Your visualizations are automatically updated with the new data.

Change the Data Source for Your Visual Exploration

To change the data source for your visual exploration:

- 1 Select **Data** ▶ **Change Data**.
- 2 From the Change Data window, select the data source that you want to use.

Note: The new data source must contain all of the data items from the current data source.

- 3 Click **Open** to open the new data source.

Managing Data Properties

Overview of Managing Data Properties

At any time after you have loaded a data source for your visual exploration, you can modify the properties of the data items. The changes that you make are stored as part of the visual exploration, but do not affect the original data source.

You can manage the data properties for all of your data items by using the [Data Properties window](#).

You can manage the data properties for individual data items by using the [Data pane](#).

Manage Data Properties by Using the Data Properties Window

To access the Data Properties window, select **Data** ▶ **Data Properties** from the menu bar.

For each of your data items, you can specify the following properties:

Name

specifies the name of the data item

Role

specifies whether the data item is a category, a measure, or a geography data item.

Model Type

specifies the data model for a measure or for a category with datetime data. Select whether the data model is continuous or discrete.

The model type that you specify determines the type of data filters that you can create for the data item.

Format

specifies the data format for the data item.

See “[Select a Numeric Format](#)” on page 83 and “[Select a Datetime Format](#)” on page 83 for details about specifying formats.

Aggregation (measures only)

specifies the method that is used to aggregate values for the measure. Select one of the following aggregation types:

Sum

provides the sum of the values in the measure.

Average

provides the average of the values in the measure.

Maximum

provides the maximum value in the measure.

Minimum

provides the minimum value in the measure.

Manage Data Properties By Using the Data Pane

You can manage data properties for a specific data item by selecting the data item in the Data pane and then interacting with the properties table at the bottom of the Data pane.

For the selected data item, you can specify the following properties:

Name

specifies the name of the data item

Role

specifies whether the data item is a category, a measure, or a geography data item.

Model Type

specifies the data model for a measure or for a category with datetime data. Select whether the data model is continuous or discrete.

The model type that you specify determines the type of data filters that you can create for the data item.

Format

specifies the data format for the data item.

See “[Select a Numeric Format](#)” on page 83 and “[Select a Datetime Format](#)” on page 83 for details about specifying formats.

Aggregation (measures only)

specifies the method that is used to aggregate values for the measure. Select one of the following aggregation types:

Sum

provides the sum of the values in the measure.

Average

provides the average of the values in the measure.

Maximum

provides the maximum value in the measure.

Minimum

provides the minimum value in the measure.

Select a Numeric Format

To select a numeric format:

- 1 In the Data Properties window or the property table in the **Data** pane, click the format for the data item that you want to manage. A window appears.
- 2 Select the basic format type from the **Format Type** list.
- 3 For the **Numeric** and **Float** formats, specify the **Width** parameter for the format. The width of the format specifies the maximum number of characters that a value can contain, including decimal points.

A preview of your format appears in the **Sample** field.

- 4 Specify the number of **Decimals** for the format. The decimals value specifies the number of digits that appear to the right of the decimal place.
- A preview of your format appears in the **Sample** field.

- 5 Click **OK** to apply your format.

Select a Datetime Format

To select a datetime format:

- 1 In the Data Properties window or the property table in the **Data** pane, click the format for the data item that you want to manage. A window appears.
- 2 Select the basic format type from the **Format Type** list.
- 3 If necessary, select a more specific format from the **Format** list. The format names also represent sample values for the format.
- 4 Click **OK** to apply your format.

Define a Geography Data Item

About Geography Data Items

A geography data item is a category whose values are mapped to geographical locations or regions. Geography data items can be used with [geo map visualizations](#) to visualize your data on a geographic map.

For example, if your data source contains a `CountryName` column that identifies countries, then you might create a geography data item for `CountryName` by using the predefined geographic role, **Countries**. See “[Define a Geography Data Item by Using a Predefined Role](#)” on page 84.

You might also create a customized geography data item to identify geographic information that is specific to your organization (for example, sales regions, warehouse locations, oil platforms, and so on). See “[Define a Geography with a Custom Role](#)” on page 85.

Define a Geography Data Item by Using a Predefined Role

To define a geography data item by using a predefined geographic role:

- 1 In the data pane or in the Data Properties window, locate a category that contains geographic information.
- 2 Change the data role for the category to **Geography**. The Geography Role window appears.
- 3 Select the predefined geographic role for your category:

Countries

specifies country names.

US States

specifies the names of states and territories in the United States.

US States Abbreviated

specifies the two-letter postal codes for states and territories in the United States.

US ZIP Codes

specifies five-digit ZIP codes for the United States.

- 4 Click **OK** to apply your changes.

Define a Geography with a Custom Role

To define a geography data item by using a custom geographic role:

- 1 In the data pane or in the Data Properties window, locate a category that contains geographic information.
- 2 Change the data role for the category to **Geography**. The Geography Role window appears.
- 3 Select **Custom**, and then specify the following:

Latitude

specifies a measure from your current data source that contains the latitude (Y) coordinate values for the geographic role that you want to define.

Longitude

specifies a measure from your current data source that contains the longitude (X) coordinate values for the geographic role that you want to define.

Coordinate Space

specifies the coordinate space (coordinate system) that is used to project the longitude and latitude coordinates.

- 4 Click **OK** to apply your changes.

11

Working with Visualizations

<i>Overview of Visualizations</i>	89
About Visualizations	89
Visualization Types	89
<i>Working with Visualizations</i>	95
Create a New Visualization	95
Manage Visualizations	95
Arrange the Visualizations in the Workspace	96
<i>Managing Visualization Data Roles</i>	99
Add a Data Item	99
Replace a Data Item	101
Remove a Data Item	101
Switch Data Roles	101
<i>Working with Filters</i>	101
<i>Working with Automatic Charts</i>	102
<i>Working with Bar Charts</i>	103
About Bar Charts	103
Specify Properties for a Bar Chart	103
Data Roles for a Bar Chart	104
Sort Data Values	105
<i>Working with Line Charts</i>	105
About Line Charts	105
Specify Properties for a Line Chart	106

Data Roles for a Line Chart	107
Sort Data Values	108
Working with Tables	108
About Tables	108
Specify Properties for a Table	108
Managing Columns	109
Working with Crosstabs	109
About Crosstabs	109
Specify Properties for a Crosstab	109
Data Roles for a Crosstab	110
Managing Rows and Columns	110
Working with Scatter Plots	110
About Scatter Plots	110
Specify Properties for a Scatter Plot	111
Data Roles for a Scatter Plot	111
Applying Data Analysis	111
Working with Bubble Plots	112
About Bubble Plots	112
Specify Properties for a Bubble Plot	113
Data Roles for a Bubble Plot	113
Using Animated Bubble Plots	114
Working with Histograms	115
About Histograms	115
Specify Properties for a Histogram	116
Data Roles for a Histogram	117
Working with Box Plots	117
About Box Plots	117
Specify Properties for a Box Plot	118
Data Roles for a Box Plot	119
Working with Heat Maps	120
About Heat Maps	120
Specify Properties for a Heat Map	120

Data Roles for a Heat Map	121
Applying Data Analysis	121
<i>Working with Geo Maps</i>	122
About Geo Maps	122
Specify Properties for a Geo Map	122
Data Roles for a Geo Map	123
Zoom a Geo Map	123
Pan (Scroll) a Geo Map	124
<i>Working with Correlation Matrices</i>	124
About Correlation Matrices	124
Specify Properties for a Correlation Matrix	124
Data Roles for a Correlation Matrix	125

Overview of Visualizations

About Visualizations

The SAS Visual Analytics explorer interface displays data by using visualizations. A visualization is an interactive visual representation of your data. A visualization can be a table, a crosstab, a chart, a histogram, or a geographic map.

Visualization Types

You can assign any of the following types to your visualizations:

Autochart



Automatically selects the chart type according to the data that is assigned to the visualization. When you are first exploring a new data set, autocharts are useful to give you a quick view of the data.

For more information, see [“Working with Automatic Charts” on page 102](#).

Table

Origin	Type	Horsepower	MPG (Highway)
Asia	SUV	265	23
Asia	Sedan	104	32
Europe	Wagon	170	31
Asia	Sedan	200	31
Asia	Sedan	124	32
Europe	Wagon	270	25
Asia	Sedan	200	29
Asia	Sedan	124	32
Europe	SUV	268	20
Asia	Sedan	270	28
Asia	Sedan	124	32
Europe	Sedan	170	29
Asia	Sedan	225	24

Displays the data as a table. Tables enable you to examine the raw data for each observation in the data source. You can rearrange the data columns and apply sorting.

For more information, see [“Working with Tables” on page 108](#).

Crosstab

Visualization 1

Options ▾

Origin		Europe		
Type	DriveTrain	Invoice	Length (IN)	MPG (City)
Sports	All	\$142,836	493	
	Front	\$108,929	350	
	Rear	\$1,249,562	3142	
Sedan	All	\$699,377	3328	
	Front	\$949,671	5379	
	Rear	\$1,447,264	5619	
Truck	All	.	.	
	Rear	.	.	
SUV	All	\$442,913	1852	
	Front	.	.	
Wagon	All	\$201,202	923	
	Front	\$140,201	912	
	Rear	.	.	

Displays the data as a crosstab. Crosstabs enable you to examine the data for intersections of category values. You can rearrange the rows and columns and apply sorting.

For more information, see [“Working with Crosstabs” on page 109](#).

Bar Chart

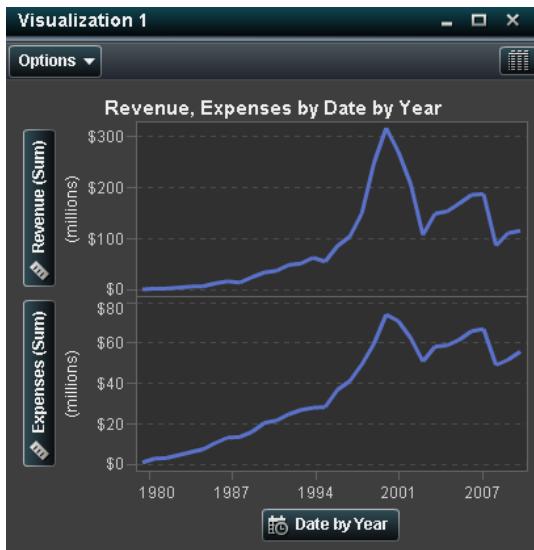


Displays the data as a bar chart. Bar charts are especially useful for comparing data according to each discrete value of a category.

You can configure a bar chart with either vertical bars or horizontal bars. You can also assign grouping to the bars and create lattices.

For more information, see [“Working with Bar Charts” on page 103](#).

Line Chart

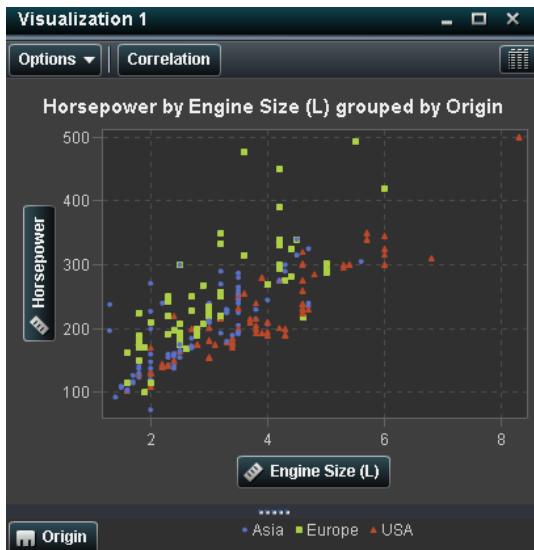


Displays the data as a line chart. A line chart is most useful for data trends over time.

You can apply grouping, and create lattices.

For more information, see [“Working with Line Charts” on page 105](#).

Scatter Plot



Displays the data as a scatter plot. Scatter plots are most useful to examine the relationship between variables.

In a scatter plot, you can apply statistical analysis with correlation and regression. Scatter plots also support grouping.

When you apply more than two measures to a scatter plot, the visualization automatically displays a scatter plot matrix to compare each pairing of measures.

For more information, see [“Working with Scatter Plots” on page 110](#).

Bubble Plot

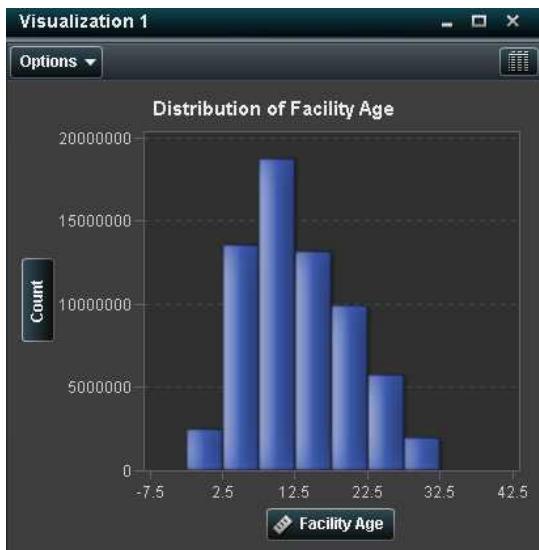


Displays the data as a bubble plot. A bubble plot displays the relationship between three measures, where two measures are represented by the plot axes and the third measure is represented by the size of the plot markers.

You can apply grouping and create lattices for a bubble plot. By assigning a datetime data item to the plot, you can animate the bubbles to display changes in the data over time.

For more information, see [“Working with Bubble Plots” on page 112](#).

Histogram

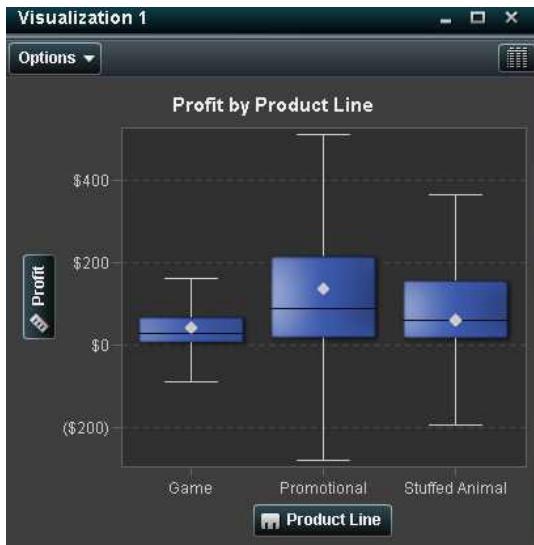


Displays the data as a histogram. A histogram displays the distribution of values for a single measure.

You can select the bar orientation, and select whether the distribution values are displayed as a percentage or as the row number of values.

For more information, see [“Working with Histograms” on page 115](#).

Box Plot



Displays the data as a box plot. A box plot displays the distribution of values for a measure by using a box. The size and location of the box indicate the range of values that are between the 25th and 75th percentile. Additional statistical information is represented by other visual features.

You can create lattices, and select whether the average (mean) value and outliers are displayed for each box.

For more information, see [“Working with Box Plots” on page 117](#).

Heat Map



Displays the data as a heat map. A heat map displays the distribution of values for two data items by using a table with colored cells. If you do not assign a measure to the color data role, then the cell colors represent the frequency of each intersection of values. If you assign a measure to the color data role, then the cell colors represent the measure value for each intersection of values.

For more information, see [“Working with Heat Maps” on page 120](#).

Geo Map



Displays the data as a geo map. A geo map displays your data as a bubble plot that is overlaid on a geographic map. Each bubble is located at a geographic location or at the center of a geographical region.

For more information, see [“Working with Geo Maps” on page 122](#).

Working with Visualizations

Create a New Visualization

To create a new visualization, click  on the toolbar, or select **View** ▶ **New Visualization** from the menu bar.

Manage Visualizations

Delete a Visualization

To delete a visualization, click the  button on the visualization or select **View** ▶ **Delete Visualization** to delete the selected visualization.

You can delete all of your visualizations by selecting **View** ▶ **Delete All Visualizations**.

Duplicate a Visualization

To duplicate a visualization, select **View** ▶ **Duplicate Visualization** from the main menu or select **Options** ▶ **Duplicate Visualization** from the visualization toolbar.

The new visualization is named “Copy of *visualization-name*”. You can assign a new name on the **Properties** tab.

Show and Minimize Visualizations

To minimize a visualization, click  for the visualization that you want to minimize. The visualization appears on the dock panel at the bottom of the workspace.

To restore a visualization, click the name of the visualization on the dock pane.

You can also select which visualizations are visible by using the  button on the dock pane. The  button opens a menu that enables you to perform the following tasks:

Show all of the visualizations

 Select **Show all visualizations**.

Minimize all of the visualizations

 Select **Minimize all visualizations**.

Show or minimize individual visualizations

 Select the name of the visualization that you want to show or minimize. A check mark appears beside the visualizations that are currently shown.

Maximize a Visualization

To maximize a visualization to fill the entire workspace, click  for the visualization that you want to maximize. To return the visualization to its normal size, click .

Arrange the Visualizations in the Workspace

Move the Position of a Visualization

To move a visualization, drag the title bar of the visualization to the location where you want to place the visualization.

When you drag a visualization onto another visualization, the placement of the visualization is determined by the edge that is closest to the drop point.

For example: Your workspace contains two visualizations named “A” and “B”. If you drag A to the left edge of B, then A is placed to the left of visualization B. If you drag A to the bottom edge of B, then A is placed below B.

Display 11.1 shows an example layout with four visualizations.

Display 11.1 Example Layout with Four Visualizations



Display 11.2 shows the effect of moving a visualization to the right edge of another visualization.

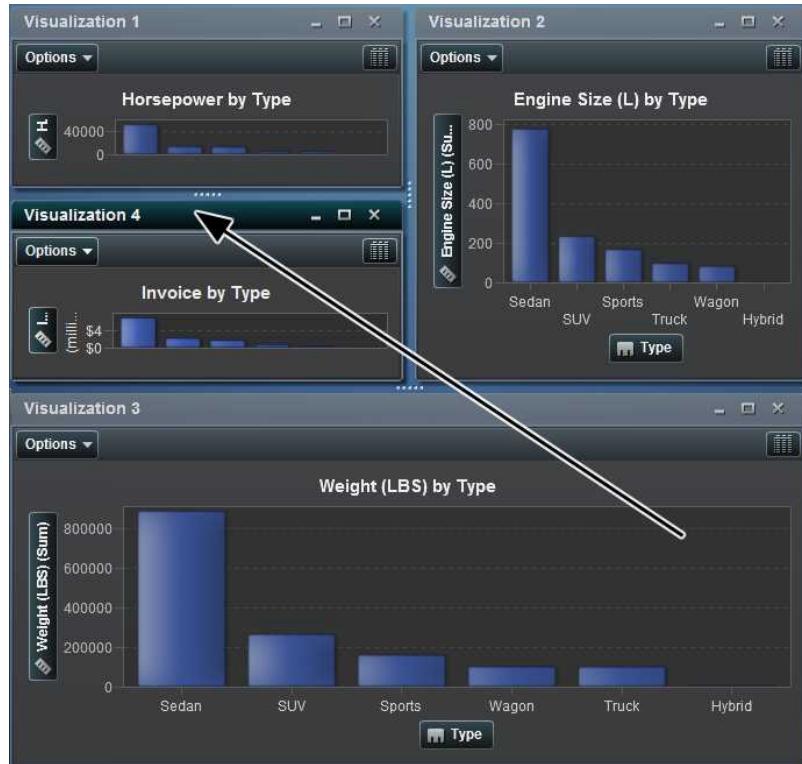
Display 11.2 Example Layout After Moving 4 to the Right Edge of 1



When your workspace contains visualization in multiple rows, if you drag to the bottom or top edge of a visualization then the target visualization and the moved visualization split the space that was previously occupied by the target visualization.

Display 11.3 on page 99 shows the effect of moving a visualization to the bottom edge of another visualization.

Display 11.3 Example Layout After Moving 4 to the Bottom Edge of 1



Resize a Visualization

To resize a visualization in your workspace, drag the resizing tab  at the bottom edge or the right edge of the visualization that you want to resize.

Managing Visualization Data Roles

Add a Data Item

In order for a visualization to display data, you must assign data items to it. You can assign data items to a visualization by performing any of the following actions:

- Drag the data item onto the center of the visualization. The data item is assigned to a role automatically. If the visualization already has data items assigned to the required roles, then you can choose how to assign the new data item.
- Drag the data item onto a **Measure** or **Category** button in the visualization to assign the new data item to a specific data role.
- From the visualization toolbar, select **Options** ▶ **Add Category** or **Options** ▶ **Add Measure**.
- Use the **Roles** tab in the right pane. Either drag a data item to a role, or select a data item from the drop-down list.

Each visualization requires a minimum number of each type of data item. The following table lists the requirements for each visualization.

Table 11.1 Required Data Items for Visualizations

Visualization Type	Requirements
Autochart	one data item of any type
Table	one measure or category
Crosstab	one category
Bar Chart	one category or hierarchy
Line Chart	one measure and one category or hierarchy
Scatter Plot	two measures
Bubble Plot	three measures
Histogram	one measure
Box Plot	one measure and one category
Heat Map	two data items (measures, categories, or hierarchies)
Geo Map	one geography

Replace a Data Item

You can replace a data item by using any of the following methods:

- Drag the new data item from the **Data Items** pane to the data item in the visualization that you want to replace.
- Right-click the data item that you want to replace from either the visualization or the **Roles** tab and then select **Replace *item-name* ▶ *new-item-name***.

Remove a Data Item

You can remove a data item by using any of the following methods:

- Drag the data item from the visualization to the **Data Items** pane.
- Select **Options ▶ Remove ▶ *item-name*** from the visualization toolbar.
- Right-click the data item in the visualization or in the **Roles** tab and then select **Remove *item-name***.

Switch Data Roles

To switch the data items that are assigned to two roles, drag one item onto another either by using the buttons in the visualization or by using the fields on the **Roles** tab.

For example, you can switch the axes of a scatter plot by dragging the measure on the X axis to the measure on the Y axis.

Working with Filters

You can use the **Filters** tab to subset the data in your visualizations. See Chapter 12, “Working with Visualization Filters,” on page 127.

Working with Automatic Charts

Automatic charts (autocharts) display different types of charts automatically based on the number of category and measure data items that are assigned to the chart.

The type of chart can be any of the following:

Table 11.2 Automatic Chart Types

Data Items	Chart Type
One measure or category	Bar chart
One datetime category and any number of other categories or measures	Line chart
One geography and zero to two measures	Geo map
Two measures	Scatter plot or heat map
Three or more measures	Scatter plot matrix or correlation matrix
One geography and three or more measures	Bar chart
One or more categories and any number of measures and geographies	Bar chart

Note: The **Roles** tab for an automatic chart always contains **Categories** and **Measures**. To set advanced data roles such as grouping and lattices, change the visualization type to select the specific chart type that you want. For example, if your automatic chart shows a bar chart, then change the type to **Bar Chart** to enable advanced data role settings.

Working with Bar Charts

About Bar Charts

A bar chart represents the values of category data by using bars. The height of each bar represents the value.

Specify Properties for a Bar Chart

On the **Properties** tab, you can specify the following options:

Name

specifies the name of the visualization.

Graph title

specifies the title that appears above the graph.

Note: The **Graph title** is disabled if you select **Generate graph titles**.

Generate graph titles

specifies that the graph title is generated automatically based on the data items in the visualization.

Show grid lines

displays grid lines for each tick on the response axis.

Bar direction

specifies whether the bars are vertical or horizontal.

Grouping style

specifies how grouped data are displayed. If you select **Stacks**, then the values of the grouping variable are displayed as segments of each bar. If you select **Colors**, then each value of the grouping variable is displayed as a separate bar.

Note: This option is disabled if no data item is assigned to the grouping role.

Measure layout

specifies whether the measures share a single response axis (**Shared axis**) or have separate response axes for each measure (**Separate axes**).

Note: The **Measure layout** option is disabled if your visualization contains exactly one measure.

Overview

specifies whether the chart overview is enabled.

Show data values

shows the data values as text in the visualization.

Note: You can always view the data values as data tips when you move the cursor over the graph.

Category label policy

specifies how the values for the category axis are displayed when there is not enough space. Select one of the following options:

Automatic

specifies that the visualization will select the best label policy according to your data and the space requirements.

Stagger

places the labels with an alternating vertical offset.

Thin

displays a subset of the labels.

Rotate

displays the labels at an angle

Note: The **Category label policy** option has no effect when you select a horizontal bar orientation.

Data Roles for a Bar Chart

The basic data roles for a bar chart are categories and measures. You can assign one category only, and the values of the category are plotted on the category axis. You can

assign many measures, and the measure values are plotted on the response axis. If a bar chart contains no measures, then the frequency of the category values is plotted on the response axis.

In addition to the basic data roles, you can assign additional roles:

Group

groups the data according to the values of the category data item that you assign.

Depending on the value that you selected for the **Grouping style** property, the group values are either shown as individual bars or as segments of each bar.

Note: Grouping is not available if you assign multiple measures to the visualization.

Lattice columns

creates a lattice of charts, with a column for each value of the category data item that you assign.

Lattice rows

creates a lattice of charts, with a row for each value of the category data item that you assign.

Sort Data Values

By default, a bar chart is sorted alphabetically by the values of the category data item. To change the sorting, right-click the data item that you want to sort on, and then select **Sort ▶ [sort-method]**.

Working with Line Charts

About Line Charts

A line chart represents the values of category data by using a line that connects the data values. If you assign multiple measures to a line plot, then you can create separate Y axes for each measure.

Specify Properties for a Line Chart

On the **Properties** tab, you can specify the following options:

Name

specifies the name of the visualization.

Graph title

specifies the title that appears above the graph.

Note: The **Graph title** is disabled if you select **Generate graph titles**.

Generate graph titles

specifies that the graph title is generated automatically based on the data items in the visualization.

Show grid lines

displays grid lines for each tick on the response axis.

Overview

specifies whether the chart overview is enabled.

Measure layout

specifies whether the measures share a single response axis (**Shared axis**) or have separate response axes for each measure (**Separate axes**).

Note: The **Measure layout** option is disabled if your visualization contains exactly one measure.

Show markers

shows markers for the data points in the visualization.

Show data values

shows the data values as text in the visualization.

Note: You can always view the data values as data tips when you move the cursor over the graph.

Category label policy

specifies how the values for the category axis are displayed when there is not enough space. Select one of the following options:

Automatic

specifies that the visualization will select the best label policy according to your data and the space requirements.

Stagger

places the labels with an alternating vertical offset.

Thin

displays a subset of the labels.

Rotate

displays the labels at an angle

Data Roles for a Line Chart

The basic data roles for a line chart are categories and measures. You can assign one category only, and the values of the category are plotted on the category axis. You can assign many measures, and the measure values are plotted on the response axis.

In addition to the basic data roles, you can assign additional roles:

Group

groups the data according to the values of the category data item that you assign. A separate line is created for each value.

Note: Grouping is not available if you assign multiple measures to the visualization.

Lattice columns

creates a lattice of charts, with a column for each value of the category data item that you assign.

Lattice rows

creates a lattice of charts, with a row for each value of the category data item that you assign.

Sort Data Values

By default, a line chart is sorted alphabetically by the values of the category data item. To change the sorting, right-click the data item that you want to sort on, and then select **Sort ▶ [sort-method]**.

Note: Sorting is not available if a datetime data item is assigned to the **Category** role.

Working with Tables

About Tables

A table shows data values as text. The values for each measure or category that is assigned to the table are displayed as columns. The data values in the table are not aggregated.

Note: For very large data sources, the table visualization displays only the first two billion (2,147,483,647) rows.

Note: If you sort the table, then the table displays only the first 5,000 sorted rows.

Specify Properties for a Table

On the **Properties** tab, you can specify the following option:

Name

specifies the name of the visualization.

Managing Columns

Sort Columns

To sort the table by a column, click the column heading. An arrow appears in the column heading to indicate the sorting. If the arrow points upward, the sort is ascending. If the arrow points downward, then the sort is descending.

Note: If you sort the table, then the table displays only the first 5,000 sorted rows.

Rearrange Columns

To rearrange your columns, drag and drop the column headings.

Resize Columns

To resize your columns, click and drag the left or right edge of the column headings.

Working with Crosstabs

About Crosstabs

A crosstab shows the intersections of category values and measure values as text. If your crosstab contains measures, then each cell of the crosstab contains the aggregated measure values for a specific intersection of category values. If the crosstab does not contain measures, then each cell contains the frequency of an intersection of category values.

Specify Properties for a Crosstab

On the **Properties** tab, you can specify the following option:

Name

specifies the name of the visualization.

Data Roles for a Crosstab

The data roles for a crosstab are columns, rows, and measures. You can assign any number of categories to the columns and rows roles. If you assign measures to the crosstab, then the measure values are displayed in the cells of the crosstab. If you do not assign measures, then the cells of the crosstab show the frequency of each intersection of values.

Managing Rows and Columns

Sort a Row or Column

By default, the crosstab is sorted alphabetically by the values of the first category that you assign to the **Rows** role. To change the sorting, right-click the heading for the row or column that you want to sort, and then select **Sort ▶ [sort-method]**.

Rearrange Rows and Columns

To rearrange your rows and columns, drag and drop the column headings.

Resize Columns

To resize your columns, click and drag the left or right edge of the column headings.

Working with Scatter Plots

About Scatter Plots

A scatter plot represents the values of measures by using markers. When you assign more than two measures, the visualization displays a scatter plot matrix. A scatter plot matrix is a series of scatter plots that displays every possible pairing of the measures that are assigned to the visualization.

If you create a scatter plot that has a very large number of points, then the plot is rendered either as a [heat map](#) (for two measures) or as a correlation matrix for three or more measures.

Specify Properties for a Scatter Plot

On the **Properties** tab, you can specify the following options:

Name

specifies the name of the visualization.

Graph title

specifies the title that appears above the graph.

Note: The **Graph title** is disabled if you select **Generate graph titles**.

Generate graph titles

specifies that the graph title is generated automatically based on the data items in the visualization.

Show grid lines

displays grid lines for each tick on the plot axes.

Marker size

specifies the size of each marker in pixels.

Note: Sizes below 4 pixels are not supported when your scatter plot is grouped.

Data Roles for a Scatter Plot

The basic data role for a scatter plot is the measures. You can assign any number of measures.

In addition to the measures, you can assign a **Group** variable. The group variable groups the data according to the values of the category data item that you assign. A separate set of scatter points is created for each value of the group variable.

Applying Data Analysis

About Data Analysis

For scatter plots, you can apply the following data analyses:

Correlation

identifies the degree of statistical correlation between the variables in the scatter plot.

Regression

plots a model of the relationship between the variables in the scatter plot.

Both correlation and regression analysis are applied together when you enable correlation analysis.

Enable Data Analysis

To apply data analysis, select **Options** ▶ **Show Correlation** or click **Correlation** from the visualization toolbar.

For a single scatter plot, the degree of correlation is indicated by a text label in the legend below the plot. To see additional information about the correlation results, click .

For a grouped scatter plot, the degree of correlation for each plot is indicated by the color of the border around the plot. To see additional information about the correlation results for all of the plots, click .

By default, a regression fit is applied to your plot when you enable correlation analysis. The regression fit appears as a red line. To see additional information about the regression results, click  in the legend.

Working with Bubble Plots

About Bubble Plots

A bubble plot represents the values of three measures by using differently sized markers (bubbles) in a scatter plot. The values of two measures are represented by the position on the graph axes, and the value of the third measure is represented by the marker size.

You can also create animated bubble plots to display changing data over time.

Specify Properties for a Bubble Plot

On the **Properties** tab, you can specify the following options:

Name

specifies the name of the visualization.

Graph title

specifies the title that appears above the graph.

Note: The **Graph title** is disabled if you select **Generate graph titles**.

Generate graph titles

specifies that the graph title is generated automatically based on the data items in the visualization.

Show grid lines

displays grid lines for each tick on the response axis.

Data Roles for a Bubble Plot

The basic data roles for a bubble plot are:

X Axis

specifies the measure that is assigned to the X axis.

Y Axis

specifies the measure that is assigned to the Y axis.

Size

specifies the measure that determines the marker size.

In addition to the basic data roles, you can assign additional roles:

Group

groups the data according to the values of the category data item that you assign. A separate set of points is created for each value.

Note: You cannot assign both the **Group** and the **Color** roles as the same time.

Color

specifies a data item that determines the color of the bubbles. If you specify a category, then each value of the category is represented by a different bubble color. If you specify a measure, then the bubble color represents the measure value.

Note: You cannot assign both the **Group** and the **Color** roles as the same time.

Lattice columns

creates a lattice of charts, with a column for each value of the category data item that you assign.

Lattice rows

creates a lattice of charts, with a row for each value of the category data item that you assign.

Animation

specifies a datetime data item that is used to animate the bubble plot.

Note: The **Animation** role is enabled only if you set assign a data item to the **Group** role.

Using Animated Bubble Plots

About Animated Bubble Plots

An animated bubble plot displays the changes in your data values over time. Each frame of the animation represents a value of the datetime data item that is assigned to the **Animation** data role.

For example, if you assign a category with the YEAR format to the **Animation** data role, then each frame of the animation displays a bubble plot of your data for a specific year.

Create an Animated Bubble Plot

To create an animated bubble plot:

- 1 Select an existing bubble plot, or create a new bubble plot.
- 2 Assign a data item to the **Group** data role.

3 Assign a data item with a datetime format to the **Animation** data role.

Display an Animated Bubble Plot

For an animated bubble plot, a set of animation controls appears at the bottom of the visualization.

Table 11.3 Animation Control Tasks

Task	Action
Start the animation	Click ►.
Go to the previous animation frame	Click ◀.
Go to the next animation frame	Click ▶.
Jump to a specific animation frame	Use the slider.
Specify whether to repeat the animation	Select or deselect Loop .
Select the animation speed	Use the Speed slider.
Track the movement of a specific bubble	Click the bubble that you want to track.

Working with Histograms

About Histograms

A histogram represents the distribution of values for a single measure. A series of bars represents the number of observations in the measure that match a specific value or value range. The bar height can represent either the exact number of observations or the percentage of all observations for each value range.

Note: The minimum and maximum values on the histogram axis might not match the actual extent of your data values.

Specify Properties for a Histogram

On the **Properties** tab, you can specify the following options:

Name

specifies the name of the visualization.

Graph title

specifies the title that appears above the graph.

Note: The **Graph title** is disabled if you select **Generate graph titles**.

Generate graph titles

specifies that the graph title is generated automatically based on the data items in the visualization.

Show grid lines

displays grid lines for each tick on the response axis.

Bar direction

specifies whether the bars are vertical or horizontal.

Scale

specifies the statistic that is used on the response axis. Select **Count** to show the data frequency, or select **Percent** to show a percentage.

Use default bin count

Specifies whether to use the default number of bins (value ranges) for the histogram. The default number of bins is determined by the number of data values in your histogram.

Bin count

Specifies the number of bins (value ranges) for the histogram.

Note: In order to create bins at a useful interval, the histogram might create additional bins beyond the number that you specify.

Data Roles for a Histogram

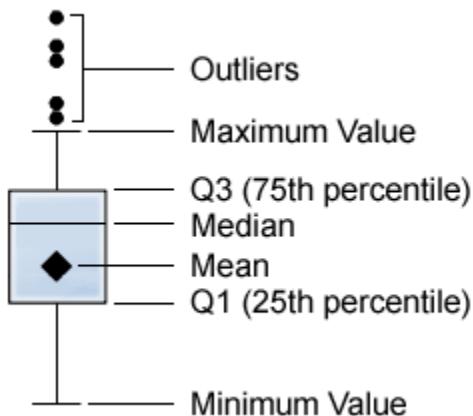
The basic data role for a histogram is a measure. You can assign only one measure to a histogram.

Working with Box Plots

About Box Plots

A box plot represents the distribution of data values by using rectangular box and lines called whiskers.

Figure 11.1 Parts of a Box Plot



[Figure 11.1 on page 117](#) shows a diagram of a box plot. The bottom and top edges of the box indicate the interquartile range (IQR). That is, the range of values that are between the first and third quartiles (the 25th and 75th percentiles). The marker inside the box indicates the mean value. The line inside the box indicates the median value.

The whiskers (lines protruding from the box) indicate the range of values that are outside of the interquartile range but are close enough not to be considered outliers.

Outliers are data points whose distance from the interquartile range is greater than 1.5 times the size of the interquartile range.

Note: If there are a large number of outlier points, then the outliers are represented by a bar. The data tip for the bar displays additional information about the outlier points.

Specify Properties for a Box Plot

On the **Properties** tab, you can specify the following options:

Name

specifies the name of the visualization.

Graph title

specifies the title that appears above the graph.

Note: The **Graph title** is disabled if you select **Generate graph titles**.

Generate graph titles

specifies that the graph title is generated automatically based on the data items in the visualization.

Show grid lines

displays grid lines for each tick on the response axis.

Box direction

specifies whether the boxes are vertical or horizontal.

Overview

specifies whether the chart overview is enabled.

Measure layout

specifies whether the measures share a single response axis (**Shared axis**) or have separate response axes for each measure (**Separate axes**).

Note: The **Measure layout** option is disabled if your visualization contains exactly one measure.

Show averages

displays the mean value as a marker inside the box.

Show outliers

shows outlier points outside of the whiskers. Outliers are data points whose distance from the interquartile range is greater than 1.5 times the size of the interquartile range.

Category label policy

specifies how the values for the category axis are displayed when there is not enough space. Select one of the following options:

Automatic

specifies that the visualization will select the best label policy according to your data and the space requirements.

Stagger

places the labels with an alternating vertical offset.

Thin

displays a subset of the labels.

Rotate

displays the labels at an angle

Note: The **Category label policy** option has no effect when you select a horizontal box orientation.

Data Roles for a Box Plot

The basic data roles for a box plot are categories and measures. You can assign one category only, and the values of the category are plotted on the category axis. You can assign many measures, and the measure values are plotted on the response axis. At least one measure is required.

In addition to the basic data roles, you can assign additional roles:

Lattice columns

creates a lattice of charts, with a column for each value of the category data item that you assign.

Lattice rows

creates a lattice of charts, with a row for each value of the category data item that you assign.

Working with Heat Maps

About Heat Maps

A heat map displays the distribution of values for two data items by using a table with colored cells. If you do not assign a measure to the **Color** data role, then the cell colors represents the frequency of each intersection of values. If you assign a measure to the **Color** data role, then the cell colors represent the measure value for each intersection of values.

Specify Properties for a Heat Map

On the **Properties** tab, you can specify the following options:

Name

specifies the name of the visualization.

Graph title

specifies the title that appears above the graph.

Note: The **Graph title** is disabled if you select **Generate graph titles**.

Generate graph titles

specifies that the graph title is generated automatically based on the data items in the visualization.

Bin count

specifies the number of value ranges that are represented as cells. **Bin count** only affects measures.

Show borders

specifies that the borders between cells are visible.

Category label policy

specifies how the values for the category axis are displayed when there is not enough space. Select one of the following options:

Automatic

specifies that the visualization will select the best label policy according to your data and the space requirements.

Stagger

places the labels with an alternating vertical offset.

Thin

displays a subset of the labels.

Rotate

displays the labels at an angle

Note: The **Category label policy** option has no effect if the heat map contains no categories.

Data Roles for a Heat Map

The basic data roles for a heat map are:

X Axis

specifies the data item that is assigned to the X axis.

Y Axis

specifies the data item that is assigned to the Y axis.

Color

specifies a measure that determines the cell color. If you do not assign the **Color** data role, then the cell color indicates frequency.

Applying Data Analysis

About Data Analysis

For heat maps, you can apply the following data analyses:

Correlation

identifies the degree of statistical correlation between the variables in the heat map.

Regression

plots a model of the relationship between the variables in the heat map.

Both correlation and regression analysis are applied together when you enable correlation analysis.

Enable Data Analysis

To apply data analysis, select **Options** ▶ **Show Correlation** or click **Correlation** from the visualization toolbar.

The degree of correlation is indicated by a text label in the legend below the heat map.

To see additional information about the correlation results, click .

By default, a regression fit is applied to your plot when you enable correlation analysis.

The regression fit appears as a red line. To see additional information about the regression results, click  in the legend.

Working with Geo Maps

About Geo Maps

A geo map displays your data as a bubble plot that is overlaid on a geographic map. Each bubble is located either at the center of geographic region or at the coordinates of a location.

In order to display a geo map, you must define one or more of your categories as a geography data item. For more information, see “[Define a Geography Data Item](#)” on [page 84](#).

Specify Properties for a Geo Map

On the **Properties** tab, you can specify the following options:

Name

specifies the name of the visualization.

Graph title

specifies the title that appears above the graph.

Note: The **Graph title** is disabled if you select **Generate graph titles**.

Generate graph titles

specifies that the graph title is generated automatically based on the data items in the visualization.

Show map controller

specifies whether the map zoom and pan controls are visible.

Note: This option does not affect your ability to zoom and pan by using the mouse or keyboard.

Data Roles for a Geo Map

The basic data roles for a bubble plot are:

Geographic Role

specifies the geography data item that identifies geographic regions for your map.

Geography data items are identified by the  icon.

Size

specifies the measure that determines the marker size.

In addition to the basic data roles, you can assign additional roles:

Color

specifies a measure that determines the bubble color.

Zoom a Geo Map

You can zoom the map by using any of the following controls:

- scroll the mouse scroll wheel

- click the zoom bar to select your zoom level
- press the + and – keys
- double-click and double right-click

Pan (Scroll) a Geo Map

You can pan (scroll) the map by using any of the following controls:

- click and drag the map
- click the arrows on the pan control
- use the keyboard arrow keys

Working with Correlation Matrices

About Correlation Matrices

A correlation matrix displays the degree of correlation between multiple intersections of measures as a matrix of rectangular cells. Each cell in the matrix represents the intersection of two measures, and the color of the cell indicates the degree of correlation between those two measures.

You cannot create a correlation matrix directly. Instead, a correlation matrix is created when you assign high-cardinality data to a scatter plot matrix.

Note: If you reduce your data by using filters, the visualization type might change to a scatter plot matrix.

Specify Properties for a Correlation Matrix

On the **Properties** tab, you can specify the following options:

Name

specifies the name of the visualization.

Graph title

specifies the title that appears above the graph.

Note: The **Graph title** is disabled if you select **Generate graph titles**.

Generate graph titles

specifies that the graph title is generated automatically based on the data items in the visualization.

Show grid lines

displays grid lines for each tick on the plot axes.

Marker size

specifies the size of each marker.

Data Roles for a Correlation Matrix

The basic data role for a correlation matrix is the measures. You can assign any number of measures.

If you reduce the number of measures to two, then the visualization type changes to a heat map.

12

Working with Visualization Filters

<i>About Filters</i>	127
<i>Create a Filter</i>	128
<i>Edit a Filter By Using Visual Filter Controls</i>	129
About Visual Filter Controls	129
Filtering Discrete Data	129
Filtering Continuous Data	130
<i>Edit a Filter By Using the Edit Filter Window</i>	131
<i>Set the Scope of a Filter (Global or Local)</i>	133
<i>Remove a Filter</i>	134
<i>Reset a Filter</i>	134

About Filters

For all visualization types, you can subset your data by using the **Filters** tab in the right pane. You can base your filters on any data item, regardless of whether the data item is assigned to the current visualization.

Your filters can be either *local filters* that apply only to the current visualization, or *global filters* that apply to all of the visualizations in the current visual exploration. For more information, see “[Set the Scope of a Filter \(Global or Local\)](#)” on page 133.

You can perform the following tasks to manage your filters:

- “Create a Filter”
- “Edit a Filter By Using Visual Filter Controls”
- “Edit a Filter By Using the Edit Filter Window”
- “Set the Scope of a Filter (Global or Local)”
- “Remove a Filter”
- “Reset a Filter”

Create a Filter

To create a filter:

- 1 Select the data item that you want to use as the base of the filter. You can select any data item, regardless of whether it is assigned to the current visualization.
- 2 For a local filter, either right-click the data item and select **Add as local filter**, or drag the data item onto the **Local Filters** area of the **Filters** tab.
For a global filter that is applied to all of your visualizations, either right-click the data item and select **Add as global filter** or drag the data item to the **Global Filters** area of the **Filters** tab.
- 3 Set the parameters for your filter. You can either create a basic filter by using [visual controls](#) or create an advanced filter by using the [Edit Filter window](#) .

Note: By default, your filter changes are applied automatically to the active visualization. To apply multiple changes together, deselect **Auto** and then click **Apply** when you are ready to apply your filter changes.

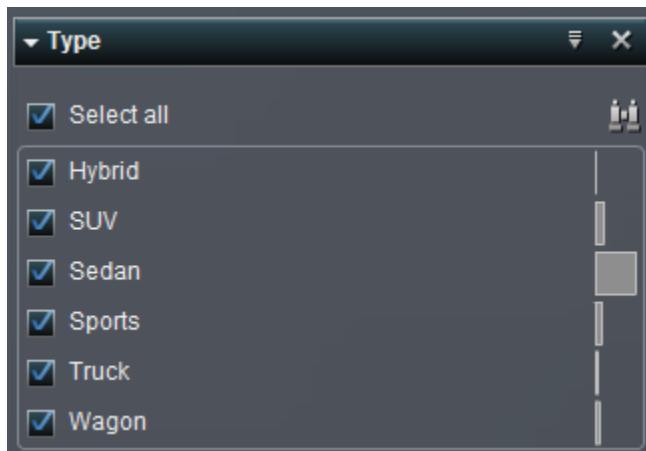
Edit a Filter By Using Visual Filter Controls

About Visual Filter Controls

Visual filter controls are accessible directly from the **Filters** tab. The visual filter controls are different depending on the model type of the filter data item.

Filtering Discrete Data

Display 12.1 A Visual Filter for Discrete Data



Note: If your data contains a large number of discrete values, then visual filtering is not available. Use the [Edit Filter window](#) instead.

For data items that use the discrete data model, the filter displays all of the distinct values for the data item. To the right of each value, a bar indicates the frequency of each value.

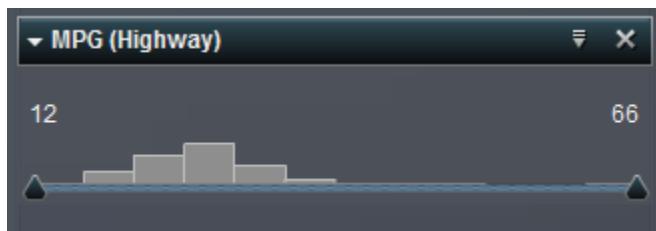
Use the check boxes next to each value to select the values for the filter. You can search for a value by clicking . The search locates values that begin with your search string.

Click **Select All** to select or deselect all of the values. To invert your selection, open the  drop-down list and then select **Invert Selection**.

By default, your filter changes are applied automatically to the active visualization. To apply multiple changes together, deselect **Auto** and then click **Apply** when you are ready to apply your filter changes.

Filtering Continuous Data

Display 12.2 A Visual Filter for Continuous Data



For measures that use the continuous data model, the filter shows the range of data values as a histogram. Sliders at each side of the histogram enable you to select the upper and lower limits of the data range that is selected by the filter.

You can also enter the limit values explicitly:

- For numeric data, click a limit values and then enter the value in the text field.
- For datetime data, click  and then select a date or time.

To exclude observations where the filter measure has a missing value, deselect **Include missing values**.

By default, your filter changes are applied automatically to the active visualization. To apply multiple changes together, deselect **Auto** and then click **Apply** when you are ready to apply your filter changes.

Edit a Filter By Using the Edit Filter Window

To edit a filter by using the Edit Filter window:

- 1 On the **Filters** tab, open the ▾ drop-down list for the filter that you want to edit and then select **Edit Filter**. The Edit Filter window appears.

Note: The **Edit Filter** window is not available for datetime data. See “Edit a Filter By Using Visual Filter Controls” on page 129 for details about using visual filter controls for a datetime data item.

- 2 Select the operator for the filter from the **Choose operator** drop-down list.

Table 12.1 Operators for Character Data

Is	Specifies that a matching value must match one of the filter values exactly.
Is not	Specifies that a matching value must not match any of the filter values exactly.
Begins with	Specifies that a matching value must contain the filter value at the start of the value.
Ends with	Specifies that a matching value must contain the filter value at the end of the value.
Contains	Specifies that a matching value must contain the filter value.
Does not contain	Specifies that a matching value must not contain the filter value.
Is missing	Specifies that a missing value matches the filter.
Is not missing	Specifies that any nonmissing value matches the filter.

Table 12.2 Operators for Numeric Data

Is	Specifies that a matching value must match the filter value exactly.
Is not	Specifies that a matching value must not match the filter value exactly.
Between	Specifies that a matching value must be greater than or equal to the first filter value and less than or equal to the second filter value.
Less than	Specifies that a matching value must be less than the filter value.
Less than or equal	Specifies that a matching value must be less than or equal to the filter value.
Greater than	Specifies that a matching value must be greater than the filter value.
Greater than or equal	Specifies that a matching value must be greater than or equal to the filter value.
Is missing	Specifies that a missing value matches the filter.
Is not missing	Specifies that any nonmissing value matches the filter.

3 Depending on the operator that you specified, select or specify a filter value:

Is, Is not

For character data, select one or more filter values from a drop-down list. For numeric data, enter the filter value in the text field.

Note: For character data, if the number of distinct values is very large then the enter a value in the text field.

Is missing, Is not missing

No filter value is needed.

Between

Enter the lower value in the first text field and the higher value in the second text field.

For all other operators, enter the filter value in the text field.

- 4 Click **Apply** to create the filter condition.
- 5 If you create multiple filter conditions, select whether a matching value must meet all of the conditions (**All**) or if a matching value can meet only some of the conditions (**Any**.)
- 6 (Optional) To delete a filter condition, click the  icon next to the condition that you want to delete.
- 7 (Optional) To preview your filter results, click **Preview Results**.
- 8 For discrete data, select whether to display the visual filter controls on the **Filters** tab for this filter (**Show results in a list of check boxes**) or to show a text description of the filter syntax (**Apply results as a dynamic rule**.)

Note: If your data contains a large number of discrete values, then visual filtering is not available.
- 9 When you are finished editing your filter, click **OK** to apply the filter.

Set the Scope of a Filter (Global or Local)

Your filters can be either local (specific to the current visualization) or global (applied to all visualizations in the visual exploration.)

To set the scope of a filter, open the  drop-down list and then select **Global Filter** or **Local Filter**.

Note: If you change a global filter's scope to local, then a local filter is created for all of your visualizations.

Note: A global filter replaces any local filters that use the same data item. The local filters are replaced in all of your visualizations.

Remove a Filter

To remove a filter, click  on the filter on the **Filters** tab.

Note: If you remove a global filter, then the filter is removed from all of your visualizations.

Reset a Filter

To restore a filter to its initial state, open the  drop-down list for the filter on the **Filters** tab and then select **Reset Filter**.

13

Exporting Data

<i>Overview of Exporting Data</i>	135
<i>Export a Visual Exploration as a Report</i>	136
<i>Export a Visual Exploration as a PDF</i>	136
<i>Save a Visualization as an Image File</i>	137
<i>E-mail a Visual Exploration as a Link</i>	138

Overview of Exporting Data

You can export your data and your visualizations from the SAS Visual Analytics explorer interface by using the following tasks:

- “Export a Visual Exploration as a Report”
- “Export a Visual Exploration as a PDF”
- “Save a Visualization as an Image File”
- “E-mail a Visual Exploration as a Link”

Export a Visual Exploration as a Report

You can export your visual exploration as a report, which can be refined in the designer interface or viewed directly on a mobile device or in the Web viewer.

To export your current visual exploration as a report:

- 1 Select **File** ▶ **Export** ▶ **Exploration as report**.
- 2 In the Export as Report window, select the visualizations that you want to include in the report.

Note: The following types of visualizations are not available to include in your report:

- visualizations that do not contain any data
- box plots, geo maps, and animated bubble plots
- depending on your administrator settings, visualizations that contain a large number of values

Click **OK**.

- 3 Select the location where you want to save the report and then click **Save**.

Export a Visual Exploration as a PDF

To export your current visual exploration as a PDF:

- 1 Select **File** ▶ **Export** ▶ **Exploration as PDF**.
- 2 In the **Export as PDF** wizard, enter a **Title** and a **Description** for the PDF document. Then, select the following options:

Page numbers

specifies the PDF document contains page numbers.

Summary data

specifies whether the PDF document contains summary data tables for each visualization.

Filter descriptions

specifies whether each visualization includes a description of its active filters.

Click **Next**.

- 3 Select the visualizations that you want to include in the PDF document.

Note: Visualizations that do not contain data are not available to include in your PDF output.

Click **Next**.

- 4 Click **Finish** to open a download window for your browser.

- 5 Select the location where you want to save the PDF document.

Note: If the legend for a visualization is too large, then it is not included in the PDF document.

Save a Visualization as an Image File

To save a visualization as an image file:

- 1 Select the visualization that you want to save as an image.
- 2 If your visualization contains hierarchies, drill to the hierarchy levels that you want to show in the image.
- 3 Select **Options** ▶ **Export Image** from the visualization toolbar.

- 4 If necessary, scroll the visualization to show the data items that you want to include in the image. If your visualization contains filters, legends, or hierarchy breadcrumbs, select whether to include these elements.
- 5 Click **Save** to open a download window for your browser.
- 6 Select the location where you want to save the image.

E-mail a Visual Exploration as a Link

To e-mail a link to your visual exploration:

- 1 Select **File** ▶ **Email**. The Email window appears.
- 2 Enter the following information:

To

specifies the e-mail destination. To enter multiple addresses, separate the addresses by using commas, semicolons, or spaces.

From

specifies the e-mail address that is shown as the sender of the e-mail. This e-mail address is also used for notifications if the e-mail cannot be delivered to the addresses in the **To** field.

Subject

specifies the subject line of the e-mail.

Message

specifies text for the e-mail message body. This field is optional.

Note: The message body of the e-mail automatically includes a link to the current visual exploration. If you specify text for the **Message** field, then your text appears before the link in the e-mail.

- 3 Click **OK** to send the e-mail.

14

Managing Hierarchies

<i>Overview of Hierarchies in the SAS Visual Analytics Explorer Interface</i>	139
What Is a Hierarchy?	139
Hierarchy Management Tasks	140
<i>Create a New Hierarchy</i>	140
<i>Edit a Hierarchy</i>	141
<i>Delete a Hierarchy</i>	141

Overview of Hierarchies in the SAS Visual Analytics Explorer Interface

What Is a Hierarchy?

A hierarchy is an arrangement of category columns that is based on parent-child relationships. The levels of a hierarchy are arranged with more general information at the top and more specific information at the bottom.

For example, you might create a hierarchy of date-time columns with Year as the top level, Month as the next level, and Day as the bottom level.

Creating hierarchies enables you to add drill-down functionality to your visualizations. For example, if you use a date-time hierarchy, you can drill down to the data for a specific year. Then, drill down to the data for a specific month.

When you drill down a hierarchy, a set of breadcrumb links at the top of your visualization enables you to drill back up the hierarchy.

Hierarchy Management Tasks

In the SAS Visual Analytics explorer interface, you can perform the following tasks to manage hierarchies:

- “Create a New Hierarchy”
- “Edit a Hierarchy”
- “Delete a Hierarchy”

Create a New Hierarchy

To create a new hierarchy:

- 1 Select **Data** ► **New Hierarchy**. The Create Hierarchy window appears.
- 2 Enter a **Name** for the hierarchy.
- 3 Select the categories that you want to include in the hierarchy, and then click ➔ to add them to the hierarchy.

Note: You can also drag and drop categories.

To change the order of the categories in your hierarchy, select the category that you want to move. Then, click  to move the category up, or click  to move the category down.

To remove a category from the hierarchy, select the hierarchy that you want to remove. Then, click .

- 4 Click **OK** to finish creating the hierarchy.

Edit a Hierarchy

To edit an existing hierarchy:

- 1 From the **Data Items** window, right-click the hierarchy that you want to edit, and then select **Edit**. The Edit Hierarchy window appears.
- 2 Enter a **Name** for the hierarchy.
- 3 To add a category to the hierarchy, select the category and then click ➔.

Note: You can also drag and drop categories.

To change the order of the categories in your hierarchy, select the category that you want to move. Then, click ↑ to move the category up or ↓ to move the category down.

To remove a category from the hierarchy, select the hierarchy that you want to remove. Then, click ←.

- 4 Click **OK** to save the changes to your hierarchy.

Delete a Hierarchy

To delete a hierarchy:

From the **Data items** window, right-click the hierarchy that you want to delete and then select **Delete**. Click **Yes** to confirm that you want to delete the hierarchy.

15

Managing Visual Explorations

<i>What Is a Visual Exploration?</i>	143
<i>Create a New Visual Exploration</i>	143
<i>Saving Your Visual Exploration</i>	144
<i>Deleting Visual Explorations</i>	144

What Is a Visual Exploration?

A visual exploration is a metadata object that contains all of the visualizations and data settings from a SAS Visual Analytics explorer interface session. You can use visual explorations to save your session for later and to share it with others.

Create a New Visual Exploration

To create a new visual exploration, select **File** ▶ **New Exploration** from the menu bar. The Select Data wizard appears. See [Chapter 9, “Opening a Data Source,” on page 73](#) for more information about opening a data source for a new visual exploration.

Saving Your Visual Exploration

To save your current visual exploration, select **File** ▶ **Save as** and then select a location and a name.

Deleting Visual Explorations

To delete a visual exploration, select **File** ▶ **Manage explorations** to open the Manage explorations window. Locate and select the visual exploration that you want to delete, and then click .

16

Performing Data Analysis

<i>Overview of Analyzing Your Data</i>	145
<i>Add Analysis to an Existing Visualization</i>	147
<i>Correlate Multiple Measures as a New Visualization</i>	148

Overview of Analyzing Your Data

SAS Visual Analytics enables you to perform two basic types of data analysis:

Correlation

identifies the degree of statistical correlation between measures.

Regression

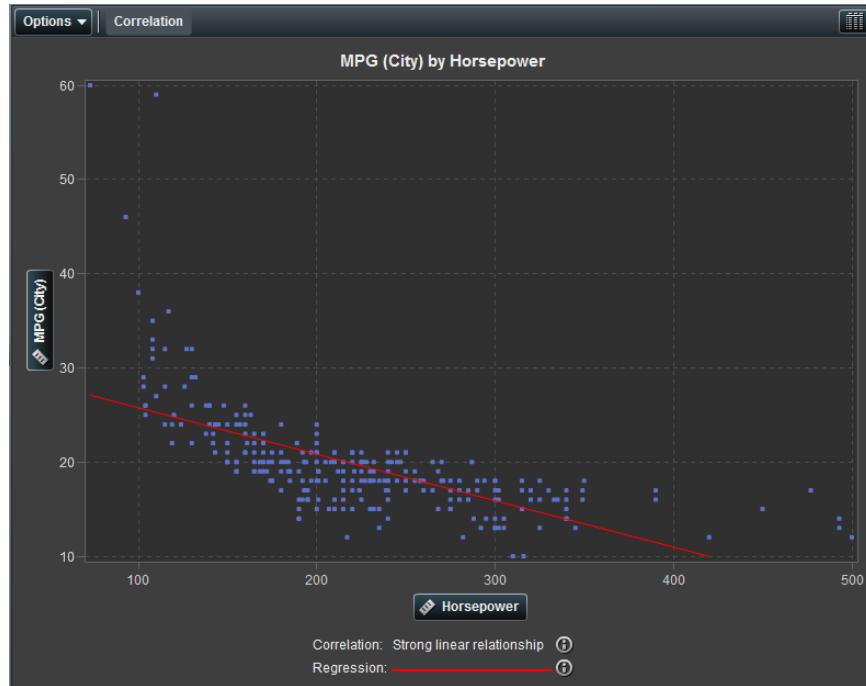
plots a model of the relationship between measures.

Both correlation and regression analysis are applied together when you enable correlation analysis.

For visualizations that contain a single plot, the correlation is identified by a text label in the legend and the regression is identified by a red line.

Display 16.1 displays a scatter plot with analysis enabled.

Display 16.1 A Basic Scatter Plot with Analysis Enabled



For visualizations that contain a plot matrix, the correlation for plot is identified by a colored border around the plot. The regression is identified by a red line in each plot.

Display 16.2 displays a scatter plot matrix with analysis enabled.

Display 16.2 A Scatter Plot Matrix with Analysis Enabled



For all visualizations with data analysis enabled, you can click  to view additional information about the regression and correlation results.

Add Analysis to an Existing Visualization

To add data analysis to a scatter plot or heat map, click **Correlation** on the visualization toolbar or select **Options** ▶ **Show Correlation** from the visualization toolbar.

Correlate Multiple Measures as a New Visualization

To create a new visualization that correlates multiple measures together, select **Analysis** ▶ **Correlate Measures**. The Correlate Measures window appears. Select the measures that you want to include in the new visualization, and then click **OK** to create the new visualization.

If you selected exactly two measures, then the new visualization is either a scatter plot or a heat map (for high-cardinality data). Correlation and regression analyses are automatically applied to the visualization.

If you selected three or more measures, then the new visualization is either a scatter plot matrix or a correlation matrix (for high-cardinality data). Correlation and regression analyses are automatically applied to the visualization.

17

Specifying Your Preferences

<i>Specifying Global Preferences</i>	149
<i>Specify Preferences That Are Specific to the Explorer Interface</i>	149

Specifying Global Preferences

See “[Specifying Global Preferences](#)” on page 15 for details about global preferences for SAS Visual Analytics.

Specify Preferences That Are Specific to the Explorer Interface

To specify preferences that are specific to the SAS Visual Analytics explorer interface:

- 1 Select **File** ▶ **Preferences** to open the Preferences window.
- 2 Select **Visual Analytics Explorer** ▶ **General**.
- 3 Select a **Graph skin** to change the visual effects that are applied to your graphs. For example, many of the **Graph skin** settings apply lighting effects to the features of your graph to create a 3-D appearance.

- 4 Select **Visualization data threshold** to specify the amount of data values that your visualizations can process. The precise amount of data that is supported varies depending on the visualization type.

Note: A large value for **Visualization data threshold** might degrade your application performance or lead to time-out errors.

- 5 To return to the default settings, select **Reset to defaults**.
- 6 Click **OK** to apply your changes.

Part 4

Designing Reports

<i>Chapter 18</i>	
<i>Overview of the Designer Interface</i>	153
<i>Chapter 19</i>	
<i>Opening Reports</i>	161
<i>Chapter 20</i>	
<i>Creating and Saving Reports</i>	163
<i>Chapter 21</i>	
<i>Selecting Data for Reports</i>	169
<i>Chapter 22</i>	
<i>Displaying Results in a Table, Graph, or Gauge</i>	185

<i>Chapter 23</i>	
<i>Sorting Data Items in Reports</i>	219
<i>Chapter 24</i>	
<i>Working with Interactions</i>	223
<i>Chapter 25</i>	
<i>Maintaining Multi-Section Reports</i>	233
<i>Chapter 26</i>	
<i>Exporting Data from Report Objects</i>	237

18

Overview of the Designer Interface

<i>About the SAS Visual Analytics Designer Interface</i>	153
<i>Your First Look at the Designer Interface</i>	154
<i>Open Window</i>	157
About the Open Window	157
Working with Folders in the Open Window	158
<i>Search Window</i>	159

About the SAS Visual Analytics Designer Interface

The designer interface enables users to easily create reports or dashboards that can be saved and viewed on either a mobile device or in a Web viewer that is provided with SAS Visual Analytics. The designer interface is part of the SAS Visual Analytics product that enables someone with either the Analysis or Administration role to view, interact with, and create reports. Simply open an existing report and interact with the information based on your current needs. Report authors can easily point and click to query central sources of data. You can drag-and-drop tables, graphs, and gauges to create a well-designed report. You can also add text and images to reports. All of this is accomplished by using the designer interface, which runs in a Web browser. You do not need to understand a programming language.

Report authors can easily create basic queries and reports based on data sources that have been provided by an administrator. They can also update reports that were created from visual explorations. You can also create reports by importing objects or visual explorations from other reports. Report authors can also define interactions (either filtering or brushing) for report objects and include SAS analytical results into a single report.

Your First Look at the Designer Interface

When you are designing reports in SAS Visual Analytics, you see the designer interface.

Here are the features of the designer interface:

Figure 18.1 The Designer Interface



- 1 The application bar enables you to return to the SAS Visual Analytics home page and access your recently created or viewed reports. You can also access recently used visual explorations in your recent history. When you select a visual exploration, the application takes you to the explorer interface.
- 2 The menu bar provides menu items that apply to the entire report or to the currently displayed report section. Actions include creating a new report, adding a new section, inserting new report objects, and adding interactions. You can also log off from SAS Visual Analytics.

- 3 The toolbar contains icons that enable you to manage your reports. You can click  to hide the left and right panes. Or, you can click  to display the left and right panes.
- 4 The **Objects** and **Imports** tabs are in the left pane.

The **Objects** tab provides a list of the tables, graphs, and gauges that can be used in a report or dashboard.

The **Imports** tab provides a list of reports and report objects that have been created in the designer interface or exported from the explorer interface. This enables you to create reports from multiple sources and create full-featured reports. You can choose objects or sections in these reports to include in either a new or an existing report.

- 5 The **Data** and **Shared Rules** tabs are in the left pane.

The **Data** tab enables you to select a data source (or data sources) and the data items for your report. You can add, change, or remove a data source using the icons above the list of data items. You can also check the details for the measures in the data set. For more information, see [“Using Data Items in a Query” on page 174](#).

The **Shared Rules** tab enables you to create a new display rule, which is used by a gauge to designate intervals and colors for ranges. You can also edit or delete an existing display rule. These display rules are shared by multiple report objects and can be created at any time.

- 6 The data item table provides information about a selected data item, including the name, role, format, and aggregation. These data item properties can also be modified, which impacts all of the report objects that use the data item. For more information, see [“Modifying Data Item Properties” on page 178](#).
- 7 The canvas is the workspace for building a report. Each report section has its own canvas, which can contain a single or multiple report objects. For more information, see [“Creating a New Report” on page 163](#).

A report can have multiple sections, which can be accessed by using the tabs at the top of the canvas. For more information, see [“Maintaining Multi-Section Reports” on page 233](#).

There is also an area at the top of the canvas where you can drop categories to create section filters.

8 The tabs in the right pane enable you to work with details about the report and report objects. The **Properties**, **Styles**, and **Display Rules** tabs are displayed by default. You can choose to display icons instead of tab names. To display the icons on the tabs, click ▾ after the last tab name and select **Show icons only**.

The **Properties** tab lists the properties for the currently selected report object. The object's name, title, and description are listed. Information that is specific to the report object type, such as the axes and legend are also included. If a report object is not selected, then the properties are for the section and the report.

The **Styles** tab enables you to specify the data styling, frame styling, text styling, and data colors for a selected report object.

The **Display Rules** tab enables you to populate or add intervals for a gauge that is currently selected in the report canvas. You can also edit or delete an existing display rule for the selected gauge.

The **Comments** tab enables you to add comments to the report after you save it.

The **Interactions** tab enables you to add or update filter or brush interactions to the report objects in a section. For more information, see “[Creating Interactions](#)” on page 224.

The **Roles** tab enables you to add or update data role assignments when you have a report object with a data source selected.

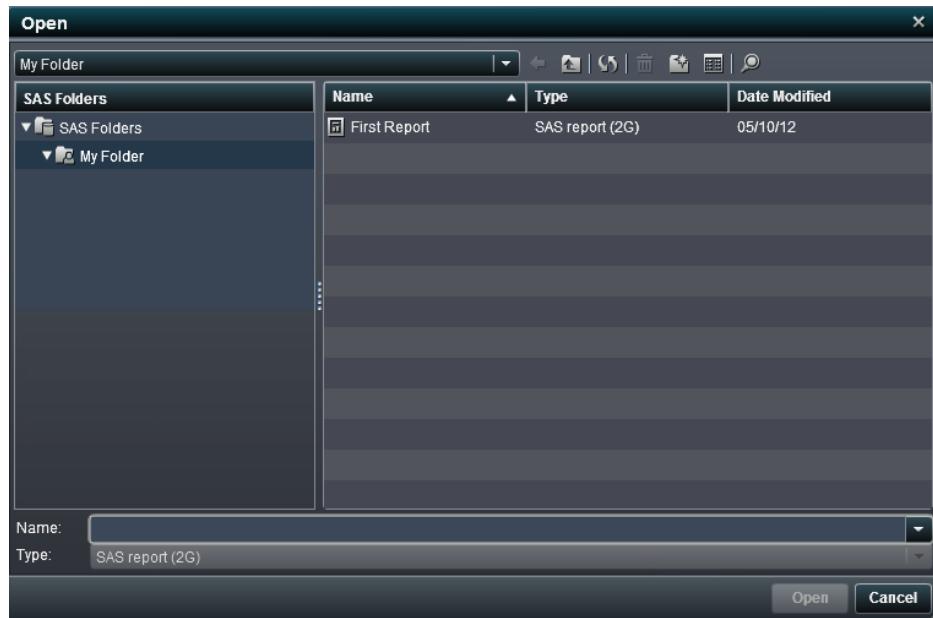
Open Window

About the Open Window

Click  on the menu bar or select **File** ▶ **Open** to display the Open window.

The Open window enables you to search for and open saved reports. However, the Open window also enables you to create new folders and perform many other tasks on selected reports and folders.

Display 18.1 Open Window



Working with Folders in the Open Window

You can work with folders in the Open window. The icons for folders are located to the right of the drop-down list of folder names, as shown in [Display 18.1 on page 158](#).

Icon	Action
←	Navigates back one level.
↑	Moves up one level.
⟳	Refreshes the items in the current location.
ⓧ	Deletes the selected item.

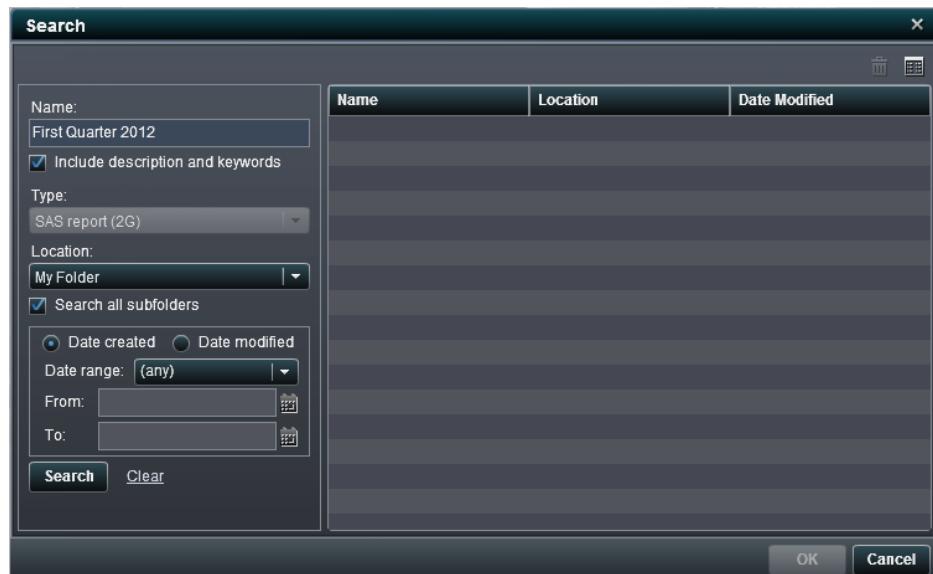
Icon	Action
	Creates a new folder.
	Specifies either details or a list.
	Opens the Search window.

You can also create a new folder when you copy a report, move a report, or save a report.

Search Window

The Search window in the designer interface enables you to navigate to find the report or exploration that you need. For example, you might need to navigate through multiple folders to find a report. In the Open window, click  to open the Search window.

Display 18.2 Search Window



To search for files:

- 1 In the **Name** field, type text for which you want to search.
- 2 (Optional) To include descriptions in your search, select the **Include description and keywords** check box.
- 3 Select a **Location**. If you want to search multiple folders, make sure that **Search all subfolders** check box is selected.
- 4 (Optional) To limit your search by time, select a **Date range**. You can also specify either **Date created** or **Date modified**.
- 5 Click **Search**.

Note: After a search, the list of report or exploration files also includes the path information for each located report.

19

Opening Reports

<i>Opening Reports Created in the Designer Interface</i>	161
<i>An Example Report</i>	162

Opening Reports Created in the Designer Interface

Reports in the designer interface are saved when you use either the Save or the Save As window. A saved report contains at least one section. Typically, that section uses a query method of data items from a data source. The section has at least one report object (a table, a graph, or a gauge) to display the results of the query.

To open a saved report:

- On the SAS Visual Analytics home page, select a report. In the details view, click **Edit**. Or you can double-click the report to open it in the designer interface.
- In the designer interface:
 - Select **File** ▶ **Recent** and then select a report name.
 - Click ▼ beside the recent content drop-down list, and then select a report name.
 - Select **File** ▶ **Open** to display the Open window.
 - Click  to navigate to a report and then select it.

An Example Report

The following display shows a report in the designer interface. The report contains a pie chart, a bubble plot, a scatter plot, a bar chart, a dual axis bar chart, and a targeted bar chart.

Display 19.1 An Example Report with Six Report Objects



20

Creating and Saving Reports

<i>Creating a New Report</i>	163
<i>Basing a New Report on Report Objects</i>	
<i>Imported from One or More Existing Reports</i>	164
<i>Saving a Report</i>	168

Creating a New Report

To create a new report in the SAS Visual Analytics designer interface:

- 1 Choose your data source and data items. For information about selecting data, see “[Selecting Data for Reports](#)” on page 169.
- 2 Select the report object (or objects) that you want to use in the report. For information about selecting report objects, see “[Displaying Results in a Table, Graph, or Gauge](#)” on page 186.
- 3 Drag and drop the report object (or objects) on the canvas.
- 4 Update the properties for the report and the report objects. For more information, see “[Specify Table Styles](#)” on page 190, “[Specify Graph Styles](#)” on page 202, or “[Specify a Name, Title, or Description for a Gauge](#)” on page 207.

- 5 Update the styles for the report objects. For more information, see “Specify a Name, Title, or Description for a Table” on page 190, “Specify a Name, Title, or Description for a Graph” on page 201, or “Specify Gauge Styles” on page 208.
- 6 If there is a gauge in the report, create or modify the display rule. For more information, see “Specify a Gauge Display Rule” on page 209.
- 7 (Optional) Add interactions between the report objects in a section. For more information, see “Overview of Interactions” on page 223.
- 8 (Optional) Add a new section (or sections) to the report. For more information, see “Overview of Report Sections” on page 233.
- 9 Save your report.
- 10 (Optional) Add comments to your saved report.

You can also create a new report based on an existing report or on report objects from a report that you have imported. For more information, see “Basing a New Report on Report Objects Imported from One or More Existing Reports” on page 164.

Basing a New Report on Report Objects Imported from One or More Existing Reports

If there is an existing saved report that closely matches the report that you want to create, then you can open the report that contains the report objects that you want to reuse. The report objects can be from either a report that has been saved in the designer interface or from a report that has been exported from the explorer interface. You can save the report under a new name or you can edit the report to make some customizations and then save it.

To base a new report on report objects imported from one or more existing reports:

- 1 Open or import the existing report by doing one of the following:

- On the SAS Visual Analytics home page, double-click the report that you want to open.
- In the designer interface:
 - Select **File** ▶ **Recent** and then select the report name.
 - Click ▾ beside the recent content drop-down list, and then select a report name.
 - Select **File** ▶ **Open** to display the Open window.
 - Click  to navigate to a report and then select it.
 - Click **Imports** in the left pane. Select a report name or click **Import another report** to display the Open window, where you can select a report.

Display 20.1 Imports Tab with a Report List

In the tree view on the left, a list of sections and report objects appears. Drag a single report object, multiple report objects, or entire section that you want to include in the new report and drop it onto the canvas.

Display 20.2 Imports Tab with a List of Report Objects

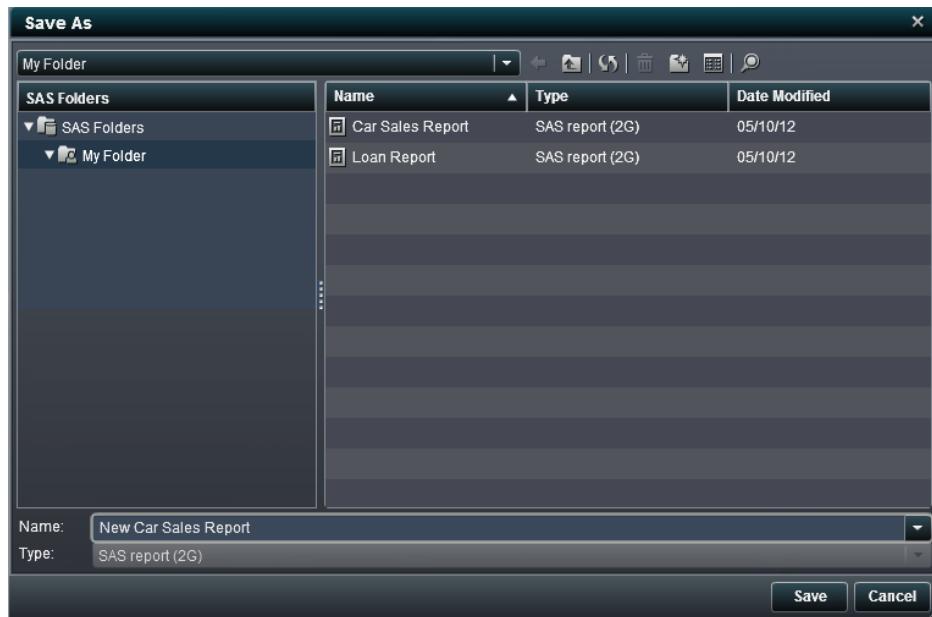
- 2 (Optional) Modify the report objects, properties, and styles. If you have a gauge (or gauges) in your report, then you can also update the display rules.
- 3 Select **File** ▶ **Save** or click , which displays the Save As window. Enter a **Name**. Report names cannot use these characters: / \
- 4 Click **Save**.

Saving a Report

To save a report:

- 1 Select **File** ▶ **Save** or click . If you are saving a new report, then the Save As window is displayed.

Display 20.3 Save As Window



- 2 For new reports, enter a **Name**. Report names cannot use these characters: / \ If you are saving an existing report, then the name of that report is listed here. You can either leave the name as it is and overwrite the existing report, or you can change the name to create a new report.
- 3 Click **Save**.

21

Selecting Data for Reports

<i>Overview of Queries Based on Data Items</i>	169
<i>Adding a Data Source or Data Sources</i>	170
<i>Changing a Data Source</i>	172
<i>Removing a Data Source</i>	173
<i>Using Data Items in a Query</i>	174
About Data Items	174
Select Data Items	175
Duplicating Data Items	176
Removing Data Items	177
Modifying Data Item Properties	178
<i>Modifying Data Role Assignments in Report Objects</i>	182
<i>Removing Data Assignments from Report Objects</i>	183

Overview of Queries Based on Data Items

Data sources are prepared by a data source administrator in a centrally managed location so that you can easily define a report. Authorized users who have an Administrative role can load tables into memory using the data preparation interface. All

data sources contain data items, which can refer to calculations or columns in physical data (tables). Reports can include query results from more than one data source.

Each data source includes one or more data items. For example, a data source named *Order Information* might include standard data items such as *Order ID*, *Product ID*, *Unit Cost*, *Order Date*, and *Order Amount*. You decide which data items to use. You can select all of the data items in the data source or a subset of the data items.

Adding a Data Source or Data Sources

You can use one or more data sources for a report in the designer interface.

To add a data source for a report:

- 1 On the **Data** tab, click ▾ beside the **Select a data source** text to display the Add Data Source window.

Display 21.1 *DataTable*

2 In the Add Data Source window, select a data source.

Display 21.2 Add Data Source Window



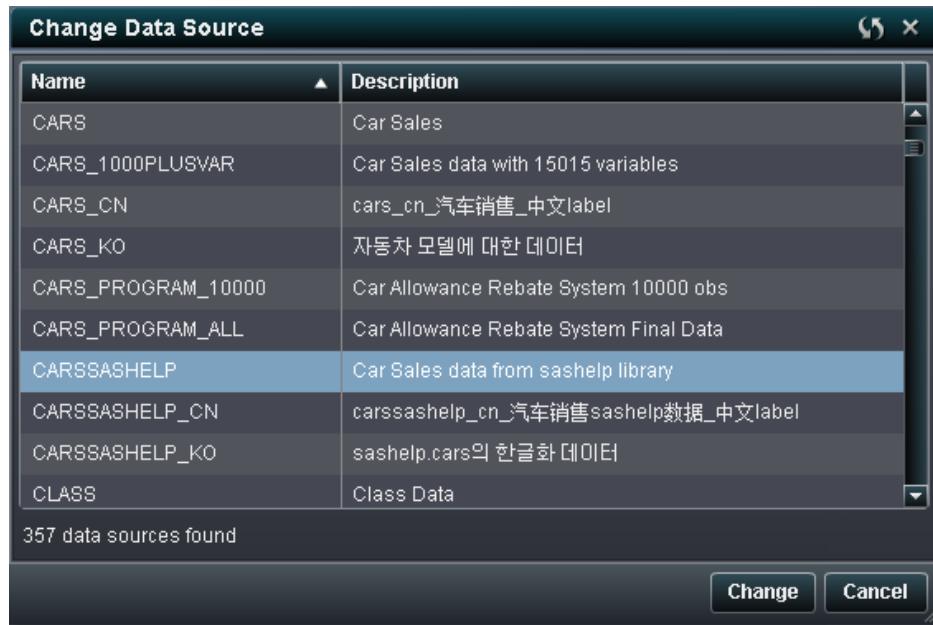
- 3 Click **Add**. The list of available data items is displayed in the **Data** tab.
- 4 To add another data source, click , which displays the Add Data Source window. Select the data source that you want, and then click **Add**. The **Data** tab is populated with a list of all of the data items that are in the data source.

Changing a Data Source

To change a data source for a report:

- 1 On the **Data** tab:
 - Click . The Change Data Source window is displayed.
 - Click  to display the Add Data Source window.
- 2 In the Change Data Source window, select a data source.

Display 21.3 Change Data Source Window



3 Click **Change**. If the data sources are compatible, then the list of available data items is displayed in the **Data** tab. This means that every data item that is defined in the data source that is being replaced has to be defined in the replacement data source. The data items must share the same names (which is not the same thing as the labels). The replacing data source can have additional data items. The additional data items are displayed in the **Data** tab when the data source is replaced.

If the data sources are not compatible, a message is displayed.

Removing a Data Source

You can remove a data source from a report. Be aware that removing a data source means that all related data items are also removed from the report.

To remove a data source for a report:

- 1 On the **Data** tab, click .

2 Click Yes in the confirmation message.

Display 21.4 The Remove Data Source Message



Using Data Items in a Query

About Data Items

Each data source includes one or more standard data items. You decide which data items to use to define a query for each report object. You can use all the data items in the data source or a subset of data items. Each data item is classified as either a category or a measure.

These data items are available in the designer interface:

Data Item	Icon	Description
Category	■	<p>A data item whose distinct values are used to group and aggregate measures. There are four types of categories: alphanumeric, date, timestamp, and time. Alphanumeric categories can be made up of all letters, all digits, or a combination of the two. Categories that have values that are all digits might be physically stored as character or numeric data. The data type affects how values are handled in relation to some functionality, such as filtering, sorting, and formatting.</p> <p>Examples of alphanumeric categories include data items such as <i>Product ID</i>, <i>Country</i>, <i>Employee Number</i>, and <i>Employee Name</i>.</p> <p>Note: If you change a measure to a category, then it also uses this category icon.</p>

Data Item	Icon	Description
Date and Time		<p>A category data item whose distinct values are used to group and aggregate measures. There are three types of date categories: date, timestamp, and time.</p> <p>Examples of date, timestamp, and time categories are <i>Order Year</i>, <i>Date of Sale</i>, and <i>Delivery Time</i>.</p>
Measure		<p>A data item whose values can be used in computations. These values are numeric. Examples of measures include <i>Sales Revenue</i>, <i>Units Sold</i>, and <i>Salary</i>.</p> <p>The designer interface assigns a default aggregation method to every measure. Almost all measures are assigned sum, but certain formats are assigned an average. You can change the aggregation method.</p>

Note: Report objects that are imported from the explorer interface use embedded data. Therefore, you have a very limited ability to make changes to these data items in the designer interface.

Select Data Items

To select data items to use in queries for the current report section:

- 1 On the **Data** tab in the left pane, click the down arrow to display a list of available data sources. Select a data source and the **Data** tab is populated with a list of all of the data items that are in the data source.

If the data source that you want is not in the list, click , which displays the Add Data Source window. Select the data source that you want, and then click **Add**. The **Data** tab is populated with a list of all of the data items that are in the data source.

If you do not want to use the data source that you originally selected, click **Remove data source**. A message box asks you to confirm that you want to remove the data source. Click **Yes**.
- 2 (Optional) To see more information about a particular data item, select it in the list. The **Name**, **Role**, **Format**, and **Aggregation** are displayed in the data item table below the list of data items.

Here is an example of the details for a data item called *Engine Size*:

Display 21.5 Details about a Selected Measure Data Item



The screenshot shows the SAS Data tab interface. At the top, there are tabs for 'Data' and 'Shared Rules', with 'Data' being the active tab. Below the tabs is a dropdown menu showing 'Car Sales data from sashelp library'. Underneath the dropdown are several icons: a star, a plus sign, a trash can, and a document. The main list area contains the following data items:

DriveTrain	3
Make	38
Model	425
Origin	3
Type	6
Cylinders	
Engine Size (L)	
Horsepower	

The 'Engine Size (L)' item is highlighted with a blue selection bar. Below this list is a table with the following properties:

Name	Engine Size (L)
Role	Measure
Format	Numeric
Aggregation	Sum

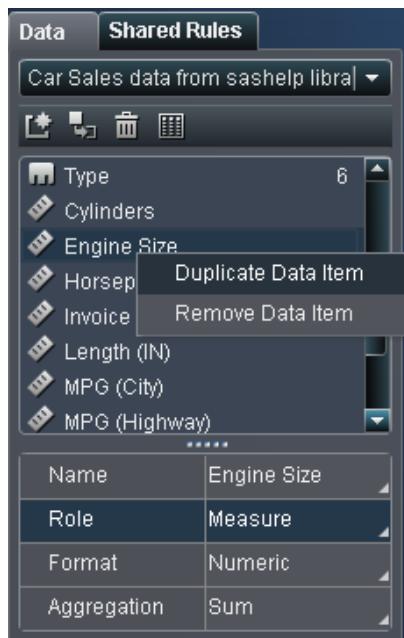
Duplicating Data Items

Duplicating data items enables you to see all the aggregations of a data item (**Sum**, **Average**, **Minimum**, **Maximum**, and **Count**) side-by-side in a table. If you save a report with duplicate data items, then those data items are available when you edit the report the next time.

To duplicate a data item:

- 1 On the **Data** tab in the left pane, right-click the data item that you want to duplicate.

Display 21.6 Duplicate Data Item Menu Selection



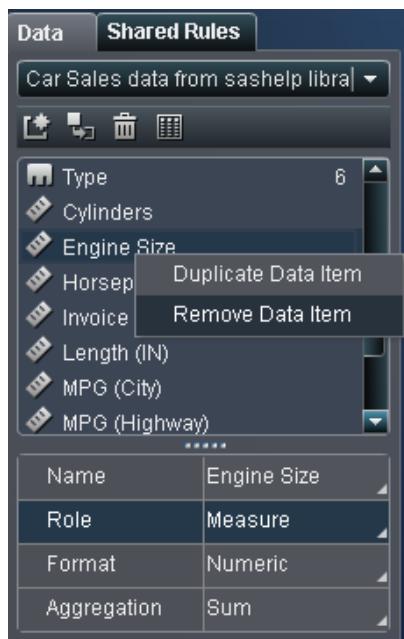
- 2 Select **Duplicate Data Item**. The data item appears in the list of data items. For example, if the original data item name is `Engine Size`, then the duplicate data item is displayed as `Engine Size (1)`. If you choose to duplicate the same data item again, then it is displayed as `Engine Size (2)`.
- 3 (Optional) Rename the duplicate data item (or data items).

Removing Data Items

You can remove data items so that they do not appear in the **Data** tab in the left pane. This means that the data item (or data items) are removed when you open that data source. If an administrator removes the data source and then adds it back, then the deleted data items appear again. If you save a report with deleted data items, then those data items are not available the next time you edit the report.

To remove a data item (or data items):

- 1 On the **Data** tab in the left pane, right-click the data item that you want to remove.

Display 21.7 Remove Data Item Menu Selection

2 Select **Remove Data Item**. The data item is removed from the list of data items.

Modifying Data Item Properties

Rename a Data Item

To rename a data item:

- 1 Select a data item on the **Data** tab.
- 2 In the data item table, select the existing name for the data item and then enter a new name. Your change is saved automatically.

Modify a Data Item's Role

To modify a data item's role:

- 1 Select a data item on the **Data** tab.

2 In the data item table, select the existing role name for the data item. In the drop-down menu, select either **Measure** or **Category**. Select one and your change is saved automatically.

Note: Categories data items cannot be converted into measure data items.

Modify the Format of a Measure or a Date-Time Data Item

To change the format of a measure or date-time data item:

- 1 Select a measure data item on the **Data** tab.
- 2 In the table, select the existing format. A list is displayed with the **Format type**, **Width**, and **Decimals**. Make your selections. A sample of your selection (or selections) is displayed under the list. Here is an example of the list:

Display 21.8 Available Formats for a Measure Data Item



Display 21.9 Available Formats for a Date–Time Data Item

3 Click **OK** to save your changes.

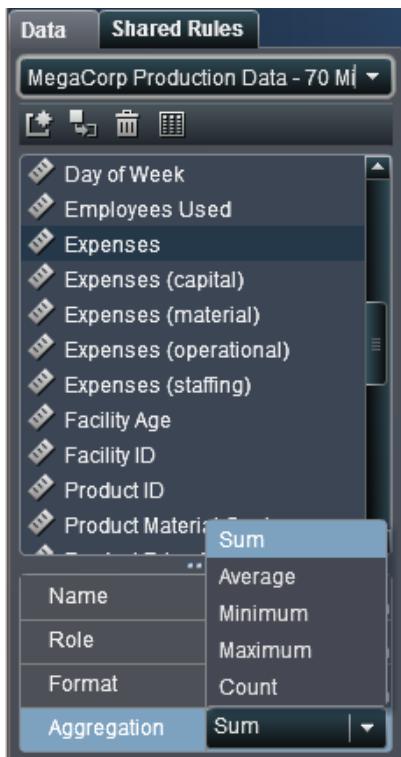
Modify How a Measure Is Aggregated

You can change the aggregation method for a measure in a data source using either the **Data** tab or in the canvas.

Note: You need to understand your data, because some aggregation methods are not always appropriate. For example, an average of an average is not valid.

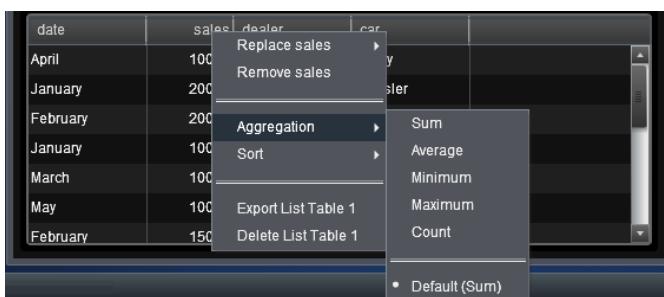
To change the aggregation method using the **Data** tab:

- 1 Select a measure data item on the **Data** tab.
- 2 In the data item table, select the existing aggregation. A drop-down list is displayed with aggregations. **Sum**, **Average**, **Minimum**, **Maximum**, and **Count** are the available aggregations. Select one, and your change is saved automatically.

Display 21.10 Available Aggregations for a Data Item

To change the aggregation method when you are working with a report object in the canvas:

- 1 Right-click a measure data item in the report object.

Display 21.11 Aggregation Menu Items

2 Select **Aggregation** ► **aggregation name**, where *aggregation name* is **Sum**, **Average**, **Minimum**, **Maximum**, and **Count**. Select one, and your change is saved automatically.

Modifying Data Role Assignments in Report Objects

After you have selected a report object and a data source, the **Roles** tab in the right pane shows which data items have been assigned to which role.

To modify data role assignments:

- 1 Select a report object in the canvas that has data items assigned.
- 2 Click the **Roles** tab in the right pane.

Display 21.12 Roles Tab



- 3 Click ▾ beside the role that you want to edit. For roles that allow multiple data items, the menu items that are available can vary.

If you select a data item in the **Roles**, then the **Aggregation**, **Add**, **Remove**, and **Replace** menu items are displayed. You can also right-click a data item, and the actions specific to that data item are displayed.

For example, for the bar chart shown in [Display 21.12 on page 182](#), you can replace or remove the **Category**. For the **Measure**, you can modify the aggregation, add another measure, replace the existing measure, or remove the measure. You can also modify the **Group**, the **Lattice columns**, or **Lattice rows**.

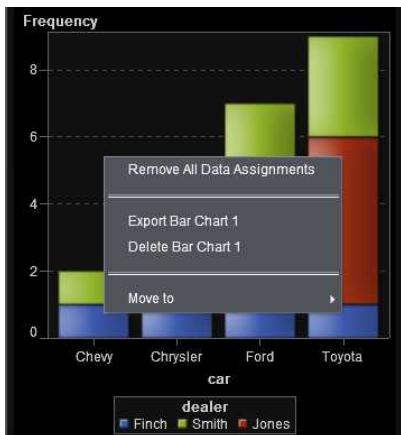
Removing Data Assignments from Report Objects

To remove data items from their assigned roles in a specific report object:

- 1 Right-click the report object in the canvas. A menu is displayed.

Here is an example of a bar chart with the object menu:

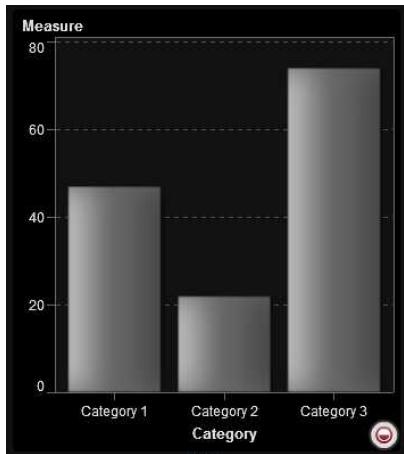
Display 21.13 Report Object Menu



- 2 Select **Remove All Data Assignments**. All data items are removed from the assigned roles and the report object turns gray. The  status icon appears in the lower right corner to let you know that the required data roles are not assigned.

Here is an example of what a bar chart looks like after all the data items have been removed from their assigned roles:

Display 21.14 Report Object with Missing Data Assignments



22

Displaying Results in a Table, Graph, or Gauge

<i>Using Tables, Graphs, and Gauges</i>	186
<i>Using Tables to Display Results</i>	187
Overview of the Table Types	187
Insert a Table into a Report	189
Specify a Name, Title, or Description for a Table	190
Specify Table Styles	190
<i>Using Graphs to Display Results</i>	191
Overview of the Graph Types	191
Insert a Graph into a Report	200
Specify a Name, Title, or Description for a Graph	201
Specify Graph Styles	202
<i>Using Gauges to Display Results</i>	204
Overview of the Gauge Types	204
Inserting a Gauge into a Report	206
Specify a Name, Title, or Description for a Gauge	207
Specify Gauge Styles	208
Specify a Gauge Display Rule	209
<i>Adding Text and Images</i>	212
Overview	212
Insert Text into a Report	212
Insert an Image into a Report	212
Specify a Name, Title, or Description for Text Objects	214

Specify a Name, Title, or Description for Images	215
<i>Adding Comments to Report Objects</i>	216

Using Tables, Graphs, and Gauges

After selecting your data source and data items, add one or more report objects to display the results. SAS Visual Analytics designer interface provides tables, graphs, gauges, text, and images for all of your reports.

You can select these table types for reports:

- list tables
- crosstabs

The following graph types are available in the designer interface:

- bar charts
- targeted bar charts
- waterfall charts
- line charts
- pie charts
- scatter plots
- time series plots
- bubble plots
- treemaps
- dual axis bar charts
- dual axis line charts
- dual axis bar-line charts
- dual axis time series plot

The following gauges are available in the designer interface:

- bullet
- slider
- thermometer
- dial
- speedometer

These objects are also available in the designer interface:

- text
- image

In the designer interface, you have access to report objects from the explorer interface. You can use the histogram or heat map report objects that have been exported from the explorer interface. However, you cannot create new histograms or heat maps in the designer interface.

Using Tables to Display Results

Overview of the Table Types

About List Tables

A *list table* is a two-dimensional representation of data in which the data values are arranged in unlabeled rows and labeled columns. List tables can use any data items from a data source.

Display 22.1 A List Table

Make	MSRP	Invoice	MPG (City)	MPG (Highway)
Acura	\$42,939	\$38,591	19.42857	26.14286
Audi	\$43,308	\$39,330	18.47368	25.78947
BMW	\$43,285	\$39,621	18.7	27
Buick	\$30,538	\$27,855	18.88889	27.77778
Cadillac	\$50,474	\$46,427	16.5	23
Chevrolet	\$26,587	\$24,061	19.66667	26.85185
Chrysler	\$27,252	\$25,270	19.86667	27.33333
Dodge	\$26,254	\$24,160	19.38462	26.23077
Ford	\$24,016	\$21,953	19.26087	25.73913
GMC	\$29,561	\$26,289	15.375	19.625
Honda	\$21,435	\$19,513	27.82353	34
Hummer	\$49,995	\$45,815	10	12
Hyundai	\$17,477	\$16,035	23	29.91667

About Crosstabs

A *crosstab* (also known as a crosstabulation table) shows an aggregate metric for the intersections of two or more categories. In a crosstab, categories are typically displayed on both the columns and the rows, and each cell value represents the aggregated measure from the intersection of the categories on the specific row and column. This type of table uses less space than a list table and is easier to read because data is grouped both horizontally and vertically.

Display 22.2 A Crosstab

Month	Jan2010		Feb2010		Mar2010		
	Country	Actual Profit	Actual Revenue	Actual Profit	Actual Revenue	Actual Profit	Actual Revenue
		1,037,176	6,234,929	5,136,199		1,508	
Africa		-506,999	449,465	903,973	3,189,596	126,664	
Argentina		0	0	-60,299	2,418,404	-755	
Australia		95,810,318	75,206,328	-593	194	829,233	
Austria		-388,328	487,969	0		88,734	
Belgium		0	0	-3,594,299	15,666,961	-1,338,338	3,453,5
Brazil		0	0	18,483,306	15,369,360	18,715	
Brunei							
Canada							
China		12,642,194	35,175,045	121,302,358	-11,034	23,764,507	-14,7
Colombia		535,038	467,240	242,377	8,967,675	971,779	
France		819,039	-1,030,269	60,615,817	343,401,650	7,818,974	
Germany		-9,928,098	-52,717,507	21,625	-45,579	-2,064,001	14,7
Great Britain		-73,317	0	0		68,201	
Greece		7,121,313	3,822,522	616,510		394,582	
Hong Kong		40,471,401	18,863,260	58,273	68,740	-104,656,418	87,027,4
India							
Indonesia							

Insert a Table into a Report

To insert a table into a report:

- Drag the table icon from the **Objects** tab in the left pane and drop it onto the report canvas.
- Select **Insert ▶ Tables** and then select the menu item for the table object that you want to insert. The table is automatically placed in the report canvas. If you want the table to appear in a different place, then drag and drop it in a new location.

The following table lists the available table objects:

Icon	Table Type
	list table
	crosstab

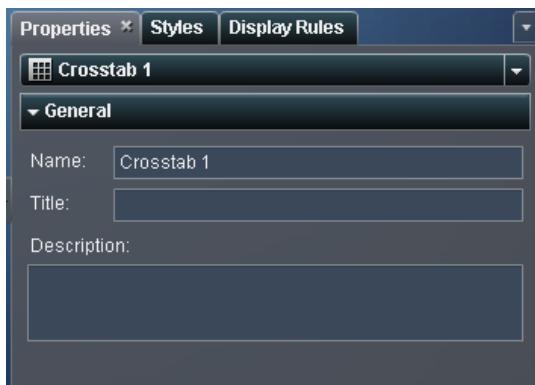
Specify a Name, Title, or Description for a Table

To specify general properties for tables:

- 1 If not already selected, select the table in the report canvas that you want to update.
- 2 In the right pane, click the **Properties** tab.
- 3 Update the general properties for the table. Your choices are **Name**, **Title**, and **Description**.

Here is an example of the properties for crosstabs:

Display 22.3 Properties for a Crosstab



Specify Table Styles

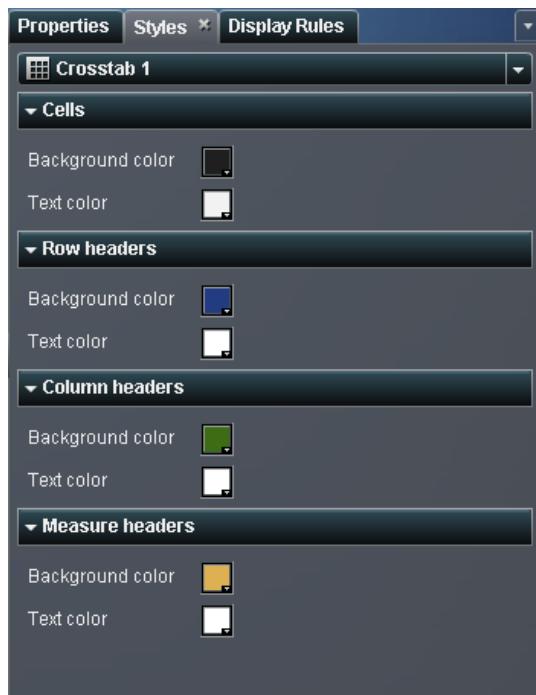
To specify styles for tables:

- 1 If not already selected, select the table in the report canvas that you want to update.
- 2 In the right pane, click the **Styles** tab.

3 Update the styles for the table. For a list table, you can customize the colors of **Cells** and **Column headers**. For crosstabs, you can customize the colors of **Cells**, **Row headers**, **Column headers**, and **Measure headers**.

Here is an example of the styles for crosstabs:

Display 22.4 Styles for a Crosstab



Using Graphs to Display Results

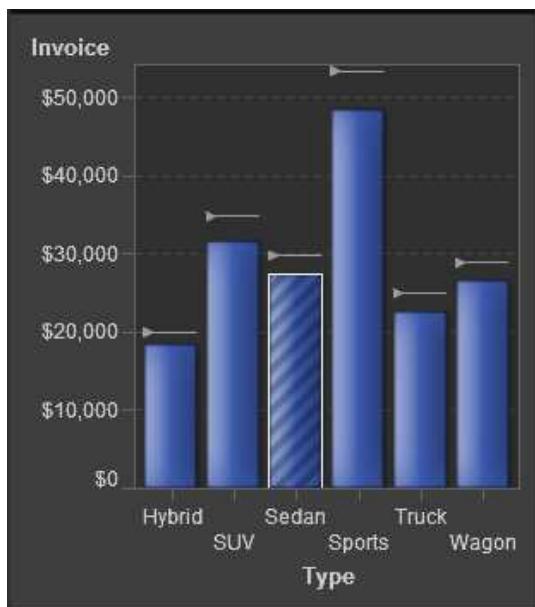
Overview of the Graph Types

About Bar Charts

A *bar chart* consists of a grid and some vertical or horizontal bars. Each bar represents quantitative data.

Display 22.5 A Bar Chart**About Targeted Bar Charts**

A *targeted bar chart* is a variation of the bar chart that has pointers to target values. In [Display 22.6 on page 192](#), the pointers appear above each bar.

Display 22.6 A Targeted Bar Chart

About Waterfall Charts

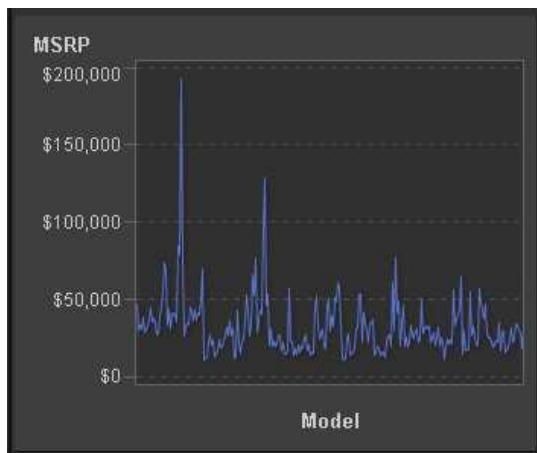
A *waterfall chart* shows how the initial value of a measure increases or decreases during a series of operations or transactions. (This chart is also called a progressive bar chart.) The first bar begins at the initial value, and each subsequent bar begins where the previous bar ends. The length and direction of a bar indicates the magnitude and type (positive or negative, for example) of the operation or transaction. The resulting chart is a stepped cascade that shows how the transactions or operations lead to the final value of the measure.

Display 22.7 A Waterfall Chart

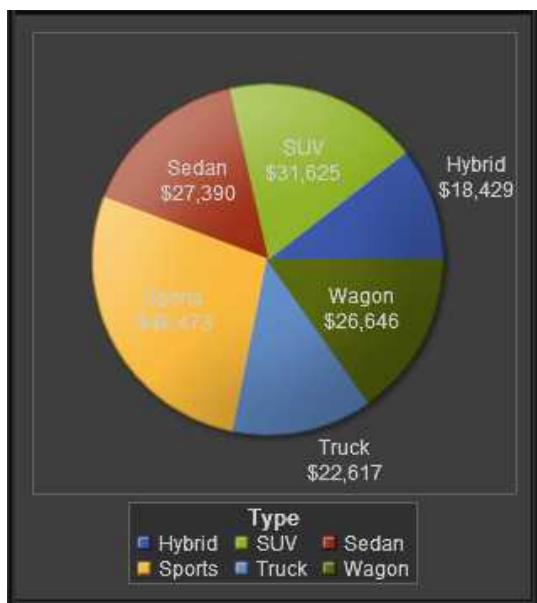


About Line Charts

A *line chart* shows the relationship of one variable to another, often as movements or trends in the data over a period of time. Line graphs summarize source data and typically are used to chart response values against discrete categorical values.

Display 22.8 A Line Chart**About Pie Charts**

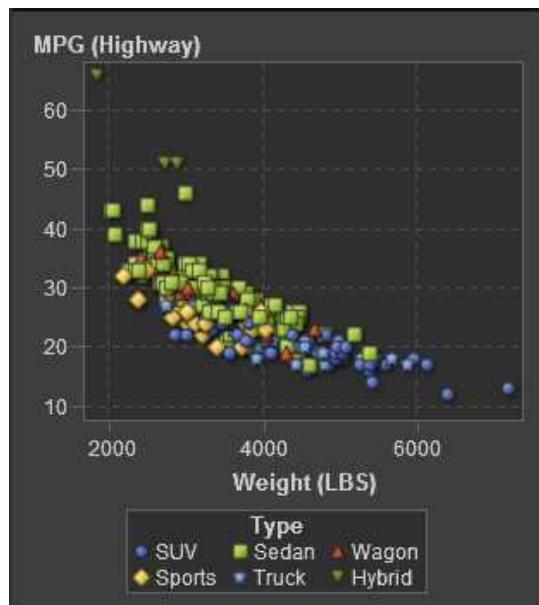
A *pie chart* is a circular chart that is divided into slices by radial lines. Each slice represents the relative contribution of each part to the whole.

Display 22.9 A Pie Chart

About Scatter Plots

A *scatter plot* is a two-dimensional plot that shows the joint variation of two data items. In a scatter plot, each marker (represented by symbols such as dots, squares, and plus signs) represents an observation. The marker position indicates the value for each observation.

Display 22.10 A Scatter Plot



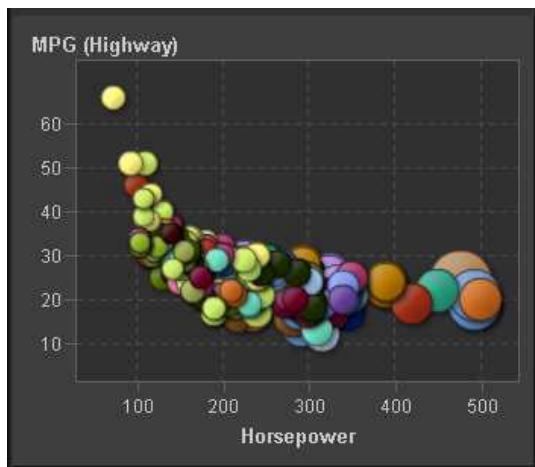
About Time Series Plot

A *time series plot* shows an ordered sequence of values of a discrete category variable that are observed at equally spaced time intervals.

Display 22.11 A Time Series Plot**About Bubble Plots**

A *bubble plot* is a variation of a scatter plot in which the markers are replaced with bubbles. In a bubble plot, each bubble represents an observation. The location of the bubble represents the value for the two measure axes; the size of the bubble represents the value for the third measure. A bubble plot is useful for data sets with dozens to hundreds of values or when the values differ by several orders of magnitude. You can also use a bubble plot when you want specific values to be visually represented by different bubble sizes.

Display 22.12 A Bubble Plot



About Treemaps

A *treemap* is divided into rectangular areas called tiles. The color of each tile represents the value of the first measure in the query. The size of each tile represents the value of the second measure in the query. For example, a treemap might be used to represent sales data where the tile sizes vary according to the number of orders invoiced and the tile colors are derived from a color gradient that represents low to high sales figures.

Display 22.13 A Treemap



About Dual Axis Bar Charts

A *dual axis bar chart* is a variation of the bar chart that has two measures, one on each axis. There can be one measure per axis for a total of two measures.

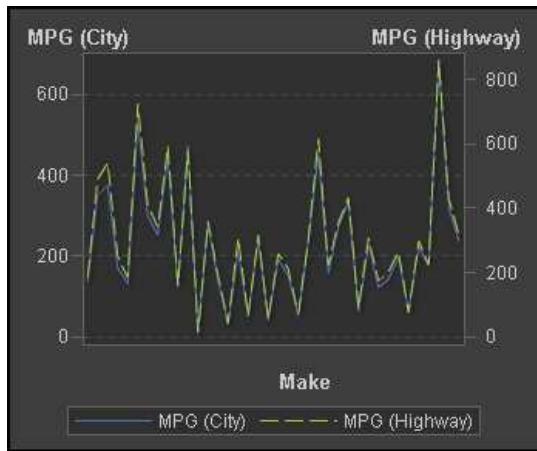
Display 22.14 A Dual Axis Bar Chart



About Dual Axis Line Charts

A *dual axis line chart* is a variation of the line chart that has two measures, one on each axis. There can be one measure per axis for a total of two measures.

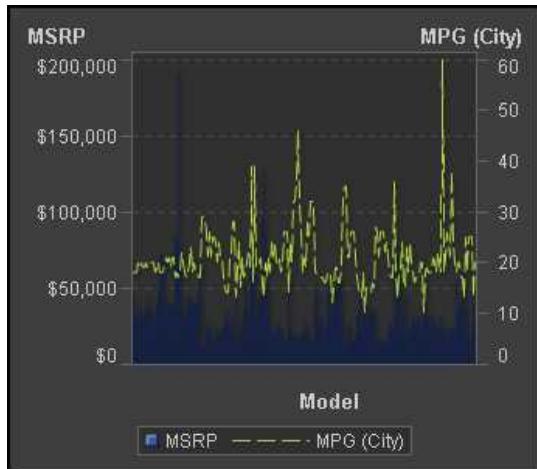
Display 22.15 A Dual Axis Line Chart



About Dual Axis Bar-Line Charts

A *dual axis bar-line chart* is a bar chart that has two measures, one on each axis, and is overlaid by a line chart. There can be one measure per axis for a total of two measures.

Display 22.16 A Dual Axis Bar-Line Chart



About Dual Axis Time Series Plots

A *dual axis time series plot* is a variation of the time series plot that has two measures, one on each axis. There can be one measure per axis for a total of two measures.

Display 22.17 A Dual Axis Time Series Plot



Insert a Graph into a Report

To insert a graph into a report:

- Drag the graph icon from the **Objects** tab in the left pane and drop it onto the report canvas.
- Select **Insert ▶ Graphs** and then select the menu item for the graph object that you want to insert. The graph is automatically placed in the report canvas. If you want the graph to appear in a different place, then drag and drop it in a new location.

The following table lists the available graph objects:

Icon	Graph Type
	bar chart
	targeted bar chart
	waterfall chart
	line chart

Icon	Graph Type
	pie chart
	scatter plot
	time series plot
	bubble plot
	treemap
	dual axis bar chart
	dual axis line chart
	dual axis bar-line chart
	dual axis time series plot

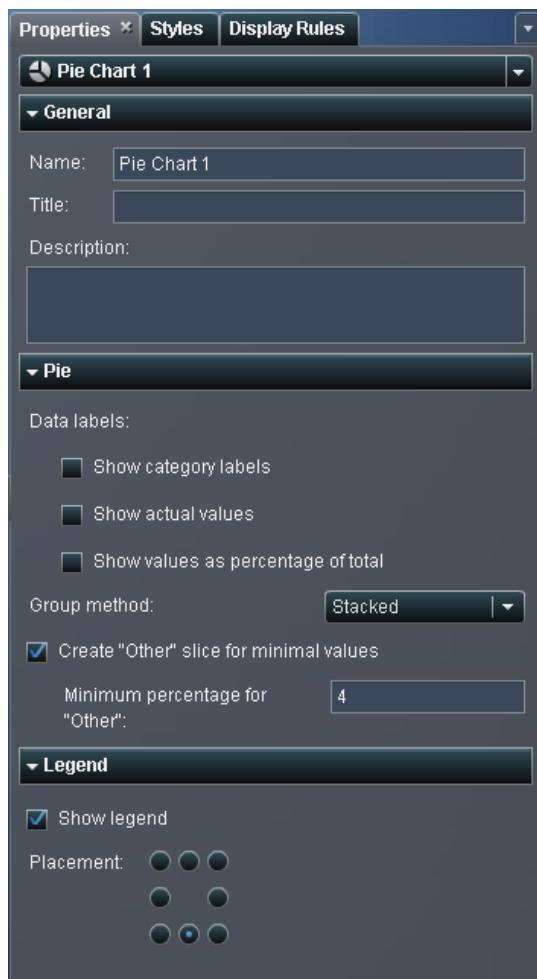
Specify a Name, Title, or Description for a Graph

To specify general properties for graphs:

- 1 If not already selected, select the graph in the report canvas that you want to update.
- 2 In the right pane, click the **Properties** tab.
- 3 Update the general properties for the graph. The available properties depend on the selected graph type. For example, for a pie chart, you can specify **General**, **Pie**, and **Legend** properties.

Here is an example of the properties for a pie chart:

Display 22.18 Properties for a Pie Chart



Specify Graph Styles

To specify styles for graphs:

- 1 If not already selected, select the graph in the report canvas that you want to update.
- 2 In the right pane, click the **Styles** tab.

3 Update the styles for the graph. The available styles depend on the selected graph type. For example, for a pie chart, you can specify **Data Styling**, **Frame Styling**, **Text Styling**, and **Data Colors** styles.

Here is an example of the styles for a pie chart:

Display 22.19 Styles for a Pie Chart



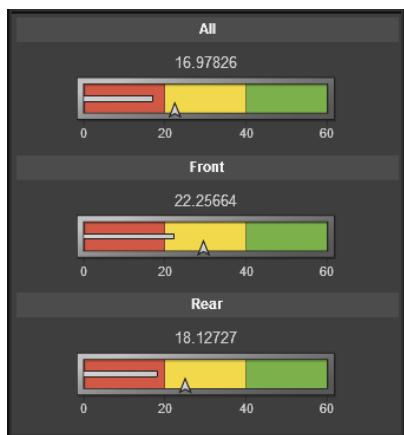
Using Gauges to Display Results

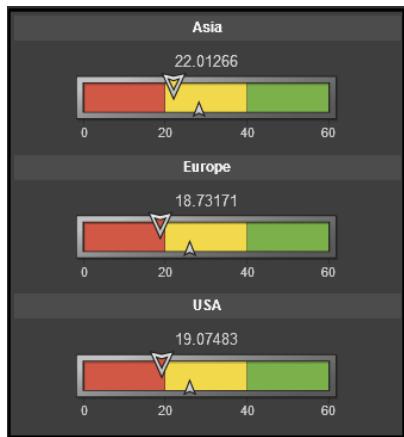
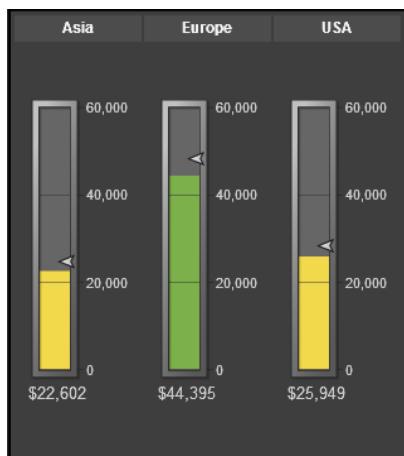
Overview of the Gauge Types

A gauge displays the status or measure of a variable or variables in relation to a target, goal, or interval. Gauges are designed to achieve this goal in a way that is familiar to users. Many real-life objects use gauges, such as cars and machines. Gauges can be used to display a quantity, range, variable, or status.

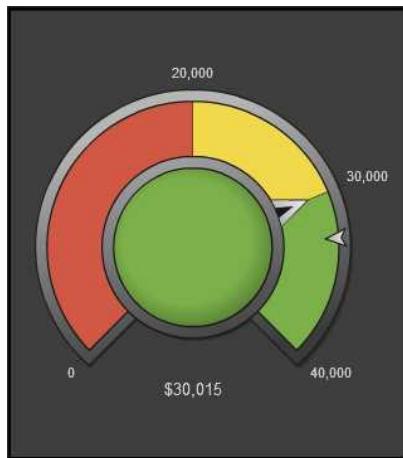
The following gauges are available in the designer interface:

Display 22.20 A Bullet Gauge



Display 22.21 A Slider Gauge**Display 22.22 A Thermometer Gauge**

Display 22.23 A Dial Gauge



Display 22.24 A Speedometer Gauge



Inserting a Gauge into a Report

To insert a gauge into a report:

- Drag the gauge icon from the **Objects** tab in the left pane and drop it onto the report canvas.
- Select **Insert ▶ Gauges** and then select the menu item for the gauge object that you want to insert. The gauge is automatically placed in the report canvas. If you

want the gauge to appear in a different place, then drag and drop it in a new location.

The following table lists the available gauge objects:

Icon	Gauge Type
	bullet
	slider
	thermometer
	dial
	speedometer

Specify a Name, Title, or Description for a Gauge

To specify general properties for gauges:

- 1 If not already selected, select the gauge in the report canvas that you want to update.
- 2 In the right pane, click the **Properties** tab.
- 3 Update the general properties for the gauge. The available properties depend on the selected gauge type. For example, for a dial gauge, you can specify **General** and **KPI Chart** properties.

Here is an example of the properties for a dial gauge:

Display 22.25 Properties for a Dial Gauge



Specify Gauge Styles

To specify styles for gauges:

- 1 If not already selected, select the gauge in the report canvas that you want to update.
- 2 In the right pane, click the **Styles** tab.
- 3 Update the styles for the gauge. The available styles depend on the selected gauge type. For example, for a pie chart, you can specify **Data Styling**, **Frame Styling**, and **Text Styling**.

Here is an example of the styles for a dial gauge:

Display 22.26 Styles for a Dial Gauge



Specify a Gauge Display Rule

A display rule is used by a gauge to designate intervals and colors for ranges.

To specify the display rule for gauges:

- 1 If not already selected, select the gauge in the report canvas that you want to update.
- 2 In the right pane, click the **Display Rules** tab.
- 3 Specify (or modify) the display rules for the gauge. You can populate intervals, edit the display rule, and specify the **Type**.

Here is an example of the display for a dial gauge:

Display 22.27 Display Rule for a Dial Gauge



To populate the intervals:

- a Click  to display the Populate Intervals window.

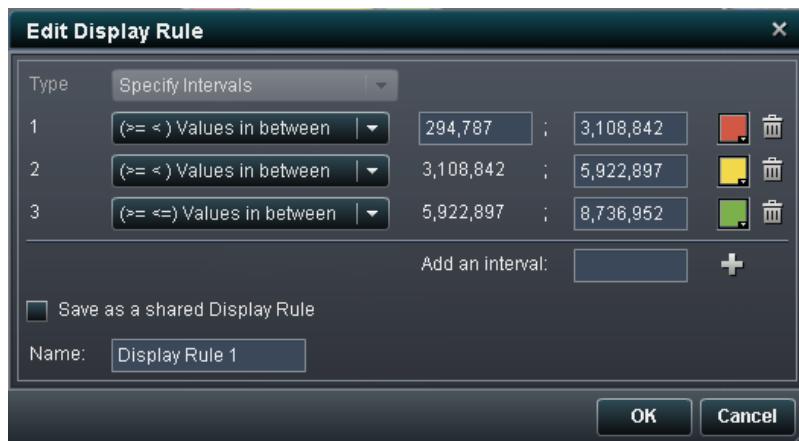
Display 22.28 Populate Intervals Window



- b Enter a number for **Target value**.
- c Click **OK**.

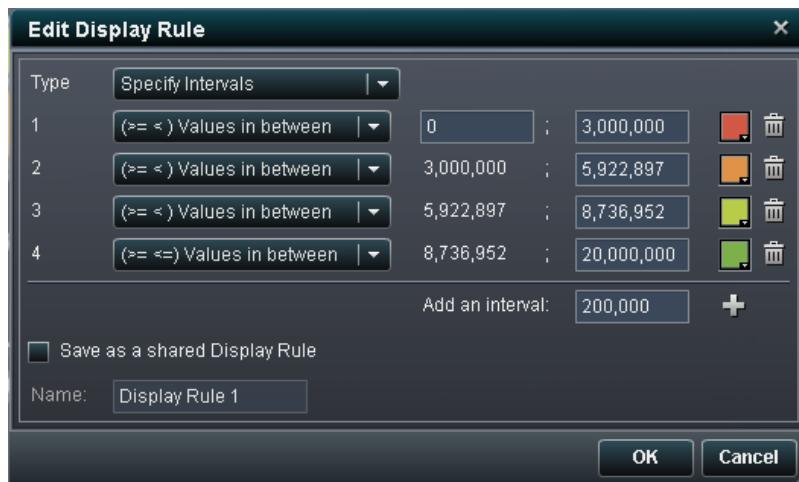
To edit the display rule:

- a Click  to display the Edit Display Rule window.

Display 22.29 Edit Display Rule Window

b Select a **Type**. Your choices are either **Use a Shared Display Rule** or **Specify Intervals**.

If you select **Specify Intervals**, then the window contents change so that you can add intervals. Here is an example:

Display 22.30 Specifying Intervals in the Edit Display Rules Window

Update the conditions for the values, the intervals, and the color for the range. Select a color to open the color picker. To add new intervals to the display rule, enter a number and click **Add**. Select the **Save as a shared Display Rule** check box and specify a **Name** if you want to save the display rule.

- c Click **OK**.

Adding Text and Images

Overview

About Text and Image Objects

Using text and images enhances the appearance of a report. You can insert images from a repository or from your local machine. If you select an image from your local machine, it is saved to the repository. You can also add tooltip text to an image. Text objects can be used to display static text.

Insert Text into a Report

To insert text into a report:

- 1
 - Drag  from the **Objects** tab in the left pane and drop it onto the report canvas.
 - Select **Insert** ▶ **Other** ▶ **Text**. The text object is automatically placed in the report canvas. If you want the text to appear in a different place, then drag and drop it in a new location.
- 2 Double-click inside the text object in the canvas and enter the text. You can use the floating toolbar to change the font, font size, text color, and text background color. You can also specify whether the text is bold, italic, or underlined and whether it is left-aligned, centered, or right-aligned.

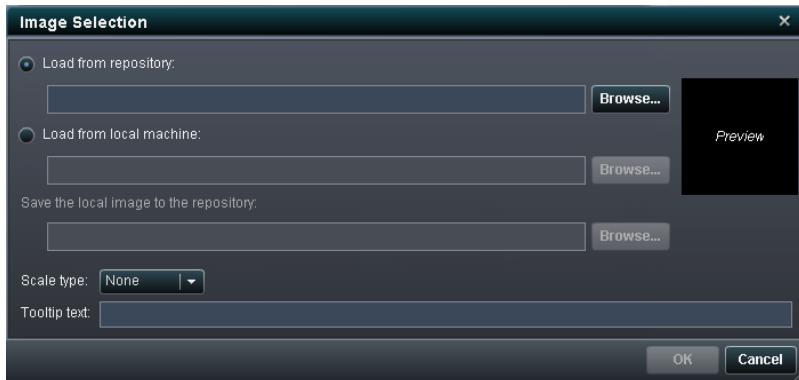
Insert an Image into a Report

To insert an image into a report:

- 1
 - Drag the  from the **Objects** tab in the left pane and drop it onto the report canvas. The Image Selection window is displayed.

- Select **Insert** ▶ **Other** ▶ **Image**. The Image Selection window is displayed. The image object is automatically placed in the report canvas. If you want the image to appear in a different place, then drag and drop it in a new location.

Display 22.31 Image Selection Window



- 2 Select the image from one of the following locations:

Load from repository

Select this option to choose an image that is stored on the same server as the reports.

Load from local machine

Select this option to choose an image from your local machine. Click **Browse** to choose a file on your local machine. Specify a repository in the **Save the local image to the repository** field. If you click **Browse**, the Save As window is displayed. Select a folder and then click **OK** to return to the Image Selection window.

A preview of the image is displayed.

- 3 (Optional) Specify the **Scale type**:

None

The actual size of the image is maintained. The image might or might not fill the entire area of the image's visual container. If the image is larger than the visual container, then scroll bars are displayed.

Stretch

The height and width of image are set to the height and width of the image's visual container. The image's original aspect ratio is not maintained.

Fit All

The image is modified to fit best into the image's visual container. The image's original aspect ratio is maintained.

Fit Width

The width of the image is set to the width of the image's visual container. The height maintains the image's original aspect ratio. Scroll bars are displayed if the set height of the image is greater than the height of the visual container.

Fit Height

The height of the image is set to the height of the image's visual container. The width maintains the image's original aspect ratio. Scroll bars are displayed if the set width of the image is greater than the width of the visual container.

Tile

The image is tiled in the visual container. The image's original size is maintained. There are no scroll bars.

4 (Optional) Specify the **Tooltip text**.

5 Click **OK**.

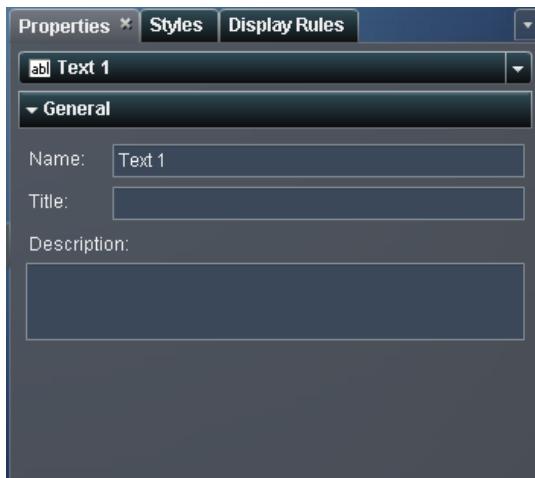
Specify a Name, Title, or Description for Text Objects

To specify general properties for text:

- 1 If not already selected, select the text in the report canvas that you want to update.
- 2 In the right pane, click the **Properties** tab.
- 3 Update the general properties for the text. Your choices are **Name**, **Title**, and **Description**.

Here is an example of the properties for a text object:

Display 22.32 Properties for a Text Object



Specify a Name, Title, or Description for Images

To specify general properties for an image:

- 1 If not already selected, select the image in the report canvas that you want to update.
- 2 In the right pane, click the **Properties** tab.
- 3 Update the general properties for the table. Your choices are **Name**, **Title**, and **Description**. Here is an example of the properties for an image:

Display 22.33 Properties for an Image Object



- 4 Update the image properties. Your choices are **Location**, **Scale type**, and **Tooltip text**.

Adding Comments to Report Objects

You must save a report before you can add comments to any report objects.

To add a comment to a report object:

- 1 Click the **Comments** tab in the right pane.

Display 22.34 Comments Tab

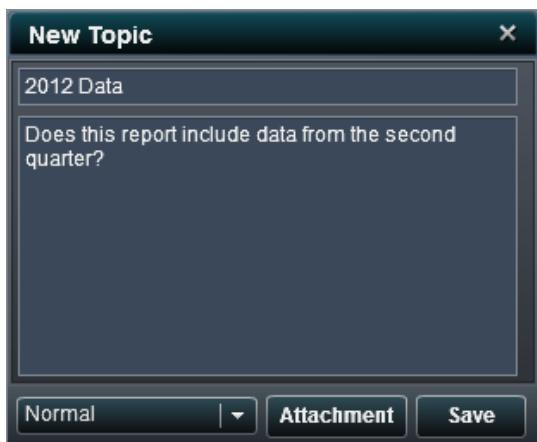


2 Click **New Topic** to open the New Topic window.

3 Enter a topic name and a comment.

Here is an example:

Display 22.35 New Topic Window for Adding Comments in the Designer Interface



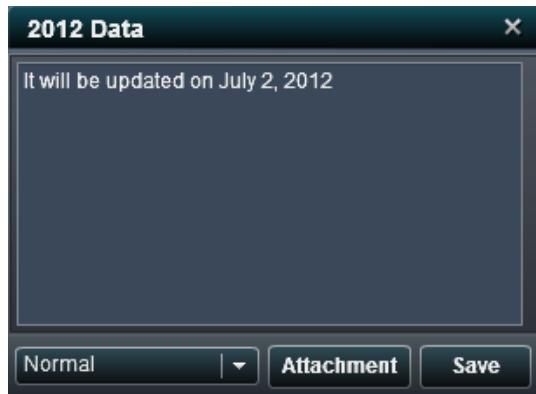
- 4 (Optional) Set a priority for the comment using the drop-down list. You can select **Normal**, **High**, or **Low**.
- 5 (Optional) Click **Attachment** to add a file or image to your comment.
- 6 Click **Save**. Your comment will be added to the **Comments** tab.

To respond to an existing comment:

- 1 Click the **Comments** tab in the right panel.
- 2 Click **Reply** to open the Comments window.
- 3 Enter a comment.

Here is an example:

Display 22.36 Window for Replying to a Comment



- 4 (Optional) Set a priority for the reply using the drop-down list. You can select **Normal**, **High**, or **Low**.
- 5 (Optional) Click **Attachment** to add a file or image to your reply.
- 6 Click **Save**. Your reply will be added to the **Comments** tab.

23

Sorting Data Items in Reports

<i>How Sorting Can Help with Analysis</i>	219
<i>Sorting Data in a List Table</i>	219
<i>Sorting Data in a Crosstab</i>	220
<i>Sorting Data in a Graph</i>	221

How Sorting Can Help with Analysis

Information can be easier to understand when it appears in an intentional order. Applying a sort order to one or more data items enables you to arrange rows and columns in tables and axis labels on charts in some order, such as alphabetically or highest to lowest numerically. Interactively changing the order of data can provide you with a different perspective that often facilitates valuable insight. For example, in a report, sales employees who are initially arranged alphabetically can be re-sorted by sales amount.

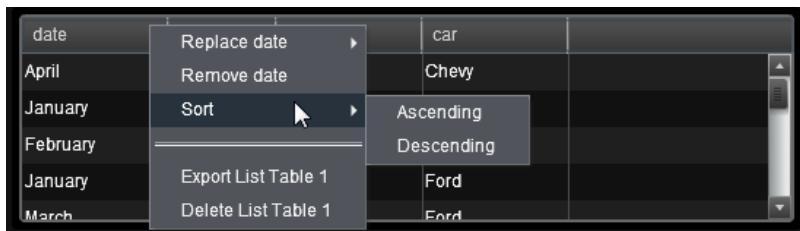
Sorting Data in a List Table

To sort values in a list table:

- 1 Select a column heading.

2 Right-click and select either **Ascending** or **Descending**.

Display 23.1 Sort Menu Item for List Tables



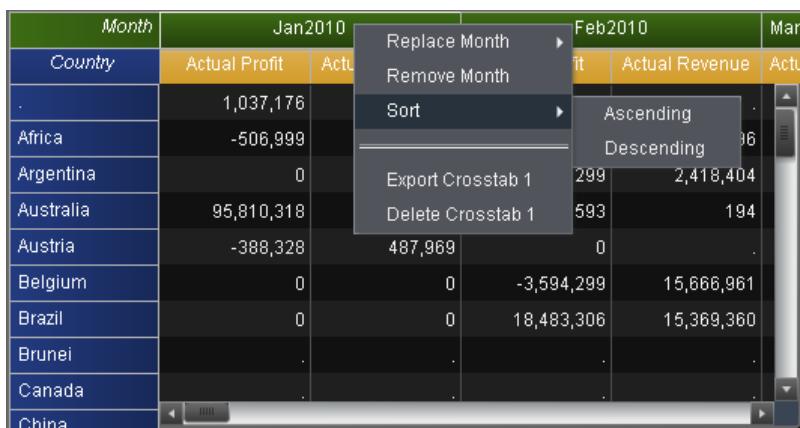
Sorting Data in a Crosstab

In crosstabs, right-click a column heading to sort data. You can also sort categories, rather than measures, by clicking the category name.

To specify a sort for values in a crosstab:

1 To sort by category values in a crosstab, right-click a column or row heading. Then select **Sort**. Then select either **Ascending** or **Descending**.

Display 23.2 Sort Menu for a Category in a Crosstab



2 To sort by measure values in a crosstab, right-click a column or row heading. Then select **Sort**. Then select either **Ascending by Value** or **Descending by Value**.

Display 23.3 Sort Menu for a Measure in a Crosstab



A screenshot of a Crosstab interface. The table has 'date' as the first column and 'Region' as the first row. The 'Region' row contains four cells with values: East (68,26), Midwest (85,32), Southeast (57,51), and West (87,06). The 'date' column has three visible cells: '31Mar2008', '30Jun2008', and '30Sep2008'. The cell for 'East' under '31Mar2008' is highlighted. A context menu is open over this cell, showing the following options: 'Replace Return_percent' (disabled), 'Remove Return_percent', 'Aggregation' (disabled), 'Sort' (selected), 'Ascending by Value', and 'Descending by Value'. Below the menu, the text 'Export Crosstab 1' and 'Delete Crosstab 1' are visible. The menu has a dark gray background with white text, and the selected 'Sort' option is highlighted with a dark blue background.

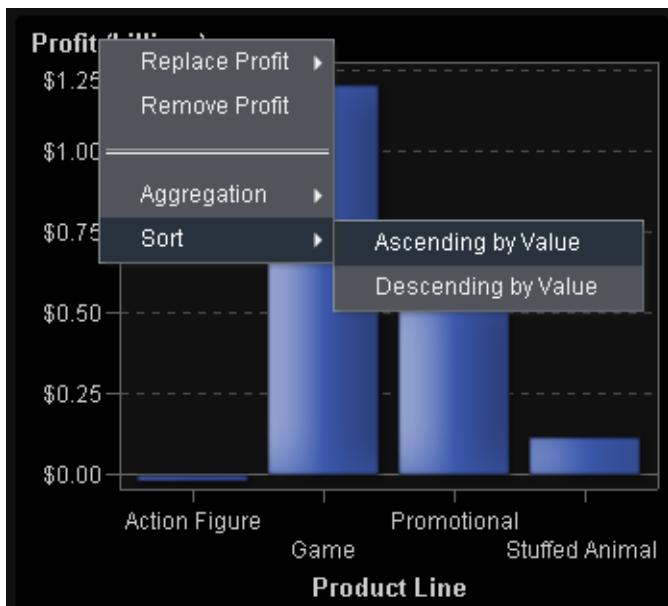
date	31Mar2008	30Jun2008	30Sep2008
Region	Return_percent	Return_dollars	Return_percent
East	68,26	3,504.88%	157,950,776.91
Midwest	85,32	6,051.25%	176,095,530.26
Southeast	57,51	2,261.67%	78,668,663.418
West	87,06	3,059.21%	176,095,530.26

Sorting Data in a Graph

To sort values in a graph:

- 1 To sort by category values in a graph, right-click a measure name and select **Sort**. Then select either **Ascending by Value** or **Descending by Value**.

Display 23.4 Sort Menu for a Category in a Graph



- 2 To sort by measure values in a graph, right-click a category name and select **Sort**. Then select either **Ascending Alphabetically** or **Descending Alphabetically**.

24

Working with Interactions

<i>Overview of Interactions</i>	223
<i>Creating Interactions</i>	224
Creating a Filter Interaction	224
Creating a Data Brushing Interaction	228
<i>Deleting Interactions</i>	229

Overview of Interactions

Interactions are used to direct a report viewer's attention to specific results in a report. Interactions allow data to be subset to reduce the amount of data, and enables users to understand it within a particular context.

The interaction view in the designer interface enables report authors to specify which interactions they would like to add to tables, graphs, and gauges in a report.

There are two types of interactions:

filter

is used to restrict the data that is returned from a query to a data source. Filters are simply a set of rules, or conditions, that you specify to subset the data that is displayed in a table or graph. The goal is to display only the data that you need to see to perform your analysis.

brush

is short for *data brushing*, which enables you to show the same data selected simultaneously in two or more tables, graphs, or both. Brushing highlights a percentage that reflects the number of shared observations in the data set. Brushing does not highlight a percentage that corresponds to the aggregated value. The brushed data has the same appearance in each object, which makes the data easily apparent to report viewers.

Typically, graphs are the source of an interaction, with the exception of scatter plots. Tables, graphs, and gauges can be the target for an interaction. However, tables, gauges, and scatter plots cannot be the source of an interaction.

Creating Interactions

Creating a Filter Interaction

You can create interactions using the interaction view or the **Interactions** tab in the right pane.

To create a filter interaction using the interaction view:

- 1 Add the report objects that you want to use to the report canvas. For example, you might have a bar chart, a line chart, and a list table.
- 2 Click  above the canvas to switch to interaction view.
- 3 In the interaction view, draw a connection between the source and target report objects. When you have created an interaction, the  is displayed between the report objects. Here is an example of an interaction between a bar chart (the source) and a list table (the target):

Display 24.1 Creating a Filter Interaction between a Bar Chart and a List Table



4 Right-click . Then select **Interaction Type** ▶ **Filter**.

Display 24.2 Interaction Type Menu

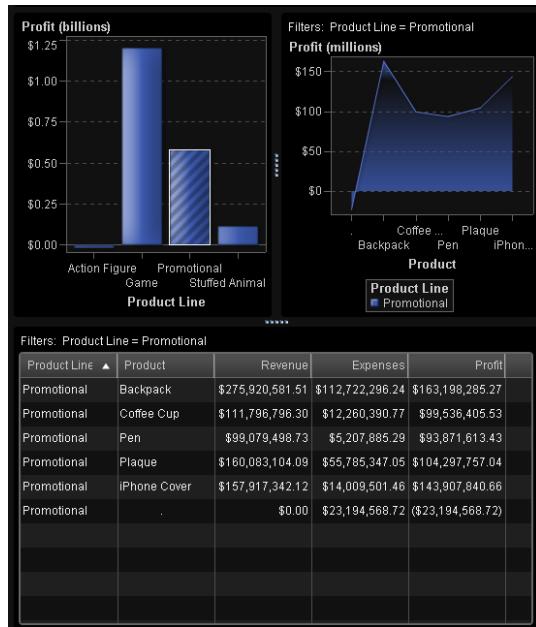


5 (Optional) Add other filter interactions.

6 Click to switch to the layout view. Select data in the source report object to filter data in the target report object (or objects). In the following example, the promotional

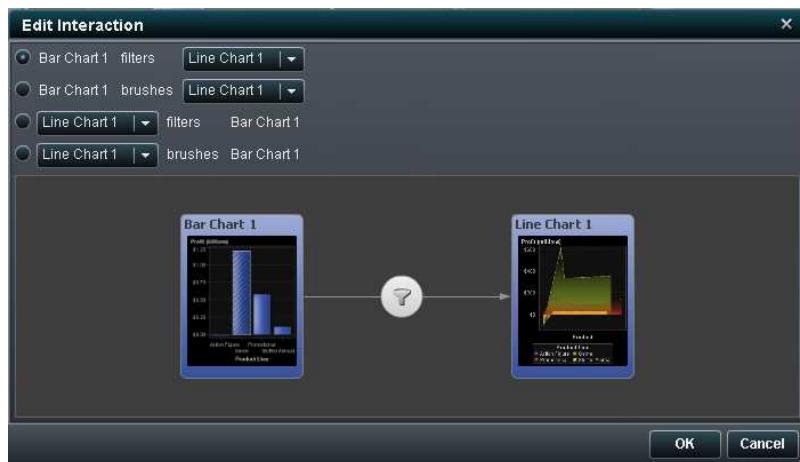
product line was selected in the bar chart. Simultaneously, the line chart and the list table show the same filter.

Display 24.3 A Filtering Example with a Bar Chart, a Line Chart, and a List Table



To create an interaction using the **Interactions** tab:

- 1 Select a report object in the canvas. For example, you might have a bar chart.
- 2 Select the **Interactions** tab.
- 3 Click . The **Edit Interaction** window is displayed.

Display 24.4 Edit Interaction Window

- 4 In the Edit Interaction window, select the report object that you want as the source of a filter or brush interaction. For example, if your report has a bar chart and you want it to filter a line chart, the first line should read *Bar Chart 1 filters Line Chart 1*.
- 5 When you are finished working with interactions, click **OK**.
- 6 The new interaction is displayed in the **Interactions** tab.

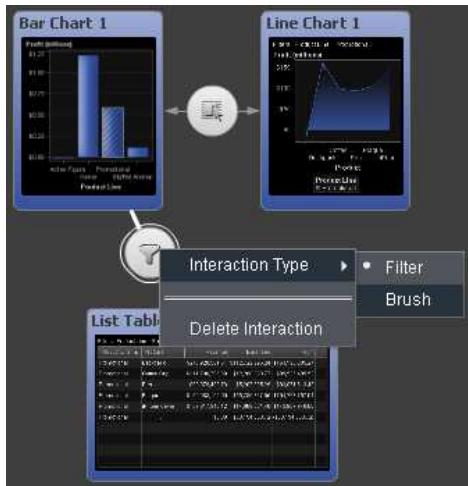
Display 24.5 A New Filter in the Interactions Tab

Creating a Data Brushing Interaction

To create an interaction using data brushing:

- 1 Add the report objects that you want to use to the canvas. For example, you might have a bar chart, a line chart, and a list table.
- 2 Click  above the canvas to switch to interaction view.
- 3 In the interaction view, draw a connection between the source and target report objects. When you have created an interaction, the  is displayed between the report objects.
- 4 Right-click . Then select **Interaction Type** ▶ **Brush**.

Display 24.6 Interaction Type Menu with Brush Item Selected



The filter icon changes to .

- 5 (Optional) Add other data brushing interactions.
- 6 Click  to switch to the layout view. Select data in the source report object to brush data in the target report object (or objects). In the following example, the stuffed animal product line was selected in the bar chart. Simultaneously, the pie chart and the crosstab highlight the same data.

Display 24.7 A Data Brushing Example with a Bar Chart, a Line Chart, and a List Table



Deleting Interactions

You can delete interactions using the interaction view or the **Interactions** tab in the right pane.

To delete either a filter or data brushing interaction using the interaction view:

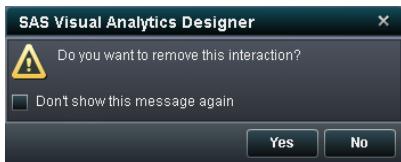
- 1 Click  above the canvas to switch to interaction view.
- 2 Right-click . Then select **Delete Interaction**.

Display 24.8 Interaction Type Menu with Delete Selected



- 3 Click **Yes** in the message that is displayed.

Display 24.9 Delete Interaction Confirmation Message



- 4 (Optional) Remove other interactions.

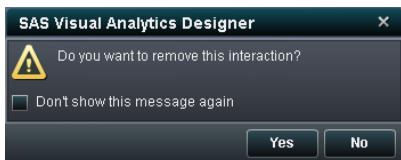
- 5 Click  to switch to the layout view.

To delete an interaction using the **Interactions** tab:

- 1 Click the **Interactions** tab.

Display 24.10 A New Filter in the Interactions Tab

- 2 Click .
- 3 Click **Yes** in the message that is displayed.

Display 24.11 Delete Interaction Confirmation Message

25

Maintaining Multi-Section Reports

<i>Overview of Report Sections</i>	233
<i>Adding a Section to a Report</i>	233
<i>Renaming a Report Section</i>	234
<i>Deleting a Section from a Report</i>	235
<i>Moving or Dragging a Report Object to Another Section</i>	235

Overview of Report Sections

Any report can have multiple sections. Multiple sections can be used to present different views of the data to the person who views the report. Each section has one or more data sources. There is no limit to the number of sections that can be included in a report.

Adding a Section to a Report

You can add sections to any report, including reports that were created in the exploration environment.

Here is an example of what the first section tab looks like.

Display 25.1 One Section Tab



To add a new section to a report, click  in the tab to the right of the first section in the report. The new tab appears to the right of the existing tab (or tabs).

Here is an example:

Display 25.2 Multiple Section Tabs

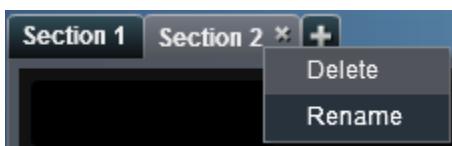


Renaming a Report Section

To rename a report section:

- 1 Right-click the section tab that you want to rename. A menu is displayed.

Display 25.3 The Menu Option for Renaming a Section



- 2 Select **Rename**. The name of the tab is highlighted.
- 3 Type a new name, and then press Enter. The new name also appears on the **Properties** tab in the right pane.

Deleting a Section from a Report

To delete a section from a report, right-click the section tab that you want to remove. When the menu is displayed, select **Delete**.

Note: The **Delete** option is not available if the report has only one section.

Moving or Dragging a Report Object to Another Section

To move a report object from one section in a report to another section, right-click the report object and select **Move to**. Then, select the name of the other section. You can also drag an object from one section and drop it onto the tab for another section to place it in that section.

26

Exporting Data from Report Objects

<i>Overview of Exporting from the Designer Interface</i>	237
<i>Exporting Data from Report Objects</i>	238
Export Data from a List Table	238
Export Data from a Crosstabulation Table	240
Export Data from a Graph	243

Overview of Exporting from the Designer Interface

The designer interface enables you to export data from report objects to Microsoft Excel format for future viewing or printing. This output can be saved locally on disk and then opened in Microsoft Excel.

When you export a graph, you are exporting the data not the visual graph representation.

Not all report objects in the designer interface support the exporting feature. For example, you cannot export data for gauges. If the export feature is not available for a particular report object, the **Export <reportObjectName>** menu item does not appear when you right-click in the object.

Exporting Data from Report Objects

Export Data from a List Table

To export data from a list table:

- 1 If not already selected, select the list table that you want to update.
- 2 Right-click in the list table and select **Export <listTableName>**, where **<listTableName>** is the name of the report object.

Here is an example of the **Export** menu option for a list table.

Display 26.1 Export Menu for List Tables



The screenshot shows a list table with four columns: Product Line, Product, Revenue, and Expense. The table contains data for various products across different categories. A context menu is open over the first row of the table, with the option 'Export List Table 1' highlighted. The menu also includes 'Delete List Table 1' and other standard list table options.

Product Line	Product	Revenue	Expense
Action Figure	Musician	\$27,364,281.52	\$21,977,144.9
Action Figure	Export List Table 1	\$13,647.13	\$22,326,021.5
Action Figure		\$0.00	\$125,851,856.0
Action Figure	Soldier	\$37,201,659.56	\$22,215,934.1
Action Figure	Firefighter	\$38,023,231.27	\$22,379,251.6
Action Figure	Police	\$37,153,375.28	\$21,788,384.6
Action Figure	Super Hero	\$37,624,002.02	\$22,351,722.9
Action Figure	Movie Star	\$37,409,043.75	\$22,347,912.7
Game	Puzzle	\$472,919,309.69	\$117,551,980.2
Game	Card	\$400,034,569.40	\$67,100,375.5
Game	Board	\$808,701,412.21	\$194,615,006.7
Game		\$0.00	\$98,927,868.2

The Export or Save As window is displayed.

Display 26.2 Export or Save As Window for List Tables



- 3 Choose one of the following export options for rows:
 - Select the **All rows** radio button to export all rows.
 - Select the **Rows** radio button to specify a range of rows to export. Enter a number in the **From** and **To** fields.
- 4 Choose one of the following export options for columns:
 - Select the **All columns** radio button to export all columns.
 - Select the **Selected columns** radio button to specify which columns to export. Then, select the check box to the left of the column (or columns) that you want to export. At least one column is required. If you do not select a column, then a message is displayed and the Export or Save As window cannot be closed.

- 5 (Optional) To choose whether the exported data is formatted, either select or clear the **Formatted data** check box. This check box is selected by default.
- 6 If you select the **Export to** radio button, then the only option is **Excel 2007 Workbook (*.xlsx)** to create a Microsoft Excel spreadsheet.
- 7 If you select the **Save as** radio button, choose one of the following options:
 - Select **Tab-Separated Values (.tsv)** to create a data file.
 - Select **Comma-Separated Values (.csv)** to create a data file.
- 8 Click **OK**.
- 9 When you are prompted, choose either to open the file or to save it.

Export Data from a Crosstabulation Table

To export detail data from a crosstab:

- 1 If not already selected, select the crosstab that you want to update.
- 2 Right-click in the crosstab and select **Export <crosstabName>**, where **<crosstabName>** is the name of the report object.

Here is an example of the **Export** menu option for a crosstab.

Display 26.3 Export Menu for Crosstabs

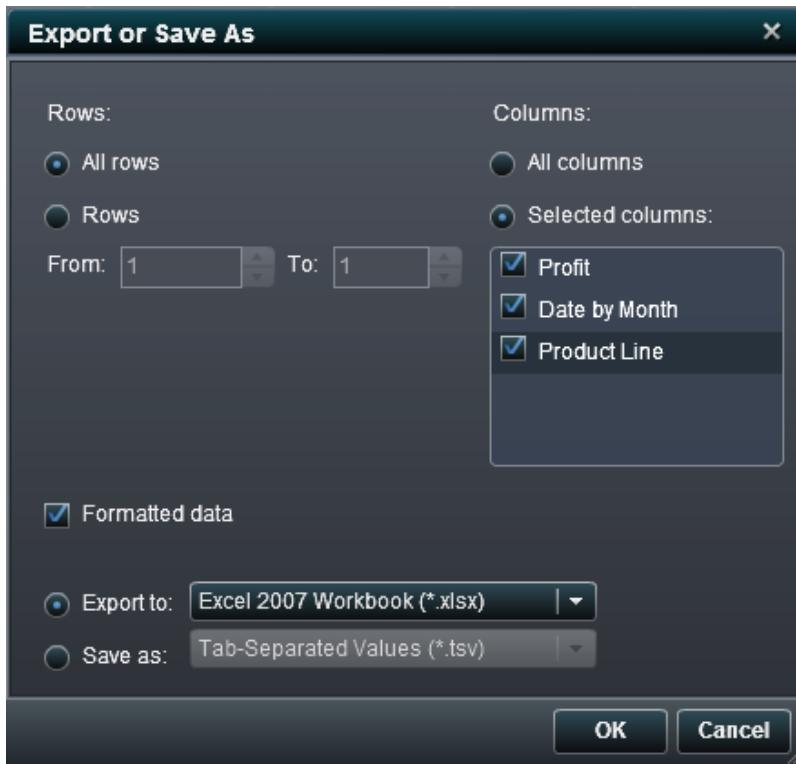


A screenshot of a Microsoft Access application showing a crosstab query. The crosstab has 'Date by Year' on the top row and 'Product Line' on the left column. The data is organized into three columns: '2010', '2011', and 'Revenue'. The 'Revenue' column contains values for four product lines: Action Figure, Game, Promotional, and Stuffed Animal. A context menu is open over the 'Revenue' cell for 'Promotional', showing options 'Export Crosstab 1' and 'Delete Crosstab 1'. The background of the application shows a dark blue ribbon and a toolbar at the bottom.

Date by Year		2010	2011	
Product Line		Revenue	Expenses	Revenue
Action Figure		\$10,151,171.85	\$17,409,398.39	\$12,270,844.
Game		\$64,299,269.35	\$21,264,482.96	\$70,327,582.
Promotional		\$28,747,506.64	\$	Export Crosstab 1
Stuffed Animal		\$7,566,435.64	\$	Delete Crosstab 1

The Export or Save As window is displayed.

Display 26.4 Export or Save As Window for Crosstabs



- 3 Choose one of the following export options for rows:
 - Select the **All rows** radio button to export all rows.
 - Select the **Rows** radio button to specify a range of rows to export. Enter a number in the **From** and **To** fields.
- 4 Choose one of the following export options for columns:
 - Select the **All columns** radio button to export all columns.
 - Select the **Selected columns** radio button to specify which columns to export. Then select the check box to the left of the column (or columns) that you want to export. At least one column is required. If you do not select a column, then a message is displayed and the Export or Save As window cannot be closed.
- 5 (Optional) To choose whether the exported data is formatted, either select or clear the **Formatted data** check box. This check box is selected by default.

- 6 If you select the **Export to** radio button, then the only option is **Excel 2007 Workbook (*.xlsx)** to create a Microsoft Excel spreadsheet.
- 7 If you select the **Save as** radio button, choose one of the following options:
 - Select **Tab-Separated Values (.tsv)** to create a data file.
 - Select **Comma-Separated Values (.csv)** to create a data file.
- 8 Click **OK**.
- 9 When you are prompted, choose either to open the file or to save it.

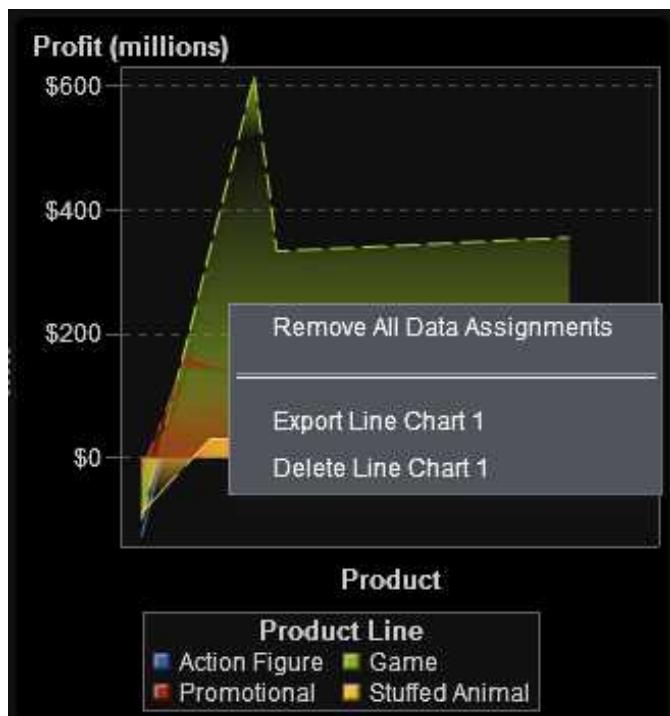
Export Data from a Graph

To export data from a graph:

- 1 If not already selected, select the graph that you want to update.
- 2 Right-click in the graph and select **Export <graphName>**, where **<graphName>** is the name of the report object.

Here is an example of the **Export** menu option for a line chart.

Display 26.5 Export Menu for Graphs



The Export or Save As window is displayed.

Display 26.6 Export or Save As Window for Graphs



- 3 Choose one of the following export options for rows:
 - Select the **All rows** radio button to export all rows.
 - Select the **Rows** radio button to specify a range of rows to export. Enter a number in the **From** and **To** fields.
- 4 Choose one of the following export options for columns:
 - Select the **All columns** radio button to export all columns.
 - Select the **Selected columns** radio button to specify which columns to export. Then select the check box to the left of the column (or columns) that you want to export. At least one column is required. If you do not select a column, then a message is displayed and the Export or Save As window cannot be closed.

- 5 (Optional) To choose whether the exported data is formatted, either select or clear the **Formatted data** check box. This check box is selected by default.
- 6 If you select the **Export to** radio button, then the only option is **Excel 2007 Workbook (*.xlsx)** to create a Microsoft Excel spreadsheet.
- 7 If you select the **Save as** radio button, choose one of the following options:
 - Select **Tab-Separated Values (.tsv)** to create a data file.
 - Select **Comma-Separated Values (.csv)** to create a data file.
- 8 Click **OK**.
- 9 When you are prompted, choose either to open the file or to save it.

Part 5

Viewing Reports

Chapter 27

Viewing Reports on a Mobile Device 249

Chapter 28

Viewing Reports in the Web Viewer 251

27

Viewing Reports on a Mobile Device

<i>Opening the App for the First Time</i>	249
<i>Connecting to Your Library Server for the First Time</i>	250

Opening the App for the First Time

You can use the SAS Mobile BI for iPad app to view SAS Visual Analytics reports on your mobile device. The SAS Mobile BI app is available in the iTunes App store.

When you open the app for the first time, the following tutorials are available in My Portfolio:

- Introduction
- Explore the Portfolio
- Explore the Library

Tap each tutorial to view information about the features provided by the app and how to use the app. Additional tutorials are available in the Library.

The app also provides help that you can view in My Portfolio. Tap **Help** ▶ **Get Help** to learn how to access more tutorials in the Library.

Connecting to Your Library Server for the First Time

To connect to a new server:

- 1 Tap **Library** in My Portfolio.
- 2 Tap **Configure and Connect** in the Welcome window.
- 3 Tap the **User ID** field and enter your user ID.

Note: Contact your SAS Visual Analytics administrator for your user ID, password, and server information.

- 4 Tap the **Password** field and enter your password.
- 5 Tap the **Server** field and enter the new server.
- 6 Tap the **Description** field and enter a description for the new connection.
- 7 Tap **Next** to verify the connection.
- 8 Tap **Save**. The connection is saved and the Connections window closes.

Note: If the connection fails, swipe the Connections window to scroll up. Make sure that the connection settings are correct. Your SAS Visual Analytics administrator can confirm these settings for you.

Information about available reports is automatically downloaded and appears in the Library.

28

Viewing Reports in the Web Viewer

<i>Overview of Viewing Reports in the Web Viewer</i>	251
<i>Opening a Report in the Web Viewer</i>	251
<i>Adding Comments to a Report in the Web Viewer</i>	253

Overview of Viewing Reports in the Web Viewer

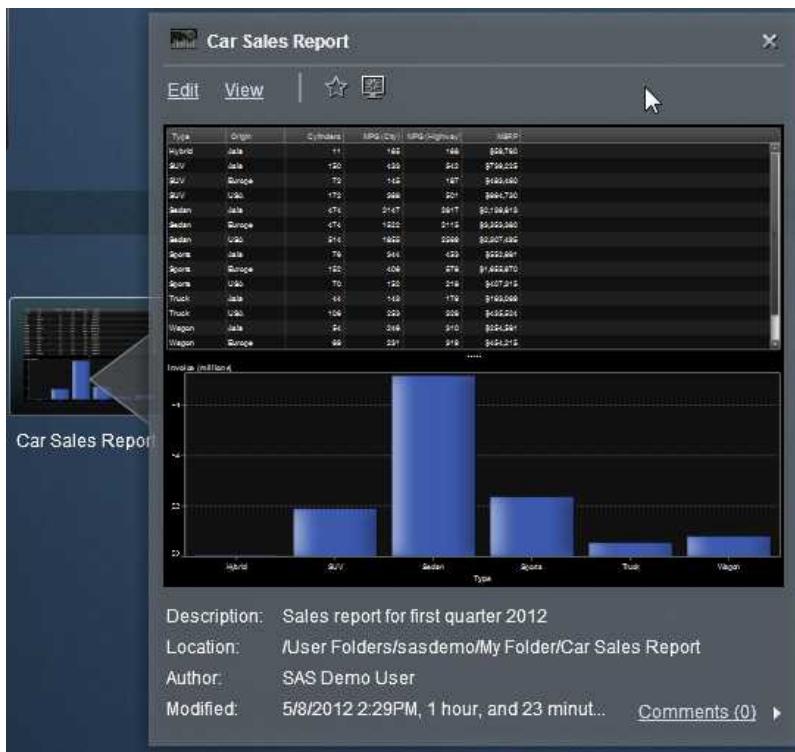
As an alternative to viewing reports on a mobile device, you can use the Web viewer. For users with a Report Viewing role, the Web viewer enables them to view report content.

Opening a Report in the Web Viewer

To open a report in the Web viewer:

- Double-click a report on the SAS Visual Analytics home page.
- Use the object inspector on the home page to view details about the report and then click **View**.

Display 28.1 The Object Inspector on the Home Page



Here is an example of a report in the Web viewer:

Display 28.2 A Report in the Web Viewer



Adding Comments to a Report in the Web Viewer

You cannot make or save changes to reports in the Web viewer. However, you can add comments to a report. The comments that you add are automatically saved with the report.

To add a comment:

- 1 Click the **Comments** tab in the right pane.
- 2 Click **New Topic** to open the New Topic window.

- 3 Enter a topic name and a comment. Here is an example:

Display 28.3 New Topic Window for Adding Comments in the Web Viewer



- 4 (Optional) Set a priority for the comment using the drop-down list. You can select **Normal**, **High**, or **Low**.
- 5 (Optional) Click **Attachment** to add a file or image to your comment.
- 6 Click **Save**. Your comment is added to the **Comments** tab in the right pane.

To respond to an existing comment:

- 1 Click the **Comments** tab in the right pane.
- 2 Click **Reply**.
- 3 Enter a comment. Here is an example:

Display 28.4 Window for Replying to a Comment



- 4 (Optional) Set a priority for the reply using the drop-down list. You can select **Normal**, **High**, or **Low**.
- 5 (Optional) Click **Attachment** to add a file or image to your reply.
- 6 Click **Save**. Your reply is added to the **Comments** tab in the right pane.

Part 6

Managing the Environment

<i>Chapter 29</i>	
<i>Working with HDFS</i>	257
<i>Chapter 30</i>	
<i>Monitoring Resources</i>	271
<i>Chapter 31</i>	
<i>Monitoring Processes</i>	275
<i>Chapter 32</i>	
<i>Managing Mobile Devices</i>	277

29

Working with HDFS

<i>What Is the HDFS Content Explorer?</i>	257
<i>What Can I Do with the HDFS Content Explorer?</i>	258
<i>Navigating HDFS Folders</i>	259
<i>Adding a Table to HDFS</i>	259
<i>Deleting a Table from HDFS</i>	260
<i>Viewing HDFS System Properties</i>	260
<i>Viewing Basic File Information</i>	262
<i>Viewing Table Information</i>	264
View Columns	264
View Row Count	265
<i>Viewing Block Information</i>	266
View Block Details	266
View Block Distribution	267

What Is the HDFS Content Explorer?

The HDFS content explorer is used to navigate through HDFS folders and to manage prepared data that is stored in HDFS. It provides a tree-based view of the HDFS folder structure and a table-based view of the details for the prepared data. The HDFS content

explorer is available for deployments that use SAS Visual Analytics Hadoop as the co-located data provider only.

The HDFS content explorer is shown in the following display:

Display 29.1 HDFS Content Explorer

Property	Value
Path	/departmentB
Copies	0
Block Size	0
Owner	admhpint
Group	deptb
Permissions	rwxrwx---
SASHDAT file?	No

The following sections provide more information about using the HDFS content explorer. It is launched from the data preparation interface by clicking **HDFS** from the toolbar.

Note: If the HDFS content explorer does not work and does not permit navigating HDFS, then contact your system administrator to make sure the SAS LASR Analytic Server Monitor is running.

What Can I Do with the HDFS Content Explorer?

A key feature of the data preparation interfaces is to prepare data for analysis. The last stage of the data preparation activity is to add the prepared data to HDFS so that it can be loaded into memory on the SAS LASR Analytic Server. Analysts can then explore the data from the SAS Visual Analytics explorer interface. For deployments that have data that is already prepared for analysis, the HDFS content explorer enables administrators to add that data to HDFS directly. The HDFS content explorer also

enables administrators to view information about the prepared data such as the row count, columns, and column information.

When tables are added to HDFS with the data preparation interface, they are stored with a SASHDAT file suffix. This is a special file format used by the SAS LASR Analytic Server. This special file format and the data redundancy provided by SAS Visual Analytics Hadoop enable the SAS LASR Analytic Server to read the data in parallel at very impressive rates. The data that is stored in HDFS is stored in blocks. The HDFS content explorer enables an administrator to view the block distribution, block redundancy, and measures of block utilization.

Navigating HDFS Folders

The HDFS content explorer provides a familiar and easy to use interface for navigating the files and folders in HDFS. You can navigate by clicking the folder icons in the folder tree and by using the following navigation buttons:

Back 

Click this button to return to the previously used folder.

Up one level 

Click this button to navigate to the parent folder.

Adding a Table to HDFS

To add a table to HDFS:

- 1 Click **HDFS** in the toolbar to access the HDFS content explorer.
- 2 Click +, select a table, and click **OK**.
- 3 On the General page, specify values for the **Job name**, **Location**, and **Description** if you intend to save this action as a job.

- 4 Select the Settings page, specify the **HDFS output path** and **Table description**, and then click **Submit**.

Note: If you refresh your view as a table is added to HDFS, the HDFS content explorer might show many files in the directory for the table. The filenames follow a **tablename-ip-address.sashdat** pattern. This is temporary and when the table is completely added to HDFS, the additional files do not appear.

Deleting a Table from HDFS

To delete a table from HDFS:

- 1 Click **HDFS** in the toolbar to access the HDFS content explorer.
- 2 Select the file to delete and click .
- 3 On the General page, specify values for the **Job name**, **Location**, and **Description** if you intend to save this action as a job.
- 4 Select the Settings page, review the HDFS filename to delete, and then click **Submit**.

Files that are not SASHDAT files are listed in the HDFS content explorer, but they cannot be deleted.

Viewing HDFS System Properties

To view the HDFS system properties:

- 1 Click **HDFS** in the toolbar to access the HDFS content explorer.
- 2 Click .

The following table describes the fields:

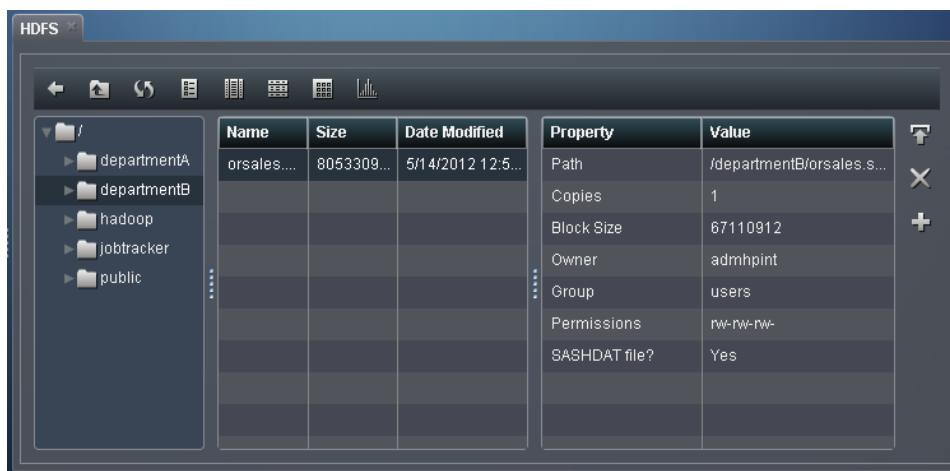
Table 29.1 HDFS System Properties

Property	Description
Command for setting permissions	Specifies the path to the setperm command. This command is used to enable security for files stored in HDFS by setting access permission.
Set permissions as root?	Specifies whether access permissions are enabled. For a SAS Visual Analytics deployment, the value is Yes .
Command for getting file information	Specifies the list command and options. For a SAS Visual Analytics deployment, the value is <code>ls -l {0}</code> .
Data directories	Specifies the directory that is used to store blocks.
NameNode	Specifies the host name of the machine that is used as the SAS Visual Analytics Hadoop NameNode.
Live DataNodes Dead DataNodes	Specifies the number of SAS Visual Analytics Hadoop DataNodes that are reachable and the number that are not available.

Viewing Basic File Information

To view basic file information, select a file in the HDFS content explorer. The following display shows an example:

Display 29.2 Basic File Information



The following information about a file is provided:

Table 29.2 Basic File Information

Field	Description
Name	Specifies the name of the file.
Size	Specifies the file size. This value includes the disk space required to store the data in blocks and metadata about the file.
Date Modified	Specifies the date that the file was created or replaced.
Path	Specifies the HDFS directory.

Field	Description
Description	Specifies the description that is stored with the data. The description is displayed beside the table name in the explorer interface.
Number of copies	Specifies the number of redundant copies of the data.
Block size	Specifies the number of bytes that are used to store each block of data.
Owner	Specifies the user account that added the data to HDFS.
Group	Specifies the primary UNIX group for the user account.
Permissions	Specifies the Read, Write, and Execute access permissions for owner, group, and other.
SASHDAT file?	Specifies that the file is in the special format that is created by the data preparation interface.

Note: The HDFS content explorer can show many files for a table as it is added to HDFS. These entries disappear once the action completes.

Viewing Table Information

View Columns

To view column information, select a table in HDFS and then click . The following display shows an example:

Display 29.3 Column Information

Columns: /departmentB/orsales.sashdat								
Column Name	Label	Type	Offset	Length	Format	Format length	Precision	Length (Formatted)
Year	Year	1	0	8	BEST	0	0	12
Quarter	Quarter	2	8	6	\$	0	0	6
Product_Line	Product Line	2	16	20	\$	0	0	20
Product_Cate...	Product Categ...	2	40	25	\$	0	0	25
Product_Group	Product Group	2	72	25	\$	0	0	25
Quantity	Number of It...	1	104	8	F	6	0	6

The following information about each column is provided:

Table 29.3 Column Information

Field	Description
Column Name	Specifies the column name from the source table.
Label	Specifies the label for the data set column when the table was added to HDFS.
Type	Numeric or Character. Numeric variables are encoded as 1.
Offset	Specifies the starting position for the variable in the SASHDAT file.
Length	Specifies the storage used by the variable.
Format	Specifies the format associated with the variable.

Field	Description
Format length	Specifies the format length of the format that existed on the variable when it was added to HDFS. This value is zero if the variable did not have a format when it was added to HDFS.
Precision	Specifies the precision portion of the format for number formats.
Length (Formatted)	Specifies the length of the variable when formatting is applied.

View Row Count

To view the row count, select a table in HDFS and then click .

Row Count: /departmentB/orsales.sashdat				
Rows	Blocks	Allocated	Used	Utilization%
4561824	12	805306368	583913472	72.508

The following information is provided:

Table 29.4 Row Count Information

Field	Description
Rows	Specifies the number of rows in the data.
Blocks	Specifies the number of HDFS blocks that are used to store the data.

Field	Description
Allocated	Specifies the number of bytes allocated to store the data. It is a multiple of the block size and the number of blocks. This value is smaller than the file size because it does not include the space needed for the SASHDAT file header.
Used	Specifies the number of bytes within the allocated blocks that are used for storing rows of data.
Utilization	Specifies the percentage of allocated space that is used for storing rows of data.

Viewing Block Information

View Block Details

To view block details, select a file in HDFS and click .

Block Details: ./departmentB/orsales.sashdat								?
Host Name	Block Name	Path	Record Length	Records	Owner	Group	Permissions	
halfpint7.unx.s...	blk_92016779...	/hadoop/hado...	128	524288	admhpint	users	-rW-rW-rW-	
halfpint2.unx.s...	blk_92016779...	/hadoop/hado...	128	524288	admhpint	users	-rW-rW-rW-	
halfpint2.unx.s...	blk_59750104...	/hadoop/hado...	128	236016	admhpint	users	-rW-rW-rW-	
halfpint5.unx.s...	blk_59750104...	/hadoop/hado...	128	236016	admhpint	users	-rW-rW-rW-	
halfpint3.unx.s...	blk_-7011788...	/hadoop/hado...	128	524288	admhpint	users	-rW-rW-rW-	
halfpint6.unx.s...	blk_-7011788...	/hadoop/hado...	128	524288	admhpint	users	-rW-rW-rW-	

Close

The following information about each block is provided:

Table 29.5 Block Details Information

Field	Description
Host Name	Specifies the machine in the cluster that stores the block of data.
Block Name	Specifies the filename for the block.
Path	Specifies the directory to the block.
Record Length	Specifies the sum of the column lengths for the variables in the data.
Records	Specifies the number of rows stored in the block. Because redundant blocks are listed in the table, the sum of the records listed does not equal the number of rows in the data.
Owner	Specifies the user account that added the data to HDFS.
Group	Specifies the primary UNIX group for the user account that stored the data.
Permissions	Specifies the Read, Write, and Execute access permissions for owner, group, and other.

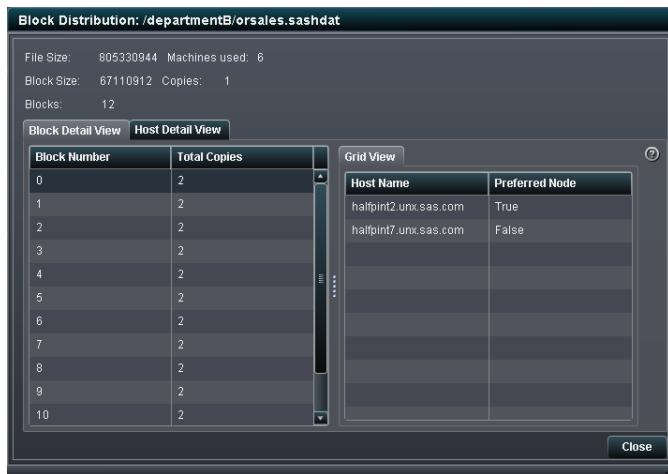
You can sort by the column headings to identify anomalies. It is normal for several blocks to be stored on the same machine. However, it is not normal for the values of **Record Length**, **Owner**, **Group**, or **Permissions** to be different from row to row.

View Block Distribution

The files added to HDFS are stored as blocks. One block is the preferred block, and additional copies of the blocks are used to provide data redundancy. The Block Distribution dialog box offers two ways to view this information. The **Block Detail View** tab enables you to select a block number and view the host names that store the

original or redundant blocks. The **Node Detail View** enables you to select a host name and view the block numbers that are stored on the machine.

To view the block distribution, select a table in HDFS and then click .



The following information about the block distribution is provided:

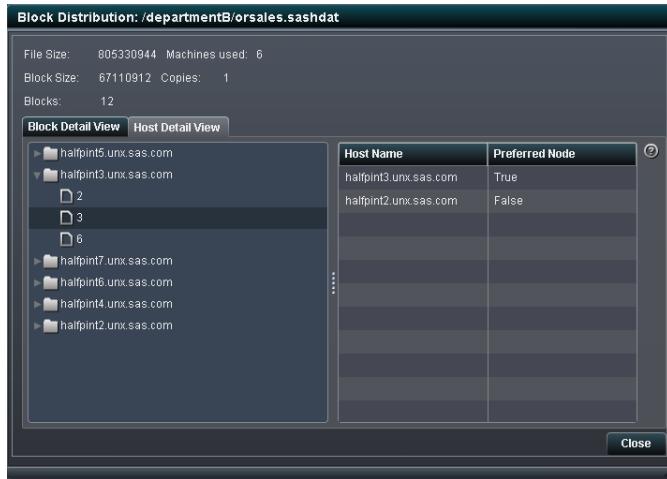
Table 29.6 Block Distribution Information

Field	Description
File size	Specifies the size of the file in bytes.
Block size	Specifies the block size for the file.
Blocks	Specifies the number of blocks used to store the original copy of the data.
Copies	Specifies the number of redundant block copies of the data.
Machines used	Specifies the number of machines in the cluster that have original or redundant blocks for the file.

On the **Block Detail View** tab, you can select a block number to view how many copies of the block exist and the host names for the machines that store the blocks. The value in the **Total Copies** column equals the number of redundant copies of the block plus

the original block. You can select the column heading to sort the rows. In an ideal distribution, the number of total copies is equal for all blocks.

On the **Node Detail View** tab, you can expand a host name node and then view the block numbers that are stored on that machine. When you select the block number, this host name and any additional machines with copies of the block are identified in the host name list. The following display shows an example:



30

Monitoring Resources

<i>Utilization History</i>	271
<i>Real-Time View</i>	272

Utilization History

Information about the resource utilization is available in the **Resource Monitor**. The resource utilization is displayed in two graphs, the Utilization History and the Real-Time View. The graphs are updated according to a sampling rate. The sampling rate is set in the application preferences.

The Utilization History graph plots the utilization for the following resources against time:

- CPU
- memory
- network input/output (I/O)

The CPU and memory utilization are plotted as a percentage of 100% capacity, though the graph scales automatically. Under periods of high demand, the upper bound can reach 100%. Under periods of low demand, the upper bound can drop below 10%.

The network I/O utilization is shown as one line plot for input and one line plot for output. The plot shows the transfer rate in units of megabytes per second.

To view the resource utilization for one sampling period, place your pointer over a line. You can also click on a line to select the same sampling period for all the resource line

plots that are shown. This way, you can place your pointer over each of the selected sampling periods to view the utilization for each resource in the same sampling period.

Figure 30.1 Resource Monitor



Note: If the Resource monitor does not display the line plot or the Real-Time View, then contact your system administrator to make sure the SAS LASR Analytic Server Monitor is running.

Real-Time View

The Real-Time View of resource utilization is provided below the Utilization History. The display is a heat map for each of the monitored resources, CPU, memory, and network I/O. Each column in the heat map represents a machine in the cluster. To view the real-time resource utilization for a resource type, position your pointer over one of the cells in the heat map.

For the section that shows CPU utilization, the tooltip shows the host name, CPU number, and the CPU utilization as a percentage. For the section that shows memory utilization, the tooltip shows the host name and memory utilization as a percentage. Both of these resources have a range of zero to 100. The color changes to represent the work load.

The bottom section shows two rows to represent network I/O. The top row shows the transfer rate for output and the bottom row shows the transfer rate for input. The tooltip shows the host name, the direction, and the rate. The color for the icons depends on the transfer rate between 0 and 25 megabytes per second. If the transfer rate exceeds 25 MB per second, then the icon color changes to purple.

31

Monitoring Processes

<i>Viewing Memory and CPU Utilization</i>	275
<i>CPU and Memory Utilization by Instance</i>	276

Viewing Memory and CPU Utilization

Information about the per-process resource utilization is available in the **Process Monitor**. The resource utilization for memory and CPU are displayed in two graphs, the Memory Utilization versus CPU Utilization and CPU and Memory Utilization by Instance. The graphs are updated according to a sampling rate. The sampling rate is set in the application preferences.

The **Selection** menu enables you to filter the display of application instances in the graphs for top resources consumers.

Each application instance is represented by a bubble in the bubble plot. The size of the bubble represents the number of processes used for the application instance. If an instance has one process for each machine in the cluster, then that ensures ideal resource utilization and performance.

The location of the bubble indicates the resource utilization for the application instance on the machines in the cluster. Place your pointer over a bubble to view information about the application instance. Select a bubble to view the detailed resource consumption for each machine that is used by the application instance.

Note: If no application instances appear in the process monitor, then contact your system administrator to make sure the SAS LASR Analytic Server Monitor is running.

CPU and Memory Utilization by Instance

The CPU and Memory Utilization by Instance graph shows a bar chart. Each application instance is represented by a bar. The lower part of the bar shows the CPU utilization for the application instance. The upper part of the bar shows the memory utilization for the application instance.

Place your pointer over a bar to view information about the application instance. Select a bar to view the detailed resource consumption for each machine that is used by the application instance.

32

Managing Mobile Devices

<i>Features in Mobile Device Administration</i>	277
<i>Managing Mobile Devices</i>	278

Features in Mobile Device Administration

This chapter addresses the use of the SAS Visual Analytics mobile device administration interface to manage mobile device users who use the SAS Visual Analytics suite of applications. This application interface includes the following administration features for managing mobile devices:

- View data about mobile devices that are set up for users to view reports or explorations that were created in SAS Visual Analytics. The displayed data includes user ID, device type, device model, device ID, OS version, application version, and timestamp of last access.
- Sort the data for mobile devices in ascending or descending order.
- Filter data for mobile devices. For example, you can filter data by the mobile device type or by the mobile device model.
- Blacklist lost or stolen mobile devices.
- Remove blacklisted mobile devices from the blacklist.

Managing Mobile Devices

The SAS Visual Analytics mobile device administration interface is easy to use. To perform tasks such as viewing, sorting, filtering, or blacklisting a mobile device:

- 1 Log on to SAS Visual Analytics.
- 2 Click the **Prepare Data** icon.
- 3 Click **Mobile Devices**.
- 4 On the main page for SAS Mobile Device Administration, select **Logon History**.
- 5 To sort the items in the columns in either ascending or descending order, click the column heading for any column on this page.
- 6 To filter and view selected data about mobile devices:
 - a From the **Filter** menu, select an item (for example, Device Type). Then, enter a valid value in the field (for example, iPad), and click **Apply**.
The value that you specified for the filter determines the type of data that is displayed.
 - b To view the entire history of events (for example, various application versions that were used by devices) associated with the item that you specified in the **Filter** field, select the **Include Device History** check box.
- 7 To add a mobile device to the blacklist:
 - a Click the device in the **Device ID** field on the Logon History page.
You are prompted by a dialog box with a question asking if you are sure that you want to blacklist this device.
 - b Click **Yes**.

The mobile device is added to the list of blacklisted devices on the Manage Blacklist page. The **Action** column indicates the status of this device (whether it is blacklisted or not), and the **Result** column indicates whether your action was successful. Blacklisted mobile devices cannot be used to access the SAS Visual Analytics suite of applications.

- 8 To remove a mobile device from the blacklist, select the device on the Manage Blacklist page. Then, click the X icon.

You are prompted by dialog box asking you to confirm your action. Click **Yes**.

- 9 To view blacklisted mobile devices or confirm that a device was blacklisted, click **Blacklist History**.

In the table on this page:

- The **Admin ID** column displays the user name of the administrator who blacklisted the device.
- The **Device ID** column displays the name of the blacklisted device.
- The **Action** column displays the type of action that was performed for the mobile device.
- The **Result** column indicates whether the attempt to blacklist a mobile device was successful.
- The **Timestamp** column displays the date and time when a specific mobile device was blacklisted.

- 10 To search for a blacklisted mobile device either by the device ID or the administrator ID, click **Manage Blacklist**. Then, select **Admin ID** or **Device ID** from the drop-down menu for **Filter**. Click **Apply**.

The mobile devices that meet the criteria that you specified are displayed.

Part 7

SAS High-Performance Computing Management Console

<i>Chapter 33</i>	
<i>Overview of the Console</i>	283
<i>Chapter 34</i>	
<i>Managing Console Users and Groups</i>	291
<i>Chapter 35</i>	
<i>Viewing Console Logs</i>	303
<i>Chapter 36</i>	
<i>Managing Users and Groups</i>	309

<i>Chapter 37</i>	
<i>Managing Database Resources</i>	325
<i>Chapter 38</i>	
<i>Gridhosts File Management</i>	331

33

Overview of the Console

<i>About the Console</i>	283
<i>Accessing the Console</i>	284
<i>Specifying Console Preferences</i>	284
<i>Configuring Console Network Settings</i>	286
<i>Configuring Console Logging</i>	287

About the Console

The SAS High-Performance Computing Management Console is a Web application that is used by system administrators to manage high-performance computing environments that use SAS software. Maintaining high-performance computing (HPC) environments is challenging because of the large number of machines that are used in the distributed computing environment. In order to ensure peak performance and reliability, whatever action is taken on one machine must be performed on all the machines. Tasks that might be simple on a single machine, such as adding a user or changing a system configuration option, become more challenging in a distributed computing environment. Most changes must be propagated to the other machines and this can be labor intensive and time consuming. The console is used to ease the administration of the environment.

The console includes its own embedded Web server. When administrators log on to the console, where authentication occurs is determined by the user name provided. If the

user name matches an operating system user account, then the operating system performs authentication. If the user name does not match an operating system user account, then the console performs authentication. In either case, the console does not perform any interaction with the SAS Metadata Server.

Accessing the Console

You can access the console interface by selecting **Tools** ▶ **SAS HPC Management Console** from the SAS Visual Analytics data preparation interface. Log on with a console user account. For more information about console user accounts, see “[Managing Console Users and Groups](#)” on page 291.

Specifying Console Preferences

To specify your preferences:

- 1 Click **Console Management** from the toolbar.
- 2 Click **Console Configuration** and then click **User Interface**.
- 3 The most important setting for a SAS Visual Analytics deployment is to make sure that the **HPC Environment being managed** menu is set appropriately.

The following table identifies the options that are available:

Table 33.1 SAS High-Performance Computing Management Console User Interface Options

Option	Description
Page background	Specify the color to use for each of these options as the RGB hexadecimal code for a color.
Normal text	
Table background	
Table header	
Link text	
Display titles as text?	Specifies whether any title is displayed as text in the console. The default value is No .
Display login and hostname (for non-framed themes)	Specify the location of the client login and host name when non-framed themes are used. Use the menu to select a choice. The default value is At bottom of browser .
Hostname to display in the Management Console	This option enables you to specify the how the host name for the SAS High-Performance Computing Management Console is displayed to clients. Use the menu to select a choice. The default value is Real hostname .
Prepend username to page titles?	
Prepend hostname to page titles?	Specify whether the user name or host name of the connected user appears in the page title bar of the Web browser.
Format for displayed dates	Specify how the date is displayed in the console. Use the menu to select a choice. The default value is dd-mon/yyyy .
File chooser size	
User chooser size	
Multiple users chooser size	
Date selector size	
Module chooser size	
Multiple modules chooser size	

Option	Description
HPC Environment being managed	<p>Specify the environment type. For SAS Visual Analytics deployments, select a menu item that includes SAS LASR Analytic Server.</p> <p>Note: For deployments that use SAS Visual Analytics Hadoop, select SAS LASR Analytic Server for Hadoop.</p>
SAS HPC MC server hostname in /etc/gridhosts	<p>Specify the host name for the machine that is used for running the SAS High-Performance Computing Management Console. Enter the host name exactly as it appears in the <code>/etc/gridhosts</code> file.</p>

- 4 Click **Save**.

Configuring Console Network Settings

By default, the server process for the SAS High-Performance Computing Management Console listens on port 10020 and on all IP addresses for the machine on which it is installed. However, the port and address settings are configurable.

Note: In a SAS Visual Analytics deployment, do not change the port number.

To specify the port and address settings:

- 1 Click **Console Management** from the toolbar.
- 2 Click **Console Configuration** and then click **Ports and Addresses**.
- 3 The following table identifies the options that are available:

Table 33.2 SAS High-Performance Computing Management Console Ports and Addresses Options

Option	Description
Listen on IPs and ports	<p>Select the Only Address option and specify an IP address to restrict the console to listen for connections on the specified IP address only. The default value is Any Address.</p> <p>Specify a port number in the Listen on port field. The default value is 10020.</p>
Accept IPv6 connections?	<p>Specifies whether the console accepts IPv6 connections. The default value is No.</p>
Listen for broadcasts on UDP port	<p>Specify whether to listen for UDP broadcasts from other console installations. Available options are Don't Listen or a specified port. The default value is to listen on port 10020. This option is not used at this time.</p>
Web server hostname	<p>Specify how the console server determines its host name. Available options are Work out from browser or specifying a host name. The default value is to determine it from the Web browser.</p>
Reverse-resolve connected IP address?	<p>Specify whether the IP addresses of the client machines are resolved to host names. The default value is Yes.</p>

4 Click **Save**.

Configuring Console Logging

To configure logging settings:

1 Click **Console Management** from the toolbar.

2 Click **Console Configuration** and then click **Logging**.

3 The following table identifies the options that are available:

Table 33.3 SAS High-Performance Computing Management Console Logging Options

Option	Description
Logging active?	Select Disable logging or Enable logging . The default value is Enable logging .
Log resolved hostnames?	Specifies whether to log host names in the logs. The default value is No .
Use combined log format (including referrer and user agent)?	Specifies whether a combined log format is used. The default value is No .
Periodically clear log files?	Specify the interval to use for clearing log files. The default value is not to clear log files.
Users to log	Specify the user accounts to record in log reports. The default is Log actions by all users .
Modules to log	Specify the modules to log. The default is Log actions in all modules .
Include Console logins and logouts in actions log?	Specify whether logon to and logoff actions from the console are logged. The default value is No .
Log changes made to files by each action?	Specify whether changes that are made to files are logged. The default value is No .
Record all modified files before actions, for rollbacks?	Specify whether modified file changes are recorded for possible rollback. The default value is No .
Permissions for log files	Specify the mode permissions for the log files. The default setting is 600 and owned by root.
Add log to syslog?	Specify whether actions are logged to syslog in addition to log files. The default value is No .

4 Click **Save**.

34

Managing Console Users and Groups

<i>Managing Console Users and Groups</i>	291
Understanding Console Users and Groups	291
Create a Console User	292
Creating a Console User from the Command Line	296
Create a Console Group	297
Best Practices for Managing Console Users and Groups	300
<i>Force Off Console Sessions</i>	300

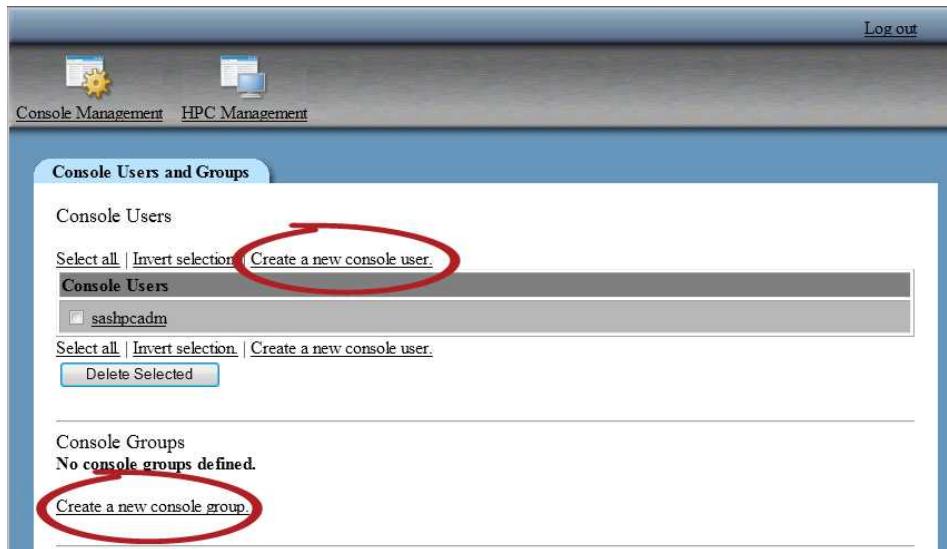
Managing Console Users and Groups

Understanding Console Users and Groups

An important concept to understand is that the SAS High-Performance Computing Management Console manages two types of users and groups. The first type is for console users and groups. This provides access control for the user accounts that can use the SAS High-Performance Computing Management Console. The second type of users and groups are the operating system users and groups. The console does assist with managing these users and groups. For information about managing operating system users and groups with the console, see “[Managing Users and Groups](#)” on page [309](#).

The following display shows the interface for managing console users and groups. The links for creating new console users and new console groups are highlighted.

Display 34.1 Console Users and Groups Management



Create a Console User

Be aware that a console user is not the same as a user that is created in the **HPC Management** section of the console. Console users are created strictly for using the console. The Users and Groups module in the **HPC Management** section manages the operating system users and groups.

TIP If the site security policy does not permit logging on the console with the root account to create the first console user, see “[Creating a Console User from the Command Line](#)” on page 296.

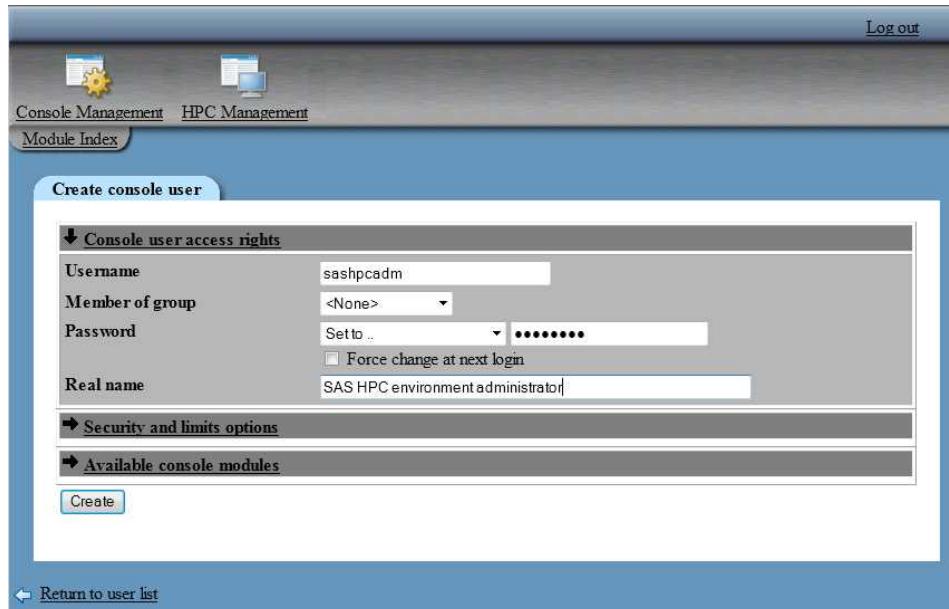
To create a console user:

- 1 Click **Console Management** from the toolbar.
- 2 Click **Console Users and Groups**.

3 Click **Create a new console user**.

The following display shows the Create console user screen.

Display 34.2 Create Console User



The screenshot shows the 'Create console user' interface. At the top, there are links for 'Console Management', 'HPC Management', and 'Module Index'. On the right, there is a 'Log out' link. The main form is titled 'Create console user'. It contains the following fields:

- Console user access rights**
 - Username:** sashpcadm
 - Member of group:** <None>
 - Password:** Set to: * * * * * * * * * * * * Force change at next login
 - Real name:** SAS HPC environment administrator
- Security and limits options**
- Available console modules**

At the bottom of the form is a 'Create' button. Below the form, there is a link to 'Return to user list'.

The following table describes the additional fields on the Create console user screen. They become available after clicking **Security and limits options** or **Available console modules**.

Table 34.1 Create Console User Field Descriptions

Field	Description
Username	Specify the user account.
Member of group	Use the menu to select a group.

Field	Description
Password	<p>Specify the password for the console user. The menu offers the following choices:</p> <p>Set to sets the password to the value entered into the text field.</p> <p>UNIX authentication passes the authentication request to the operating system. This is the most commonly used choice.</p> <p>No password accepted locks the user out of the system.</p>
Real name	(Optional) Specify an alternative description for the console user.

4 Click **Create**.

The following table describes the additional options for creating a user.

Table 34.2 Create Console User Options

Field	Description
Inactivity logout time	Specify an inactivity time-out. The default is no time-out.
Minimum password length	Specify the minimum number of characters for the user's password. The default is no minimum length.
IP access control	Select the option that meets the security requirements for the environment. If Deny from listed addresses is selected, then be aware that the console user can log on from any IP address that is not listed. The default value is Allow from all addresses .
Allowed days of the week	Specify the days of the week that the console user is permitted to log on. The default value is Every day .

Field	Description
Allowed times of the day	Specify the range of hours that the console user can log on. The default value is Any time .
Console Configuration	Select this check box to enable the console user to configure console settings such as the user interface, network settings, and console logging.
Console Users and Groups	Select this check box to enable the console user to configure console users and groups. A best practice is to limit access to this feature. For more information, see “ Best Practices for Managing Console Users and Groups ” on page 300.
Console Logs	Select this check box to enable the console user to view and search the console logs. The console logs contain information about the actions performed with the console.
Active Queries and Connections	Select this check box to enable the console user to view and terminate database queries and connections. Note: This option is available for environments that are configured to manage Greenplum Database deployments only.
Gridhosts File Management	Select this check box to enable the console user to modify the <code>/etc/gridhosts</code> file. This file is used to identify the machines in the HPC environment and is very important to the operation of SAS High-Performance Analytics software and the console itself.
Users and Groups	Select this check box to enable the console user to manage the operating system user accounts in the HPC environment.
SSH Lockout	Select this check box to enable the console user to use the SSH lockout feature.

Field	Description
Temporary Table Maintenance	Select this check box to enable the console user to view and drop temporary database tables. Note: This option is available for environments that are configured to manage Greenplum Database deployments only.

Check boxes that are marked with an 'x' identify features that are enabled through group membership.

Creating a Console User from the Command Line

The root user account has access to log on to the console as soon as the console is installed and started. Typically, the root user account logs on to the console and creates a console administrator account that is then used to perform the administrative tasks. However, the security policy at some sites might not permit sharing the password for the root user account. In these environments, it is possible to add a console user from the command line if you have permission to use the `sudo` command. The files to modify are owned by root and only the root user account can restart a service. The console user to add must have a user account on the system.

To add a console user from the command line:

- 1 Use the `sudo` command to edit the `miniserv.users` file:

```
sudo vi /opt/webmin/etc/miniserv.users
```

- 2 Add a line that is similar to the following example:

```
username:x:0:::::::0:0
```

Save and close the file.

- 3 Use the `sudo` command to edit the `webmin.acl` file:

```
sudo vi /opt/webmin/etc/webmin.acl
```

- 4 Add a line that is similar to the following example:

```
username: webmin webminlog init cron inittab proc acl
webmincron net initsetup system-status useradmin tmptblmaint
theme-xpstyle mscstyle3 actpsgqrys sshlock gridhosts
```

Note: Enter this all on one line. It is split onto more than one line for display purposes only.

Save and close the file.

- 5** Restart the SAS High-Performance Computing Management Console:

```
sudo service sashpcmc restart
```

- 6** Log on to the console with the user name.

Create a Console Group

To create a console group:

- 1** Click **Console Management** from the toolbar.
- 2** Click **Console Users and Groups**.
- 3** Click **Create a new console group**.

The following display shows the Create Console Group screen.

Display 34.3 Create Console Group

The screenshot shows a web-based interface for managing console groups. At the top, there are navigation links for 'Console Management', 'HPC Management', and 'Module Index'. The 'Console Management' link is underlined, indicating it is the active module. The main content area is titled 'Edit Console Group' and contains a form for a group named 'consoleadmins'. The form includes fields for 'Group name' (set to 'consoleadmins'), 'Group description' (set to 'Administrators for the SAS High-Performance Computing Manager'), and 'Storage type' (set to 'Local files'). Below the form is a section titled 'Available console modules' with two columns of checkboxes. The left column contains 'Console Configuration', 'Console Users and Groups', 'Gridhosts File Management', and 'Users and Groups'. The right column contains 'Console Logs', 'SSH Lockout', and 'Select all' and 'Invert selection' buttons. At the bottom of the screen are 'Save', 'Clone', and 'Delete' buttons.

The following table describes the fields on the Create Console Group screen.

Table 34.3 Create Console Group Field Descriptions

Field	Description
Group name	Specifies the console group name.
Group description	Specifies a description for the console group.
Member of users and groups	When a console group that has members is edited, this field lists the members of the group.
Console Configuration	Select this check box to enable the members of the group to configure console settings such as the user interface, network settings, and console logging.

Field	Description
Console Users and Groups	<p>Select this check box to enable members of the group to configure console users and groups. A best practice is to limit access to this feature. For more information, see "Best Practices for Managing Console Users and Groups" on page 300.</p>
Console Logs	<p>Select this check box to enable members of the group to view and search the console logs. The console logs contain information about the actions performed with the console.</p>
Active Queries and Connections	<p>Select this check box to enable members of the group to view and terminate database queries and connections.</p> <p>Note: This option is available for environments that are configured to manage Greenplum Database deployments only.</p>
Gridhosts File Management	<p>Select this check box to enable members of the group to modify the <code>/etc/gridhosts</code> file. This file is used to identify the machines in the HPC environment and is very important to the operation of SAS High-Performance Analytics software and the console itself.</p>
Users and Groups	<p>Select this check box to enable members of the group to manage the operating system user accounts in the HPC environment.</p>
SSH Lockout	<p>Select this check box to enable members of the group to use the SSH lockout feature.</p>
Temporary Table Maintenance	<p>Select this check box to enable members of the group to view and drop temporary database tables.</p> <p>Note: This option is available for environments that are configured to manage Greenplum Database deployments only.</p>

- 4 Click **Create**.

Best Practices for Managing Console Users and Groups

Console groups can be used to grant identical access rights to groups of users that share the same administration responsibilities. The **Console Users and Groups** module enables an administrator to manage the access permissions for console users. Limit access to this module because a user can change his access permissions and dilute the security of the SAS High-Performance Computing Management Console.

Force Off Console Sessions

The console enables administrators to force off other console users. To force off a console user:

- 1 Click **Console Management** from the toolbar.
- 2 Click **Console Users and Groups**.
- 3 Click **View Login Sessions** near the bottom of the page.
- 4 Select the session ID from the table to force off the console user. There is no confirmation prompt.

Display 34.4 SAS High-Performance Computing Management Console Current Login Sessions

The screenshot shows the SAS HPC Management Console interface. At the top, there are three icons: a gear, a document, and a computer monitor. Below the icons, the menu bar includes 'Console Management', 'HPC Management', and 'Module Index'. On the right side of the header is a 'Log out' link. The main content area is titled 'Current Login Sessions'. A sub-instruction below the title reads: 'Current console session logins are listed below. To cancel an existing session and force the user to login again, click on its session ID.' A table lists three sessions:

Session ID	Console user	IP address	Logged in at	Action
vPa7j5l8g4CYAyC4jVHh/1	sashpcadm		08/May/2012 10:02	View logs...
wjBw07XYURMlarWq8m5w1	sashpcadm		08/May/2012 04:34	View logs...
ACCT3L4d7cjVdIV79IbSi/	sashpcadm		03/May/2012 10:08	View logs...

At the bottom left, there is a link 'Return to user list' with a back arrow icon.

Your session ID is shown in bold text.

35

Viewing Console Logs

<i>View Logs for an Active Session</i>	303
<i>Search Console Logs</i>	304
<i>Understanding Log Details</i>	305

View Logs for an Active Session

Follow the steps described in “[Force Off Console Sessions](#)” on page 300. Instead of selecting the session ID, select **View logs** for the console user to monitor. The console uses the session ID to perform a search of the console logs. The following display shows an example of the search results:

Display 35.1 Console Logs by User and Session ID

Action	Module	User	Client Address	Date	Time
Added Users Success	SSH Lockout	sashpcadm		08/May/2012	10:20
Add Lock Success	SSH Lockout	sashpcadm		08/May/2012	10:20

Select the action to review and then review the details. For information about the details, see “[Understanding Log Details](#)” on page 305.

Search Console Logs

To search console logs:

- 1 Click **Console Management** from the toolbar.
- 2 Click **Console Logs**.
- 3 Specify the search criteria in the **Search the console log for actions** fields and click **Search**.

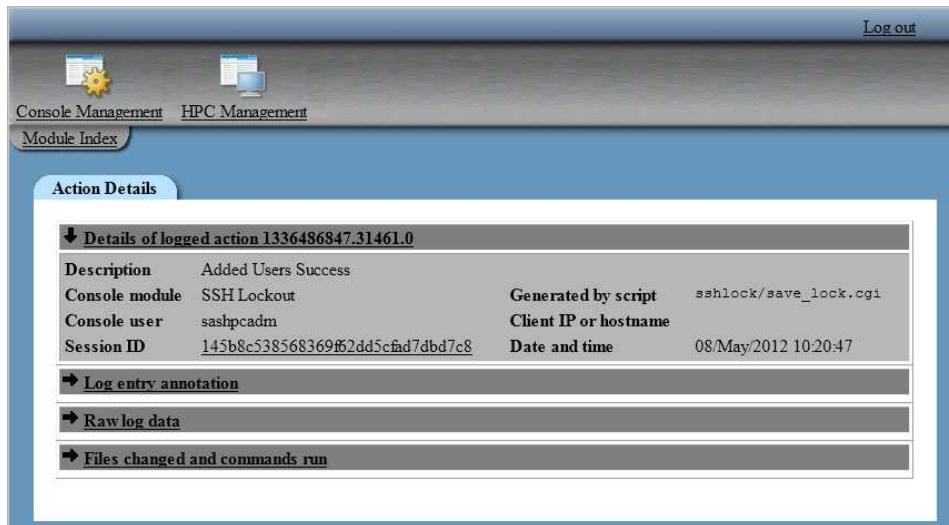
Note: The **By non-console user** option does not work in a SAS Visual Analytics deployment.
- 4 Select the action to review and then review the details. For information about the details, see “[Understanding Log Details](#)” on page 305.

Understanding Log Details

The console logs record information such as the action that was performed, the module that was used, the date and time, and the parameters associated with the action. The information in these logs is valuable when reviewing the actions that were performed by a console user. In addition to viewing logs, the console enables an administrator to annotate the log record with comments.

The following display shows a console log example for starting an SSH lockout. The information that the console shows in most of the fields, and especially in the **Raw log data** section, differs according to which action is performed.

Display 35.2 Console Log Action Details



The screenshot shows the Oracle Linux 6.3 Console Management interface. At the top, there are icons for Console Management, HPC Management, and Module Index. The Module Index tab is currently selected. In the center, a box titled "Action Details" displays the following information for a logged action (ID 1336486847.31461.0):

Details of logged action 1336486847.31461.0	
Description	Added Users Success
Console module	SSH Lockout
Console user	sashpcadm
Session ID	145b8c538568369f62dd5cfad7dbd7c8
Generated by script	sshlock/save_lock.cgi
Client IP or hostname	
Date and time	08/May/2012 10:20:47

Below this, there are three expandable sections:

- Log entry annotation
- Raw log data
- Files changed and commands run

The **Details of logged action** section includes the following information:

Table 35.1 Console Log Details of Logged Action

Field	Description
Description	Specifies the action that was performed.

Field	Description
Console module	Specifies the console module that was used to perform the action.
Console user	Specifies the user name that performed the action.
Session ID	Specifies the Web session that was used when the action was performed. You can select the link to view the console logs for all actions performed during the Web session.
Generated by script	Specifies the console script that was used to perform the action.
Client IP or hostname	Specifies the IP address or host name that was used by the console user to perform the action.
Date and time	Specifies the date and time for when the action was performed. The date and time format can be changed by specifying a different format for displayed dates as described in "Specifying Console Preferences" on page 284 .

The **Log entry annotation** section enables an administrator to save a comment about the action. Click **Save** after adding a comment.

The information in the **Raw log data** section differs according to the action that was performed. The information is shown as a two-column table of parameter names and parameter values. Most of the parameters are related to the descriptions provided in the previous table. However, the appearance of some fields is different. The following list describes those fields:

sid

This is the value for the Web session ID that was in use when the action was performed.

time

This value is a UNIX time representation for when the action was performed. When the data and time are shown in the **Details of logged action** section, the time is formatted for display from this value.

id

This value is the identifier for the log record.

The **Files changed and commands run** section shows the operating system commands that were executed to perform the action and the files that were used. This section is populated when the console is configured to log such information. By default, the console is not configured to log commands and files because the list of commands can be large and can contribute to a full system disk.

36

Managing Users and Groups

<i>About Managing Users and Groups</i>	309
<i>Managing Users and Groups</i>	311
Understanding Users and Groups	311
Configuring the Middle-Tier Shared Key	312
Create a User	313
About Editing and Deleting Users	316
SSH Key Management Features	317
Create a Group	317
About Editing and Deleting Groups	318
<i>Managing User Sessions</i>	319
Features in User Session Management	319
Review Recent SSH Logon Activity	319
Show Currently Logged On Users	320
<i>Perform an SSH Lockout</i>	321
Understanding the SSH Lockout Feature	321
Perform an SSH Lockout	323

About Managing Users and Groups

The SAS High-Performance Computing Management Console enables two types of management for users and groups. The first type, for console users and groups, is described in “[Managing Console Users and Groups](#)” on page 291. This section

describes the second type of user and group management, system users and groups. This management is performed with the High-Performance Computing (HPC) management section of the interface. It is also used to perform an SSH lockout. A lockout is used to limit access to machines in the cluster.

The management operations that are described in the previous paragraph apply to a deployment that configures the console to manage an environment for SAS LASR Analytic Server. For information about configuring the console environment, see “[Specifying Console Preferences](#)” on page 284.

The following figure shows the interface for an administrator who has permission to perform HPC management only. The **Active Queries and Connections** and **Temporary Table Maintenance** modules apply to a SAS LASR Analytic Server with Greenplum environment only. The options that are used to perform console management are not available for this administrator.

Display 36.1 HPC Management Interface



Managing Users and Groups

Understanding Users and Groups

The users and groups module is used to create, manage, and propagate operating system user accounts and groups throughout the machines in the cluster. It also enables an administrator to generate and distribute SSH keys for each user account. The key generation can be performed as the account is propagated to the machines in the cluster, or it can be performed on existing accounts.

An important requirement for SAS Visual Analytics deployments is to append the SSH public key for the user account that runs JBoss to the authorized_keys files for operating system user accounts. This task can be performed automatically when creating user accounts with the HPC management interface. The public key can also be appended to the authorized_keys file for existing accounts.

The following display shows the interface for managing operating system users and groups.

Display 36.2 Managing Users and Groups

	Username	User ID	Group	Real name	Home directory	Shell	SSH Keys?	Last login
<input type="checkbox"/>	sasdemo	32241	users	SAS Demo Account_9999999,0425	/home/sasdemo	/bin/ksh	Yes	Tue May 8 11:06
<input type="checkbox"/>	sas	32237	sas		/home/sas	/bin/bash	Yes	Tue May 8 04:33
<input type="checkbox"/>	sassrv	501	sas		/home/sassrv	/bin/bash	No	

The **HPC Users** tab shows the following information for each user account:

Table 36.1 User Accounts Field Descriptions

Field	Description
Username	Specifies the user account name.
User ID	Specifies the UID for the user account.
Group	Specifies the primary group for the user account.
Real name	Specifies details about the user that make the user account more identifiable.
Home directory	Specifies the home directory for the user account.
Shell	Specifies the UNIX shell to use for the user account.
SSH Keys?	This field is set to Yes if the <code>\$HOME/.ssh/id_rsa</code> file exists. For environments that use Network File System (NFS) to manage home directories, this field is set to NFS . The console does not attempt to detect the existence of the <code>id_rsa</code> file.
Last login	Specifies the date of the last logon.

Configuring the Middle-Tier Shared Key

SAS Visual Analytics deployments require appending the SSH public key for the user account that runs JBoss to the `authorized_keys` files for user accounts that use the explorer and designer interfaces. Perform this step before you create users.

To configure the middle-tier shared key:

- 1 Access the `.ssh/id_rsa.pub` file for the user account that is used to run JBoss. This file contains the SSH public key. Copy the contents of the file to your clipboard.

- 2** Click **HPC Management** from the toolbar.
- 3** Click **Users and Groups**.
- 4** Click **Midtier Shared Key** and specify values for the following fields:

Table 36.2 *Middle-Tier Shared Key Field Descriptions*

Field	Description
TKlasrkey location	Specify the fully qualified path to the tklasrkey file. The default location is <code>/opt/TKGrid/bin/tklsrkey</code> .
Shared Public Key	Paste the contents of the <code>id_rsa.pub</code> file from your clipboard.
Mid Tier Hostname	(Optional) If you specify the host name for the machine that is used to run JBoss, then the host name is included in the <code>authorized_keys</code> file. This provides an additional measure of security.

- 5** Click **Save**.

Create a User

To create a user:

- 1** Click **HPC Management** from the toolbar.
- 2** Click **Users and Groups**.
- 3** Click **Create a new user** and specify values for the following fields:

Table 36.3 *Create User Field Descriptions*

Field	Description
Username	Specify the user account name, such as <code>madupr</code> or <code>team1usr</code> .

Field	Description
User ID	<p>Select an option:</p> <p>Automatic the operating system selects an unused UID.</p> <p>Calculated (rarely used) the UID is created based on a Berkeley CRC and mkuid. The mkuid command assumes a standard naming convention for user names.</p> <p>Specified specify the UID to use. By default, this field shows the UID that is assigned if the Automatic option is used.</p> <p>The default option is Automatic.</p>
Real name	<p>Specify details about the user that make the user account more identifiable.</p>
Home directory	<p>Select an option:</p> <p>Automatic <code>/home/\$username</code> is the home directory value.</p> <p>Directory specify the fully qualified path to use as the home directory or click the Browse button to select a location.</p> <p>The default is Automatic.</p>
Shell	<p>This field is not selectable or configurable. The default is <code>/bin/ksh</code>.</p>
Password	<p>Select an option:</p> <p>No password required sets the password to null.</p> <p>Normal password enter the plain-text password in the field.</p> <p>Pre-encrypted password enter a password in encrypted form.</p>

Field	Description
Password changed	Specifies the last time the password was changed.
Expiry date	Specify the date on which the password should expire. You can enter the date or use the calendar. The default is no date.
Minimum days	Specifies the minimum number of days between password changes. The default is zero.
Maximum days	Specifies the maximum number of days between password changes. The default is 99999.
Warning days	Specifies the number of days that a password expiration warning is generated before a password expires. The default is zero.
Inactive days	Specifies the number of days that a user must be inactive before the user account is locked. The default is zero.
Force change at next login?	Select Yes to force the user to change the password after the next logon. The default is No .
Note: This option is not available for deployments that use SUSE Linux Enterprise Server. In this case, edit the user account after it is created to force the change.	
Primary group	Select a group name from the menu. This list is filtered to show GID values that are greater than or equal to 100.
Propagate User	Select Yes to add the user to each machine in the environment. The default is No .
Generate and Propagate SSH Keys	Select Yes to generate SSH keys and propagate them when the user is created. The default is No .

Field	Description
Add Shared Midtier Key	Select Yes to include the information from the Midtier Shared Key tab in the authorized_keys file for the user. The default is No .
Create home directory?	Select Yes to create the user's home directory. Select No if the directory already exists. The default is Yes .
Copy template files to home directory?	Select Yes to copy standard environment files to the user's home directory at creation time. The default is Yes .

4 Click **Create**.

If **Propagate User** was set to **Yes**, then the progress of adding the user to the machines in the environment is shown.

If a mismatch is detected on a machine, such as a UID already in use, then the change fails and the mismatch is reported.

About Editing and Deleting Users

You can edit a user by selecting the user name on the **HPC Users** tab. Unlike creating a user account, very few fields are available for edit. The following list identifies the fields that can be changed:

- **Password**
- **Expiry date**
- **Minimum days**
- **Maximum days**
- **Warning days**
- **Inactive days**
- **Force change at next login?**

The field descriptions for these options are provided in [Table 36.3 on page 313](#).

You can also delete users individually or in groups by selecting the check box for the user on the **HPC Users** tab and then clicking **Delete Selected Users**.

SSH Key Management Features

As described in the section about creating users, the console can generate and propagate SSH keys. In addition, the console can append a middle-tier shared key to the authorized_keys file for the user. These two features, SSH key generation and propagating the middle-tier shared key, can also be performed on existing user accounts.

To use these features, edit the user by selecting the user name on the **HPC Users** tab, select the radio button for the features that you want to use, and click **Save**.

Create a Group

To create a group:

- 1 Click **HPC Management** from the toolbar.
- 2 Click **Users and Groups**.
- 3 Select the **HPC Groups** tab and then click **Create a new user**. Specify values for the following fields:

Table 36.4 Create Group Field Descriptions

Field	Description
Group name	Specify the name for the group, such as <i>finance</i> or <i>team1</i> .

Field	Description
Group ID	<p>Select an option:</p> <p>Automatic the operating system selects an unused GID.</p> <p>Calculated (rarely used) the GID is created based on a Berkeley CRC and mkgid.</p> <p>Specified specify the GID to use. By default, this field shows the GID that is assigned if the Automatic option is used.</p> <p>The default option is Automatic.</p>
Members	<p>Select users from the list and use the buttons to specify the members of the group.</p>

4 Click **Create**.

If a mismatch is detected on a machine, such as a GID already in use, then the change fails and the mismatch is reported.

About Editing and Deleting Groups

You can edit a group by selecting the group name on the **HPC Groups** tab. Users can be added as members of the group. This action sets secondary group membership for the user account.

You cannot delete a group if it is used as the primary group for any user accounts.

Managing User Sessions

Features in User Session Management

The HPC management interface of the SAS High-Performance Computing Management Console enables administrators to monitor and manage user logon sessions. The following features are available to administrators:

- review recent SSH log on activity for one user or all users
- display a list of users currently logged on through SSH
- limit SSH logon actions by performing an SSH lockout

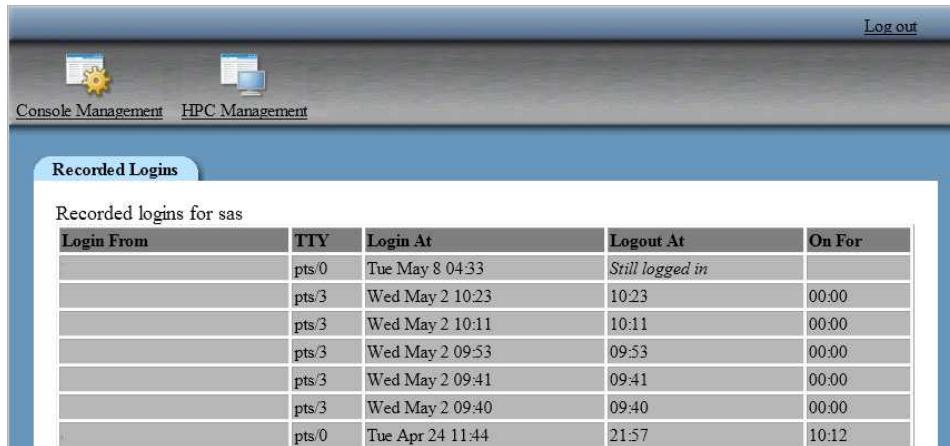
Review Recent SSH Logon Activity

To review recent SSH logon activity:

- 1 Click **HPC Management** from the toolbar.
- 2 Click **Users and Groups**.
- 3 Scroll to the bottom of the table of user accounts and then select either **All users** or **Only user**. If you select **Only user**, then enter the user account to review or click the button to select the user account from a list.
- 4 Click **Display Logins By**.

The recent logon activity is listed in the **Recorded Logins** table.

Display 36.3 Recent SSH Logon Activity



The screenshot shows a software interface with a toolbar at the top. The toolbar includes icons for a sun, a computer monitor, and a gear, followed by the text 'Console Management' and 'HPC Management'. On the far right of the toolbar is a 'Log out' button. Below the toolbar, the title 'Recorded Logins' is displayed in a blue header bar. The main content area is a table titled 'Recorded logins for sas'. The table has columns: 'Login From', 'TTY', 'Login At', 'Logout At', and 'On For'. The data in the table is as follows:

Login From	TTY	Login At	Logout At	On For
	pts/0	Tue May 8 04:33	<i>Still logged in</i>	
	pts/3	Wed May 2 10:23	10:23	00:00
	pts/3	Wed May 2 10:11	10:11	00:00
	pts/3	Wed May 2 09:53	09:53	00:00
	pts/3	Wed May 2 09:41	09:41	00:00
	pts/3	Wed May 2 09:40	09:40	00:00
	pts/0	Tue Apr 24 11:44	21:57	10:12

The **Login From** field is used to show the IP address or host name that was used for the logon session.

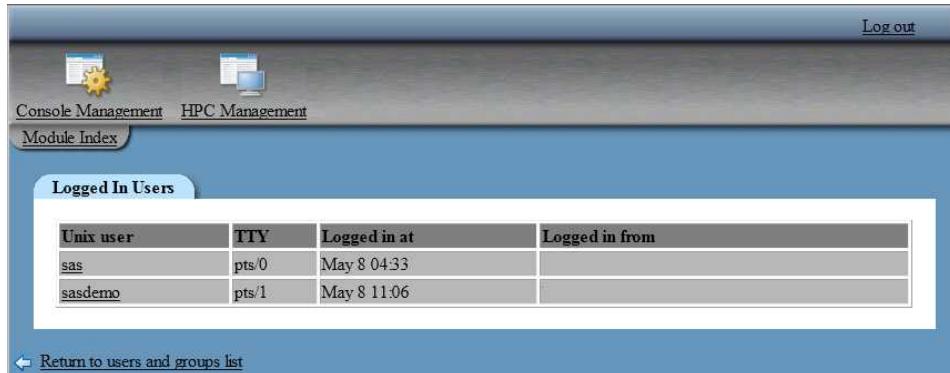
Show Currently Logged On Users

To show the users that are currently logged on with SSH:

- 1 Click **HPC Management** from the toolbar.
- 2 Click **Users and Groups**.
- 3 Scroll to the bottom of the table of user accounts and click **Show Logged In Users**.

The currently logged on users are listed **Logged In Users** table. You can select the user name to review the recent SSH logon activity for the user account.

Display 36.4 Logged In Users



The screenshot shows a software interface with a blue header bar. On the left, there are icons for 'Console Management' (a sun), 'HPC Management' (a computer monitor), and 'Module Index' (a list icon). The 'Module Index' icon is highlighted with a red box. In the center, a table titled 'Logged In Users' is displayed. The table has four columns: 'Unix user', 'TTY', 'Logged in at', and 'Logged in from'. The first row shows 'sas' at 'pts/0' on 'May 8 04:33'. The second row shows 'sasdemo' at 'pts/1' on 'May 8 11:06'. At the bottom of the table, there is a link 'Return to users and groups list'.

Unix user	TTY	Logged in at	Logged in from
sas	pts/0	May 8 04:33	
sasdemo	pts/1	May 8 11:06	

The **Logged in from** field is used to show the IP address or host name that was used for the logon session.

Perform an SSH Lockout

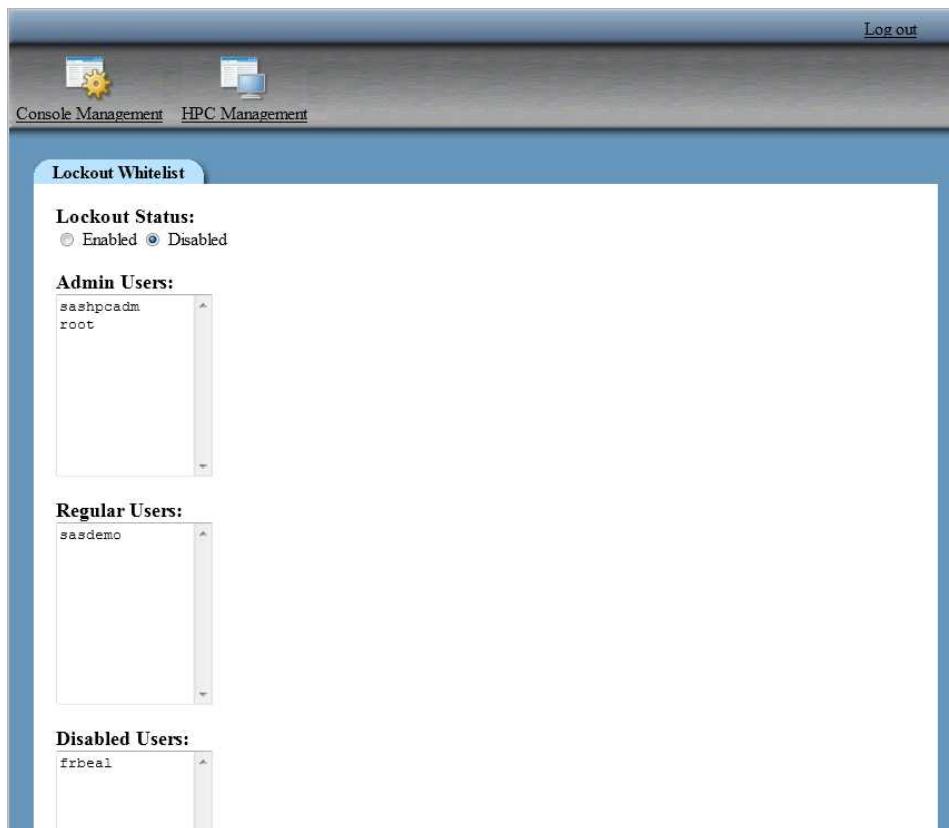
Understanding the SSH Lockout Feature

The SSH Lockout module enables an administrator to limit SSH logon access to the environment. This feature provides a mechanism for ensuring that processing resources are available for high-priority tasks as well as to assist with managing scheduled maintenance. Be aware that existing SSH logons are not affected by a lockout. A lockout prevents new SSH logons only. Access to software that does not rely on SSH is unaffected.

The feature provides user-level controls for denying SSH logons broadly while permitting SSH logons from specific user accounts. This level of control is available

based on how you set the lockout controls. The following figure shows the fields that are available on the **Lockout Whitelist** tab.

Display 36.5 SSH Lockout



The screenshot shows the SAS Management Console interface. At the top, there are icons for 'Console Management' and 'HPC Management', and a 'Log out' link. The main content area is titled 'Lockout Whitelist'. It contains three sections: 'Admin Users' (listing 'sashpcadm' and 'root'), 'Regular Users' (listing 'sasdemo'), and 'Disabled Users' (listing 'frbeal'). Each section has a scroll bar on the right.

Lockout Status:
 Enabled Disabled

Admin Users:
sashpcadm
root

Regular Users:
sasdemo

Disabled Users:
frbeal

The following table provides a description for each of the fields on the **Lockout Whitelist** tab and describes how the settings interact with each other. The lockout status is not changed until the **Commit** button is clicked.

Table 36.5 Lockout Whitelist Field Descriptions

Field	Description
Lockout Status	<p>This field shows the current lockout status and is used to set the lockout status:</p> <p>Enabled only the user accounts listed in the Admin Users and Regular Users are permitted SSH logon access.</p> <p>Disabled all users, except the users listed in the Disabled Users are permitted SSH logon access.</p> <p>Note: User accounts in the Disabled Users are denied SSH logon access at all times.</p>
Admin Users	Specifies the user accounts that are permitted SSH logon access when the lockout is active.
Regular Users	Specifies the user accounts that are permitted SSH logon access when the lockout is active.
Disabled Users	Specifies the user accounts that are denied SSH logon access at all times. This list can be used to deny access to single users even when a lockout is not active.

Perform an SSH Lockout

To perform an SSH lockout:

- 1 Click **HPC Management** from the toolbar.
- 2 Click **SSH Lockout**.

- 3** Select **Enabled** on the **Lockout Status**.
- 4** Review the **Admin Users**, **Regular Users**, and **Disabled Users** lists.

For information about the lists, see [Table 36.5 on page 323](#).

- 5** Click **Commit**.

To disable an SSH lockout, set the Lockout Status to **Disabled** and then click **Commit**.

37

Managing Database Resources

<i>Managing Active Queries and Connections</i>	325
Understanding the Active Queries and Connection Feature	325
Configuring the Greenplum Database Connection	326
Restrict the Database User Account	327
Viewing Active Queries and Connections	328
<i>Managing Temporary Tables</i>	329
About Temporary Tables	329
Viewing Temporary Tables	329

Managing Active Queries and Connections

Understanding the Active Queries and Connection Feature

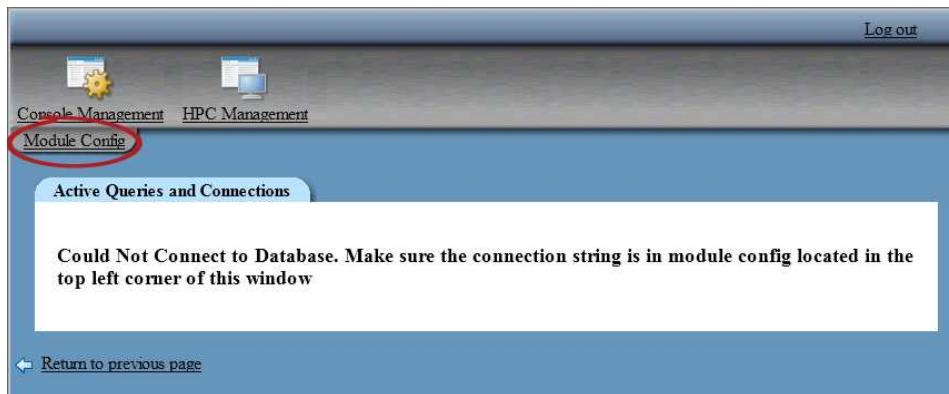
For deployments that use Greenplum Data Computing Appliance (DCA) as a co-located data provider, this feature enables an administrator to monitor long running queries and terminate database connections.

In order to monitor queries and terminate connections, a database user account is needed to connect to the database with administrative privileges. In order to list the active queries for all connections and to be able to terminate connections, the database user account should be granted the superuser role.

Configuring the Greenplum Database Connection

When you first access the **Active Queries and Connections** module, the console cannot connect to the Greenplum database because it is not configured with the connection parameters. The following display shows an example of the console before it is configured.

Display 37.1 Unconfigured Active Queries and Connections



To configure the module:

- 1 Click **HPC Management** from the toolbar.
- 2 Click **Active Queries and Connections**.
- 3 Select **Module Config**. (Shown in the preceding display.)
- 4 Specify the following parameters:

Table 37.1 Active Queries and Connections Configuration

Field	Description
Account Name for DB Connections	Enter the name of a database user account that has been granted the superuser role. For information about securing this user account, see “ Restrict the Database User Account ”.
Account Password	This field is used for storing a password. This field is also used to avoid changing a stored password: Don't change Select this option to avoid changing a stored password. Set to Select this option and specify the password for the database user account. The password is stored in <code>/opt/webmin/etc/actpsqqrays/config</code> .
Database Name	Specifies the database name to use.
MDW Hostname	Specifies the host name for the Greenplum Database Master Host.

5 Click **Save**.

Restrict the Database User Account

The database user account that is used for the Greenplum Database connection must have superuser privileges in order to monitor queries and terminate connections. You can increase the security for the account by configuring Greenplum Database to accept connections for the account from the machine that hosts the console.

To restrict the database user account:

1 Edit the `$MASTER_DATA_DIRECTORY/pg_hba.conf` file.

2 Add a line to the file that is similar to the following example:

```
# permit the 'sashpcadm' role access to any database from
# IP address 192.168.1.1 only and use md5 encrypted password
host    all    sashpcadm  192.168.1.1/32  md5
```

3 Save and close the file.

4 Run `gpstop -u` so that Greenplum Database rereads the file and applies the change.

The same database user account can be used for the active queries and connections module and the temporary table maintenance module. For more information about editing the `pg_hba.conf` file, see *Greenplum Database Administrator Guide*.

Viewing Active Queries and Connections

Once the database connection information has been configured, the active queries and connections module shows the active query and connection information. Each row provides information about a database server process.

Display 37.2 Active Queries and Connections

PID	Username	Current State or Query	Last Query Started	Client IP Address
7478		Idle	2012-05-08 13:11:44.408076-04	
7467		Idle	2012-05-08 15:18:20.520483-04	

Select the check box for a server process and click **Kill Selected** to terminate the queries, connection, and server process.

Managing Temporary Tables

About Temporary Tables

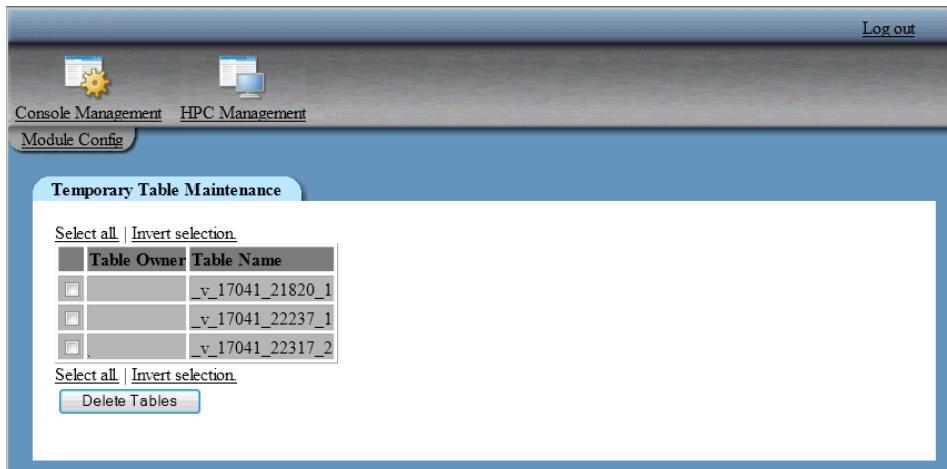
Temporary tables are a feature of Greenplum Database and several other third-party DBMS vendors. In some circumstances temporary tables are created by processes but are not dropped. The console provides the temporary table maintenance module for listing temporary tables and dropping them to restore system resources.

Before using the temporary table maintenance module, you must configure the Greenplum Database connection. Select **HPC Management** ▶ **Temporary Table Maintenance** ▶ **Module Config**. The configuration parameters are identical to those described in “[Configuring the Greenplum Database Connection](#)” on page 326.

Viewing Temporary Tables

To view temporary tables:

- 1 Click **HPC Management** from the toolbar.
- 2 Click **Temporary Table Maintenance**.

Display 37.3 Temporary Table Maintenance

The screenshot shows a software interface for managing temporary tables. At the top, there are icons for 'Console Management' and 'HPC Management', and a 'Log out' link. Below these are tabs for 'Module Config' and 'Temporary Table Maintenance', with the latter being the active tab. A sub-header 'Temporary Table Maintenance' is displayed. Under this, there are two rows of buttons: 'Select all' and 'Invert selection'. A table is shown with three rows, each containing a checkbox, the 'Table Owner' (which is 'v_17041'), and the 'Table Name' (which are '_21820_1', '_22237_1', and '_22317_2' respectively). At the bottom of the table are the same 'Select all' and 'Invert selection' buttons, followed by a large blue 'Delete Tables' button.

Table Owner	Table Name
<input type="checkbox"/>	_v_17041_21820_1
<input type="checkbox"/>	_v_17041_22237_1
<input type="checkbox"/>	_v_17041_22317_2

3 Select the check box for a table to drop, and click **Delete Tables** to drop the tables.

Note: It is possible to drop a table that is in use. Make sure that the table that you have selected to drop is not in use. Perform this action during scheduled maintenance or after using some other method to ensure that the table is not in use.

38

Gridhosts File Management

<i>What Is the Gridhosts File?</i>	331
<i>Managing the Gridhosts File</i>	332
Creating the File	332
Editing Host Names in the Gridhosts File	333

What Is the Gridhosts File?

Identifying all the machines in the environment is critical to ensuring that the SAS High-Performance Computing Management Console can perform operations in parallel or serially on the machines in the cluster. The `/etc/gridhosts` file is used to identify the machines. It is a list of the machines, with one host name on each line. The `/etc/gridhosts` file is not the same as the `/etc/hosts` file.

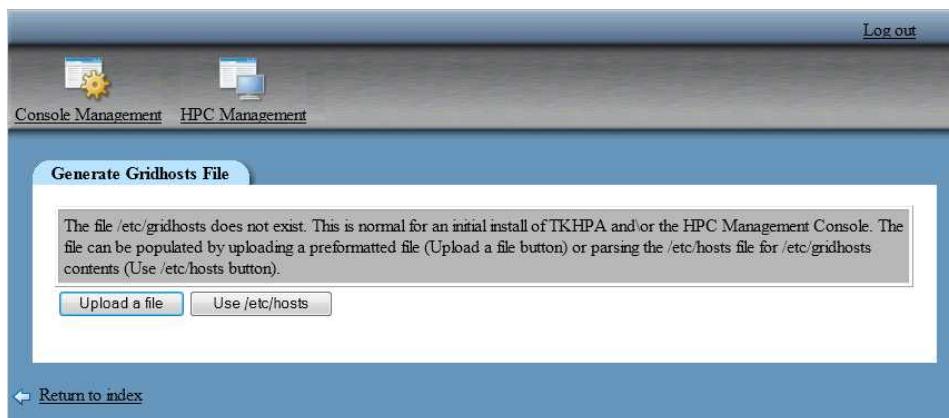
The `/etc/gridhosts` file is created immediately after installing SAS High-Performance Computing Management Console. If the `/etc/gridhosts` file has not been created already, then follow the steps described in “[Creating the File](#) on page [332](#)”. The console is also used to manage the host names that are listed in the file. The following sections provide more detail.

Managing the Gridhosts File

Creating the File

The contents of the file (host names) are typically created in a text editor or with shell scripting to avoid repetitive typing. If the file is not present on the machine that is hosting the console, then the following page is seen after accessing the **Gridhosts File Management** module of the console.

Display 38.1 Generate Gridhosts File



Note: The **Use /etc/hosts** button cannot be used in a SAS Visual Analytics deployment to create the **/etc/gridhosts** file.

If you can access the file, then you can upload it to the machine that is hosting the console. To upload a file:

- 1 Click **Upload a file**.
- 2 Click **Browse** and then navigate to the file that contains the host name of the machines in the cluster.

Click **Upload**.

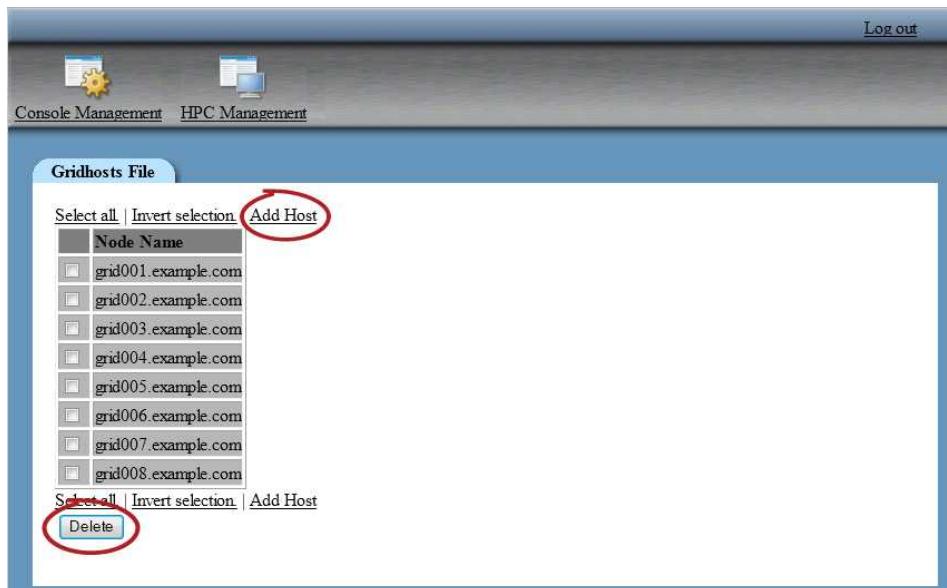
3 Review the host names that are listed in the table. If the list of host names is accurate, click **Select all** at the top of the table. Click **Submit**.

After the **Submit** button is clicked, the **/etc/gridhosts** file is created on the machine that is hosting the console.

Editing Host Names in the Gridhosts File

You can use the console to add and delete host names from the **/etc/gridhosts** file. The following display shows the link for adding host names and the button for deleting them.

Display 38.2 Gridhosts File



The screenshot shows a web-based interface for managing the Gridhosts File. At the top, there are navigation links for 'Console Management' and 'HPC Management'. On the right, there is a 'Log out' link. The main content area is titled 'Gridhosts File'. At the top of this area, there are buttons for 'Select all' and 'Invert selection', followed by a red circle around the 'Add Host' button. Below this is a table with a header row 'Node Name' and eight data rows, each containing a checkbox and a host name: 'grid001.example.com', 'grid002.example.com', 'grid003.example.com', 'grid004.example.com', 'grid005.example.com', 'grid006.example.com', 'grid007.example.com', and 'grid008.example.com'. At the bottom of the table, there are buttons for 'Select all', 'Invert selection', 'Add Host', and 'Delete', with the 'Delete' button circled in red.

To add a host name:

- 1 Click **Add Host**.
- 2 Enter the host name in the **Hostname** field and click **Add**.

Note: The console does not verify that the specified machine is reachable.

To delete a host name:

- 1 Select the check box beside the host name to remove.
- 2 Click **Delete**.

Glossary

Apache Hadoop

a framework that allows for the distributed processing of large data sets across clusters of computers using a simple programming model.

bar chart

a chart that consists of a grid and some vertical or horizontal columns (bars). Each column represents quantitative data.

box plot

See box-and-whisker plot

box-and-whisker plot

a graphical display of five statistics (the minimum, lower quartile, median, upper quartile, and maximum) that summarize the distribution of a set of data. The lower quartile (25th percentile) is represented by the lower edge of the box, and the upper quartile (75th percentile) is represented by the upper edge of the box. The median (50th percentile) is represented by a central line that divides the box into sections. The extreme values are represented by whiskers that extend out from the edges of the box.

capability

an application feature that is under role-based management. Typically, a capability corresponds to a menu item or button. For example, a Report Creation capability might correspond to a New Report menu item in a reporting application. Capabilities are assigned to roles.

crosstab

See crosstabulation table

crosstabulation table

a two-dimensional table that shows frequency distributions or other aggregate statistics for the intersections of two or more category data items. In a crosstabulation table, categories are displayed on both the columns and rows, and each cell value represents the data result from the intersection of the categories on the specific row and column.

data item

an item in a data source that is either a logical view of a data field or a calculation. The author of a report decides which data items to use in a particular section of a report. There are three types of data items: hierarchies, categories, and measures.

data source

a table, view, or file from which you will extract information. Sources can be in any format that SAS can access, on any supported hardware platform. The metadata for a source is typically an input to a job.

filter

specified criteria that are applied to data in order to identify the subset of data for a subsequent operation, such as continued processing.

heat map

a graphical representation of data where the values taken by a variable in a two-dimensional map are represented as colors.

list table

a two-dimensional representation of data, in which the data values are arranged in rows and columns.

pie chart

a circular chart that is divided into slices by radial lines. Each slice represents the relative contribution of each part to the whole.

progressive chart

See waterfall chart

report

output that is generated by running custom SAS code against the data in your project.

role

a set of capabilities. In some applications, certain actions are available only to users or groups that have a particular role.

scatter plot

a two- or three-dimensional plot that shows the joint variation of two (or three) variables from a group of table rows. The coordinates of each point in the plot correspond to the data values for a single table row (observation).

scatter plot matrix

a grid of scatter plots showing pairwise combinations of multiple numeric variables.

source

See data source

time series

an ordered sequence of values of a variable that are observed at equally spaced time intervals.

treemap

a graph that represents the relative values of data by using rectangular areas. The color of each area represents the value of one measure in the query. The size of each area represents the value of the another measure in the query. (In SAS/GRAF a treemap is created with the GTILE procedure.)

user role

See role

visual exploration

a metadata object that contains visualizations and data settings that are saved from a session of the SAS Visual Analytics explorer.

visualization

an interactive visual representation of data. A visualization can be a table, a chart, or a geographic map.

waterfall chart

a form of data visualization that is used to understand or explain the cumulative effect on an initial value of sequentially introduced positive or negative values. Usually the initial and the final values are represented by whole columns, while the intermediate values are denoted by floating columns.

Index

A

aggregation, modifying for measures [180](#)
 aggregations [81](#)
 alphanumeric categories [174](#)
 analysis
 sorting and [219](#)
 autocharting [102](#)
 automatic charts [102](#)

B

bar charts [103](#), [191](#)
 box plots [117](#)
 bubble plots [112](#), [196](#)
 bullet gauges [204](#)

C

calculated columns
 data preparation interface [46](#)
 capabilities, access to [7](#)
 category data items [174](#)
 changing data sources [172](#)
 charts
 See also [bar charts](#)

dual axis bar charts [198](#)
 dual axis bar-line charts [199](#)
 dual axis line charts [198](#)
 line charts [193](#)
 pie charts [194](#)
 targeted bar [192](#)
 waterfall charts [193](#)
 comments
 adding in designer interface [216](#)
 adding in Web viewer [253](#)
 connecting to library server for
 mobile viewer [250](#)
 correlation matrix [124](#)
 creating reports [163](#)
 crosstabs [188](#)
 exporting data from [240](#)
 sorting data in [220](#)

D

data
 exporting from crosstabs [240](#)
 data formats [81](#)
 data items
 See also [standard data items](#)
 category data items [174](#)
 data items in queries [174](#)

- duplicating 176
- measure data items 175
- modifying format 179
- modifying role 178
- queries based on 169
- removing 177
- renaming 178
- selecting 175
- data properties 80
- data sources
 - changing 172
 - modifying measure
 - aggregation 180
 - removing 173
 - selecting 170
- date categories 175
- deleting
 - report sections 235
- designer interface 153
 - adding comments 216
 - an example report 162
 - exporting data 237
 - gauges 204
 - graphs 191
 - images 212
 - Open window 157
 - opening reports 161
 - Search window 159
 - sorting data 219
 - tables 187
 - tables, graphs, and gauges 186
 - text 212
- dial gauges 204
- display rule
 - gauges 209
- dual axis bar chart 198
- dual axis bar-line charts 199
- dual axis line chart 198
- dual axis time series plot 199
- duplicate data items 176

E

- exporting 237
- code 34
- data from crosstabs 240
- graph data 243
- list table data 238
- report data 238
- report objects 238

F

- files, searching for 159
- filtering
 - continuous data 130
 - discrete data 129
 - in Select Data wizard 74
- filters 127, 223
 - creating 128
 - Edit Filter window 131
- folders
 - in Open window 158
- formats
 - calculated column 47
 - modifying for data items 179

G

gauges
 bullet 204
 dial 204
 display rule 209
 inserting into a report 206
 properties 207
 slider 204
 speedometer 204
 styles 208
 thermometer 204
 geo maps 122
 geographic maps 122
 graph types 191
 graphs
See also charts
 bubble plots 196
 displaying results 191
 exporting data from 243
 inserting into a report 200
 pie charts 194
 properties 201
 scatter plots 195
 styles 202
 time series plots 195
 treemaps 197

H

heat maps 120
 histograms 115
 home page

adding comments to reports 22
 discovering details about objects 20
 first look 11
 managing my content 14

I

images 212
 displaying in reports 212
 inserting into a report 212
 properties 215
 importing
 report objects 164
 reports 164
 interactions 223
 adding data brushing 228
 brush 223
 creating filters 224
 filter 223
 removing 229
 interface 11
 designer 153
 home page 11

J

jobs
 data preparation 35
 editing data preparation 35
 SAS Data Integration Studio 35

joining tables
inner join 45

L

line charts 105, 193
list tables 187
 exporting data from 238
 sorting data in 219
logging off 9
logging on 8

M

measure data items 175
measures
 modifying aggregation of 180
missing values, excluding 130
mobile devices
 blacklisting 277
 managing 277
 sorting 277
 viewing 277
mobile viewer 249
 connecting to library server 250
 opening the app 249
modifying role
 data items 178

O

object inspector 20

adding comments 22
discovering details 20
Open window 157
 folders in 158
opening
 reports 161
opening the mobile viewer 249

P

passwords 9
pie charts 194
plots
 bubble plots 196
 dual axis time series plots 199
 scatter plots 195
 time series plots 195
preferences 15
 general 17
 global for all SAS applications 15
initial page 18
process monitoring
 resource utilization 275
properties
 gauges 207
 graphs 201
 images 215
 tables 190
 text 214

Q

queries
 based on data items [169](#)
 data items in [174](#)
 long running [325](#)
 query results
 displaying in tables [187](#)

R

remove data items [177](#)
 removing data sources [173](#)
 renaming
 data items [178](#)
 report sections [234](#)
 report data
 exporting [238](#)
 report objects
 exporting [238](#)
 report sections [233](#)
 adding [233](#)
 deleting [235](#)
 renaming [234](#)
 reports
 adding comments [22](#)
 adding sections [233](#)
 an example [162](#)
 basing on existing report [164](#)
 creating [163](#)
 deleting sections [235](#)
 importing [164](#)
 inserting gauges [206](#)
 inserting graphs [200](#)

inserting images [212](#)
 inserting tables [189](#)
 inserting text [212](#)
 naming [168](#)
 opening [161](#)
 replying to comments [22](#)
 saving [168](#)
 viewing in the Web viewer
[251](#)

resource utilization
 by process [275](#)
 monitoring [271](#)

results
 displaying in graphs [191](#)

roles [7](#)

S

SAS statements
 logging [34](#)
 SAS Visual Analytics [3](#)
 access to capabilities [7](#)
 benefits of [4](#)
 home page [11](#)
 how it works [4](#)
 logging off [9](#)
 logging on [8](#)
 preferences [15](#)
 roles [7](#)
 Web viewer [251](#)

saved reports [161](#)
 saving reports [168](#)
 scatter plots [110, 195](#)
 Search window [159](#)

searching
 for descriptions 159
 for files 159
 for keywords 159
 for report names 159
 security
 limit logon access 321
 select data items 175
 select data sources 170
 Server configuration files 56
 Server run time 56
 slider gauges 204
 sorting
 and analysis 219
 data in crosstabs 220
 data in graphs 221
 data in list tables 219
 sorting data 219
 speedometer gauges 204
 SSH keys 312
 generating and propagating 315
 SSH lockout 321, 324
 standard data items 174
 styles
 gauges 208
 graphs 202
 tables 190

T

table types 187
 tables 108
 See also crosstabs

displaying query results in 187
 inserting into a report 189
 list tables 187
 properties 190
 styles 190
 targeted bar chart 192
 text 212
 displaying in reports 212
 inserting into a report 212
 properties 214
 thermometer gauges 204
 time categories 175
 time series plot 195
 timestamp categories 175
 treemaps 197
 troubleshooting
 HDFS content explorer 258
 Process monitor 276
 Resource monitor 272

U

UNIX groups
 creating 317
 members 318
 UNIX user accounts
 creating 313
 deleting 317
 secondary group membership 318
 user interface 11
 designer 153
 home page 11

user management
audit SSH logon activity [319](#)
creating users [313](#)
deleting users [317](#)
group membership [318](#)
prevent SSH logons [321](#)
show logged on users [320](#)

V

visual filters [129](#)
visualizations [89](#)
arranging [96](#)

creating [95](#)
managing [95](#)
overview [89](#)
types of [89](#)

W

waterfall charts [193](#)
Web viewer
adding comments [253](#)
opening reports [251](#)
overview [251](#)

