SAS® Visual Analytics 7.4: User's Guide

SAS® Documentation
January 8, 2019

The correct bibliographic citation for this manual is as follows: SAS Institute Inc. 2017. *SAS® Visual Analytics 7.4: User’s Guide*. Cary, NC: SAS Institute Inc.

SAS® Visual Analytics 7.4: User’s Guide

SAS Institute Inc., SAS Campus Drive, Cary, NC 27513-2414

January 2019

7.4-P2:vaug

Contents

PART 4 Exploring Data 105

Using This Book

Audience

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

- 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.

Persons responsible for managing SAS servers and managing the SAS Visual Analytics environment should refer to *SAS Visual Analytics: Administration Guide*.

The content of this document is also applicable to other SAS solutions that integrate with and use SAS Visual Analytics features.

Prerequisites

Here are the prerequisites for using SAS Visual Analytics:

- A user ID and password for signing in to SAS Visual Analytics.
- A supported web browser installed on your desktop client.
- A supported version of the Adobe Flash player installed on your desktop client.
- Access to data sources that can be used to obtain data for exploration or reports.

Note: SAS Visual Statistics is visually and functionally integrated with SAS Visual Analytics Explorer. SAS Visual Statistics requires a separate license.

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

Documentation Conventions

This book uses short forms of the following phrases where the meaning is clear from context:

Long Form	Short Form	User Interface Labels*
SAS Home SAS Visual Analytics Hub	the home page	SAS Home Home
SAS Visual Analytics Administrator	the administrator	Administrator
SAS Visual Analytics Explorer**	the explorer	Data Explorer
SAS Visual Analytics Designer	the designer	Report Designer (Create Report)
SAS Visual Analytics Graph Builder	the graph builder	Custom Graph Builder
SAS Visual Analytics Viewer	the viewer	Report Viewer
SAS Visual Data Builder	the data builder	Data Preparation (Prepare Data) (Create Data Query)

* Labels in parentheses are used only in the classic (Flash) presentation mode.

** Not all SAS Visual Analytics orders include the explorer.

What's New

What's New in SAS Visual Analytics 7.4

General Enhancements to SAS Visual Analytics

The general enhancements include:

- SAS Visual Analytics Designer (the designer) has dynamic text, improved calculations, and enhanced prompted filter controls, parameters, report links, and section links.
- The modern SAS Visual Analytics Viewer (the viewer) now contains most of the features that the classic viewer had, so many users can switch to using the modern viewer.
- Enhanced printing features enable you to have page breaks in the PDF for list tables and to show the filter context for filter controls. Another new feature enables the same footer to be printed on every page of a report. For example, you can have the same legal statement on each page in all of your company's reports. The customized footer is also displayed in reports that are distributed using the designer. (This feature must be enabled by a SAS administrator using SAS Management Console.) For more information, see the *SAS Visual Analytics: Administration Guide*.
- Importing data from Teradata no longer requires a user name and password.

SAS Visual Analytics Designer

New and enhanced features in the designer include:

- A new dynamic text feature, which enables you to display a measure value in a large font. The dynamic text can display the last date that an in-memory table was updated. It can display the value of a parameter or the value of a category data item associated with a filter control.
- Calculations have been enhanced. The distinct count option ignores missing values. You can specify a reset-to-zero offset for time-based calculations,

which enables you to adjust for a fiscal calendar or another calendar that does not start on January 1. Time period calculations are not restricted to the applied filter view of data.

- Parameters have been enhanced. You can now have parameters with multiple values, which enables you to use them in a list control. And, parameters can now be based on date and datetime formats.
- There are now cascading prompts. Filter controls placed at the report level and the section level can now include dependencies.
- You can now synchronize prompt values and parameters across linked reports.
- The CumulativePeriod operator for aggregated measures now enables you to customize the starting month for each year (for example, to calculate year-to-date based on the fiscal year for your company).
- Periodic operators for aggregated measures have a new parameter that enables them to be calculated before datetime filters are applied.
- Slider controls have two new properties that enable you to set dynamic minimum and maximum values for the slider. The dynamic minimum and maximum values automatically adjust to the current data query.
- Custom shape files for geo maps are now supported. For more information about custom polygon data, see the *SAS Visual Analytics: Administration Guide*.

SAS Visual Analytics Viewer

New and enhanced features in the modern viewer include:

- Data can be exported to Microsoft Excel or to a CSV file.
- An image of a graph can be exported to a PNG file.
- The context of a report filter can be displayed.
- There is now an option to expand all hierarchies in crosstabs.
- Pages (or sections) in a report can be navigated using tabs.
- You can select all or clear all selections for list prompt controls using the right mouse button.
- You can sort on graphs.
- There is a redesigned landing page, which includes new features. For example, you can e-mail and print a report from the landing page without actually opening the report.

SAS Visual Analytics Administration

For information about changes and enhancements in the administration of SAS Visual Analytics, see the *SAS Visual Analytics: Administration Guide*.

Accessibility

For information about the accessibility of this product, see Accessibility Features of SAS Visual Analytics 7.4.

Part 1

Introduction to SAS Visual Analytics

1

About SAS Visual Analytics

What Is SAS Visual Analytics?

SAS Visual Analytics is an easy-to-use, web-based product that leverages SAS high-performance analytic technologies. SAS Visual Analytics empowers organizations to explore huge volumes of data very quickly to identify patterns, trends, and opportunities for further analysis. SAS Visual Data Builder (the data builder) enables users to summarize data, join data, and enhance the predictive power of their data. Users can prepare data for exploration and mining quickly and easily. The highly visual, drag-and-drop data interface of SAS Visual Analytics Explorer (the explorer), combined with the speed of the SAS LASR Analytic Server, accelerate analytic computations and enable organizations to derive value from massive amounts of data. This creates an unprecedented ability to solve difficult problems, improve business performance, predict future performance, and mitigate risk rapidly and confidently. SAS Visual Analytics Designer (the designer) enables users to quickly create reports or dashboards, which can be viewed on a mobile device or on the web.

Starting in the 7.2 release, the explorer enables you to create, test, and compare models based on the patterns discovered during exploration of the data. The explorer enables you to explore, discover, and predict using your data. You can export the score code, before or after performing model comparison, for use with other SAS products and to put the model into production.

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 important, reports can be displayed on a mobile device or in the SAS Visual Analytics Viewer (the viewer).

Benefits of Using SAS Visual Analytics

Using SAS Visual Analytics, users can enhance the analytic power of their data, explore new data sources, investigate them, and create visualizations to uncover relevant patterns. Users can then easily share those visualizations in reports. In traditional reporting, the resulting output is well-defined up-front. 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 its relationships. Then, when useful visualizations are created, you can incorporate those visualizations into reports that are available on a mobile device or in the viewer.

SAS Visual Analytics provides users with the following benefits:

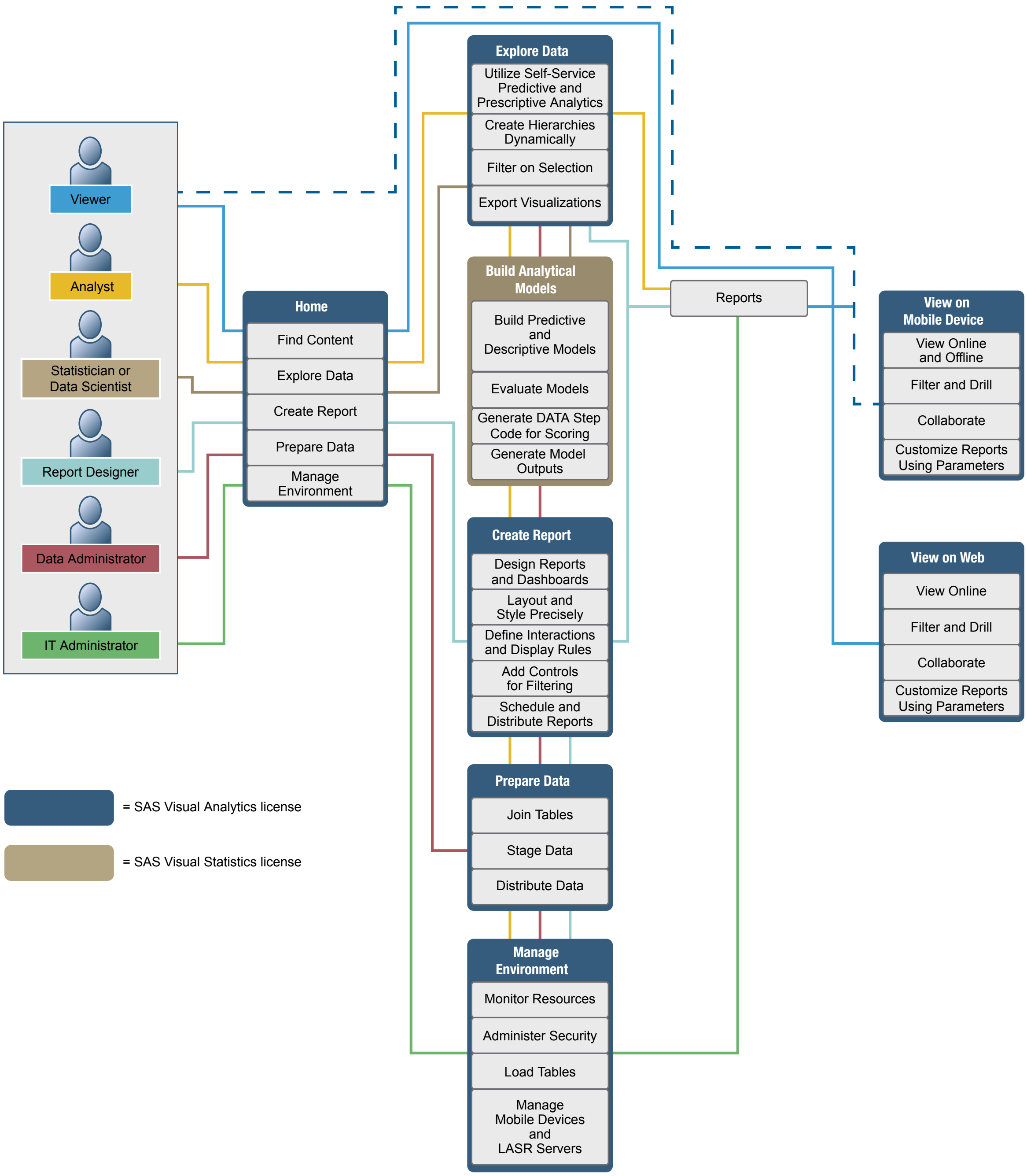
- enables users to apply the power of SAS analytics to massive amounts of data
- empowers users to visually explore data, based on any variety of measures, at amazingly fast speeds
- enables users to quickly create powerful statistical models if SAS Visual Statistics is licensed at your site
- enables users to quickly create reports or dashboards using standard tables, graphs, and gauges
- enables users to quickly create customized graphs
- 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 using a native mobile app or on the web. You can explore your data by using interactive visualizations such as charts, histograms, and tables. Report designers can easily point and click to query central sources of data. You can add filters and design the layout using tables, graphs, and gauges. You can use drag and drop to create a well-formatted report.

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

Figure 1.1 Overview of SAS Visual Analytics



SAS Visual Analytics Community

The SAS Visual Analytics community is dedicated to users who are focused on exploratory visualization and analytical techniques, data preparation, dashboard reporting, and mobile reporting. You can share your experiences, discuss topics and ideas, seek help from your peers, and share information about upcoming events. You can access the user community at communities.sas.com/visual-analytics.

For questions requiring immediate technical assistance, contact SAS Technical Support at support.sas.com.

2

Accessing SAS Visual Analytics

About SAS Visual Analytics Users

Authenticated Users

SAS Visual Analytics uses the standard sign-in window for SAS applications. To display the sign-in window, use the URL that is supplied by your system administrator. For example, you might enter: http://*host*/SASVisualAnalyticsHub

Click **Sign Out** in the upper right corner of the user interface to sign out of SAS Visual Analytics. When you click **Sign Out**, you are signed out of all SAS web applications. For example, suppose that you have SAS Home (the home page), the explorer, and the designer open, and then you click **Sign Out** when you finish working on a report in the designer. In this case, you have also signed out of the home page and the explorer.

Guest Access

SAS Visual Analytics system administrators can configure support for guest access. Users with guest access can access only the home page and SAS Visual Analytics Viewer (the viewer). Guest access uses a shared account, so it does not provide individualized features, such as history or alerts. If provided by the system administrator, favorites and preferences are read-only features.

Accessing SAS Visual Analytics as a guest is useful if you do not have a metadata identity. This enables you to view reports that are widely available under a generic, shared account. You can also view reports that are available to the public on the Internet.

Use SAS Home to Access SAS Visual Analytics

After you sign in to SAS Visual Analytics using the standard sign-in window for SAS applications, you will see SAS Home (the home page). The home page enables you to create new content in SAS Visual Analytics. In addition, it enables you to access content that you and others have created. For more information, refer to the online Help that is available for the home page.

Overview of 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.

Note: Access to functionality depends on how SAS Visual Analytics is installed at your site. For example, a site might not have the explorer.

SAS Visual Analytics provides five predefined roles—Basic, Report Viewing, Analysis, Data Building, and Administration. A predefined set of capabilities is available for each role. A system administrator can modify these roles and specify the capabilities for each role 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 the 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.

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 a Data Building role to prepare data.
- your location in SAS Visual Analytics. For example, some application features 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*.

About Application Themes in SAS Visual Analytics

An application theme is the collection of colors, graphics, and fonts that appear in the application. SAS Visual Analytics provides the following themes: SAS Corporate (default theme), SAS Blue Steel, SAS Dark, SAS High Contrast, and SAS Light. To change the application theme, see "Specifying Your Preferences" on page 9.

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

Report themes are available in the designer. The designer provides the following report themes: SAS Snow, SAS Light, SAS Dark, or SAS High Contrast. SAS Snow is the default report theme. Your site might also have custom report themes. For more information, see "About Report Themes" on page 304.

Specifying Your Preferences

All of your preferences and settings persist between sessions. Settings and preferences are not available for guest users.

Specify Preferences for SAS Visual Analytics

You can specify preferences for SAS Visual Data Builder (the data builder), SAS Visual Analytics Explorer (the explorer), SAS Visual Analytics Designer (the designer), SAS Visual Analytics Graph Builder (the graph builder), SAS Visual Analytics Viewer (the viewer), or SAS Visual Statistics. For example, you can specify a default scheduling server for the data builder, a default map provider mode for the explorer, or your preferred report theme for the designer.

Preferences are saved on a per-user basis.

If you are in the data builder, the explorer, the designer, the viewer, or SAS Visual Statistics, then select **File** ▸ **Preferences** to open the Preferences window.

For information about the specific preferences that are available, see the following:

- Preferences for the data builder on page 41
- Preferences for the explorer on page 111

- Preferences for SAS Visual Statistics on page 236
- Preferences for the designer on page 305
- Preferences for the graph builder on page 508
- Preferences for the classic viewer on page 545

Note: Settings are available for the modern viewer. For more information, refer to the online Help that is available for the modern viewer.

Specify Settings Using SAS Home

You can specify settings for the modern home page (or preferences for the classic home page). For example, you can specify the initial screen when the home page is displayed.

Settings for the home page affect SAS Visual Analytics. For example, you can specify which piece of SAS Visual Analytics opens a report. Suppose that you specify **Edit - Report Designer** as the first action in the list of actions for the **SAS report (2G)** content type. Then, you receive an e-mail message with a link to a report. When you click the link to the report, it is displayed in the designer instead of the viewer.

To specify settings:

1. On the modern home page, select your name, and then click or tap **Settings**.
2. Click or tap **Home** in the side menu.
3. Specify any of the following settings:

   Default Appearance
   enables you to specify the appearance when the home page is displayed.

   Initial Screen
   enables you to specify the initial screen when the home page is displayed. The available options depend on your role and capabilities.

   Application Shortcuts
   enables you to specify the order of your application shortcuts.

   Tiles
   enables you to specify which tiles are displayed and the order in which they are displayed.

   For more information, refer to the online Help that is available for the modern home page.
4. Click or tap **Done** to apply your changes.

For information about the classic home page, see “Specify Your Preferences for the Classic Home Page” on page 638.

Specify Global Settings Using SAS Home

You can specify global settings that are applied to all SAS web applications. These settings are set by each user.

To specify global settings:

1. On the modern home page, select your name, and then click or tap **Settings**.

2 Click or tap **Global** in the side menu.

3 Specify any of the following settings:

General
enables you to specify the **User locale** and **Theme**.

Note: If you change the **User locale**, then you must sign out and sign in to SAS Visual Analytics for the change to take effect.

Side Menu
enables you to hide or change the order of the SAS applications that are displayed in the side menu.

Accessibility
enables you to specify your preferences for assistive technologies.

For more information, refer to the online Help that is available for the modern home page.

4 Click or tap **Done** to apply your changes.

For information about the classic home page, see “Specify Global Preferences Using the Classic Home Page” on page 637.

Personalize SAS Visual Analytics Using SAS Home

SAS provides accessibility features that you can use to personalize the user interface to make it easier to use. Accessibility features are part of the global settings, which are applied to all SAS web applications, including SAS Visual Analytics. You can specify global settings using the modern home page. For more information, see “Specify Global Settings Using SAS Home” on page 10.

The following accessibility features are available:

- **Themes:** You can change the colors, graphics, and fonts that appear in the application. Several themes are available, including the SAS High Contrast theme.

  You can also change the themes that are used in reports. For more information, see “About Report Themes” on page 304.

- **Configurable focus indicator:** The interface provides an indicator of the current location of the focus. You can configure the focus indicator to make it easier to see.

- **Color inversion:** You can invert the colors in the interface to improve readability.

- **Zoom support:** You can make your screen content larger and smaller by using the zoom in (Ctrl+plus sign) and zoom out (Ctrl+minus sign). You can reset the zoom state (Ctrl+0) keyboard shortcuts.

- **Landmarks in the user interface:** Landmarks are references to the primary areas of an application’s interface. They provide a quick and easy way for keyboard users to navigate to these areas. Press Ctrl+F6 to open the Landmarks window.

- **Keyboard shortcuts:** You can use keyboard shortcuts as a quick and easy way to perform tasks or navigate the user interface. For more information, see "Keyboard Shortcuts for SAS Visual Analytics" on page 549.

For more information, see Accessibility Features of SAS Visual Analytics 7.4.

For information about the classic home page, see "Specify Global Preferences Using the Classic Home Page" on page 637.

Part 2

Accessing Data

3

Overview of Data Flow in SAS Visual Analytics

Data Flow in SAS Visual Analytics

All reporting and exploration of data in SAS Visual Analytics is performed against data that is in memory on a SAS LASR Analytic Server. In addition to the in-memory tables, you can use your own data with self-service data access if you have been granted the capability to import data.

If your site has data that must be prepared before it is ready for analysis, then the data builder can perform basic data preparation and load data to memory on a SAS LASR Analytic Server. The following list summarizes the data builder features that enable the flow of data from source systems into SAS Visual Analytics:

- Joining tables, such as fact and dimension tables.
- Working with SAS/ACCESS engines to read data from operational systems.
- Scheduling data queries for basic data flow automation.

SAS Visual Analytics Administrator provides features that are most commonly associated with managed data access. The administrator enables control of whether tables are loaded to memory and secures access to in-memory tables.

Self-Service Data Access

SAS Visual Analytics offers a variety of ways for non-administrative users to bring data into the SAS Visual Analytics environment, which enables you to work with your data quickly. The following list summarizes the different ways:

Import data from a file
: When you are creating data queries, explorations, or reports, you can import data from a Microsoft Excel spreadsheet, a delimited text file (CSV), or a SAS data set. After you import the data into the SAS LASR Analytic Server, it is ready to use.

Import SAS data set on a server
: You can direct the SAS Application Server to import a SAS data set that is already on the server. The import process is optimized and you can import large files when a SAS data set is already on the server.

Import a database table
: After providing connection information, you can transfer a table from a database to a SAS LASR Analytic Server. Afterward, you can use the in-memory table as is for reports and explorations, or you can prepare it for analysis using the data builder.

  Note: Your site must license and configure the related SAS/ACCESS engine to use this feature.

Import data from Facebook, Google Analytics, or Twitter
: After authenticating with Facebook, Google Analytics, or Twitter and providing search criteria, you can import data into memory on a SAS LASR Analytic Server. You can then use the unstructured data with the explorer. For example, a typical exploration of Twitter data can perform text analysis to look for patterns and trends in the tweets.

  Note: Your access to, and use of, social media data through a social media provider's public APIs is subject to the social media provider's applicable license terms, terms of use, and other usage terms and policies.

TIP The self-service import feature keeps track of your most recently used values for each import type to simplify repeated import actions.

Managed Data Access

Many sites prepare data sources to be used by business analysts. This information can come from data sources such as data warehouses and transactional systems. A data administrator enables access by registering the tables and libraries in SAS metadata. A data administrator might use additional SAS products to provide advanced data governance, data quality, and data management support.

SAS/ACCESS engines can be used to access data in operational systems, transactional systems, or data warehouses. SAS offers a variety of SAS/ACCESS engines for accessing data from operational systems. These engines must be licensed and configured at your site to connect to the data. After a connection is set up, you can use SAS Management Console and SAS Visual Analytics Administrator for registering the libraries and tables from these data sources.

After a SAS LASR Analytic Server Restart

When an administrator restarts a SAS LASR Analytic Server, all the tables on the server are removed from memory as the server stops. Tables that you import with the self-service features remain in memory so long as the server is running. However, administrators cannot interactively reload imported tables for you. If you want to use a table that you imported after a server restart, then you must repeat the import action.

TIP Your administrator can configure a server to automatically reload tables that were imported from local files after a restart.

Requirements for Importing Data

Importing data requires starting a SAS session on the SAS Application Server. Typically, this requirement is met by each user who has a host account.

For deployments on Microsoft Windows, the host account must have the **Log on as a batch job** Windows privilege. For deployments on Linux that use a distributed SAS LASR Analytic Server, the host account must be configured for passwordless SSH. For more information about working with the server, see *SAS LASR Analytic Server: Reference Guide*.

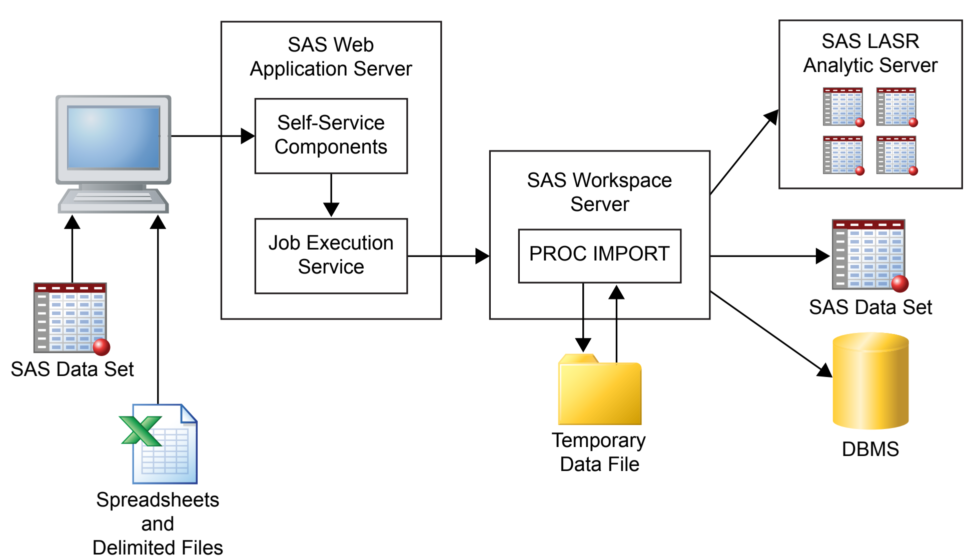
4

Importing Local Data Files

Import a Local Data File

When you import a local data file from your desktop, such as a spreadsheet, a delimited text file, or a SAS data set, the file is transferred as data to SAS LASR Analytic Server. This enables you to access data without needing assistance from an administrator or information technology group.

The following figure shows how a file is accessed from your PC, transferred to the SAS Workspace Server, and then stored in an output table.



Note: When you import a SAS data set, it is not processed with PROC IMPORT. SAS data sets are transferred to output with a DATA step.

Note: Only the data builder can output a SAS data set or a DBMS table. The explorer and designer can import data to SAS LASR Analytic Server only.

You can import data files that are available from the file systems on your PC. This includes local file systems such as `C:\` on Windows machines and paths such as `/home/$USER` on UNIX machines. Network file systems and shared folders are included, such as UNC paths like `\\nas\spreadsheets`.

To import data from a Microsoft Excel spreadsheet, delimited text file, ZIP file, or SAS data set:

1 In the Import Data window, click the link for the type of data file that you want to import, and then select the file.

2 Specify the following input file options:

Note: For a SAS data set, there are no options to specify.

Spreadsheet options
The following options are available when you import a Microsoft Excel spreadsheet:

Select worksheet
Select **All** or select the check boxes for the worksheets to import.

Begin import on row
The default is to import data from the first row of the spreadsheet. If the data begins on a different row, select the row.

Includes column names
Select this check box when the row on which to begin the import has the column names.

Text file options
The following options are available when you import a delimited text file or a ZIP file:

Delimiter
Select the delimiter that is used in the file that you want to import. You can specify a single character to use as a user-defined delimiter.

First row contains column names
Select this check box when the file has the column names on the first row.

Data records begin on row
The default is to import data records from the second row. If you deselect **First row contains column names**, then this value indicates the first row.

Number of rows to scan
The default is to read up to the first 500 rows in the file to determine the data type and length for each column. A smaller value causes the import to complete quickly, but you risk the chance of determining a value too short for character columns. Larger values reduce the chance of truncating character columns, but they increase processing time.

Encoding
Select the encoding of the file. If you are importing UTF-8 or UTF-16 data, then make sure that the SAS Web Application Server is a Unicode server or that all of the file contents can be successfully transcoded to the encoding of the SAS Web Application Server.

3 (Optional) Click **Preview** to view the data. Preview displays up to 500 rows from the file.

> **TIP** Previewing the data can help you determine whether you specified the correct encoding.

4 (Optional) In the **Output Table** section, enter the name for the output table and a description. The description is limited to 256 characters. Review the library and location settings by clicking **Advanced**. Make any necessary changes.

You do not have access to the **Advanced** section if you are using the explorer or designer and do not have the Build Data capability. Instead, you can import the data to a general-purpose area or select **Store the table in a private location** to prevent other users from accessing the data.

5 Click **OK**.

Limitations and Restrictions for Importing Local Data Files

Large Data Files

When you import data files (spreadsheets, SAS data sets, or delimited text files) from your desktop, you are limited to files that are 4 GB or less. This constraint is set by the web browser. Because importing large data files through the web browser impacts overall performance and because you can experience long wait times, alternate approaches are recommended for importing large files.

Note: Administrators can specify a limit that is less than 4 GB.

As an alternative to importing data files through your web browser, you can use autoload. You can autoload data from files that are larger than 4 GB. In general, FTP and network file systems transfer data files faster than web browsers.

Another alternative for importing a large data file is to compress it and import it as a ZIP file. You can use this option to import comma-separated values (CSV) text files only. To use this option, compress the file that you want to import, and then click **Text Files** in the Import Data window. Here are some key points about importing ZIP files:

- Make sure that the ZIP file contains only one file. If the ZIP file contains more than one file, then only the first file is imported.
- A ZIP file must contain only comma-separated values (CSV) files. Other file formats are not supported.
- A CSV file must have the file extension .csv. Other file extensions are not supported.

Importing Data from Spreadsheets

The following list identifies considerations for importing data from spreadsheets:

- You can import Excel workbook (XLSX, XLSM, and XLSB) files and Excel 97-2003 workbook (XLS) files. You cannot import XLST or other Excel file types.

  Here are some key points about importing XLSB files:

  - The machine for the SAS Workspace Server must have the following provider software installed: Microsoft Data Access Components (MDAC) and Microsoft Jet (Joint Engine Technology) or Microsoft Access Database Engine (formerly know as Microsoft Access Connectivity Engine or ACE) for 2007 and later.
  - The bit version of Microsoft Access Database Engine must be the same as the bit version of SAS.
  - You can import XLSB files only if the SAS Workspace Server is running on Microsoft Windows.

  TIP If your spreadsheet is from an unsupported Excel file type, then try saving it as an XLSX file before importing it.

- When you import a spreadsheet (from your PC) that has multiple worksheets, by default, all of the worksheets are imported. A table is created for each worksheet. You can clear the check boxes for the worksheets that you do not want to import.
- Importing pivot tables is not supported.

Table Names, Column Names, and Special Characters

In general, you can import files that use blanks and special characters in the filenames and column names. The following list identifies how table names are handled:

- For text files (such as CSV files), the table name is initially set from the filename.
- When you import a spreadsheet, table names are handled as follows:
  - If the spreadsheet contains a single worksheet, then the table name is initially set from the filename.
  - If the spreadsheet contains multiple worksheets, then each table name (for each worksheet) is initially set as a combination of the filename, an underscore, and the name of the worksheet.
- Some special characters can be used, including spaces. Unsupported special characters include `/ \ * ? " < > | :` - and period (`.`). After the initial table name is determined, any unsupported special character in the name is replaced with an underscore.
- Table names are shortened to 32 characters because that is the table name length that is supported by SAS. The entire name appears as the label.

If you clear the **Includes column names** check box or the **First row contains column names** check box, then the column names are generated for you as follows:

- **Spreadsheets** Column names are assigned A, B, C, and so on.

- **Text files** Column names are assigned VAR1, VAR2, and so on.

The following table identifies how a column name that begins with a number, such as `2014sales`, or that uses numbers only, such as `2014`, will appear after being imported:

Source File Column Name	Imported Column Name
`2014sales`	`_2014sales` for Microsoft Excel files and text files.
`2014`	`_2014` for XLSX, XLSM, XLSB, and text files. For Excel 97-2003 workbook (XLS) files, a letter such as `A`, `B`, `C`, and so on, is substituted for the column name, depending on the column position in the file.

Usage Notes

Review the following notes if you have trouble importing data:

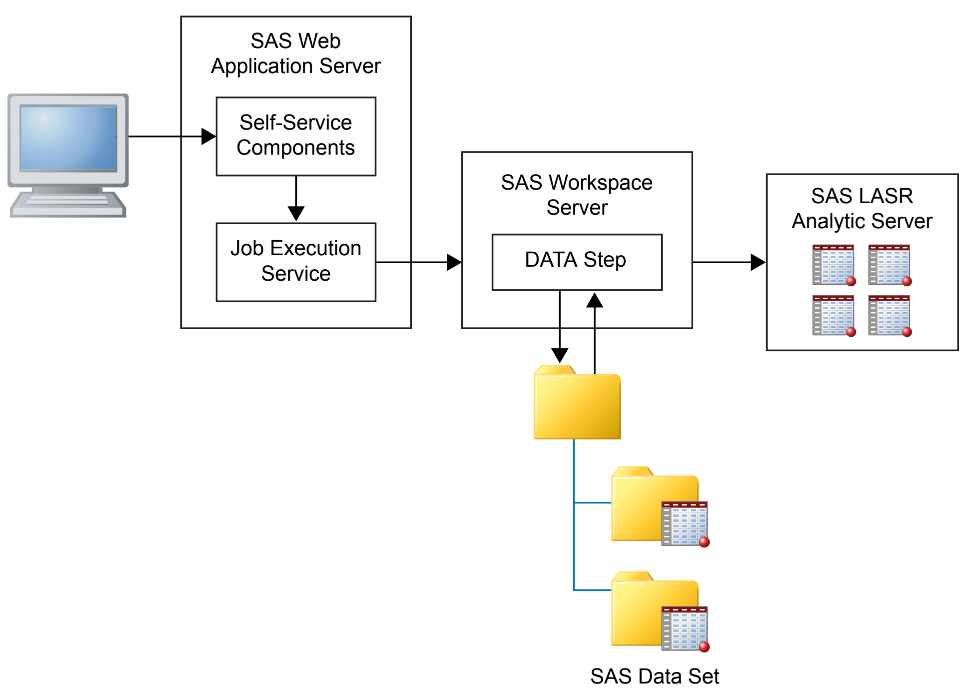
- Before you click **OK** to import the data, click **Preview**. Preview shows an accurate representation of the column names and data values that will be available after the import.
- If SAS is configured as a Unicode server at your site, then you have the most flexibility for importing data. Specifically, SAS as a Unicode server helps with using column names or filenames (that are used as table names) that have double-byte characters.
- When you import a delimited text file (CSV file), you must specify the encoding of the text file. In some cases, the import reports success, even though the data might be corrupted. It is important to verify the imported data.
- If you import a SAS data set that uses user-defined formats, then you must ensure that the custom format catalog is available to the SAS Application Server. For more information, see "Working with User-Defined Formats" on page 39.
- If importing large data files at the same time is common for your deployment, then you should be aware that large data files are written to temporary disk space on the server. In extreme cases, this can cause temporary disk space to become full. Systems that run out of disk space can become unresponsive and difficult to troubleshoot.
- If you import data from text files and plan to append the data, then you must verify that the column data types and lengths match the table that you want to append to.
- When you import data, a SAS LASR Analytic Server does not maintain pre-existing sort orders. You must re-sort the data after you import it.
- Importing indexed SAS data sets is not supported.

5

Importing Data from Servers

Import a SAS Data Set on a Server

The following figure depicts how your PC can be used to specify a SAS data set on the SAS Application Server machine (represented as the SAS Workspace Server), and then load the data set into memory on a SAS LASR Analytic Server.



You can direct the SAS Workspace Server to access the file systems on the server. For example, if you have a large data set, you can use FTP or another method to copy it to a directory on the server, and then use the server to import it.

To import a SAS data set that is accessible from your SAS Application Server:

1 Select **SAS Data Set** from the list of **Server** data types, navigate to the SAS data set, and click **OK**.

   Note: Remember that the data files and directories are on the remote machine, not on your PC.

2 (Optional) In the **LASR Table** section, enter the name for the table and a description. The description is limited to 256 characters. Review the library and location settings by clicking **Advanced** and make changes if necessary.

   You do not have access to the **Advanced** section if you are using the explorer or designer and do not have the Build Data capability. Instead, you can import the data to a general-purpose area or select **Store the table in a private location** to prevent other users from accessing the data.

3 Click **OK**.

Import a Database Table

To import database tables, the SAS/ACCESS product for the database must be licensed and configured for the SAS Workspace Server. You can import data from the following databases:

- Server databases — SAS Data Set, Aster, DB2, Greenplum, MySQL, Netezza, ODBC, Oracle, PostgreSQL, Salesforce, SAP HANA, SQL Server, Teradata, Vertica
- Hadoop databases — BigInsights, Cloudera, Cloudera Impala, Pivotal HAWQ, Hortonworks, MapR, Pivotal HD

Note: BigInsights, Cloudera, Hortonworks, MapR, and Pivotal HD all use the SAS/ACCESS Interface to Hadoop, even though they have different menu selections.

To import a database table:

1 From the Import Data window, select the database name from the list of server or Hadoop databases. These lists include only the data sources for which a SAS/ACCESS product is licensed and configured and that your administrator has granted you the capability to use.

   If you want to import SAS Data Set on a server, see “Import a SAS Data Set on a Server” on page 25 for more information.

2 Specify the connection information. Here are some key points about specifying connection information:

   - The **Server** field corresponds to the host name for the server. Some databases connect using a data source name instead of the combination of server and port.

- In the **DBMS table names** field, you can import multiple tables at the same time. To do this, hold down the Ctrl key while selecting the table names in the Choose Tables window. During the import, an icon in the **Status** column indicates whether the table was successfully imported, if it failed, or if you chose to cancel. By clicking the link in the **Remarks** column, you can view additional information, such as log or error messages.

For connection details for specific databases, see the “Database Connection Tips” on page 28 topic.

Most fields are case sensitive. For example, specifying a value of **products** in the **Database** field might not be the same as specifying **PRODUCTS**. Case sensitivity depends on the database vendor. Furthermore, some databases use schemas. Some databases automatically use the user ID as the schema if a schema is not explicitly specified. Be aware that the **User ID** and **Schema** fields can be case sensitive. Check with your database administrator if you are unsure.

3 Click **Browse** to select the table to import.

4 (Optional) Expand **Options** to indicate additional connection options. Here are some key points about specifying additional connection options:

- For more information about valid values for the ***Database* options** field, see *SAS/ACCESS for Relational Databases: Reference*. View the Data Set Options topic for the type of database that you are working with (for example, Data Set Options for ODBC).
- You can use the **SAS system options** field to specify environment variables such as the following:

```
set=SAS_HADOOP_JAR_PATH="/path/to/files"
```

The **options** keyword is submitted with any options that you specify in this field.

5 (Optional) Review the library and location settings by clicking **Advanced**. Make any necessary changes.

You do not have access to the **Advanced** section if you are importing data from the explorer or designer and do not have the Build Data capability. Instead, you can import the data to a general-purpose area or select **Store the table in a private location** to prevent other users from accessing the data.

6 Click **OK**.

After you have successfully imported a table, the connection information is saved, except for the password. This enables you to import additional tables quickly or to reload the table as needed.

If you want to reload data for an existing table using the same table name, you must do one of the following:

- Use the same library and output folder that were used when the data for the table was originally imported.
- Indicate both a different library and a different output folder than when the data for the table was originally imported.

If the table fails to reload, the log might not contain any error information. In this case, the log is most likely displaying information about the last successful action on the table.

Note: If you are importing a single table, then you cannot cancel it. You can choose to cancel if you are importing multiple tables. However, the table that is being processed when the cancel was submitted cannot be canceled and will finish loading.

Database Connection Tips

Additional Options for Importing Hadoop Tables

SAS Visual Analytics offers self-service options for importing data from BigInsights, Cloudera, Cloudera Impala, Pivotal HAWQ, Hortonworks, MapR, and Pivotal HD. Each of these databases requires separate setup by your administrator.

A common connection type for all of these databases is to connect to Hive or HiveServer2, and then import tables.

If the Hadoop cluster is configured with the SAS Embedded Process, then you can perform parallel loading to the SAS LASR Analytic Server. In this case, indicate one of the following in the Import Data window:

- In the **SAS system options** field, specify the SAS_HADOOP_CONFIG_PATH environment variable. This is not necessary if your administrator has already specified the values.
- In the **Configuration** field, specify the path to a Hadoop configuration file.

Specifying more options might be necessary for your site. For information about setting up parallel loading from Hadoop, see “Where Do I Locate My Analytics Cluster” in *SAS Visual Analytics: Installation and Configuration Guide (Distributed SAS LASR)*.

Additional Options for Importing ODBC Tables

The **Specify connection options** field provides you with an additional way to connect to an ODBC database instead of using a data source name. For more information about possible options, see the LIBNAME Statement Specifics for ODBC topic in *SAS/ACCESS for Relational Databases: Reference*.

Additional Options for Importing Oracle Tables

The value for the **Path** field is related to the net service name in the tnsnames.ora file. The tnsnames.ora file is generated during the Oracle client installation on the machine for the SAS Web Application Server. The file is typically stored in an Oracle installation directory such as `/opt/oracle/app/oracle/product/10.2.0/db_1/network/admin/tnsnames.ora`. The net service name for the connection information is in this file. See the following figure:

```
# tnsnames.ora Network Configuration File:
C:\oracle\product\10.2.0\client_1\network\admin\tnsnames.ora
# Generated by Oracle configuration tools.

NEWSERVER10G =
  (DESCRIPTION =
    (ADDRESS_LIST =
      (ADDRESS = (PROTOCOL = TCP)(HOST = server.na.sas.com)(PORT = 1521))
    )
    (CONNECT_DATA =
      (SERVICE_NAME = server10G)
    )
  )
```

Additional Options for Importing PostgreSQL Tables

The **Schema** field is not case sensitive when you browse for tables, but it is case sensitive when the import is performed. As a result, if you specify a schema in the wrong case, you can successfully browse for a table, and then select it in the **Choose Tables** window. However, the import fails. In this case, contact your database administrator for assistance with the schema name.

Additional Options for Importing Teradata Tables

The **Teradata Management Server** field is used to determine whether the SAS LASR Analytic Server is co-located on the same data appliance. If the SAS LASR Analytic Server and the Teradata database are on the same data appliance, then make sure that the **Teradata Management Server** field includes the host name that the SAS LASR Analytic Server uses.

SAS Visual Analytics and the Teradata database can be configured to transfer data in parallel when they are not co-located on the same data appliance. For information about setting up parallel loading, see the Where Do I Locate My Analytics Cluster topic in *SAS Visual Analytics: Installation and Configuration Guide (Distributed SAS LASR)*.

6

Importing Data from Other Sources

Import Data from Facebook

To import data from Facebook:

1 Click **Facebook** in the Import Data window.

2 Click **OK** in the Import Facebook Data window to accept the terms and conditions.

  If you would like to view the terms or remove authorization after completing this step, click **Clear Authorization** on the lower left-hand side of the Import Facebook Data window.

3 Indicate the Facebook fan page that you want to import. Valid values include the full URL (for example, `http://www.facebook.com/SASsoftware`) or the page name (for example, `SASsoftware`).

4 Select a date range.

5 Enter the maximum number of posts, comments, and replies to return.

6 (Optional) In the **LASR Table** section, modify the table name and description. Review the library and location settings by clicking **Advanced**. Make any necessary changes.

  You do not have access to the **Advanced** section if you are using the explorer or designer and do not have the Build Data capability. Instead, you can import the data to a general-purpose area or select **Store the table in a private location** to prevent other users from accessing the data.

7 (Optional) Review the **Proxy Server** section. Make changes if necessary.

8 Click **OK**.

A Facebook limitation can cause the table to be created with only partial data. If this occurs, you will receive a warning message. To get a complete data set, try making one or more of the following adjustments before rerunning the import:

- Indicate a smaller date range.
- Limit the number of posts, comments, or replies.
- Run the import at a lower traffic time.

Import Data from Google Analytics

To begin importing data from Google Analytics, your administrator must give you access. Contact your administrator and provide your Google Analytics account information. Administrators must give access using the Google Analytics website. The drop-down menu options for the **Analytics account**, **Property**, and **View (Profile)** fields in the Import Google Analytics window will be based on the type of access that your administrator assigns to you. To import data from Google Analytics:

1 Click **Google Analytics** in the Import Data window.

2 Click **Obtain access code** in the Import Google Analytics window. The Google Analytics service opens in a new window.

3 After signing in with your Google user name and password, click **Accept** to accept the terms and conditions.

4 Highlight and copy (Ctrl+C) the access code.

5 Navigate back to SAS Visual Analytics. Paste the code (Ctrl+V) in the **Paste access code here** field.

6 Click **OK**. The Import Google Analytics window appears.

If you would like to change users or remove authorization after completing this step, click **Clear Sign In Information** on the lower left-hand side of the Import Google Analytics window.

7 (Optional) Modify the fields and date range based on the data that you would like to import.

Note: The default value for the maximum number of rows that you can import is 100,000. For more information, contact your administrator.

8 (Optional) In the **LASR Table** section, modify the table name and description. Review the library and location settings by clicking **Advanced**. Make any necessary changes.

You do not have access to the **Advanced** section if you are using the explorer or designer and do not have the Build Data capability. Instead, you can import the data to a general-purpose area or select **Store the table in a private location** to prevent other users from accessing the data.

9 (Optional) Review the **Proxy Server** section. Make any necessary changes.

10 Click **OK**.

Note: As an attempt to decrease the processing time of your import, Google Analytics sometimes returns sampled data. If this occurs, you will receive a warning message. A column that indicates that the data is sampled will appear in the table. To increase your chance of getting a complete data set, try rerunning the import with a smaller date range.

Note: If you want to reload data for an existing table using the same table name, you must do one of the following:

- Use the same library and output folder that were used when the table was originally imported.
- Indicate both a different library and a different output folder than when the table was originally imported.

Note: It is recommended that your SAS server be configured to use UTF-8 encoding while importing data from Google Analytics. Otherwise, it could impact your ability to import data that includes nonstandard ASCII characters.

For more information about dimensions and metrics in Google Analytics, see https://support.google.com/analytics/answer/1033861?hl=en and https://developers.google.com/analytics/devguides/reporting/core/dimsmets.

Import Tweets from Twitter

To search for tweets and import them:

1. Click **Twitter** in the Import Data window.

   The first time you import tweets, you are directed to the Twitter website to log on to your account and authorize SAS Visual Analytics. After you enter your logon information and click **Authorize app**, the SAS product page opens. Close this page and navigate back to SAS Visual Analytics.

   After the initial logon, SAS Visual Analytics uses authorization tokens for accessing Twitter instead of requiring you to log on each time. If you would like to change users or remove authorization after completing this step, click **Clear Sign In Information** on the lower left-hand side of the Import Twitter Data window.

2. Enter a search term and the maximum number of tweets to return.

   The search operators that you can use are described at https://dev.twitter.com/rest/public/search.

   Note: SAS does not support the following:

   - searches in languages other than English. Requests in other languages (such as specifying `lang=fr`) return no tweets.
   - search terms cannot include double-byte characters.

3. (Optional) In the **LASR Table** section, enter the name for the table and a description. Review the library and location settings by clicking **Advanced**. Make any necessary changes.

   You do not have access to the **Advanced** section if you are using the explorer or designer and do not have the Build Data capability. Instead, you can import the data to a general-purpose area or select **Store the table in a private location** to prevent other users from accessing the data.

4. (Optional) Review the **Proxy Server** section. Make any necessary changes.

5 Click **OK**.

The search results from a Twitter import in SAS Visual Analytics and the search results from Twitter's own search interface do not match exactly. Each uses a different mechanism to download tweets. A Twitter import in SAS Visual Analytics uses Twitter's public search API. There are limits on what data and how much data that SAS can download using Twitter's public search API. These limits might not apply to Twitter's own search interface. For more information, see https://dev.twitter.com/rest/reference/get/search/tweets.

For information about the data structure of imported tweets, see Appendix 10, "Schema for Imported Tweets," on page 621.

Note: Rate limits apply to the Twitter service. Such limits are beyond the control of SAS Visual Analytics. You are required to follow all applicable terms of use that Twitter and others might promulgate for Twitter data.

Note: Certain functionality in SAS Visual Analytics enables you to invoke external third-party resources. Be aware that use of these resources might result in disclosure and transmission of information that you submit to these resources.

Part 3

Preparing Data

7

Overview of SAS Visual Data Builder

What Is SAS Visual Data Builder?

The data builder enables analysts and data administrators to perform basic data preparation. You can create data queries to perform joins, add calculated columns, and subset and sort data. Several productivity features accelerate the creation of columns based on common aggregation functions.

Once you create your data queries, you can reuse them as subqueries for more sophisticated data queries, export them as jobs for scheduling, or schedule them directly from the user interface.

The data builder has self-service data import features that enable you to access data from the following data sources:

- Microsoft Excel spreadsheets
- delimited text files
- SAS data sets
- database tables
- Facebook, Google Analytics, and Twitter

After you import the data, you can prepare it for analysis or join it with existing data.

The data builder provides a series of features that take advantage of the in-memory tables in SAS LASR Analytic Server.

You can perform the following operations to add data to memory in the server:

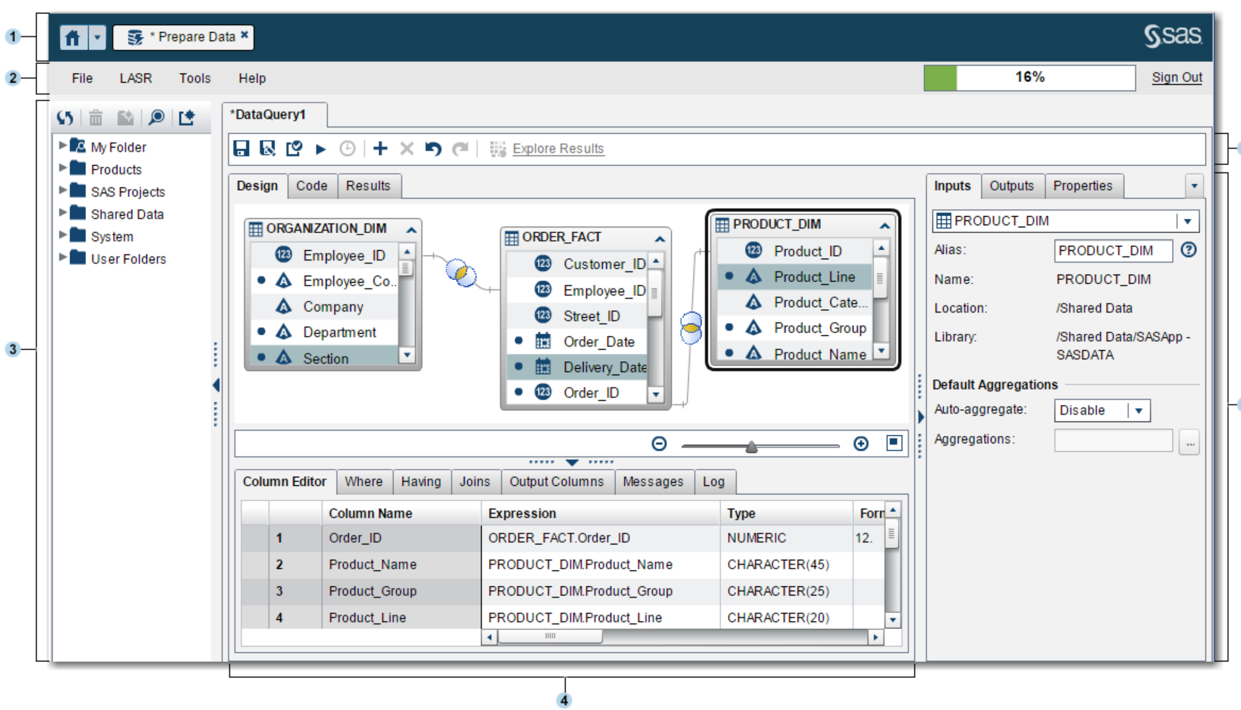
- load an existing table directly into memory
- load the results of a data query into memory (or stage the data and then load it into memory)
- append rows to an in-memory table

After the data is in memory, you can perform the following operations with in-memory tables:

- join in-memory tables to form a LASR star schema
- append entire in-memory tables to another in-memory table
- save in-memory tables to SASHDAT for persistence and fast reloads

Your First Look at SAS Visual Data Builder

Here are the features of the data builder:



1. The application bar enables you to return to the home page and to access other parts of SAS Visual Analytics and other SAS applications that integrate with the home page. You can access your recently created or viewed reports, explorations, stored processes, data queries, or other objects in your recent history. Buttons are displayed for each open application.
2. The menu bar contains menus that enable you to perform tasks such as creating new data queries and LASR star schemas. The right side of the menu bar has a memory gauge that displays the memory utilization for a distributed SAS LASR Analytic Server. You can also sign out of SAS Visual Analytics.
3. The navigation pane displays a tree (the SAS Folders tree) of tables and data queries.
4. The center of the screen contains the workspace. When you create a new object, such as a data query, it is represented as a tab on the top of the workspace.

   The bottom of the workspace contains a series of tabs that enable you to create column expressions, joins, and filter data.

5 The right pane enables you to manage the properties of the item that is selected in the workspace.

6 The toolbar contains icons that enable you to manage, run, and schedule data queries.

Importing Data

The data builder has self-service data import features. After you select **File** ▸ **Import Data**, you can click the link for the type of data that you want to import.

For more information, see Chapter 4, “Importing Local Data Files,” on page 19, Chapter 5, “Importing Data from Servers,” on page 25, and Chapter 6, “Importing Data from Other Sources,” on page 31.

About Managed Access to DBMS Data

The data builder can be used to read source tables from third-party vendor databases and to write tables to them as well. In order to use this feature, your site must have a SAS/ACCESS Interface product license for the database that you want to use. The user ID and password that you use to log on to the data builder might not be valid for a third-party vendor database. If this is the case, then you are prompted for credentials to the DBMS when you access a registered table from a library with a Read operation or a Write operation. As an alternative to being prompted, you can store a login in metadata that has valid DBMS credentials. For more information, see "How to Store Passwords for a Third-Party Server" in *SAS Intelligence Platform: Security Administration Guide*.

If you are prompted for credentials and supply an invalid user ID or password, then you are denied access to the data. In this case, you can select **File** ▸ **Clear Credentials Cache** to remove the invalid credentials from your session. The next time you access the data source, you are prompted again.

Working with User-Defined Formats

A format is a set of instructions that SAS uses to write data values. Formats are used in the data builder to control the written appearance of data values. User-defined formats are specialized formats that are not supplied by SAS. These formats are stored in a custom format catalog.

The preferred method for making user-defined formats available to a SAS Application Server is to name the custom 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*.

When a user-defined format is permanently associated with a variable, the data builder uses the format and shows it in the **Format** column. However, if you select the **Format** menu and choose a different format, you cannot use the

Format menu to go back to the original user-defined format. You can click to undo the change to the format if you have not saved your work yet, or you can remove and add the column back again.

Even if the custom format catalog is made available to the SAS Application Server correctly, you cannot associate a user-defined format to a variable with the data builder. The menu in the **Format** column does not enable you to specify a user-defined format.

8

Specifying Preferences for SAS Visual Data Builder

Specifying Global and General Preferences

To specify global SAS preferences, see “Specifying Your Preferences” on page 9. To specify general preferences, see “Specify Settings Using SAS Home” on page 10.

Specify Your Preferences for the Data Builder

To specify preferences that are specific to SAS Visual Data Builder, perform the following steps after you log on to the data builder:

1 Select **File** ▸ **Preferences** to open the Preferences dialog box.

2 Select **SAS Visual Data Builder** ▸ **General**.

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

3 Select **SAS Visual Data Builder** ▸ **Scheduling**.

Specify a default scheduling server, batch server, and deployment directory. For more information, see “Scheduling Preferences” on page 95.

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

9

Creating Data Queries

What Is a Data Query?

A data query is your primary method for selecting and formatting data that is used with data exploration and creating reports.

A data query is a metadata object that manages the references to input tables, output tables, staging tables, joins, and summarizations from a SAS Visual Data Builder session. You can save data queries and open them later to edit the data preparation operations that are performed in the data query. You can use saved data queries as subqueries when you create a new data query.

About Creating Data Queries

You can use the data builder to create data queries to prepare data for analytics. You can use a data query to subset, sort, join, and add calculated columns to tables.

As you create your data query, click to validate your data query. Use the button to preview or run the data query.

The preview option uses a temporary table for the output table. Clicking the **Results** tab shows only the first 100 rows. When you run the data query, it uses the specified output table instead of a temporary table.

Save Your Data Query

When you click , you clear the undo and redo history.

The default location for saving data queries is initially set to `/My Folder`. If you save the data query in a different folder, then the complete length of the path and name is limited to 128 characters.

Note: If the data query uses a SAS LASR Analytic Server library for the output table (this is the default), then the server does not need to be running when you save the data query. However, the server must have been started at least once in the past. When you save it, the data builder checks the metadata authorization for the library and table before it performs the save. If the server has never been started, then there is no security key to use for checking authorization. For more information, see "Security Keys" in *SAS Visual Analytics: Administration Guide*.

Save a Data Query as a New Data Query

Open an existing data query from the SAS Folders tree that closely matches the data query that you want to create. Click , and then enter a new name and select a location. Perform any customizations.

In many cases, you want to use the same input tables, joins, and so on, but you want a different output table. In this case, click **Clear** on the **Outputs** tab, and then specify new output table information. This step ensures that you use a different output table for the copied data query.

Note: If you change only the name of the output table in the copied data query, then the name of the original output table in the original data query is also changed.

For example, an analyst creates a data query that summarizes sales data and includes several geographic regions. The data query is saved with an output table that is named `Sales`. A regional sales manager wants to perform a similar summarization, but he wants to filter the data for a single geographic region. In this case, the regional sales manager performs the following steps:

1. Opens the analyst's data query, and clicks to save a copy of the data query with a new name.
2. Filters the data on the **Where** or **Having** tab for the geographic region.
3. Clicks **Clear** on the **Outputs** tab, and then specifies new output table information. This ensures that the original output table, `Sales`, for the original data query is not overwritten.
4. Saves and runs the copied data query.

See Also

"Output Tables" on page 50

Using the Design Tab

The **Design** tab is the default view for working with data queries. This tab provides an easy-to-use interface for creating a data query. You can perform the following tasks with the **Design** tab:

- Drag and drop tables or data queries from the SAS Folders tree onto the workspace.
- Join tables by using your pointer to select the source column, and then drag the pointer to the corresponding column in the joined table.
- Add columns to the **Column Editor** tab by clicking the column name from the table in the workspace or by right-clicking on the table and selecting **Add All Columns**.
- Use the **Column Editor** tab to specify column expressions, aggregations, and sort. You can set the attributes for a column, such as the type, format, and label.
- View the **Output Columns** tab to see the number of output columns and attributes. When you specify aggregations and pivot by columns, the number of output columns can increase dramatically.
- Click to check that the data query is valid.
- Check the **Messages** tab for information about warnings and errors such as invalid column expressions.
- Check the **Log** tab to view the SAS log. A SAS log is generated when you preview, run, or validate a data query.
- Use the **Outputs** tab in the right pane to specify output table information.

See Also

- Chapter 10, "Working with Tables in Data Queries," on page 49
- Chapter 11, "Working with Columns in Data Queries," on page 55
- Chapter 12, "Working with Filters in Data Queries," on page 63
- Chapter 13, "Working with Joins in Data Queries," on page 67

Adding a Data Source

Add a Table

To add a table to a data query, use the SAS Folders tree to locate the table ( ) and then drag and drop the table onto the workspace.

TIP You can also click and search for the table by name and location.

Add a Subquery

After you have created a data query and saved it, it can be used as an input data source to another data query. To add a subquery, use the SAS Folders tree to locate the data query ( ), and then drag and drop the data query onto the workspace.

The subquery is represented in the workspace by the columns that are selected for output in the subquery.

Support for Special Characters

In most cases, you can use table names and column names that contain special characters, including blank characters. When you use a column in your data query, the data builder applies the n-literal syntax, such as `'table-name'n.'column-name'n`, so that SAS can use the column.

Rules for SAS names apply. For more information, see “Names in the SAS Language” in *SAS Language Reference: Concepts*.

The data builder does not apply the n-literal syntax to code that you enter manually on the **Where** tab or **Having** tab or in column expressions. For example, if your table has a column that is named **`quantity ordered`**, then you must add the n-literal syntax (similar to the following example):

```
AVG(table.'quantity ordered'n)
```

Best Practices for Adding Data Sources

When you plan to join data sources (tables or subqueries), the order in which you add the data sources to the workspace matters. The first data source that is added to the workspace is automatically assigned as the left table for any joins that you add to the data query.

If you are creating a data query that uses a fact table and dimension tables, then the simplest approach is to drag and drop the fact table onto the workspace first. You can perform left, right, or full joins with the dimension tables faster because you must specify only the join type. However, if you drag and drop a dimension table first, then you can easily use the button on the **Joins** tab to switch the left and right tables in the join.

If you are not using a fact table or dimension tables, then the sequence for adding tables to the workspace might not be very important. Just remember that the first table dropped onto the workspace is assigned as the left table, and you can switch the left table and right table on the **Joins** tab. The data builder takes advantage of vendor-specific features in SQL processing whenever possible. If the source tables are from a third-party vendor database, then the SAS/ACCESS Interface engine can optimize the performance of the data query by passing the SQL statements through to the database.

Specifying Properties for a Data Query

On the **Properties** tab in the right pane, you can view or specify the following properties:

Name
displays the name for the data query. The initial value is DataQuery1. You can specify a different name when you save the data query.

Location
displays the metadata folder location for the data query object. The initial value is `/My Folder`. You can specify a different location when you save the data query.

Description
specifies a description of the data query.

Create an SQL query view
specifies to create a view for either the work table or the output table. For more information, see "Create SQL Query Views" on page 53.

Unique values
specifies whether the SQL keyword DISTINCT is applied to the SELECT statement that is used to generate the result set for the data query.

Append data
specifies whether the result set for the data query is appended to the output table. If a staging table is used, then the staging table is replaced with the result set before appending to the output table.

On the **Outputs** tab, you can view or specify the following properties:

Table
specifies the table name for either the staging output or the final output. For more information, see "Specify an Output Table" on page 51.

Compress data
specifies to compress the output table. For more information, see "Output and Staging Table Interactions" on page 52.

Location
specifies the metadata folder location to use for registering the staging table metadata or output table metadata. Click to select a different location.

Library
specifies the library to use for the staging table or output table. Click to select a different library.

Partition by
specifies the column to use for partitioning the output table. This property only applies when the output table is in a SAS LASR Analytic Server library or SASHDAT library. For more information, see "Distributed Server: Partition Tables" on page 83.

See Also

"Specifying Properties for a Source Table" on page 50

10

Working with Tables in Data Queries

Source Tables

When you drop a table onto the workspace, the data builder connects to the SAS Metadata Server to determine the column names and data types for the table. When you drop a subquery onto the workspace, the data builder makes the same request, but determines the column names and data types for the output table of the subquery.

You (or an administrator) must register a table in metadata before using it as a source table for data preparation. When you import data from a file, a database, Facebook, Google Analytics, or Twitter, the data builder automatically registers the data in metadata as a table. Source tables can also be registered using SAS Visual Analytics Administrator or SAS Management Console.

Data queries reference input tables by their unique metadata IDs. For this reason, do not delete the metadata registrations for input tables that are used in your data queries. If you delete the input tables from metadata, their unique metadata IDs are also deleted from metadata. If you register the same input tables again, they are assigned new metadata IDs. As a result, SAS Visual Data Builder cannot locate the tables because of their new metadata IDs. If the input tables cannot be located, the query is corrupted, and you must re-create it.

See Also

- “Adding a Data Source” on page 45

- *SAS Visual Analytics: Administration Guide*

- *SAS Intelligence Platform: Data Administration Guide*

Specifying Properties for a Source Table

When you create a data query and select a table in the workspace or when you select it from the menu on the **Inputs** tab in the right pane, you can specify the following:

Alias
displays the SAS table name that is stored in metadata. You can specify a new value to use as a table alias. The alias name is stored with the data query, and it does not affect the metadata information for the table.

Name
displays the metadata object name. You can change the metadata object name in the SAS Folders tree by selecting it, right-clicking, and selecting **Rename**.

Location
displays the table's metadata folder.

Library
displays the table's library.

Auto-aggregate
specifies whether to apply aggregations to the columns for this table when the columns are added to the query.

Aggregations
specifies the aggregations to apply to the columns for this table when the columns are added to the data query.

See Also

"Use the Auto-Aggregate Functions" on page 59

Output Tables

About Output Tables

When you create a data query, you specify an output table on the **Outputs** tab in the right pane. When you save the data query, the output table is registered (or updated) in the metadata. Registering the table in metadata enables you to use it as a source table for another data query or another SAS application. When you run the data query, the physical output table is created, and the table is updated in metadata.

Subqueries do not require an output table. If you do not need to view the results of a data query that you intend to use as a subquery, then you can click **Clear** on the **Outputs** tab and still save the subquery. You can then drag and drop the subquery into a data query.

When you create a data query, the default output table name is OutputTable.

Specify an Output Table

Every data query must have an output table in order to save the results in a table. How you specify the output table affects whether metadata is updated or created.

You can specify an output table name, location, and library on the **Outputs** tab in the right pane.

The following table shows alternative ways to specify an output table.

Action	How To
Change the name of the output table. *	Enter the new name in the **Table** field, and save the data query. All data queries, reports, explorations, and so on, still reference the same table object.
Create a new output table. *	Click **Clear**, and then specify a table name, location, and library. This action disassociates the previously used output table.
Reuse an existing table.	Click , and then browse for the table to use. This action disassociates the previously used output table.

* If you replace the default table name, OutputTable, with another name before you save the data query, the data builder registers a new output table and uses it with the data query.

You can enter a name that is up to 32 characters as the output table name. If a third-party vendor database product is used for the output table, then the number of characters might be less.

If you select a SAS LASR Analytic Server or SASHDAT library, the **Partition by** menu becomes available.

Staging Tables

About Staging Tables

Staging data is a best practice because you can use the data builder to access and transfer data from operational systems once, rather than frequently interfere with the operational systems and reduce their performance. Using the data builder to stage data can provide the advantage of adding calculated columns when you stage the data.

Like the output table, the staging table is registered in metadata when you save the data query. The physical table for the staging table is created when you run the data query.

You cannot specify the name for a staging table. The name of the output table is used, and an `_STG` suffix is applied to the name. The suffix is used for the table name in metadata. The physical name of the staging table does not include the suffix.

Specify a Staging Table

To use a staging table, perform the following steps on the **Outputs** tab:

1. Select the **Use a staging table** check box.
2. Specify a library.

   Note: The data builder is initially configured to use the Visual Analytics Public LASR library and the Visual Analytics Public HDFS pair of libraries. If you specify different libraries, then make sure that you understand how the path is related to the server tag for the SAS LASR Analytic Server library. For more information, see *SAS Visual Analytics: Administration Guide*.

Output and Staging Table Interactions

The physical table is always replaced with the results of the data query. When you use a data query to append data and stage the data too, the staging table holds the data to append. As a result, the output table that the data query appends to is typically much larger than the staging table.

In the right pane, if you select the **Compress data** check box on the **Outputs** tab, then the tables are compressed as follows:

- If the output table is in the SAS LASR Analytic Server library or SASHDAT library, then the output table is compressed.
- If the output table is in the SAS LASR Analytic Server library and the staging table is in the SASHDAT library, then the staging table is compressed. The staging table will automatically be compressed when it is loaded to memory on the SAS LASR Analytic Server.

The following table identifies the supported combinations for output tables and staging tables. In addition, whether you can append data to tables is indicated.

Table 10.1 *Output Table and Staging Table Interactions*

Output Table	Staging Table	Append Data
SAS or DBMS *	None	Supported
SASHDAT	None	Not supported
Co-located HDFS or NFS-mounted MapR	None	Supported
SAS LASR Analytic Server	None	Supported
SAS or DBMS *	SAS or DBMS *	Supported
SAS LASR Analytic Server	SAS or DBMS *	Supported **

Output Table	Staging Table	Append Data
SAS LASR Analytic Server	SASHDAT **Note:** When the staging table is in SASHDAT, SAS LASR Analytic Server is the only choice for the output table.	Not supported
SAS LASR Analytic Server	Co-located HDFS or NFS-mounted MapR	Supported **

* The SAS or DBMS value represents data stored in SAS data sets or a third-party vendor database, respectively.

** Appending data is performed by the SAS LASR Analytic Server engine. Appends are not performed by having the server read data in parallel.

The information about appending data in the previous table applies to data queries. See “Append In-Memory Tables” on page 79 if you are working with in-memory tables on SAS LASR Analytic Server exclusively.

See Also

Chapter 15, “Working with SAS LASR Analytic Server,” on page 77

Create SQL Query Views

You can influence whether the work table or output table is a view or a table. To specify view, select the **Create an SQL query view** check box on the **Properties** tab. The option is enabled by default. In most cases, this option improves performance by reducing data movement and storage requirements.

When working with tables in a database, if the source tables and output tables are in the same library, the **Create an SQL query view** option can enable you to pass a CREATE VIEW statement directly to the database. When this happens, the data query runs almost instantaneously because there is no data movement to create a view.

The following table summarizes the interactions between library types and views:

Library Type	Interactions
Base SAS	Base SAS libraries always support views for output tables or staging tables. Source tables can be from different libraries because the connection information is stored in the view.

Library Type	Interactions
DBMS	If the source tables and the output tables or staging tables are in the same library, then the view is created in the database. If they are in different libraries, then the check box is not enabled and the output table or staging table is created as a physical table.
SAS LASR Analytic Server or SASHDAT	These libraries types do not support views. Selecting the check box applies to the work table, rather than the output table or staging table. If the source tables and the output tables are in the same SAS LASR Analytic Server library, then the check box is not enabled, and the work table must be a physical table.

Note: In some cases, the check box is enabled, but using this option can result in an error when running the data query. For example, if you use a DBMS library for the output table, you must have permission to create a view. You can clear the check box to create the output table as a physical table in these cases.

11

Working with Columns in Data Queries

Adding Columns to a Data Query

When you add a data source (a table or subquery) to the workspace, the columns from the data source are not automatically added as output columns. You must add the columns to the data query that you want to use.

After the columns are added, you can specify column expressions and aggregations and use the sort and pivot by features.

Note: As an exception, the auto-aggregate feature requires that you set the default aggregations for the table before they are added to the data query.

When a data source is dropped onto the workspace, the column types are represented by the following icons:

Table 11.1 *Icons for Data Types*

Icon	Description
	This icon represents numeric data.
	This icon represents character data. **Note:** Date, time, and datetime data use this icon. After the column is added, the **Type** and **Format** columns are updated with information about the new column.

You can add columns to the data query in the following ways:

- Select the table in the workspace, right-click, and select **Add All Columns**.
- To add one column from a table, select the column name with your pointer.

  TIP If you select the column name an additional time, then the column is added to the data query again. This can be helpful if you want to use a column for both numeric and character data. When you add a column more than once, a number is added to the column name. If you change the column name, then you must make sure that you do not have more than one column with the same name.

- To create a new column, click the **Column Editor** tab, and then click + next to the last column that is listed. Enter a column name, expression, and type. The remaining fields are optional.

Remove Columns

To remove a column from a data query:

1. Click the **Column Editor** tab.
2. Select the column to remove, right-click, and select **Remove Column**.

Specify a Column Expression

To specify a column expression:

1. On the **Column Editor** tab, specify a name for the column.
2. Select `table-name.column-name` from the **Expression** column. If you added a new column manually, then you must make sure that you specified a table name and a column name.
3. Click to open the expression builder. The table name and column name are added automatically as the default SQL expression.

   The expression is limited to 1024 characters.
4. On the **Fields** tab, select columns from the source tables that you have added to the data query.
5. On the **Functions** tab, select the functions to apply to the source column.
6. Enter arithmetic operators and expressions such as CASE statements directly in the **SQL expression** area.
7. Click **Apply** to save the column expression.

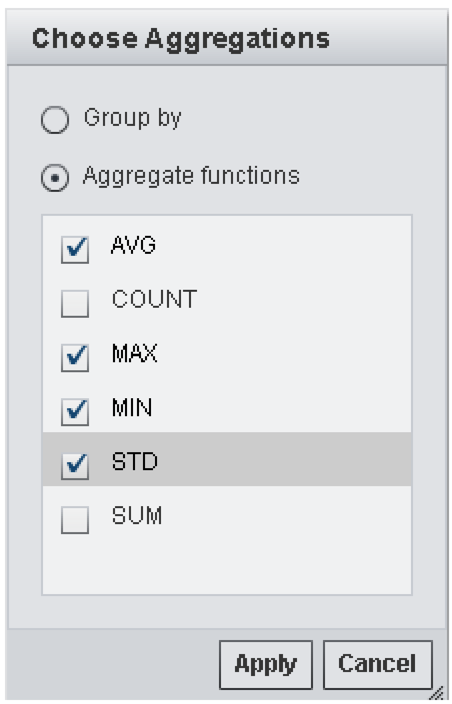
Note: If you click after you click **Apply**, then your last change to the expression is undone. However, because the expression window is closed, you might not notice the undo.

Specify Aggregations

To specify an aggregation for a column:

1 On the **Column Editor** tab, place your pointer in the **Aggregations** cell for the column and click. Click the ellipsis button to select the aggregations to use.

2 In the Choose Aggregations dialog box, select the **Aggregate functions** radio button.



3 Select the check boxes for the aggregate functions to use. Click **Apply**.

Note: After you click **Apply**, all of the other columns are automatically specified as `GROUP BY` columns. You can change a column from `GROUP BY` to an aggregate function by repeating this step for the column.

The following display shows how adding aggregations results in additional output columns. The aggregate function is appended to the column name.

Column Editor | Where | Having | Joins | **Output Columns** | Messages | Log

#	Column Name	Type	Format
7	State	CHARACTER(25)	$25.
8	Quantity_STD	NUMERIC	
9	Quantity_MIN	NUMERIC	
10	Quantity_MAX	NUMERIC	
11	Quantity_AVG	NUMERIC	
12	Total_Retail_Price_STD	NUMERIC	
13	Total_Retail_Price_MIN	NUMERIC	
14	Total_Retail_Price_MAX	NUMERIC	
15	Total_Retail_Price_AVG	NUMERIC	
16	CostPrice_Per_Unit_STD	NUMERIC	
17	CostPrice_Per_Unit_MIN	NUMERIC	
18	CostPrice_Per_Unit_MAX	NUMERIC	

Remove All Aggregations

To remove all the aggregate functions and group by settings:

1 On the **Column Editor** tab, select all the columns.

2 Right-click, and select **Remove Aggregations**.

> **TIP** This menu option is available only when all of the columns are selected and at least one aggregation is defined.

Use Group By Variables

When you add an aggregation to a column, the remaining columns are automatically used as group by variables. The **Aggregations** column displays `GROUP BY` for these variables.

You can use a column as a group by variable by following the steps in “Specify Aggregations” and selecting the **Group by** radio button.

Use the Auto-Aggregate Functions

The auto-aggregate feature is a productivity feature that enables you to specify a set of aggregations to apply as default aggregations to numeric columns for a specific table. A typical use is to automatically aggregate some of the columns in a fact table.

To use the auto-aggregate feature:

1. Select a table on the **Design** tab.
2. On the **Inputs** tab, select **Enable** for **Auto-aggregate**.
3. Click the ellipsis button next to **Aggregations** to open the Choose Aggregations window.
4. Select the check boxes for the aggregate functions that you want to apply, and then click **Apply**.

Whenever you add a column to the data query, the selected aggregate functions are automatically applied.

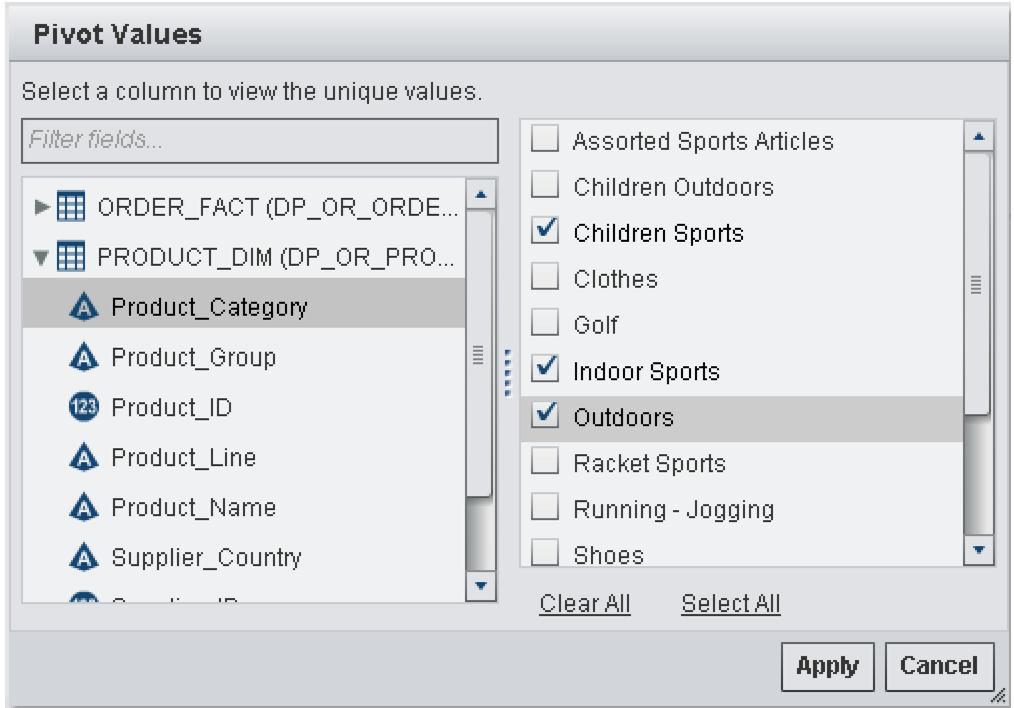
Use the Pivot By Feature

The pivot by feature provides an easy and powerful way to summarize data for analytics. You can specify a column to use as a categorical variable and the unique values to use. When the data query is run, the output table is summarized with the aggregations that you apply.

To use the pivot by feature:

1. On the **Column Editor** tab, place your pointer in the **Pivot By** cell for the column to use as the pivot column. Click the ellipsis button to select the pivot column and values.
2. In the Pivot Values dialog box, select the pivot by column. You can enter search criteria in **Filter fields** to filter the column names.

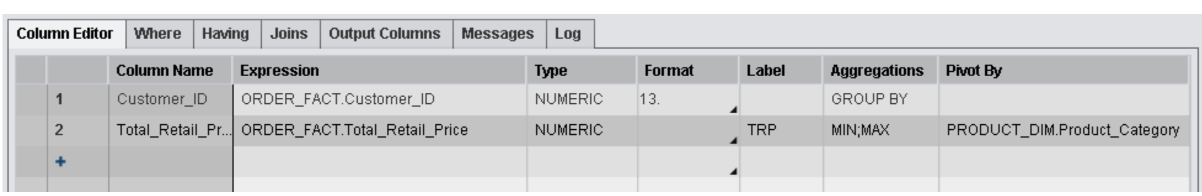
The following display shows an example of pivoting by three values in the Product_Category column.



3 After the unique values for the column are loaded, select the check boxes for the values to use in the summarization. Click **Apply**.

The following display shows an example of the **Column Editor** tab when a pivot by column is used. The minimum and maximum Total_Retail_Price are calculated for each Customer_ID and are then pivoted by (transposed by) three values of the Product_Category column.

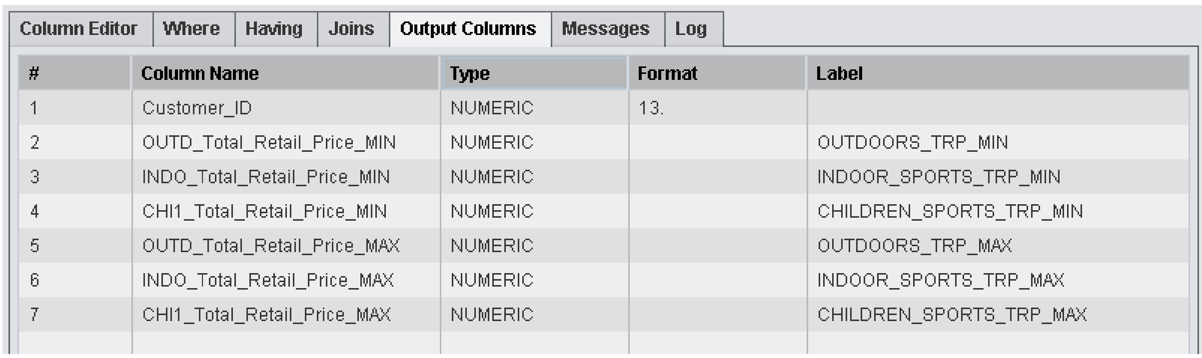
Figure 11.1 Column Editor Tab with a Pivot By Column



TIP TRP is specified as the label for the Total_Retail_Price column. Look at the next display to see how the label is used to create labels for the new columns.

The following display shows how pivoting the Customer_ID column by three values of the Product_Category column results in additional output columns. A substring of the pivot by values is used as a prefix to each column name and the aggregate function is used as a suffix. The pivot by column label and aggregate function are used in the output column label.

Figure 11.2 *Output Columns Tab with Pivot By Values*



#	Column Name	Type	Format	Label
1	Customer_ID	NUMERIC	13.	
2	OUTD_Total_Retail_Price_MIN	NUMERIC		OUTDOORS_TRP_MIN
3	INDO_Total_Retail_Price_MIN	NUMERIC		INDOOR_SPORTS_TRP_MIN
4	CHI1_Total_Retail_Price_MIN	NUMERIC		CHILDREN_SPORTS_TRP_MIN
5	OUTD_Total_Retail_Price_MAX	NUMERIC		OUTDOORS_TRP_MAX
6	INDO_Total_Retail_Price_MAX	NUMERIC		INDOOR_SPORTS_TRP_MAX
7	CHI1_Total_Retail_Price_MAX	NUMERIC		CHILDREN_SPORTS_TRP_MAX

12

Working with Filters in Data Queries

About Filtering Data

You can use the **Where** and **Having** tabs to filter data in SAS Visual Data Builder.

The SQL expression on the **Where** tab is applied to the input data. This SQL expression is often used to subset data on the columns in the source tables.

The SQL expression on the **Having** tab can be applied to either the input data or the calculated columns. This SQL expression is typically used to subset data on calculated columns in the output table.

The maximum length for a WHERE or HAVING clause is 4096 characters.

Specify a WHERE Clause

To add a WHERE clause to a data query:

1 Click the **Where** tab.

2 On the **Fields** tab, expand the table node, and select the column to use for filtering data.

> **TIP** You can enter a value in the **Filter fields** field to locate the column.

3 Double-click the column or drag and drop it in the **SQL expression** area.

4 For character variables, you can click ***column_name*** **Values** to view the values for the column. In the Filter Values window, select the values that you want to include in the filter.

   Click **Apply** to add the filter values to the **SQL expression** area.

5 (Optional) You can click the **Functions** tab and select the functions to use with filtering.

6 (Optional) You can add a subquery to the filter using a table from a different library. On the **Libraries** tab, click + to add a library.

After you select a library, expand the library node, and select the table that you want to use for the subquery.

Note: In the SQL expression, you must enclose the subquery in parentheses.

7 Edit the WHERE clause in the **SQL expression** area as follows:

Expression	Edit	Example
A single character value	Add an equal sign between the column name and the unique value.	`CARS.Make = 'Acura'`
More than one character value	Specify an IN operator and enclose the unique values in parentheses.	`CARS.Make IN ( 'Acura', 'Audi')`
Numeric comparison	Specify a numeric operator and a constant, or specify a numeric operator and another column name.	`CARS.Cylinders >= 6` `PRDSALE.Actual > PRDSALE.Estimate`

8 Click the save icon to save the filter.

Specify a HAVING Clause

To add a HAVING clause to a data query:

1 Click the **Having** tab.

2 On the **Fields** tab, expand the table node, and select the column to use for filtering data. The **Output Columns** node includes the calculated columns.

TIP You can enter a value in the **Filter fields** field to locate the column.

3 Double-click the column or drag and drop it in the **SQL expression** area.

4 For character variables, you can click ***column_name*** **Values** to view the values for the column. In the Filter Values window, select the values that you want to include in the filter.

Click **Apply** to add the unique values to the **SQL expression** area.

5 (Optional) You can click the **Functions** tab and select the functions to use with filtering.

6 (Optional) You can add a subquery to the filter using a table from a different library. On the **Libraries** tab, click + to add a library.

After you select a library, expand the library node, and select the table that you want to use for the subquery.

Note: In the SQL expression, you must enclose the subquery in parentheses.

7 Edit the HAVING clause in the **SQL expression**. The syntax is identical to that shown in step 6 of “Specify a WHERE Clause”.

8 Click to save the filter.

Best Practices for Filters

SAS Visual Data Builder provides the **Where** and **Having** tabs for filtering data. Follow these best practices:

- Filter on the **Where** tab first because the WHERE clause reduces the rows to consider for further subsetting.
- When you specify a filter on a column that is an index or primary key, avoid using a function whenever it is possible. Using a function, such as `CAST(order_id as DOUBLE)`, risks performing a full-table scan rather than using the index.
- In some cases, you can improve performance by filtering on the table that has the smallest number of rows first.
- If you must remove a table that is used on the **Where** or **Having** tabs, then remove the reference to the table from the filters before you remove the table. Otherwise, if you remove a table that is referenced in a filter, then all of the filtering conditions are cleared.
- When adding a subquery using the **Libraries** tab, an additional step is needed if the column names or table names are written in a language that uses a double-byte character set. After you finish adding the columns, you must manually enter an n-literal to the string on the **Code** tab for each column name or table name that uses a double-byte character set. For example:

  ```
  'table-name'n.'column-name'n
  ```

  For more information about editing code manually, see Chapter 18, “Customizing Code,” on page 89.

13

Working with Joins in Data Queries

About Joins

SAS Visual Data Builder supports joins for tables and subqueries. You can join tables to each other, including self joins. You can join subqueries to tables and join subqueries to subqueries. When you use a subquery in a join, the join condition is made against the output table for the subquery. The data builder supports joining up to 256 tables.

When you drop a table or subquery onto the workspace, the data builder attempts to determine a join condition automatically. When the data builder creates a join automatically using foreign keys or by matching columns, the join is added as an inner join. The data builder also supports left, right, and full joins. You must specify the join type manually if you do not want an inner join.

In addition to easily adding joins to a data query, the data builder generates an SQL statement with all of the joins declared explicitly. For example, you can specify an inner join in a WHERE clause, such as `WHERE t1.order_id=t2.order_id`. However, mixing inner joins in a WHERE clause and outer join types in a single data query can be complex to read and understand.

SAS Visual Data Builder always generates an SQL statement with inner joins declared explicitly. For example, see the following code sample:

```
LEFT JOIN
        LIB1.TRANSACTION_TYPE_DIM TRANSACTION_TYPE_DIM
            ON CASH_FLOW_FACT.TRANSACTION_TYPE_KEY =
                TRANSACTION_TYPE_DIM.TRANSACTION_TYPE_KEY
    INNER JOIN
        LIB1.TRANSACTION_DIM TRANSACTION_DIM
            ON PARTY_DIM.SEGMENT_ID = TRANSACTION_DIM.SEGMENT_ID
            AND PARTY_DIM.TRANSACTION_KEY = TRANSACTION_DIM.TRANSACTION_KEY
```

How Does the Automatic Join Feature Work?

Feature Overview

SAS Visual Data Builder attempts to join tables and data queries automatically as you add them to the workspace. When you drag and drop a table or data query onto the workspace, information about the table or data query is retrieved from the SAS Metadata Server. For subqueries, the metadata for the subqueries' output table is retrieved.

Using Foreign and Primary Keys

If primary key or foreign key information is registered in the metadata for the table that you drag and drop onto the workspace, then the data builder retrieves the foreign key and primary key information.

The data builder then iterates over each of the tables that are already in the workspace in the same sequence in which they were added to the workspace. The data builder retrieves the foreign key and primary key information for the table, and compares the length, type, and name with the key columns for the newly added table. If a match is found, then the tables are used in the join, and the columns are added as a join condition. The data builder continues to search for matches between the two tables, and it adds join conditions when possible. After a set of join tables is identified, the data builder does not continue iterating over the tables that are already in the workspace.

Matching by Name

If there is no foreign key or primary key information for the table that is dragged and dropped onto the workspace, then the data builder does not use foreign key or primary key information for the tables already in the workspace.

The data builder retrieves the column information for all of the columns in the newly added table. The data builder then iterates over each of the tables that are already in the workspace in the same sequence in which they were added to the workspace. The data builder compares the length, type, and name for each column with each column in the newly added table. If a match is found, then the tables are used in the join, and the columns are added as a join condition. The data builder does not continue to search for matches between the two tables, and it does not continue iterating over the tables that are already in the workspace.

Selecting the Join Type

When the data builder finds the first set of matching columns by comparing keys or matching names, it sets the join type for the two tables. The data builder checks the metadata for the columns to determine whether the columns are nullable. (Many third-party vendor databases support **NOT NULL** as a constraint for a column when the table is created in the database with SQL. The constraint ensures that there are no missing values for the column.) The data builder performs the following steps to set the join type:

1. If the column for the existing table is not nullable but the column for the newly added table is nullable, then the tables use a left join.
2. If the column for the existing table is nullable but the column for the newly added table is not nullable, then the tables use a right join.
3. If both columns for the existing table and newly added table are nullable, then the tables use a full join.
4. If none of the previous conditions are met, then the tables use an inner join.

You can specify the join type by right-clicking on the join icon () and selecting the join type from the menu.

Adding a Join

Drag and Drop Join Lines

To add a join by selecting a column name and dragging your pointer to another column:

1. Place your pointer on one of the columns in a table to use, and then click and drag your pointer to a column in the other table to use in the join.
2. If you want a join type other than an inner join, then select the join, right-click, and select the join type.

TIP The first table that is added to the workspace is always set as the left table. You can switch the right table and left table by using the Joins tab.

Use the Joins Tab

To add a join manually:

1. Click the **Joins** tab.
2. Click +.
3. Use the menus to replace `Not Selected` with the table names to use in the join. Click **Save**.

4 The default join type is an inner join. Use the menu in the **Join Type** column to select a different join type.

5 Click + in the **Join conditions** area.

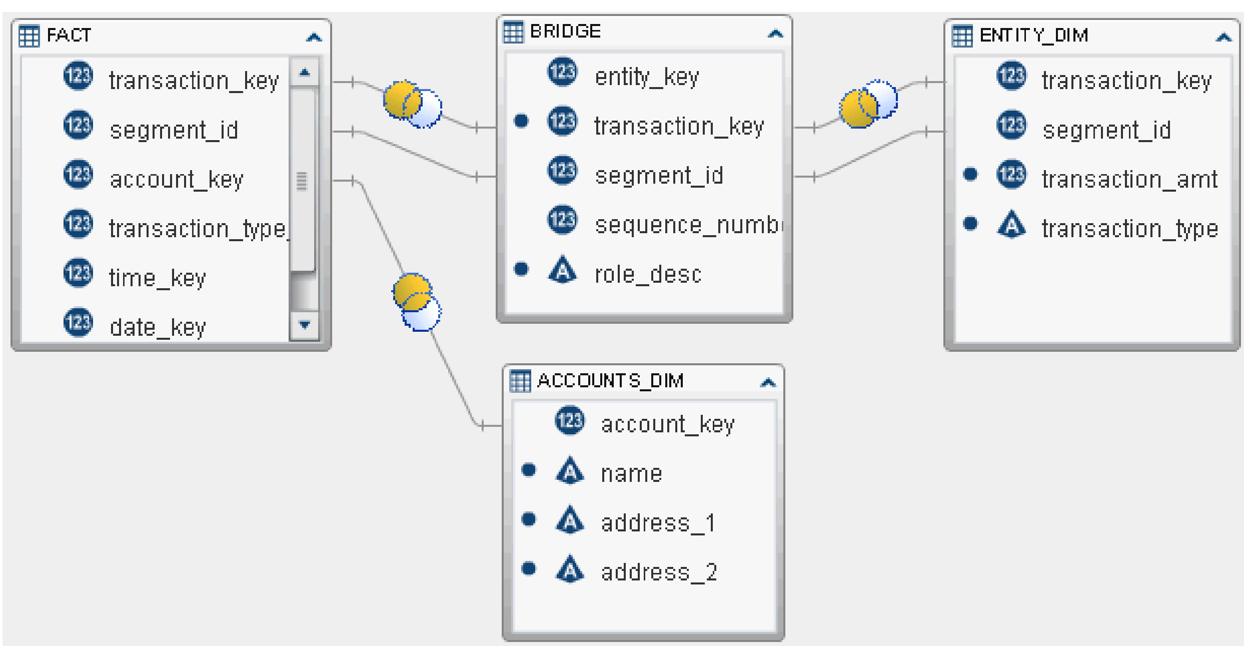
6 Use the menus to replace `Not Selected` with the column names to use in the join condition. Click **Save**.

7 Click [save icon] to save the join condition with the data query.

Example: Joins with a Junction Table

By default, the data builder considers the first table that you drop in the workspace as the fact table. As you add tables to the data query, the data builder attempts to join the tables with the first table, the fact table. This strategy works well for straightforward fact table and dimension table arrangements. However, in many cases, you need to join additional tables to a junction table. Junction tables are also known as bridge tables or link tables.

In the following figure, the Bridge and Accounts_Dim tables are joined to the Fact table. The Accounts_Dim table has a join on the Account_Key column. The Bridge table is joined on the compound key (represented as two lines), which consists of the Transaction_Key and Segment_ID columns.



To retrieve the transaction amount and transaction type from the Entity_Dim table, it must have a join that uses the Bridge table instead of the first table (fact table) that was added to the workspace.

To create a join to a junction (or bridge) table:

1 Add the fact table to the data query first.

2 Add the dimension tables, such as the Accounts_Dim table and junction tables. The data builder attempts to join the tables automatically.

Review the join as follows:

a Check that the correct columns are used.

b Change the join type, such as from inner to left, as needed.

3 Add the tables that use the junction table, such as the Entity_Dim table. Initially, the data builder attempts to join the tables to the fact table.

Correct the join as follows:

a Select the automatic join to the fact table, right-click, and select **Remove Join Condition**.

b Use your pointer to connect the keys from the junction table to the dimension table.

c Change the join type as needed.

Remove a Join

You can remove a join by using either of the following methods:

- Select the join in the workspace, right-click, and select **Remove Join Condition**.
- Click the **Joins** tab, select the row in the table, and click .

Managing Joins in a Data Query

Once a join is added to a data query, you can change the join by selecting it in the workspace, right-clicking, and changing the join type or removing the join condition.

You can also change a join by clicking the **Joins** tab, and then selecting the row in the table. You can make the following changes:

- add and remove tables from the join list
- reorder the sequence of joins by moving them up or down
- switch the left table and right table assignments for a join
- add, remove, and change the columns that are used in the join condition

The **Joins** tab shows the join condition for the entire data query. Make sure that you select a row in the upper table to set the columns in the **Join conditions** area.

The workspace shows a link between the tables that are used in a join. If you specify a left join or right join, then the icon reflects which table provides the bulk of the data. If you rearrange the tables in the workspace (switching the left table and right table), then the icon continues to reflect which table provides the bulk of the data. In order to change the data relationship, use either of the following methods:

- Switch the left table and right table by selecting a row from the table, and clicking .
- Change the join type from left to right or from right to left.

Best Practices for Managing Joins

In most cases, the sequence in which joins are specified on the **Joins** tab does not matter. However, it is possible that the query optimizer for the data source might perform the joins in a sequence that reduces performance. In these rare cases, you can change the sequence in which joins are specified on the **Joins** tab by selecting the join in the list, and clicking or . This join sequence is still determined by the query optimizer, but you can control how the SQL for the join is presented to the query optimizer.

14

Creating LASR Star Schemas

What is a LASR Star Schema?

A LASR star schema is very similar to a typical star schema in a relational database. Imagine a single fact table that is surrounded by dimension tables. Each dimension table is joined to the fact table using a dimension key.

Here are some key facts about LASR star schemas:

- Typically, the dimension key and the corresponding column in the fact table are the same data type and length. If your tables do not use the same data type and length, then there are two ways to correct the data:
  - Change the data definition in the system with the original data.
  - Create a data query that modifies the columns and outputs the table to an in-memory table.
- Single-level star schemas are supported. Snowflake schemas are not.
- A LASR Star Schema is limited to a single join condition. It does not support multiple key columns between fact and dimension tables.

Create a LASR Star Schema

To create a LASR star schema:

1. Select **LASR** ▸ **Create a Star Schema**.
2. Drag and drop tables or data queries from the SAS Folders tree onto the workspace. Keep the following suggestions in mind:

- Drag and drop the fact table first and then the dimension tables.
- Any input tables that are not already in a SAS LASR Analytic Server library are loaded to memory when the star schema is run.
- If a dimension table shows the icon, then it usually indicates that the data builder could not determine the join condition for the dimension table. Position your pointer over the icon for information about how to correct the incomplete table status.
- If a dimension table shows the icon, then it usually indicates that the table is from a different SAS LASR Analytic Server library than the output table. You must use one SAS LASR Analytic Server library only.

3 (Optional) If you do not want a column from the dimension table in the output table, then select the column in the table in the workspace, right-click, and select **Remove Column**.

   Note: All of the columns from the fact table are automatically selected and required to be included in the output table.

4 (Optional) Click the **Joins** tab to review the columns that are used to generate the output table.

5 (Optional) Specify a different column prefix. For more information, see "Column Prefixes".

6 Click the **Output** tab, and specify a name for the output table.

7 Click , and specify a name and location for the LASR star schema.

8 Click to generate the output table.

How Are Tables Used?

Input Tables

The fact table and dimension tables are used in the star schema after you drag and drop them onto the workspace. By default, the data builder considers the first table that you drop in the workspace as the fact table. Click from the toolbar to select a different fact table.

Because the star schema is formed in memory by the server, the first step is to make sure that the tables are loaded to memory. Be aware that the tables are transferred each time the star schema runs and this can reduce performance. Conversely, if the tables are already loaded to memory on the server, the star schema runs and generates the output faster.

When creating a star schema using tables from the same SASHDAT library, the star schema will fail to run if one or more of the tables in the star schema is encrypted, but the SASHDAT library is not encrypted. In this case, you will receive an error message indicating that a password is needed for the encrypted tables. This can occur when your administrator turns encryption off for a previously encrypted library. Any table that was in the library beforehand will remain encrypted. Your administrator can resolve the issue by turning encryption back on for the library. Contact your administrator for more information.

Column Prefixes

The first 15 characters of a dimension table's name and the underscore character are initially set as a prefix for the column names from the dimension table. Column names for the output table are a combination of the prefix and the original column name.

However, you can specify a different value for the prefix after you select the table name from the menu at the top of the **Inputs** tab.

Column names from the fact table are not modified with a prefix.

TIP A column name is limited to 32 characters. If you have a long column name, then reducing the prefix can help you keep more of the original column name.

Output Table and Conserving Memory

To use memory efficiently, the default output for the schema is a view. The **Create output as a view** check box on the **Output** tab is selected by default to create a view. If you clear the check box, then the output is a table.

When the output is a view, the rows are created from the original tables when the view is accessed. This is the main advantage of a view because it does not create the entire output table and hold it all in memory.

Note: If you choose to output as a view, the memory usage size will display the value as if the star schema was output as a table.

If you clear the **Create output as a view** check box to create a table from the star schema, then the system must have enough available memory to store the table. If the system runs out of memory while running the star schema, then the memory that was used for the output table is freed and you receive an error message.

If you use SAS LASR Analytic Server tables for input to the schema, then the output table for the star schema must use the same library.

When using tables from an SASHDAT library to create a star schema, the star schema will fail to save if the SASHDAT library path and the tag of the LASR output library do not match. In this case, you will receive an error message stating that the library path does not match the SAS LASR Analytic Server tag. Try changing the output table to a LASR library whose tag matches the SASHDAT library path. You can view the tag for the LASR library in SAS Management Console. For more information, contact your administrator.

Performance Considerations

Creating a LASR star schema as a view can be more convenient than creating a table, but accessing data through the view can impact performance negatively. The scale of the impact depends on the size of the dimension tables. However, the number of passes through the data is even more important. For example, requesting percentiles or box plots or fitting statistical models requires passing through the data multiple times. These requests are impacted more than a request for summary statistics.

Keep the following considerations in mind:

- The initial creation of a view is faster than forming a table. However, accessing the data in a view is slower than accessing the data from a table.
- Creating a table requires more physical memory than creating a view. If your system has sufficient memory capacity, then creating a table provides the best performance for accessing the data.

15

Working with SAS LASR Analytic Server

Using SAS LASR Analytic Server Libraries

Default Library

During installation, the SAS Deployment Wizard registers a predefined library for SAS LASR Analytic Server. This library is available for use in the SAS Folders tree, and it is located in `/Shared Data/SAS Visual Analytics/Public/Visual Analytics Public LASR`.

This library is intentionally configured as a general-purpose library with extremely limited restrictions for securing data access.

Input Libraries

When you select a SAS LASR Analytic Server table as an input table for a data query, be aware of the following best practices if the table is large. If the table is not large, then using it for input requires no special considerations.

Here are the considerations for using a large SAS LASR Analytic Server table as an input table:

- A WHERE clause is processed in memory by the server if no aggregations or joins are used. Specify a filter on the **Where** tab so that you use only the rows that you want.
- Instead of using a data query (which transfers the data from the server to the SAS Workspace Server and then performs the join), you can create a LASR star schema. The LASR star schema performs the join in memory with other in-memory tables.

Output Libraries

A SAS LASR Analytic Server library is initially configured as the default output library. When you run a data query, click **Explore Results** to explore the table immediately.

You can select the **Append data** check box on the **Properties** tab to add rows to an in-memory table. However, this option is not available if you use a SASHDAT library as a staging library. Instead, you can use the data query to append rows, and then save the in-memory table as a SASHDAT table.

Load a Table (As Is) to SAS LASR Analytic Server

The data builder is typically used to perform basic data preparation. However, there are cases when a table is already prepared for analysis. In these cases, you can load a table to memory on SAS LASR Analytic Server without any modification.

To load a table into memory on a SAS LASR Analytic Server:

1 Use the SAS Folders tree to locate the table.

> **TIP** You can also click , and search for the table by name and location.

2 Select the table, right-click, and select **Load a Table**.

3 The fields in the **Source Table** section are filled automatically.

Specify the following settings in the dialog box:

Table 15.1 Load a Table Dialog Box Properties

Field	Description
Name	Enter the name to use for the table. This field is initially set to the same name as the source table. This name is registered in metadata with the SAS LASR Analytic Server library.
Description	Enter a description of the data. This description is displayed by SAS clients such as SAS Visual Analytics Explorer.
Location	This field is initially set to a default folder. Click **Browse**, and select a folder to use for the table metadata.
Library	This field is initially set to a default library. Click **Browse**, and select the SAS LASR Analytic Server library to use with the table.

4 Click **OK**.

Appending In-Memory Tables

How Does Appending In-Memory Tables Work?

The data builder provides a very convenient way to add incremental data into a single table. For example, if sales data is loaded into memory on a daily schedule, you can append the numerous daily tables into a separate table that contains the cumulative sales data.

For the purpose of comparison, you can create a data query to append to an in-memory table with a data query. One reason to append entire tables is that you can also access the tables individually.

Append In-Memory Tables

To append in-memory tables:

1 Select **LASR** ▸ **Append Tables**.

2 Select the table that you want to append data to from the **Available tables** list, and click to move it to the **Base table** field.

If the table that you want to use is not listed, click to view a different source library. If a table is added to the library after you click , then the list does not automatically refresh to show the new table. Select the same library again to refresh its contents.

3. Add tables to the **Source tables** list. These are the tables that have the rows to append to the base table.

4. (Optional) If you select the **Unload source tables after appending them to the base table** check box, then the source tables are dropped from memory after the append completes.

   Select this option only if you do not want to access the source tables individually.

5. Click , and specify a name and location.

6. Click to append the source tables.

Note: If an administrator reloads the base table, any data that you previously appended to the table will not be saved. The base table will go back to its original content and size.

Note: If the append fails and you get an error message that says that the file is read-only, in some cases it could mean that your administrator has indicated the `fullcopyto=` option for the table. You cannot append to a table that includes this option. Contact your administrator for more information.

How Are Tables Used?

Base and Source Tables

The following list identifies the considerations for source tables:

- The source tables and base table must be in-memory and in the same library.
- The source tables can be LASR star schema views. However, the base table cannot be a view.
- If you append tables with character columns of different lengths, then the length of the column in the base table takes precedence. For example, if you have a column in the base table that is 15 characters and a column in the source table is 20 characters, the data from the source table’s column is truncated to 15 characters.
- If the base table is partitioned, then the source tables must be partitioned on the same variables.
- If the base table has a column that is not present in a source table, then the rows in the source table receive a missing value for the column that is not present in the source table.
- If a source table has columns that are not present in the base table, then the columns are dropped and are not appended to the base table. The base table always maintains the same number of columns.
- The order of the columns in a source table does not need to match the order of the columns in the base table.
- The data builder does not prevent you from appending the same source table more than once. Review the source table list to ensure you did not accidentally use a source table more than once.

Conserving Memory

After you append the source tables, the base table grows to accommodate the additional rows. This consumes memory on the SAS LASR Analytic Server.

If you do not need to access the source tables after appending them, then you can select the **Unload source tables after appending them to the base table** check box. This option conserves memory, but you must ensure that the source tables are loaded to memory again (often with newer data) before you can run the append again.

Delete Table Metadata

You can use the **Delete** menu option in the data builder to delete the metadata for a table that has not yet been loaded to SAS LASR Analytic Server.

If you want to delete a table that has been loaded to SAS LASR Analytic Server, you must first unload the table from SAS LASR Analytic Server, and then delete the metadata for the table. For more information, contact your administrator, or see **Loading Data ▸ Administer LASR Tables** in *SAS Visual Analytics: Administration Guide*.

To delete the metadata for a table:

1 Right-click on the table in the navigation pane.

2 Click **Delete** on the drop-down menu.

3 Click **Yes** in the confirmation message that is displayed.

Note: If the **Delete** option on the drop-down menu is grayed out, it means that you do not have permission to delete the metadata for the table.

Distributed Server: Using SASHDAT Libraries

Default Library

When your deployment uses Hadoop as a co-located HDFS or NFS-mounted MapR, the SAS Deployment Wizard registers a predefined library for it. This library is available for use in the SAS Folders tree, and it is located in `/Shared Data/SAS Visual Analytics/Public/Visual Analytics Public HDFS`.

Staging Library

You can specify a SASHDAT library as a staging library. This is a common use because the rows for the output table are distributed among the machines in the cluster. The server can then read the data in parallel when it loads the table to memory.

You must specify a SAS LASR Analytic Server library for the output library when you use a SASHDAT library for staging.

Output Library

You can specify a SASHDAT library as an output library. The engine distributes the rows for the table to the machines in the cluster. Afterward, you can select the table from the SAS Folders tree, right-click, and select **Load a Table**. This menu option loads the table to memory on a SAS LASR Analytic Server.

You can partition SASHDAT tables when they are used in an output library. You can select a column to use from the **Partition by** menu. Partitioning the table ensures that all of the rows with the same formatted value as the selected column are distributed to one machine in the cluster. The rows are also placed in the same block. When you load a partitioned table to memory, the partitioning information is retained, and the result is a partitioned in-memory table.

See Also

"Distributed Server: Partition Tables" on page 83

Restrictions

The following restrictions apply to using SASHDAT libraries with SAS Visual Data Builder:

- You cannot specify a SASHDAT library as an input library because the SASHDAT engine is a Write-only engine.
- The **Append data** check box on the **Properties** tab is disabled. The SASHDAT engine does not support appending data.
- If you specify a SASHDAT library as an output library, you cannot view the results on the **Results** tab because the SASHDAT engine is a Write-only engine.

Distributed Server: Save an In-Memory Table to SASHDAT

Note: This task is applicable if you use a distributed server with co-located HDFS or NFS-mounted MapR.

To save an in-memory table to SASHDAT:

1. Select **LASR** ▸ **Save a SASHDAT Table**.
2. Next to the **Name** field, click to select the SAS LASR Analytic Server table.
3. In the **SASHDAT Table** section, make any necessary changes to the **Location** and **Library** fields.

   Note: The data builder is initially configured to use the Visual Analytics Public LASR library and Visual Analytics Public HDFS library. If you specify different libraries, then make sure that you understand how the SASHDAT path is related to the server tag for the SAS LASR Analytic Server library. For more information, see *SAS Visual Analytics: Administration Guide*.

4 (Optional) Specify options in the **Properties** panel. By default, the data builder replaces the table if it already exists. You can specify the number of redundant copies to save.

5 Click , and make any necessary changes.

6 Click to save the table.

Note: The table is saved with the user ID of the user that started the SAS LASR Analytic Server.

Distributed Server: Legacy Co-located Providers

One of the most powerful benefits of SAS LASR Analytic Server is the ability to read data in parallel from a co-located data provider. In this configuration, the SAS LASR Analytic Server software is installed on the same hardware as the data provider. The following legacy data providers are supported, and their default library names and locations are as follows:

Teradata Data Warehouse Appliance
: **`/Shared Data/SAS Visual Analytics/Public/Visual Analytics Public Teradata`**

Greenplum Data Computing Appliance
: **`/Shared Data/SAS Visual Analytics/Public/Visual Analytics Public Greenplum`**

The data builder excels at accessing data from a variety of data sources, performing basic data preparation, and then staging the prepared data. After the data is staged, the server can load the data to memory for exploration, reporting, and further in-memory data preparation.

Distributed Server: Partition Tables

When you specify a SAS LASR Analytic Server or SASHDAT library as an output library, you can specify a partition key for the table. You can select a column to use from the **Partition by** menu.

Partitioning uses the formatted values of the partition key to group rows that have the same value for the key. All of the rows that have the same value for the key are loaded to a single machine in the cluster. For SAS LASR Analytic Server libraries, this means that the rows that have the same value for the key are in memory on one machine. For SASHDAT libraries, all of the rows that have the same value for the key are written to a single file block on one machine. (The block is replicated to other machines for redundancy.) When the partitioned table is loaded onto a server, the partitioning remains when it is in memory.

If you select a partition key and also specify sort options for columns on the **Column Editor** tab, the sort options are passed to the current engine in an ORDERBY= option. This enhancement applies to SAS LASR Analytic Server and SASHDAT libraries and can improve performance once the data is in memory.

When you specify a partition key, avoid using a variable that has few unique values. For example, partitioning by a flag column that is Boolean results in all rows on two machines because only two values are available. At the other end of the spectrum, partitioning large tables by a nearly unique key results in many partitions that have few rows.

Determining the optimal partition key can be a challenging task. However, as an example, if you tend to access data based on a customer ID, then you might improve performance by partitioning the data by customer.

See Also

SAS LASR Analytic Server: Reference Guide

Monitoring Memory Usage

A memory gauge is visible in the menu bar that shows the amount of physical memory that is used on the server.

The server rejects requests to add tables or append rows when 80% or more of the memory is used. (The default value for the server is 75%. SAS Visual Analytics deployments set the value at 80%.) System administrators can specify a different value for the server with SAS Management Console.

For non-distributed deployments, the 80% threshold applies to the single machine. For distributed deployments, the 80% threshold is evaluated for each machine in the cluster.

If the threshold is crossed, the operation fails, and the **Messages** tab includes an error that is similar to the following example:

```
ERROR: A server-side limit on the consumption of memory resources
has been reached. These limits can be adjusted by the owner or by
the administrator of the LASR Analytic Server.
```

In some deployments with limited network bandwidth, the data builder cannot communicate with the SAS LASR Analytic Server monitor, and the memory gauge fails to appear.

See Also

SAS Visual Analytics: Administration Guide

16

Importing SAS Information Maps

About SAS Information Maps

A SAS Information Map is descriptive information that is used by many SAS solutions and Business Intelligence products. An information map is applied to the data sources in your data warehouse to describe the structure and content of data. An information map does not contain any physical data. Information maps provide business users with a user-friendly way to understand data so that they can perform data queries and get results for themselves.

SAS Information Maps are created with either SAS Information Map Studio or the INFOMAPS procedure. SAS Visual Data Builder cannot create, edit, or save an information map. However, the data builder can import the business logic from a relational information map and represent it as an SQL query.

See Also

Base SAS Guide to Information Maps

Import a SAS Information Map

To import a SAS Information Map:

1 Use the SAS Folders tree to locate the map ().

2 Select the map, right-click, and select **Import Query**.

   A new data query is created with the same name as the map. The output table is automatically set with the same name as the map.

The business logic is imported from the map and represented as an SQL query on the **Code** tab.

TIP If you want to join a table, then save the data query, create a new data query, and add it as a subquery.

See Also

Chapter 18, "Customizing Code," on page 89

Limitations and Restrictions

SAS Visual Data Builder has the following limitations and restrictions for working with information maps:

- The data builder can use information maps that are based on relational tables. Information maps that are based on OLAP cubes cannot be used.
- The information map cannot use more than 50 physical tables. A table that is used more than once in a self join counts as one physical table.
- The information map cannot use more that 5000 physical columns. A column that is used in more than one data item counts as one physical column.
- Prompts are not supported. Even if a prompt has a default value, the default value is not included.
- Data items that are based on business data are not supported. For example, the equation `Dataitem1 = Year + 2` (where Year is a data item) is not included in the data query.
- Data items that are based on physical columns are not supported. For example, the equation `Dataitem2 = FirstName || LastName` (where FirstName and LastName are columns) is not included in the data query.
- The output table is automatically named with the same name as the map. The name field is limited to 32 characters.

17

Supporting Text Analysis

Features Overview

SAS Visual Analytics can perform text analysis with the explorer. To enable text analysis, perform the following data administration tasks with the data builder:

- Load a stop list to filter out commonly used words.
- Add a numeric key to existing tables that have unstructured text.
- Acquire Twitter data to analyze tweet trends. For more information about Twitter, see Chapter 6, “Importing Data from Other Sources,” on page 31.

Load a Stop List

A stop list is a table of words that you want to ignore in your text analysis. By eliminating some commonly used words, such as the various forms of the verb “to be,” you can filter out noise from your analysis. SAS Visual Analytics supports text analysis in English and German, and includes stop lists for both languages.

To load a stop list:

1 Select **Tools** ▸ **Load Text Analytics Stop List**.

2 Make changes to the location and library if necessary. Each stop list must be stored in its own metadata folder.

   A table named **ENGSTOPL** or **GRMSTOPL** is registered in the location and library that you specified.

3 Click **OK**.

SAS Visual Analytics supports one stop list for each SAS LASR Analytic Server. You load the stop list (which is a table) to memory by performing the previous steps. If more than one library is registered for SAS LASR Analytic Server, then you can use any one of them. If you load a stop list more than once or use more than one library, then the server uses the last stop list that was loaded to memory.

Add a Unique Numeric Key to a Table

In the case where your data already includes unstructured text that you want to analyze, the data builder can add a unique numeric key as a column in the table. For example, many data tables include a text column for comments. Using text analysis, you can look for trends or themes within these comments. The text analysis in SAS Visual Analytics requires a unique numeric key to reference the data sources.

To add a unique numeric key to a table:

1 Create a new data query.

2 Add the columns to include in the text analysis by clicking the column name in the table in the workspace.

3 In the **Column Editor**, right-click on the first row in the table, and select **Add Generated Key Column**.

   A column that is named `GENERATED_ID` is added to the table.

4 Specify an output table, library, and location.

5 Save and run the data query.

When you run the data query, the `GENERATED_ID` column begins at zero and increments for each row in the source table. If you select the **Append data** check box on the **Properties** tab, then the data query increments from the current maximum value.

If the data that you want to use is already in memory on SAS LASR Analytic Server, then performing the previous steps results in the following sequence of data movements:

- The data is read from memory on SAS LASR Analytic Server and transferred to the SAS Application Server.
- The data query runs and adds the generated key column.
- The data is transferred back to SAS LASR Analytic Server in a new table.

You can avoid these data movements by adding the generated key column before loading the table to memory on SAS LASR Analytic Server.

18

Customizing Code

Use the Code Tab

You can click the **Code** tab to view the SQL statements that are generated by the data preparation expressions on the **Design** tab.

You can also use the **Code** tab to enter custom code.

- The **Preprocess** and **Postprocess** views enable you to enter SAS statements that are run before and after the data query runs.
- The **All Code** view shows the generated SQL statements. Click to unlock the view so that you can manually edit all of the SAS statements for the data query. After you have unlocked the view, the button icon changes to .

After you have unlocked the view, you cannot continue using the **Design** tab to edit the data query, and the **Preprocess** and **Postprocess** views are disabled. However, if you have not saved your changes to the code, you can click to be able to use the **Design** tab.

Preprocess and Postprocess Code

In some cases, you might want to assign SAS options, load format catalogs, use LIBNAME statements, or run macros before running a data query. You can unlock the code to enter the statements, but that prevents you from using the **Design** tab. However, the **Preprocess** view or **Postprocess** view might enable you to supplement the data query while using the code generation features of the **Design** tab.

Any SAS statements that you enter in the **Preprocess** view or **Postprocess** view are included in the **All Code** view.

Considerations for Manually Editing Code

Before You Manually Edit Code

The code generation features in SAS Visual Data Builder can generate a lot of SAS statements with a few simple clicks. Depending on your needs, you might benefit by using the **Design** tab for adding input tables, performing joins, and filtering data before you customize the generated code.

Before you unlock the code, you must perform the following steps:

1. On the **Outputs** tab, specify the information for the output table, including the table name, location, and library. If you do not specify the table name or the table name is already in use, then you will not be able to save the data query.
2. (Optional) Specify the following optional settings, if needed:
   - On the **Properties** tab, select **Unique values** to determine whether the DISTINCT keyword is included in the generated PROC SQL statement. After you unlock the code, **Unique values** cannot be selected to include the keyword.
   - On the **Properties** tab, select **Append data** to enable you to use the generated SAS statements for appending data to the output table.
   - On the **Outputs** tab, select **Use a staging table** and specify the library for the staging table.
3. Save the data query by clicking . When you save the data query, the metadata for the table is registered. If you do not save the query before unlocking the code, you will receive empty columns when you run the query.

See Also

SAS Language Interfaces to Metadata

Table and Column Metadata

After unlocking the code and saving it, the columns in the output table are removed from metadata. This is because the data builder uses the columns selected on the **Design** tab to create column metadata, and those column selections are no longer available after you unlock the view on the **Code** tab.

When you run the unlocked data query, the column metadata is registered as part of running the code. However, the columns might not be visible in the data builder immediately because the table metadata (without any columns) is cached in the application.

If you did not save the query at least one time before unlocking the code, you will receive empty columns when you run the query. Make sure that you save the query before you unlock the code.

Input and Output Libraries

When you unlock the code for manual editing, you must specify the libraries to use for your tables. You can do this with LIBNAME statements. If you dragged and dropped the tables onto the workspace before you unlocked the code for manual editing, then the LIBNAME statements for the tables are automatically included.

See Also

- *SAS Statements: Reference*
- *SAS/ACCESS for Relational Databases: Reference*

Validate Custom Code

If you entered SAS statements in the **Preprocess** or **Postprocess** view, then those statements are not validated when you click .

If you unlock the code in the **All Code** view, then the button is disabled. You cannot validate customized code.

19

Scheduling

About Scheduling Data Queries

How Does the Scheduling Feature Work?

Note: By default, you cannot edit a schedule that was created by a different user. To edit a schedule that was created by a different user, your administrator must grant you Read and Write permissions on the .sas file, located in the `/Lev1/SASApp/SASEnvironment/SASCode/Jobs` directory for your operating system.

When you have a data query open in the workspace, you can click the clock icon to schedule the data query. When you schedule a data query, the data builder performs the following operations:

1 Creates a *job* that performs the data query operations.

2 Creates a *deployed job* from the job.

3 Places the job into a new *deployed flow*.

4 Schedules the flow on a *scheduling server*.

You can schedule the data query based on specified conditions (for example, run immediately or run whenever a trigger condition is met).

The job, deployed job, and deployed flow are metadata objects. The data builder stores them in the same metadata folder with the data query. The metadata objects are named based on the following pattern:

`vdb_`*`name_timestamp`*

Up to 42 characters from the data query name are used as the name.

When the specified conditions are met, the data query is run with the user ID that scheduled it. This is the behavior for the Operating System Services Scheduler.

Edit Data Queries That Are Already Scheduled

If you edit a data query that is already scheduled, you must click again so that the SAS statements for the data query are regenerated and saved. Otherwise, the schedule will not continue to run.

Stop Scheduled Data Queries

To stop a scheduled data query and prevent it from running, you must delete the trigger that is associated with it. To stop a scheduled data query by deleting the associated trigger, perform the following steps:

1 Open the data query that you want to edit.

2 In the workspace, click to open the Schedule window.

3 Select the trigger that you want to delete, and then click **Delete**.

4 If you deleted the last trigger in the list, you must perform one of the following steps in order to proceed:

- Create a new trigger by clicking **New Time Event**.
- Run the query immediately by selecting **Run now**.

5 Click **OK**.

6 In the workspace, click to save.

7 Click to run.

Note: If you deleted all of the triggers and then selected **Run now**, the scheduled data query no longer runs but it will still exist in the SAS Management Console. To permanently delete the scheduled data query, you must use the SAS Management Console. For more information, see the Deleting Jobs and Flows topic in *Scheduling in SAS*.

Caution about Scheduling Data Queries to Run Now

When you schedule a data query, one of the options is to run it immediately. Select **Run now** in the Schedule window.

Performing the following steps results in an error condition:

1 Use a SAS data set for the output table of the data query.

2 Run the data query.

3 Click the **Results** tab to look at the output.

4 Schedule the data query by selecting **Run now**.

These steps result in an error condition because SAS unlocks a SAS data set when it is opened for reading. When step 3 is performed, the output table is locked, and no other process can overwrite the output table. The following message is included in the SAS log:

Locked Error Message

```
ERROR: A lock is not available for OUTPUTTABLE.

ERROR: Lock held by process xxxx.
```

You can avoid this error condition. If you want to schedule the data query to run now, then close it, open it again, and schedule it to run now. Alternatively, you can schedule the data query to run in the future and then close the data query.

Scheduling Preferences

Default Scheduling Server

By default, your deployment includes a server that is named `Operating System Services -` *hostname.example.com*. This server is used as the default scheduling server.

Use the Server Manager plug-in to SAS Management Console to identify the scheduling servers that are included in your deployment. You can specify a different scheduling server in your application preferences. Any data queries that you schedule after you specify a different scheduling server will use the new scheduling server.

Note: The Distributed In-process scheduling server is not supported.

Some deployments include the Platform Suite for SAS server. To use this server, change the scheduling server. The default name is `Platform Process Manager`.

In all cases, when you schedule a new data query, the data builder retrieves your default scheduling server, and uses that value to look up the scheduling server in SAS metadata. The data builder uses the first server that matches the value in SAS metadata. Including the host name, such as `Operating System Services -` *hostname.example.com* ensures that the data builder uses the server that you specify.

Default Batch Server

By default, your deployment includes a server that is named `SASApp - SAS DATA Step Batch Server`. This server is used as the default batch server.

You can specify a different batch server in your application preferences. Consider the following before you change the default batch server:

- The batch server must be registered in metadata as a component of a SAS Application Server that you can access.

- You must specify the same SAS Application Server as your default application server in your preferences.

As with the default scheduling server, the data builder retrieves your default batch server, and uses that value to look up the batch server in SAS metadata the first time you schedule the data query. The data builder uses the first server that matches the value in SAS metadata.

If you have multiple server contexts and each context has its own batch servers, then you should specify in preferences the servers that you want to use.

Default Deployment Directory

A deployment directory is a SAS metadata object that represents the following items:

- the name of the SAS Application Server with which the deployment directory is associated. The default value is `SASApp`.
- a name for the deployment directory. The default value is `Batch Jobs`.
- the path to the deployment directory. The default value is ***SAS-config-dir*/Lev1/SASApp/SASEnvironment/SASCode/Jobs**.

When you schedule a data query, the SAS statements for the data query are saved in a file. The file is saved in the path that is associated with the deployment directory. The file is named based on the same pattern that is described in "How Does the Scheduling Feature Work?" on page 93.

The data builder looks up the SAS Application Server in the SAS Metadata Server using your scheduling server preference setting. The initial value is `SASApp`. If a matching server name is not found, then the data builder uses the first application server that is returned. After the server is determined, the data builder looks up the deployment directory in that server context that matches your scheduling server preference setting. If a matching deployment directory is not found, then the data builder uses the first deployment directory that is returned.

You can specify a different name for the default deployment directory. For more information about deployment directories and using the Server Manager plug-in to SAS Management Console, see *Scheduling in SAS*.

When Are the Scheduling Preferences Used?

Any of the preferences that you change are used the next time you create a data query and schedule it. If you edit an existing data query that is already scheduled, then the existing settings for the scheduling server, batch server, and deployment directory are not updated with the changes. To change the settings for existing data queries that are already scheduled, use SAS Management Console to redeploy the deployed job for the data query.

Creating Events

Why Use Events?

Events specify conditions that must be met before a step in the flow can take place. You can use SAS Visual Data Builder to create two types of events.

- *Time events* are evaluated based on a specified time being reached.

  Note: You can specify multiple time events. However, the Operating System Services Scheduler (the default scheduler) can use only one of the time events.

- *File events* are evaluated based on the state of a specified file.

You can create file events if your deployment includes a scheduling server that supports them, and the flow is deployed to that scheduling server. Time and file events can be used as triggers (conditions that must be met in order to run a flow on the scheduling server).

Create a Time Event

You can create a time event and use it as a trigger.

To create a time event:

1 In the Schedule window, select **Select one or more triggers for this data query**, and then click **New Time Event**.

2 In the New Time Event window, specify whether the time event should happen one time only or more than once. If the time event should happen one time only, then specify the date and time for the time event.

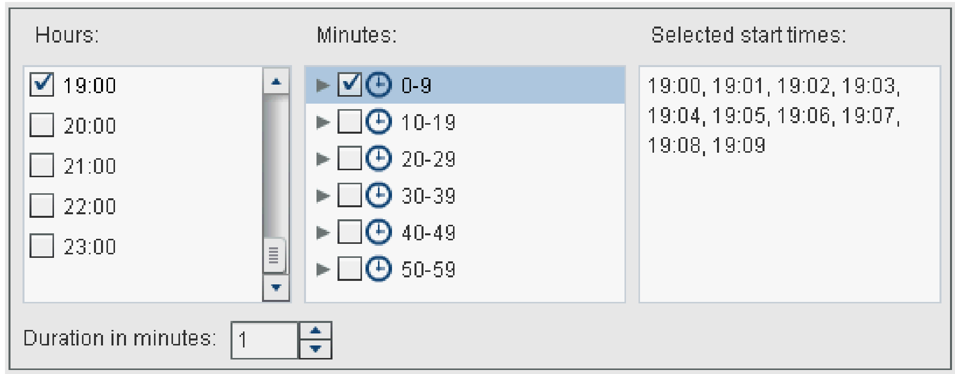
3 If the time event should happen more than once, then select **More than once** and then select a radio button for how often the time event should repeat (such as hourly, weekly, or yearly).

4 Specify the details of when the time event should repeat. The specific fields that are available depend on the recurrence interval that you select.

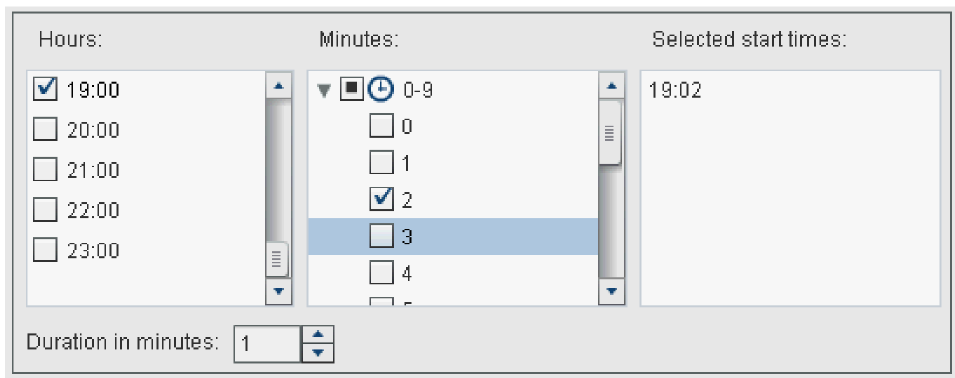
   If you select **Hourly**, then the time is calculated from hour zero on a 24-hour clock. For example, if you leave **Interval hour** set to 1, then the data query runs at hour zero (midnight), and runs each hour. If you set **Interval hour** to 2, then the data query runs every other hour, and the **Offset hour** specifies whether to run on the even hours (when set to 0) or on the odd hours (when set to 1). In this case, the **Offset hour** specifies the offset from hour zero (midnight).

   If the recurrence interval requires you to select start times, then use the **Hours** and **Minutes** check boxes to select the times. The **Minutes** area contains groupings of 10-minute intervals. Selecting a check box for a minute grouping selects all of the minutes in that grouping.

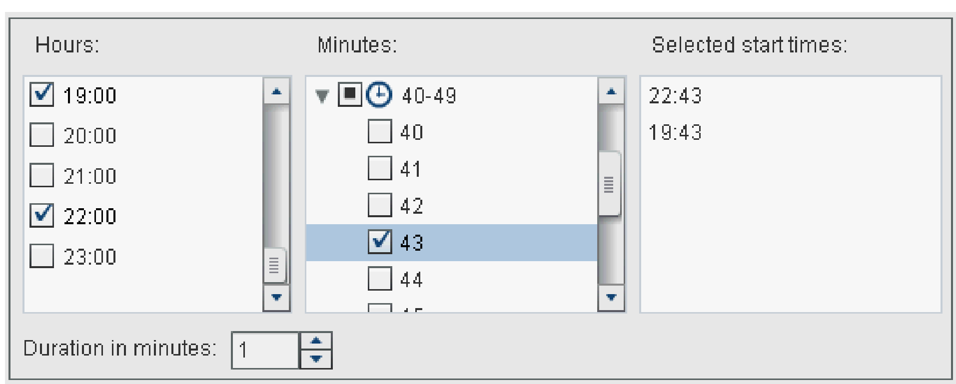
   TIP The following display is available after selecting **More than once** and then **Daily** on the New Time Event window.



To select individual minutes, expand the grouping.



If you select multiple values for **Hours**, then all of the selections for **Minutes** apply to all of the selected values. For example, if you select 19:00 (07:00 PM) and 22:00 (10:00 PM) for **Hours**, and 43 for **Minutes**, then the time event is scheduled for 19:43 and 22:43.



5 The **Duration in minutes** field is used to specify the maximum number of minutes after the specified time has been reached that the dependency is kept in an open state. This value is used when a job has multiple dependencies. It specifies a window of time in which the time event remains open so that other events can also be met.

For example, a query is specified to run only if a file event and a time event are both met. The time event has a value for **Duration in minutes** of 10. When the time specified for the time event is reached, the query runs if the file event is met at any time up to 10 minutes later.

6 If needed, specify the start date and end date for the time event. The default is to start at the current date and time and not to have an end date.

Create a File Event

You can create a file event and use it as a trigger. The file event can check for various file conditions, such as its existence, size, or age.

Note: Not all scheduling servers support file events. Platform Suite for SAS servers do support file events. The **New File Event** button that is used to create a file event is available only when the scheduling server supports file events.

To create a file event:

1 In the Schedule window, select **Select one or more triggers for this data query**, and then click **New File Event**.

2 In the New File Event window, specify or select the file to use for the file event.

3 Select the condition to evaluate the file to make the file event true. For example, if you selected **not exist** for the condition, the file event would be true only if the selected file was not in the specified location.

4 If needed, specify the details (such as size or age) about the condition.

Export Data Queries as Jobs

For deployments that include SAS Data Integration Studio and prefer to deploy jobs, create flows, and schedule flows manually, you can export a data query as a job, and then perform the deployment steps.

This feature enables you to work with a data query interactively while you create it, and then you can export the data query as a job for automating the operations. After the job is exported, you can deploy the job for scheduling with SAS Data Integration Studio. For more information about deploying jobs, see *Scheduling in SAS*.

After a job is stored in metadata, you can open the job with SAS Data Integration Studio and edit it. This might be necessary if you need to add or remove columns or change the column expression in a calculated column.

If the deployment does not include SAS Data Integration Studio, then you can modify the data query and export it as a job again. However, you cannot overwrite job objects in metadata with the data builder. If you need to delete objects from metadata (such as jobs, libraries, or tables), use SAS Management Console.

To export a data query as a job:

1 Use the SAS Folders tree to locate the data query.

2 Select the data query, right-click, and select **Export as a job**.

3 In the Export as a Job window, enter a name and specify a location. Click **Export**.

Additional Scheduling Resources

SAS Visual Data Builder provides an easy-to-use method for taking a data query and making it available for scheduling as a flow. In addition, the Schedule Manager plug-in to SAS Management Console provides additional resources for managing deployed flows, job dependencies, and scheduling servers.

For users that are familiar with the Schedule Manager plug-in, be aware that the data builder does not support scheduling a data query without a trigger. In the Schedule Manager plug-in, this option is specified as the **Manually to the scheduling server** condition. If this is your preferred method for scheduling, you can use the Schedule Manager plug-in to modify the flow for the scheduled data query. If you use the data builder to schedule the data query afterward, the setting is lost and you must repeat your steps with the Schedule Manager plug-in.

See Also

Scheduling in SAS

20

Using the Results Tab

About the Results Tab

After you have created your data query and clicked ▶, it runs, and you are prompted to determine whether you want to view the results.

If you click **Yes**, then the data builder retrieves the data and shows you the results on the **Results** tab.

Note: For large data sets, retrieving and displaying the data can take a long time.

Data Pages

When you view data, a SAS server retrieves the data from the data set. Instead of returning all of the data, the data viewer requests a page of data. You can set the page size to between 20 and 2000 rows.

A slider is provided at the bottom of the **Results** tab so that you can navigate through the pages and browse the entire data set.

Navigate within the Data

Navigate to a row in the table by entering the row number in the **Go to row** field. The **Go to first row** and **Go to last row** buttons provide navigation to the first and last rows of the table.

Find

Search for text or numbers in the data set by clicking , typing in the **Find** field, and pressing Enter. The find feature searches through the rows in the data set for the value, and it highlights the first row with the value. You can navigate through the results by clicking **Find next** and **Find previous**.

Click to set options. The following list identifies the options:

- **Options** tab
  - **Exactly matches the specified string**
  - **Contains the specified string** (default setting)
  - **Starts with the specified string**
  - **Case sensitive**
  - **Trim leading and trailing spaces** (default setting)
- **Columns** tab

  Enable and clear check boxes to specify the columns to search.

Filter and Sort

Sort the data by clicking , and then choosing the columns and sort order. The **Generated SQL statement** field shows the sort criteria.

Filter the data by clicking , and setting the following options:

- The **Sample** tab is used to limit the number of rows to return or to select distinct values.
- The **Row Filter** table is used to subset data with a WHERE clause. You can filter on the values in multiple columns, set ranges for numeric values, and use IN and NOT IN criteria for character values. If the number of distinct values is less than 50, then you can select check boxes for the values. If there are more than 50 distinct values, then you must enter the values to use for filtering.
- The **Column Filter** tab is used to select the columns to display.

Export Data

You can save the data that you view as a comma-separated values file. After you click , you can choose to export the rows on the current data page, all of the rows, or a range of rows. The export feature has a limit of exporting 200,000 rows.

If you choose to export all of the data or a range of rows, then you must click **Retrieve Data** before you can click **Export as CSV**.

Column Headings

Use the **Headings** menu to control the appearance of the column headings. You can view the column name from the data set, the column label from the data set, or a combination of the two. If a column does not have a label in the data set, then the column name is used.

Part 4

Exploring Data

21

Overview of SAS Visual Analytics Explorer

What Is SAS Visual Analytics Explorer?

SAS Visual Analytics Explorer (the explorer) 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 also apply data analysis such as forecasting, correlation, and fit lines.

Your work in the explorer is saved as a metadata object called an *exploration*. An exploration (sometimes called a visual exploration) contains all of the visualizations, data settings, and filters from your explorer session.

You can use the explorer to explore in-memory tables from the SAS LASR Analytic Server. You can either open a table directly or you can open a saved exploration.

If SAS Visual Statistics is licensed at your site, then you can perform statistical modeling tasks in the explorer. For more information, see Chapter 32, “Getting Started with SAS Visual Statistics,” on page 237.

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

You can save your explorations as PDF documents, share your saved explorations via e-mail, and export your visualizations as image files. You can export the data from your visualizations to a data file that you can use in other tools.

Note: Stored processes are not supported in the explorer.

The Welcome Window

Overview of the Welcome Window

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

The Welcome window enables you to perform the following tasks:

- "Create a New Exploration" on page 108
- "Open an Existing Exploration" on page 108

Create a New Exploration

To create a new exploration, click **Select a Data Source**. The Open Data Source window appears.

To open a data source that is already loaded, select the data source from the **Data Sources** pane and then click **Open**. In the **Search** field, you can enter a short string to find all of the data sources whose names, descriptions, or locations begin with that string.

Note: You can also perform a search that matches a string anywhere in the data source name, description, or location. Select the **Search** field, and then press Ctrl + Down to switch the search to **Search (contains)**.

To load a new data source, click one of the links in the **Import Data** pane. This task is available only if you have the Import and Load Data capability.

For information about importing data, see Chapter 4, "Importing Local Data Files," on page 19, Chapter 5, "Importing Data from Servers," on page 25, and Chapter 6, "Importing Data from Other Sources," on page 31.

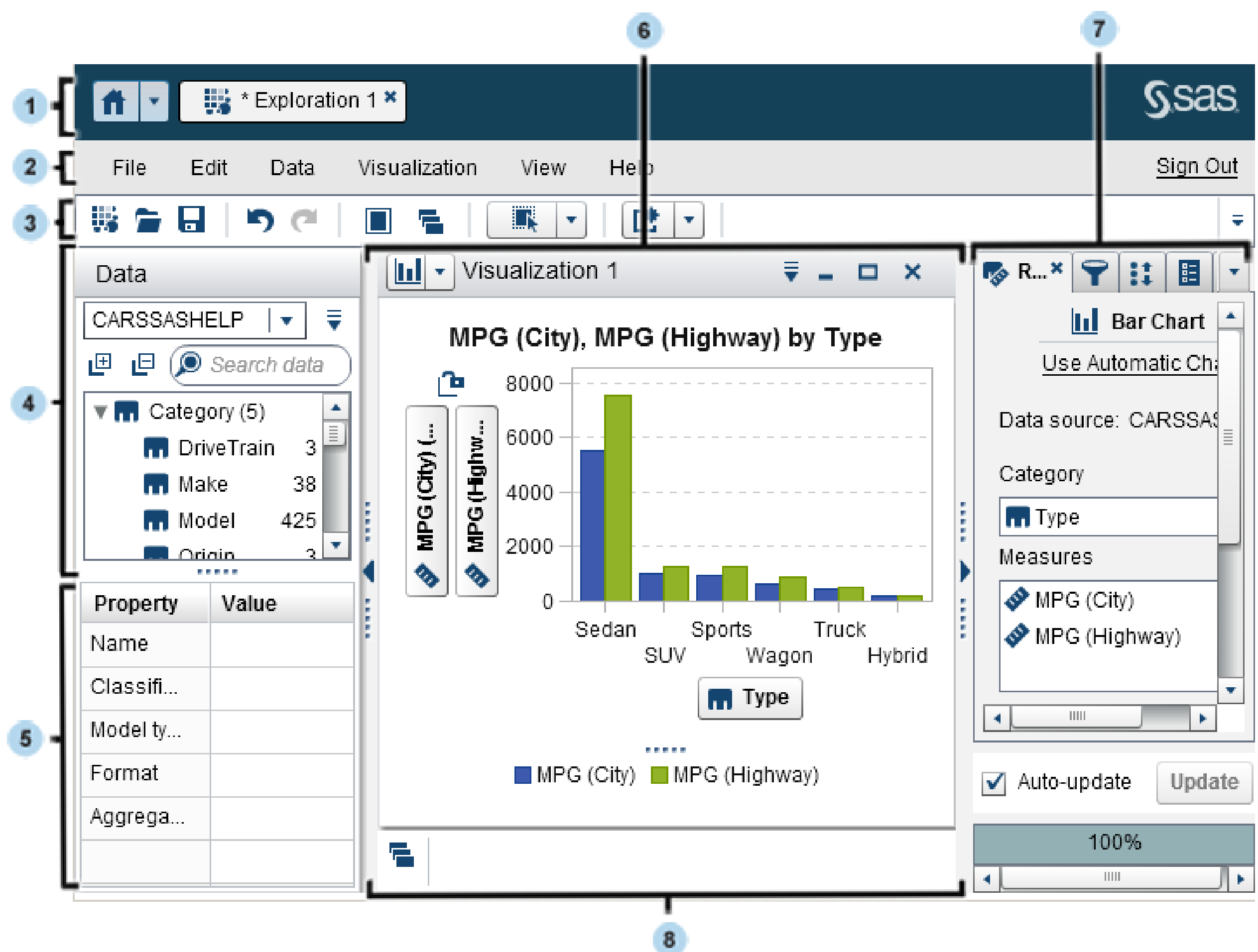
Open an Existing Exploration

To open an existing exploration, you can either select from your recent explorations, or select any exploration by clicking **Browse** to locate the exploration.

Your First Look at the Explorer

Here are the main parts of the explorer user interface:

Figure 21.1 The Explorer User Interface



1 The application bar enables you to return to the home page and to access other parts of SAS Visual Analytics and other SAS applications that integrate with the home page. You can access your recently created or viewed reports, explorations, stored processes, data queries, or other objects in your recent history. Buttons are displayed for each open application.

2 The menu bar offers common tasks, such as creating a new exploration.

3 The toolbar enables you to manage your 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 set data item properties.

6 The workspace displays one or more visualizations.

7 The right pane's tabs enable you to set properties and data roles, create filters and ranks, set global parameter values, and use comments.

8 The dock contains any minimized visualizations.

Manage Tabs in the Right Pane

In the right pane, the following tabs can be displayed:

- **Roles**
- **Filters**
- **Ranks**
- **Properties**
- **Comments**
- **Parameters**

To hide a tab, click ✕ beside the tab label.

To show a tab, select **View** ▸ ***tab-name*** from the main menu.

To display the full label name for each tab, select the ▾ drop-down list, and then deselect **Show icons only**.

You can reorder the tabs in the right pane by dragging and dropping them.

22

Specifying Preferences for SAS Visual Analytics Explorer

Specify Global Preferences

See “Specifying Your Preferences” on page 9 for details about global preferences for SAS Visual Analytics.

Specify Your Preferences for the Explorer

Here are the steps to specify preferences that are specific to SAS Visual Analytics Explorer (the explorer):

1. Select **File** ▸ **Preferences** to open the Preferences window.
2. Select **SAS Visual Analytics Explorer** ▸ **Geographic Mapping**.
3. Select the default **Map provider mode** for geographic maps. The map provider creates the background map for geo maps and for network diagrams that display a map.

   Note: You can change the map provider for an individual visualization by editing the **Map service** value on the **Properties** tab.

4. If your **Map provider mode** is **Esri**, then you must select the specific Esri map service that you want to use.
5. Select **SAS Visual Analytics Explorer** ▸ **Visualizations**.
6. 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.
7. 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. See Appendix 7, “Data Limits,” on page 609.

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

8 Select **Update automatically** to specify whether new visualizations apply your changes automatically.

9 To return to the default settings, select **Reset to Defaults**.

10 Click **OK** to apply your changes. Your preferences persist between sessions.

23

Managing Explorations

What Is an Exploration?

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

Create a New Exploration

To create a new exploration, select **File** ▸ **New Exploration** from the menu bar. The Open Data Source window appears. Select your data source, and then click **Open**.

Save Your Exploration

To save your current exploration, select **File** ▸ **Save As**, and then select a location and a name.

Delete Explorations

To delete an exploration, use the SAS Visual Analytics home page. See "Manage Content on the Classic Home Page" on page 628.

24

Managing Data

Managing Data Properties

Overview of Managing Data Properties

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

You can manage the properties for individual data items by using the **Data** pane or manage the properties for all of your data items by using the Data Properties window.

Manage Data Properties for Data Items

You can manage data properties by using the Data Properties window or by using the properties table in the **Data** pane.

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

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

Name
specifies the name of the data item.

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

Note: Changing the classification from a category to a measure or vice versa does not change the data type of the data. You can change the data type by using the **Format** and **Parse** operators in a calculated data item.

Model Type

specifies the type of 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.

Note: If you change the model type for a data item that is used in a filter, then the filter is removed.

Format

specifies the data format for the data item.

See “Select a Numeric Format” on page 117 and “Select a Date or Time Format” on page 117 for information about specifying formats.

Aggregation (for measures only)

specifies the method that is used to aggregate values for the measure. For a list of the aggregations that you can choose, see “Aggregations for Measures” on page 581.

Select a Numeric Format

Here are the steps to select a numeric format:

1 In the Data Properties window or in the property table in the **Data** pane, click the format for the data item. A window appears.

2 Select the basic format type from the **Format** list.

3 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.

   Note: You can restore the default format by clicking **Reset to Default**.

Select a Date or Time Format

Here are the steps to select a date or time format:

1 In the Data Properties window or in the property table in the **Data** pane, click the format for the data item. A window appears.

2 Select the basic format type from the **Format** list.

3 If necessary, select a more specific format from the **Sample** list. The format names also represent sample values for the format.

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

Note: You can restore the default format by clicking **Reset to Default**.

Note: Time and datetime values in the explorer ignore daylight-saving time.

Show and Hide Data Items

If your exploration contains a large number of data items, you might want to hide the items that you are not interested in.

Note: Hiding data items on the **Data** pane does not hide them in the entire exploration or in the exploration results when you export them as a report. This feature is not a way to implement column-level security.

Here are the steps to select which data items that are shown in the **Data** pane:

1 Select **Data** ▸ **Show or Hide Items** from the menu bar. The Show or Hide Items window appears.

2 Click **Select All** to select or deselect all of the data items, or click the check box beside each data item to select whether that data item is shown.

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

Sort Data Items

To sort your data items in the **Data** pane, click and then select **Sort Items**. Select **Ascending by Name** or **Descending by Name**.

Note: The explorer uses the locale of the SAS LASR Analytic Server to sort data items. For example, if you want to sort data items in the Swedish language, then the table needs to be loaded into a LASR Analytic Server that is initialized for the Swedish locale (sv_SE). Contact your system administrator for additional assistance.

Group Data Items

To group your data items in the **Data** pane, click , and then select **Group Items**. Select one of the following grouping types:

By First Letter
groups by the first letter of the data item name.

By Data Type
groups by data type (date, numeric, or character).

By Model Type
groups by model type (continuous or discrete).

By Classification
groups by data classification (category, geography, hierarchy, measure, or aggregated measure).

By Format
groups by data format (for example, $CHAR, Numeric, Percent, or MMMYYY).

By Aggregation
groups measures by the default aggregation.

Note: The default grouping is **By Classification**.

Assigning Colors for Category Values

About Colors

By default, the explorer assigns colors dynamically to category values for grouped visualizations. You can assign specific colors to your category values to ensure that the category values are always represented by a specific color. Because the same assigned colors are used in every visualization, you can easily compare values between visualizations.

Assign Colors for a Category

Here are the steps to set the colors for a category:

1. Right-click on a category from the **Data** pane, and then select **Colors**. A color selection window appears.
2. For each category value to which you want to assign a color, select the value from the **Choose** drop-down list, and then click the color box to select a color. Click **Apply** to save your selection.
3. (Optional) To automatically assign persistent colors for all of the unassigned category values, click **Assign remaining**. Colors are assigned to all of the remaining category values. These colors are used consistently in all of your visualizations.
4. Click **OK** to apply your changes.

Reset Colors for a Category

Here are the steps to reset the colors for a category:

1. Right-click on a category from the **Data** pane, and then select **Colors**. A color selection window appears.
2. Click **Reset all** to clear all of the color assignments for the category.
3. Click **OK** to apply your changes.

Assign a Color for a Specific Category Value

Here are the steps to assign the color for a specific category value:

1. Within a grouped visualization, right-click on a data value, and then select **Change Color**.
2. Click the color box to select the color.
3. Click **OK** to apply the new color. This color value is used consistently for the selected category value in all of your visualizations.

Add Additional Data Sources to Your Exploration

At any time, you can add additional data sources to your exploration. To add a data source to your exploration, select **Data** ▸ **Add Data Source**. In the Open Data Source window, select or import the data source that you want to use, and then click **Open**.

For information about importing data, see Chapter 4, “Importing Local Data Files,” on page 19, Chapter 5, “Importing Data from Servers,” on page 25, and Chapter 6, “Importing Data from Other Sources,” on page 31.

The new data source is selected automatically in the **Data** pane. You can use the drop-down list in the **Data** pane to select which data source you want to use.

Note: Each visualization in your exploration is associated with a specific data source. By default, a new visualization uses the data source that is selected in the **Data** pane when the visualization is created. For an empty visualization with no ranks or filters, you can add data items from any data source to change the data source.

Replace a Data Source in Your Exploration

At any time, you can replace a data source in your exploration with a different data source. All of the visualizations, filters, and settings that used the previous data source are applied to the new data source.

Here are the steps to replace a data source in your exploration:

1. In the **Data** pane, select the data source that you want to replace from the drop-down list.
2. Click , and then select **Change Data Source**.
3. In the Change Data Source window, select or import the data source that you want to use.

   Note: If any data items in the previous data source do not exist in the new data source, then a message appears. Any data items that do not exist in the new data source are removed from your exploration. Any filters, ranks, or calculated items that are based on the removed data items are also removed.
4. Click **Change** to add the new data source.

Remove a Data Source from Your Exploration

Here are the steps to remove a data source from your exploration:

1. On the **Data** pane, select the data source that you want to delete from the drop-down list.
2. Click ▼, and then select **Delete Data Source**.

   Note: You cannot delete the last remaining data source from an exploration.

3. Confirm that you want to delete the selected data source.

   Note: Any visualizations, filters, ranks, or calculated items that are based on the deleted data source are removed from your exploration.

Refresh Your Data Sources

At any time, you can refresh your data sources to include the latest changes.

To refresh your data sources, select **Data** ▸ **Refresh All Data Sources**.

Your visualizations are automatically updated with the refreshed data.

Creating Custom Categories

About Custom Categories

A custom category is a category data item that enables you to assign labels for groups of values in an existing category or measure.

For example, if your data source contains a category with country names, you might create a custom category to group the country names by continent. For a measure, you might label three value ranges as *Low*, *Optimal*, and *High*.

Create a Custom Category for a Measure

Here are the steps to create a custom category for a measure:

1. Select **Data** ▸ **New Custom Category**.
2. Select the measure that you want to use for the custom category.
3. Specify the **Name** for the custom category.
4. Select whether the custom category uses **Intervals** or **Distinct Values**.
5. Specify the label groups for the custom category:

a. As needed, click **New label** to add a new label group to the custom category.

b. For each label group, click the label group heading (for example, **Label1**) to specify the label text.

   Note: The label text must use characters that are compatible with the locale of the data source. If the data source uses Unicode, then your labels can contain characters from any locale.

c. For each label group, specify the values that are associated with the label:

   - For intervals, select + and specify the minimum and maximum values for the interval.
   - For distinct values, drag and drop values from the **Values** list onto the label group.

6. Select one of the following options:

   Group remaining values as
   assigns a label to all values that are not included in your label groups. The default label text is “Other.”

   Show as missing
   assigns all values that are not included in your label groups as missing values.

7. Click **OK**. The new custom category is displayed in the **Data** pane.

Create a Custom Category for a Category or a Date

Here are the steps to create a custom category for a category or a date:

1. Select **Data** ► **New Custom Category**.
2. Select the category or date that you want to use for the custom category.
3. Specify the **Name** for the custom category.
4. Specify the label groups for the custom category:

   a. As needed, click **New label** to add a new label group to the custom category.

   b. For each label group, click the label group heading (for example, **Label1**) to specify the label text.

      Note: The label text must use characters that are compatible with the locale of the data source. If the data source uses Unicode, then your labels can contain characters from any locale.

   c. For each label group, specify the values that are associated with the label.

      Drag and drop values from the **Values** list onto the label group.

5. Specify the **Options for remaining values**.

Select one of the following options:

Group remaining values as
assigns a label to all values that are not included in your label groups. The default label text is “Other.”

Show as missing
assigns all values that are not included in your label groups as missing values.

Show as is
retains the existing category value for all values that are not included in your label groups.

6 Click **OK**. The new custom category is displayed in the **Data** pane.

Working with Global Parameters

About Global Parameters

Global parameters enable you to specify a variable whose value can be changed globally at any time. You can use global parameters in filters, calculated items, and aggregated measures.

If your selected visualization contains a data item or filter that uses a global parameter, then the parameter appears on the **Parameters** tab. The same value for the parameter is used globally in your exploration for all filters, calculated items, and aggregations that include the parameter.

Note: Global parameters are shown on the **Parameters** tab only if a data item or filter that uses the global parameter is used in the selected visualization.

Create a Global Parameter

You can create a global parameter during the process of editing a filter, calculated item, or aggregated measure.

To create a global parameter, follow these steps:

1 In the expression editor, right-click the **Parameter** heading in the **Data Items** list, and then select **Create Parameter**.

2 Specify the **Name** of the parameter.

3 Specify the **Type** for the parameter. You can select either **Numeric** or **Character**.

4 For character parameters, specify the **Current value** for the parameter.

For numeric parameters, specify the following:

Minimum value
specifies the minimum possible value for the parameter.

Maximum value
specifies the maximum possible value for the parameter.

Current value
specifies the current value for the parameter.

Format
specifies the data format for the parameter.

5 Click **OK** to save the parameter.

6 Finish editing your item. If you close the expression editor window without saving, then your new parameter is not saved.

Set the Value for a Global Parameter

If your selected visualization contains a global parameter, then the parameter appears on the **Parameters** tab.

For a numeric parameter, either drag the slider to adjust the value or click the value below the slider to edit the value.

For a character parameter, enter the value in the text field.

Edit a Global Parameter

To edit the settings for a parameter, follow these steps:

1 On the **Parameters** tab, select the drop-down list for the parameter, and then select **Edit**.

2 For character parameters, specify the **Current value** for the parameter.

For numeric parameters, specify the following:

Minimum value
specifies the minimum possible value for the parameter.

Maximum value
specifies the maximum possible value for the parameter.

Current value
specifies the current value for the parameter.

Format
specifies the data format for the parameter.

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

Delete a Global Parameter

To delete a global parameter, follow these steps:

1 Edit a filter, calculated item, or aggregated measure.

2 In the **Data Items** list, right-click the parameter that you want to delete, and then select **Delete Parameter**.

Note: You can delete a global parameter only if it is not used by any filters, calculated items, or aggregated measures.

3 Finish editing your item, and then click **OK**.

Note: If you close the expression editor window without saving, then the parameter is not deleted.

Creating Calculated Data Items

About Calculated Data Items

The explorer enables you to calculate new data items from your existing data items by using an expression.

Note: All calculations are performed on unaggregated data. The calculation expression is evaluated for each row in the data source before aggregations are performed. To perform calculations on aggregated data, see "Creating Aggregated Measures" on page 126.

In addition to performing mathematical calculations on numeric values, you can use calculated data items to create character values and date and time values. For example, if your data contains separate categories for month, day, and year, then you can calculate a date value from each category.

Create a Calculated Data Item

Here are the steps to create a calculated data item:

1. Select **Data** ▸ **New Calculated Item**. The New Calculated Item window appears.

2. Enter a **Name** for the calculated data item.

3. Select the data type for the calculated data item from the **Result type** drop-down list.

   Note: The data type is updated automatically based on the outermost operator of your expression.

4. Build the expression for your calculated data item by dragging and dropping data items, global parameters, and operators onto the expression in the right pane. For each field in the expression, you can insert a data item, a global parameter, an operator, or a specific value.

   Note: You can double-click a data item in the **Data Items** pane to add it to the first available parameter in the expression.

   Note: Aggregated data items and derived data items are not supported for calculation expressions.

   When you drag and drop data items, global parameters, and operators onto the expression, the precise location of the cursor determines where and how the new element is added to the expression. As you drag the new element over the expression, a preview appears, which displays how the expression would change if you drop the element at the current location.

   For example, if your current expression is `( Profit / Revenue )`, and you drag the **x - y (subtract)** operator over the open parenthesis symbol, then the expression changes to `( [number] - ( Profit / Revenue ))`. If you drag the operator over the division symbol, then the expression changes to `( Profit - Revenue )`.

There are a large number of operator types available to perform mathematical functions, process datetime values, and evaluate logical processing such as IF clauses. See Appendix 5, “Operators for Data Expressions,” on page 583.

Note: You can also edit your expression as text. See “Editing a Data Expression in Text Mode” on page 577.

5 When you are finished creating your expression, select the **Default aggregation** for the calculated data item, and then click **Select** to choose the data format.

6 Click **Preview** to see a preview of the calculated data item as a table. The table displays the values of the calculated item and any data items that are part of the calculation expression.

7 Click **OK** to create the new calculated data item. The new data item appears in the **Data Items** pane.

Creating Aggregated Measures

About Aggregated Measures

Aggregated measures enable you to calculate new data items by using aggregated values. For example, you might want to calculate a company’s profits by subtracting expenses from revenues.

Note: To calculate data items by using unaggregated values, see “Creating Calculated Data Items” on page 125.

Aggregations are evaluated as part of the calculation expression. For each data item in your expression, you can select the aggregation type and the context for the aggregation.

Create an Aggregated Measure

Here are the steps to create an aggregated measure:

1 Select **Data** ► **New Aggregated Measure**. The New Aggregated Measure window appears.

2 Enter a **Name** for the aggregated measure.

3 Build the expression for your aggregated measure by dragging and dropping data items, global parameters, and operators onto the expression in the right pane. For each field in the expression, you can insert a data item, a global parameter, an operator, or a specific value.

Note: You can create a new calculated item to use in the aggregated measure expression. Click **New Calculated Item** to create the new calculate item.

When you drag and drop items, global parameters, and operators onto the expression, the precise location of the cursor determines where and how the new element is added to the expression. As you drag the new element over the expression, a preview appears, which displays how the expression would change if you drop the element at the current location.

There are a large number of operator types available to perform mathematical functions and evaluate logical processing such as IF clauses. See Appendix 5, “Operators for Data Expressions,” on page 583.

Note: You can also edit your expression as text. See “Editing a Data Expression in Text Mode” on page 577.

4 For each data item in your expression, select an aggregation type. By default, **Sum** is used for measures and **Distinct** is used for categories. To select a new aggregation type, drag and drop an aggregated operator from the **Operators** list onto the aggregation type in the expression. See Appendix 5, “Operators for Data Expressions,” on page 583 for a list of the aggregated operators that are available.

5 For each aggregated operator in your expression, select the aggregation context. A drop-down list beside each aggregation enables you to select one of the following context values:

ByGroup
calculates the aggregation for each subset of the data item that is used in a visualization. For example, in a bar chart, an aggregated measure with the **ByGroup** context calculates a separate aggregated value for each bar in the chart.

ForAll
calculates the aggregation for the entire data item (after filtering). For example, in a bar chart, an aggregated measure with the **ForAll** context uses the same aggregated value (calculated for the entire data item) for each bar in the chart.

By using the **ForAll** and **ByGroup** contexts together, you can create measures that compare the local value to the global value. For example, you might calculate the difference from mean by using an expression like the following:

```
Avg ByGroup(X) - Avg ForAll(X)
```

6 When you are finished creating your expression, click **Select** to choose the data format.

7 Click **OK** to create the new aggregated measure. The new data item appears in the **Data Items** pane.

Support for Aggregated Measures

Aggregated measures can be used in the following visualization types:

- Automatic Chart
- Crosstab
- Bar Chart
- Bubble Plot (grouped bubble plots only)

- Line Chart
- Treemap
- Geo Map

Creating Derived Items

About Derived Items

For the categories and measures in your exploration, you can create derived data items. Derived data items are aggregated measures that perform calculations for your data.

Derived Item Types for Categories

For categories, you can create the following types of derived items:

Distinct Count
displays the number of distinct values for the category that it is based on.

Note: If your category contains missing values, then distinct count is increased by one.

Count
displays the number of nonmissing values for the category that it is based on.

NMiss
displays the number of missing values for the category that it is based on.

Derived Item Types for Measures

For measures, you can create the following types of derived items:

Difference from Previous Parallel Period
Displays the difference between the value for the current time period and the value for the previous parallel time period within a larger time interval. For example, you might derive the difference between sales for the current month and sales for the same month of the previous year.

Note: This derived item is not available if your data source does not contain a date data item that includes the year.

Difference from Previous Period
Displays the difference between the value for the current time period and the value for the previous time period. For example, you might derive the difference between sales for the current month and sales for the previous month.

Note: This derived item is not available if your data source does not contain a date data item that includes the year.

Percent Difference from Previous Parallel Period
Displays the percentage difference between the value for the current time period and the value for the previous parallel time period within a larger time interval. For example, you might derive the percentage difference between

sales for the current month and sales for the same month of the previous year.

Note: This derived item is not available if your data source does not contain a date data item that includes the year.

Percent Difference from Previous Period

Displays the percentage difference between the value for the current time period and the value for the previous time period. For example, you might derive the percentage difference between sales for the current month and sales for the previous month.

Note: This derived item is not available if your data source does not contain a date data item that includes the year.

Percent of Subtotals

For crosstabs only, displays the percentage of the subtotal value for the measure on which it is based.

You can select whether to calculate the percentage of the subtotal value for the rows in the crosstab or for the columns in the crosstab.

For example, you might calculate the percentage of the row subtotal value for a measure that contains revenue values. The derived item displays the percentage of the subtotal of revenues for each row in the crosstab.

Note: **Percent of Subtotals** can be created only from within a crosstab visualization.

Percent of Total

Displays the percentage of the total value for the measure on which it is based. For example, you might derive the percentage of the total value for a measure that contains revenue values. If you create a bar chart by using the derived item and a category that contains products, then the bar chart shows the percentage of total revenue for each product.

Note: The percentage of the total value is relative to the subset of data that is selected by your filters and ranks.

Period to Date

Displays the aggregated value for the current time period and all of the previous time periods within a larger time interval. For example, you might derive the year-to-date total for each month.

Note: This derived item is not available if your data source does not contain a date data item that includes the year.

Year over Year Growth

Displays the percentage difference between the current time period and an equivalent time period from the previous year. For example, you might derive the difference in sales between the current month and the same month of the previous year.

Note: This derived item is not available if your data source does not contain a date data item that includes the year.

Year to Date

Displays the aggregated value for the current time period and all of the previous time periods within the year. For example, you might derive the year-to-date total for each month.

The year-to-date calculation subsets the data for each year using today's date (where today is evaluated each time you view the exploration). To use all data for every period, use a **Period to Date** item or edit the expression for the derived item.

Note: This derived item is not available if your data source does not contain a date data item that includes the year.

Year to Date Growth
Displays the percentage difference between the year-to-date value for the current time period and the year-to-date value for the same time period of the previous year. For example, you might derive the difference in year-to-date sales between the current month and the same month of the previous year.

The year-to-date calculation subsets the data for each year using today's date (where today is evaluated each time you view the exploration). To use all data for every period, use a **Period to Date** item or edit the expression for the derived item.

Note: This derived item is not available if your data source does not contain a date data item that includes the year.

Note: The derived item types that aggregate values over time use periodic operators. For information about using periodic operators, see "Periodic Operators" on page 591.

Create a Derived Item for a Measure

You can create a derived item for a measure by using the **Data** pane or by selecting a measure in a visualization.

To create a derived item from the **Data** pane, right-click the data item that the derived item is based on, and then select **Create** ▸ ***item-type***. If you select a type that uses a date, select the date data item that is used for date calculations.

To create a derived item in a visualization, right-click a measure heading in the visualization, and then select **Create and Add** ▸ ***item-type***. If you select a type that uses a date, select the date data item that is used for date calculations. The derived item is added to the visualization automatically.

For a list of item types, see "Derived Item Types for Measures" on page 128.

The new derived item appears in the **Data** pane.

Create a Derived Item for a Category

To create a derived item for a category from the **Data** pane, right-click the category, and then select **Create** ▸ ***item-type***.

The new derived item appears in the **Data** pane.

Support for Derived Items

Derived items can be used in the following visualization types:

- Automatic Chart
- Crosstab
- Bar Chart
- Line Chart

- Bubble Plot (grouped bubble plots only)
- Network Diagram
- Geo Map
- Treemap

Explore Distinct Values in a Visualization

For a bar chart, line chart, treemap, or geo map that contains a distinct count derived item, you can explore the distinct values that are associated with a data point in the visualization. To explore the distinct values, right-click on a data point in the visualization, and then select **Explore Distinct Values**. When you explore the distinct values, a new bar chart is created. The new bar chart displays the values and the frequency of each value.

For example, you might use a distinct count derived item to create a bar chart that displays the number of cities where each product line is produced.

Then, you might want to explore the distinct values for City where Promotional items are produced. By right-clicking on the bar for Promotional, and then selecting **Explore Distinct Values**, you get a new bar chart that contains the cities where promotional items are produced and the frequency of data for each city.

Editing a Calculated, Aggregated, or Derived Data Item

Open the Editor Window

To edit a duplicate data item, an aggregated measure, a calculated data item, or a derived data item, right–click the item in the **Data** pane, and then select **Edit**.

Note: You cannot edit a Percent of Total or Percent of Subtotals derived item.

Specify the Item Name

To specify the name for the item, enter a name in the **Name** field.

Edit a Parameter Value

To edit a parameter value for an operator, select the parameter, and enter a new value. Or, right-click the parameter field, and select **Replace with** to select a data item or a global parameter.

To specify a missing value, enter a period (.) character.

Add an Operator

To add an operator to the expression:

1 From the **Operators** list, select the operator that you want to add. For a list of the operators that are available, see Appendix 5, “Operators for Data Expressions,” on page 583.

2 Drag and drop the operator onto the expression.

3 For any required parameters, select the parameter, and enter a value. Or, right-click the parameter field, and select **Replace with** to select a data item or a global parameter.

Replace an Operator

To replace an operator, drag and drop a new operator onto the existing operator in the expression. You can also right-click an operator in your expression, and select **Replace Operator with** ▸ ***operator***.

Add a Global Parameter

To add an existing global parameter to your expression, drag and drop the parameter from the **Data Items** list onto your expression.

To create a new global parameter, in the expression editor, right-click the **Parameter** heading in the **Data Items** list, and then select **Create Parameter**. For more information about creating global parameters, see “Working with Global Parameters” on page 123.

Delete Part of an Expression

To delete a part of an expression, highlight the part of the expression that you want to delete, right-click, and then select **Delete** or **Clear**.

Subset an Expression

To subset an expression and delete the remainder:

1 Highlight the part of the expression that you want to keep.

2 Right-click, and select **Keep Operand**. The parts of the expression that were not selected are removed.

Edit an Expression as Text

On the **Text** tab of the expression editor, you can edit an expression as text code. See “Editing a Data Expression in Text Mode” on page 577.

Note: You can use the **Text** tab to copy and paste expressions between explorations and reports.

Display the Current Expression as Text

To display the current expression as text on the **Display Text** tab, click the drop-down list, and then select **Show display text**.

Manage the Scratch Area

The scratch area of the expression editor enables you to store elements of an expression, and then add them back into the expression as needed.

You can see the contents of the scratch area on the **Scratch** tab.

To move items to the scratch area, right-click part of the expression, and select **Move to Scratch Area**. You can also drag and drop items from your expression onto the **Scratch** tab.

To move an item from the scratch area to your expression, drag and drop the item from the **Scratch** tab onto your expression.

To delete an item in the scratch area, click next to the item.

Preview an Expression (Calculated Data Items Only)

For calculated data items, you can preview the results of your expression by clicking **Preview Results**.

You can preview the results of a subset of your expression by right-clicking a part of your expression, and then selecting **Preview Subexpression Results**.

Previewing the results of subsets of your expression is useful for troubleshooting if a complex expression does not return the expected result.

Delete a Calculated, Aggregated, or Derived Data Item

To delete a duplicate data item, an aggregated measure, a calculated data item, or a derived data item, right-click on the data item in the **Data** pane, and then select **Delete**.

Note: If the deleted data item is used by a visualization, then a confirmation message appears. The deleted item is removed from the visualization.

Note: You cannot delete a data item that is included in a hierarchy. You must remove the data item from any hierarchies before you can delete it.

Duplicate a Data Item

Duplicate data items enable you to use multiple copies of a measure that have different data formats or default aggregations. For example, you might want to use the Minimum and Maximum aggregations for a data item in the same visualization.

Here are the steps to duplicate a data item:

1 Right-click on a measure in the **Data** pane, and then select **Duplicate Data Item**. The New Duplicate Item window appears.

   Note: You can also invoke the New Duplicate Item window by dragging and dropping the same data item onto a visualization more than once.

2 Enter a **Name**, a **Format**, and a **Default aggregation** for the duplicate data item.

3 Click **OK** to create the duplicate data item.

Defining 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 maps and network diagrams 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, **Country or Region Names**. See "Define a Geography Data Item By Using a Predefined Geographic Role" on page 134.

You might 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). For each unique combination of category values, latitude values, and longitude values, a point is plotted on the geographic map. See "Define a Geography Data Item By Using a Custom Geographic Role" on page 135.

Note: For predefined geographic roles, the values of your geography data items must match the lookup values that are used by SAS Visual Analytics. To view the lookup values, see http://support.sas.com/va72geo.

Define a Geography Data Item By Using a Predefined Geographic Role

Here are the steps to define a geography data item by using a predefined geographic role:

1 In the **Data** pane, locate a category that contains geographic information.

2 Right-click the category, select **Geography**, and then select the role type. The role type can be any of the following:

   Country or Region Names
      specifies countries or regions by name.

Country or Region ISO 2-Letter Codes
specifies countries or regions by using two-letter country codes from the ISO 3166-1 standard.

Country or Region ISO Numeric Codes
specifies countries or regions by using three-digit numeric country codes from the ISO 3166-1 standard.

Country or Region SAS Map ID Values
specifies countries or regions by using the two-letter codes that are used with the MAPSGFK library that is included with SAS/GRAPH.

Note: The two-letter SAS Map ID country values are identical to the ISO 3166–1 two-letter country codes.

Subdivision (State, Province) Names
specifies country subdivisions (for example, states and provinces) by using the subdivision names.

Subdivision (State, Province) SAS Map ID Values
specifies country subdivisions (for example, states and provinces) by using the two-letter codes that are used with the MAPSGFK library that is included with SAS/GRAPH.

US State Names
specifies states and territories in the United States by using the state and territory names.

US State Abbreviations
specifies states and territories in the United States by using two-letter postal codes.

US ZIP Codes
specifies five-digit ZIP codes for the United States.

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

Define a Geography Data Item By Using a Custom Geographic Role

Here are the steps to define a geography data item by using a custom geographic role:

1 In the **Data** pane, locate a category that contains geographic information.

2 Right-click the category, and then select **Geography** ▸ **Custom**. The Custom Geography window appears.

3 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 coordinate values.

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

Define Data Items for Text Analytics

To enable text analytics for the word cloud visualization:

1 Specify a data item as the unique row identifier for the exploration. You can choose a category, a measure, or a calculated item. The unique row identifier must have a unique value for each row of the data source.

> **TIP** To determine whether the values of a category are unique, assign the category to a bar chart with no measures. If the frequency of any value is greater than one, then the category values are not unique. To determine whether the values of a measure are unique, open the **Measure Details** window, and then compare the **Total Rows** and **Distinct Count** statistics. If the statistics are identical, then the measure values are unique.

Right-click the data item in the **Data** pane, and then select **Set as Unique Row Identifier**.

Note: A nonunique row identifier can produce unreliable results.

2 Assign one or more categories to the **Document Collection** data role. Right-click a category, and then select **Document Collection**.

Note: A document collection can be used in your visualizations the same way a category is used.

25

Working with Visualizations

Overview of Visualizations

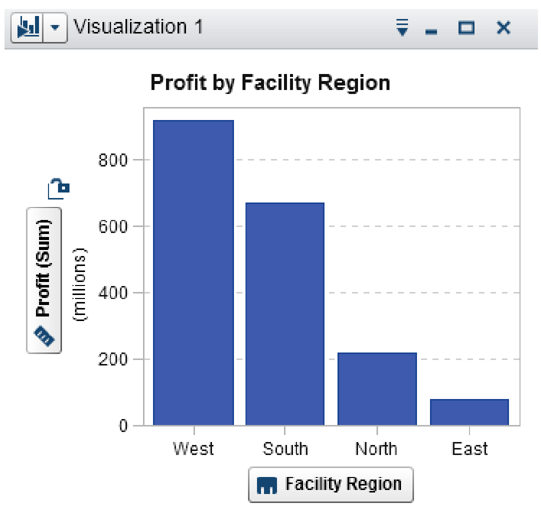
About Visualizations

SAS Visual Analytics Explorer (the explorer) displays data by using visualizations. A visualization is an interactive visual representation of your data.

Visualization Types

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

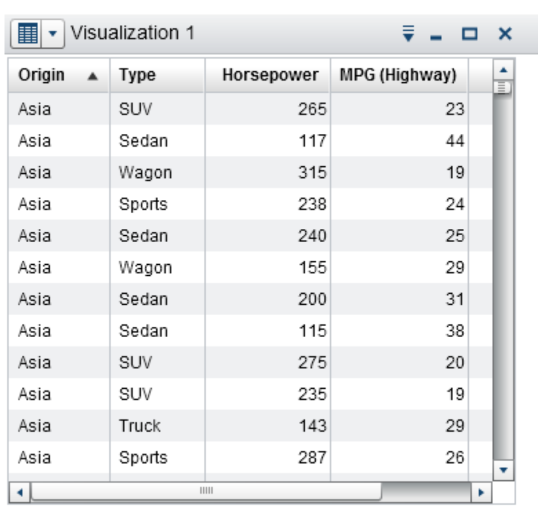
Automatic Chart



Automatically selects the chart type based on the data that is assigned to the visualization. When you are first exploring a new data set, automatic charts give you a quick view of the data.

For more information, see “Working with Automatic Charts” on page 157.

Table

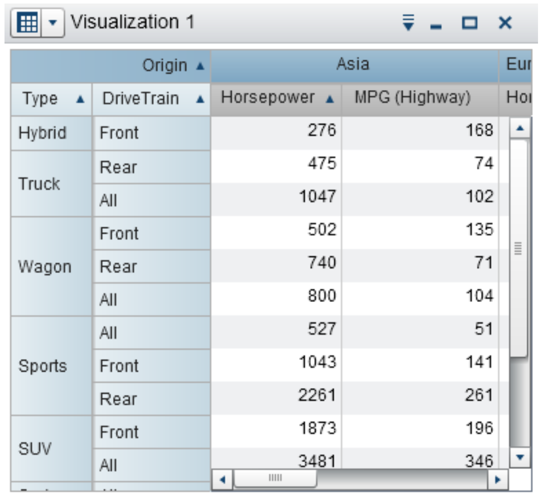


Origin	Type	Horsepower	MPG (Highway)
Asia	SUV	265	23
Asia	Sedan	117	44
Asia	Wagon	315	19
Asia	Sports	238	24
Asia	Sedan	240	25
Asia	Wagon	155	29
Asia	Sedan	200	31
Asia	Sedan	115	38
Asia	SUV	275	20
Asia	SUV	235	19
Asia	Truck	143	29
Asia	Sports	287	26

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 158.

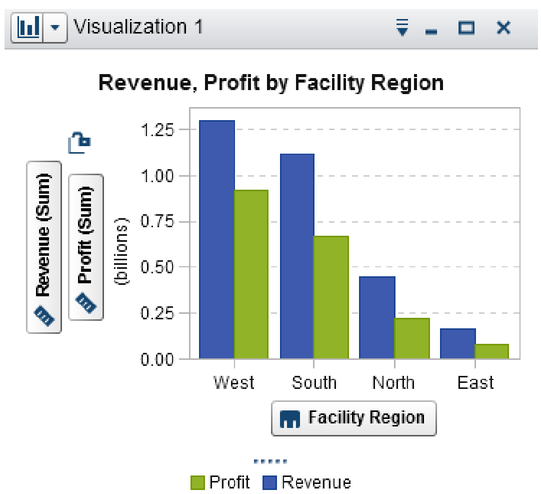
Crosstab



Displays the data as a crosstab. Crosstabs enable you to examine the data for intersections of hierarchy nodes or category values. You can rearrange the rows and columns and apply sorting. Unlike tables, crosstabs display aggregated data.

For more information, see “Working with Crosstabs” on page 159.

Bar Chart

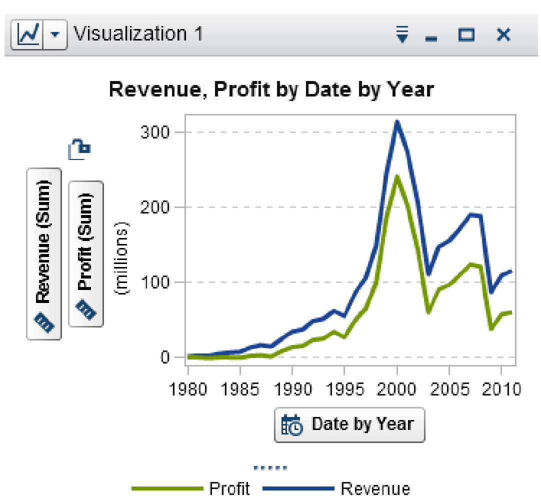


Displays the data as a bar chart. Bar charts are useful for comparing data that is aggregated by the distinct values of a category.

A bar chart consists of vertical bars or horizontal bars. You can apply grouping and create lattices.

For more information, see “Working with Bar Charts” on page 160.

Line Chart

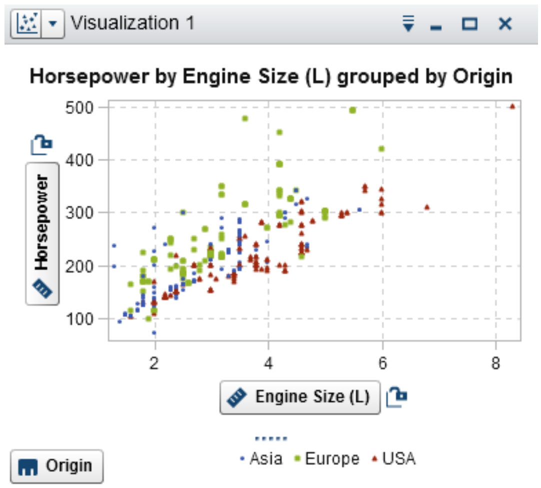


Displays the data as a line chart. A line chart is useful for data trends over time. Line charts support forecasting for predicting future values.

You can apply grouping and create lattices.

For more information, see “Working with Line Charts” on page 162.

Scatter Plot



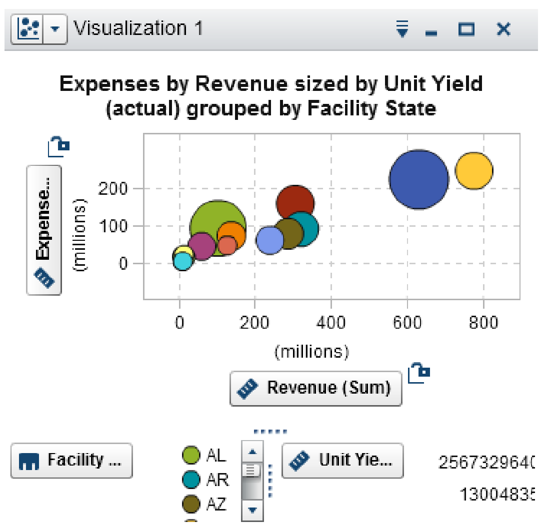
Displays the data as a scatter plot. Scatter plots are useful to examine the relationship between numeric data items.

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

When you apply more than two measures to a scatter plot, a scatter plot matrix compares each pair of measures.

For more information, see "Working with Scatter Plots" on page 165.

Bubble Plot

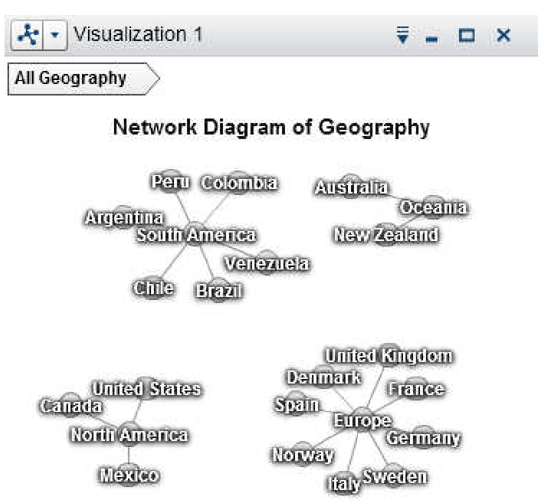


Displays the data as a bubble plot. A bubble plot displays the relationships among at least three measures. Two measures are represented by the plot axes, and the third measure is represented by the size of the plot markers.

You can apply grouping and create lattices. 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 167.

Network Diagram

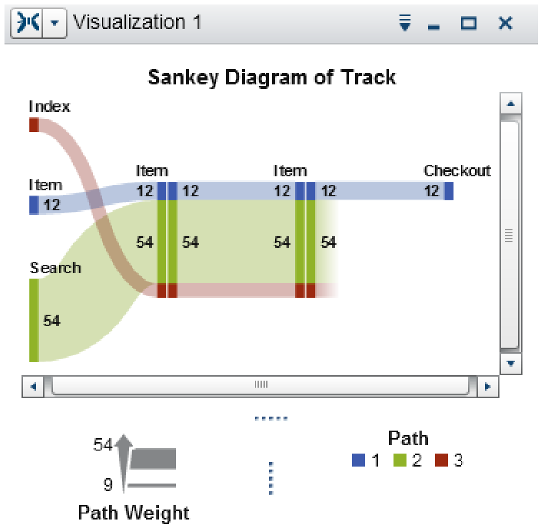


Displays a series of linked nodes. A network diagram displays the relationships between the values of categories or hierarchy levels.

You can indicate the values of measures through the sizes and colors of the nodes and the node links.

For more information, see "Working with Network Diagrams" on page 170.

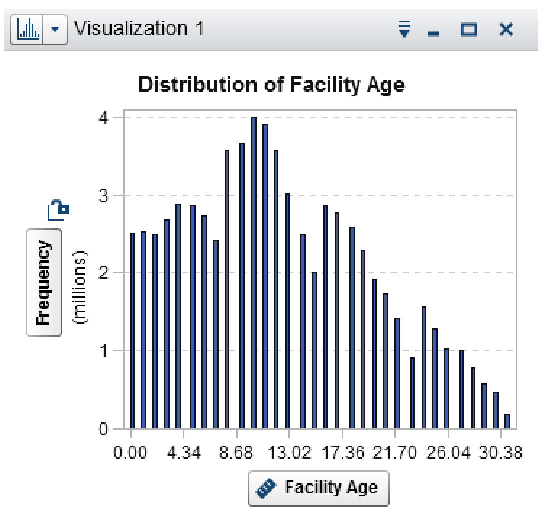
Sankey Diagram



Displays a series of linked nodes, where the width of each link indicates the frequency of the link or the value of a measure. A Sankey diagram enables you to perform path analytics. Path analytics displays flows of data from one event (value) to another, as a series of paths.

For more information, see "Working with Sankey Diagrams" on page 174.

Histogram

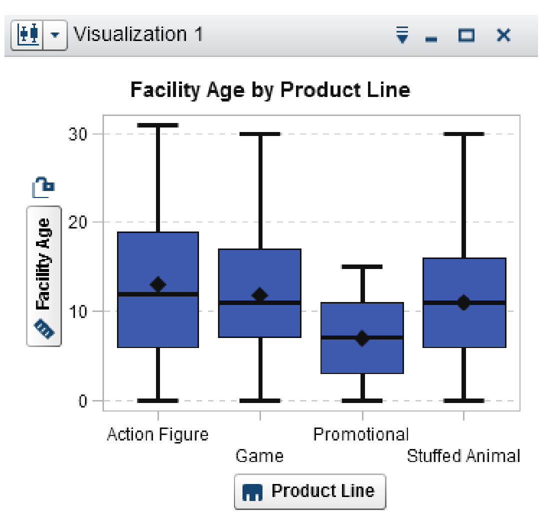


Displays the data as a histogram. A histogram displays the distribution of values for a single measure.

You can select the bar orientation, and you can select whether the values are displayed as a percentage or as a count.

For more information, see "Working with Histograms" on page 181.

Box Plot

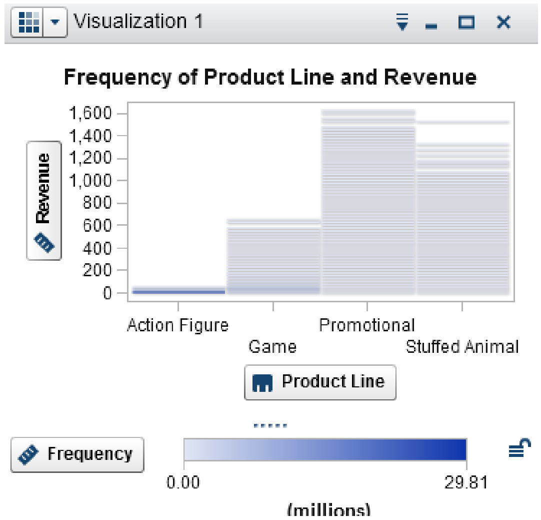


Displays the data as a box plot. A box plot displays the distribution of values for a single measure using a box and whiskers. The size and location of the box indicate the range of values that are between the 25th and 75th percentile. Additional statistical information is represented by other visual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

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

For more information, see "Working with Box Plots" on page 182.

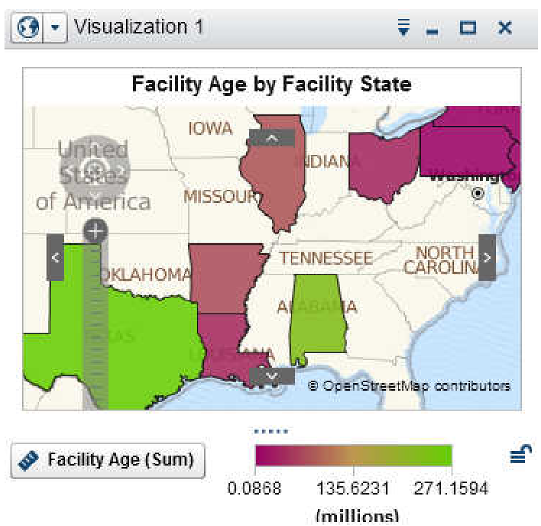
Heat Map



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

For more information, see "Working with Heat Maps" on page 184.

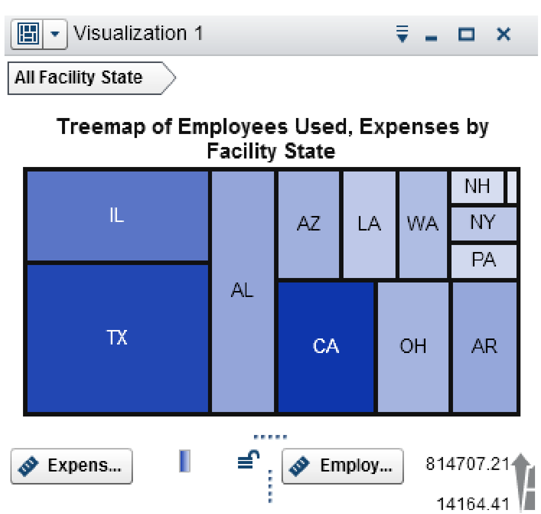
Geo Map



Displays the data as a geo map. A geo map displays your data as an overlay on a geographic map. You can display your data either as bubbles, as a scatter plot, or as colored regions.

For more information, see "Working with Geo Maps" on page 186.

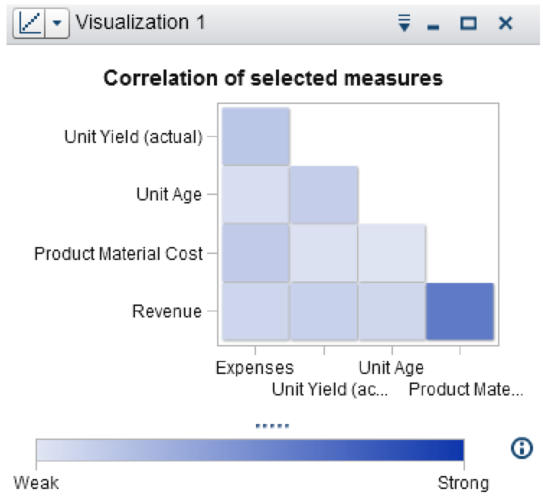
Treemap



Displays the data as a treemap. A treemap displays your data as a set of rectangles (called tiles). Each tile represents a category value or a hierarchy node. The size of each tile represents either the frequency of the category or the value of a measure. The color of each tile can indicate the value of an additional measure.

For more information, see "Working with Treemaps" on page 188.

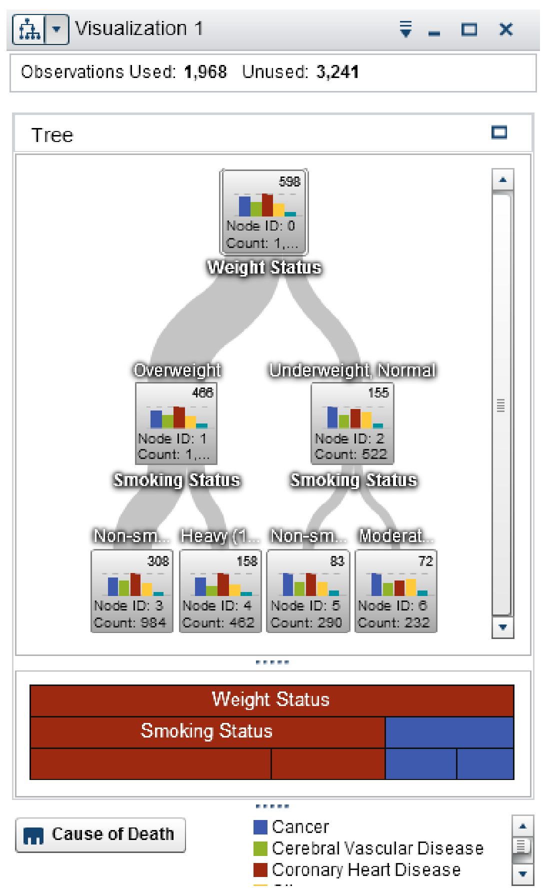
Correlation Matrix



Displays the data as a correlation matrix. A correlation matrix displays the degree of correlation between measures as a series of colored rectangles. The color of each rectangle indicates the strength of the correlation.

For more information, see "Working with Correlation Matrices" on page 190.

Decision Tree

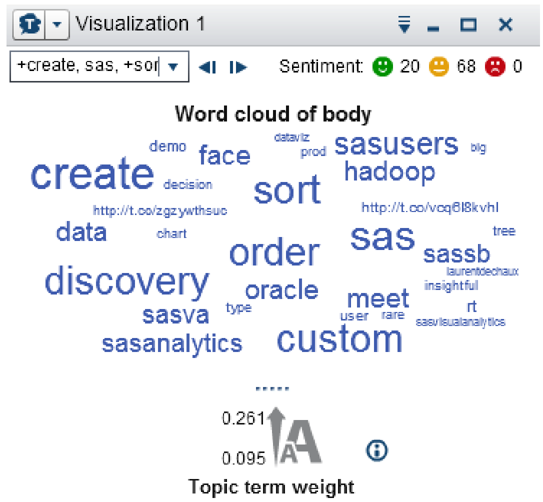


Displays the data as a decision tree. A decision tree displays a series of nodes as a tree, where the top node is the response data item, and each branch of the tree represents a split in the values of a predictor data item.

The splits enable you to see which values of the predictor data item correspond to different distributions of values in the response data item.

For more information, see "Working with Decision Trees" on page 191.

Word Cloud



Displays a set of words from a character data item. Depending on the type of word cloud and your data roles, the size of each word in the cloud can indicate the relevance of the word to a topic, the frequency of the word in a category, or the value of a measure.

You can use text analytics in a word cloud to identify topics and terms that appear together in your data and to analyze the sentiment of the documents in a topic.

For more information, see "Working with Word Clouds" on page 196.

If SAS Visual Statistics is licensed at your site, then the following modeling visualizations are available:

- Linear Regression
- Logistic Regression
- Generalized Linear Model
- Cluster

For more information, see Chapter 33, "Modeling Information," on page 249.

Working with Visualizations

Create a New Visualization

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

Manage Visualizations

Delete a Visualization

To delete a visualization, click the button on the visualization, or select **Visualization** ▸ **Delete** from the menu bar 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 **Visualization** ▸ **Duplicate** from the main menu, or select the drop-down list from the visualization toolbar, and then select **Duplicate**.

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

Show and Minimize Visualizations

To minimize a visualization, click in the visualization that you want to minimize. The visualization appears in the dock pane at the bottom of the workspace.

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

You can select which visualizations are visible by using the Manage Visualizations window.

To minimize all visualizations, select **View** ▸ **Minimize All Visualizations** from the menu bar.

To show all visualizations, select **View** ▸ **Show All Visualizations** from the menu bar.

Maximize a Visualization

To maximize a visualization to fill the entire workspace, click in 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 and drop the title bar of the visualization on the location where you want to place the visualization.

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

When your workspace contains visualizations in multiple rows, if you drag and drop a visualization onto the bottom edge 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.

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.

Use the Manage Visualizations Window

The **Manage Visualizations** window enables you to manage all of your visualizations together. Each visualization is represented by a thumbnail image or by an icon of the visualization type if the visualization has not been displayed in the current session.

To open the Manage Visualizations window, select **View** ▸ **Manage Visualizations** from the main menu bar.

The Manage Visualizations window enables you to perform the following tasks:

Task	Action
Add visualizations to the selection list. (Restore visualizations.)	Select a visualization from the **Available** list, and then click , or click to add all of the available visualizations.
Remove visualizations from the selection list. (Minimize visualizations.)	Select a visualization from the **Selected** list, and then click , or click to remove all of the visualizations.
Change the order of the selected visualizations.	In the **Selected** list, drag and drop the visualizations in the order that you want, or select a visualization and click the up and down arrows to move it.
Search the available visualizations.	In the **Search** field, enter a string to search for. A visualization matches the search if the string appears anywhere in the visualization name. Only the matching visualizations appear in the **Available** list.
Filter the available visualizations.	Click , and then select your filter parameters. You can filter either on the visualization type or on the data items that are used in each visualization. Only the matching visualizations appear in the **Available** list.
Rename a visualization.	Right-click on any visualization, and then select **Rename**. Enter a new name, and then click **OK**.
Delete a visualization.	Right-click on any visualization, and then select **Delete**.

Display Detailed Data for a Visualization

For all visualization types besides tables and crosstabs, you can display the detailed data for the visualization in the details table. To display the details table for a visualization, click the drop-down list from the visualization toolbar, and then select **Show Details**.

Change the Data Source for a Visualization

Each visualization in your exploration is associated with a specific data source.

For an empty visualization with no data items, ranks, or visualization filters, you can add a data item from any data source. The new data source is assigned to the visualization automatically.

To change the data source for a visualization that is not empty, follow these steps:

1 If the data source that you want to use is not part of the exploration, then add the new data source. See “Add Additional Data Sources to Your Exploration” on page 120.

2 Select the visualization for which you want to change the data source.

3 Remove any ranks from the **Ranks** tab and any visualization filters from the **Filters** tab.

4 On the **Roles** tab, remove all data items from the data roles.

5 From the **Data source** drop-down list, select the data source for the visualization.

Control Visualization Data Updates

By default, the explorer applies changes to your visualizations automatically when you change data roles, filters, or ranks.

To disable automatic updates, deselect the **Auto-update** check box at the bottom of the right pane. When you are ready to apply your updates, click **Update** at the bottom of the right pane.

Note: You can change the default behavior for new visualizations by deselecting **Update automatically** in the Preferences window.

Manage Visualization Comments

You can use the **Comments** tab to view and create comments for the current visualization. See “Sharing Comments in the Explorer” on page 231.

Managing Visualization Data Roles

Add a Data Item

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 and drop 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 and drop the data item onto a **Measure** or **Category** button in the visualization to assign the data item to a specific data role.
- From the visualization toolbar, select the drop-down list, and then select **Add Category** or **Add Measure**.
- Use the **Roles** tab in the right pane. Either drag and drop a data item onto a role, or expand the drop-down list beside the role and then select a data item.

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

Table 25.1 Required Data Items for Visualizations

Visualization Type	Requirements
Automatic Chart	one data item of any type
Table	one data item of any type (except aggregated measure)
Crosstab	one data item of any type
Bar Chart	one category or hierarchy
Line Chart	one category or hierarchy
Scatter Plot	one measure
Bubble Plot	three measures
Network Diagram	one category or hierarchy
Sankey Diagram	one category, one data item of any type (except aggregated measure), and one datetime or measure
Histogram	one measure
Box Plot	one measure
Heat Map	two data items of any type (except aggregated measure)
Geo Map	one geography
Treemap	one category or hierarchy
Correlation matrix	two measures
Decision tree	two data items of any type (except aggregated measure)
Word Cloud	one category or document collection

Replace a Data Item

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

- Drag and drop the new data item from the **Data** pane onto the data item in the visualization that you want to replace.
- Right-click on the data item that you want to replace in either the visualization or on 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 and drop the data item from the visualization onto the **Data** pane.
- Select the drop-down list from the visualization toolbar, and then select **Remove** ▸ ***item-name*** .
- Right-click on the data item that you want to delete in either the visualization or on the **Roles** tab, and then select **Remove *item-name***.

Switch Data Roles

To switch the data items that are assigned to two roles, drag and drop one data item onto another data item 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.

Work with Filters

You can use the **Filters** tab to subset the data in your visualizations. See Chapter 26, “Working with Filters,” on page 201.

Ranking Data

Overview of Ranking

You can use the **Ranks** tab to create ranks to subset the data in your visualizations. A rank selects either the top (greatest) or the bottom (least) aggregated value for a category.

A rank selects values for a category based on either the frequency of the category values or the aggregated value of a measure.

For example, you might create a rank of the top 10 countries by frequency to select the 10 countries that are most represented in your data source. As another example, you might create a rank of the top 10 countries by population to select the 10 countries with the greatest population.

Note: If the category for the rank is part of a hierarchy that is used in the current visualization, then the rank is applied only when the hierarchy is drilled to the level of the rank category.

Create a New Rank

To create a rank:

1. From the **Data** pane, select the data item that you want to use as the base of the rank. You can select any category or geography data item, regardless of whether it is assigned to the current visualization.
2. Either right-click on the data item, and select **Add as Rank on Visualization**, or drag and drop the data item onto the **Ranks** tab.
3. Set the parameters for the rank:
   a. Select the type of rank from the drop-down list. **Top** specifies that the rank selects the greatest value. **Bottom** specifies that the rank selects the least value.
   b. Specify the number of values for the rank. For example, if you specify `5`, and you select **Top** as the rank type, then the rank selects the five greatest values.
   c. From the **By** drop-down list, select the measure that is used to create the rank. You can either select any measure, or select **Frequency** to use the frequency of the rank category.

      Note: If you select a measure that is used in the current visualization, then the rank uses the same aggregation for the measure that is used by the current visualization.

      Note: Only the Sum, Average, Count, Minimum, and Maximum aggregation types can be used in a rank.
   d. (Optional) Select **Ties** to include ties in the rank.

      If you select **Ties**, then the rank selects as many values as necessary to include all of the ties. If you do not select **Ties**, then the rank selects only the number of values that are specified by the rank parameters.

      For example, if your rank selects the top three values, but there are five values tied for the greatest value, then the number of values that are selected by the rank depends on the **Ties** option. If you select **Ties**, then the rank includes all five of the tied values. If you do not select **Ties**, then the rank includes only three of the tied values.

      Note: If the ranking does not select all of the tied values, or if the number of tied values exceeds the maximum that is set by your administrator, then a message appears.

   Note: By default, your rank changes are applied automatically to the current visualization. To apply multiple changes together, deselect **Auto-Update**, and then click **Update** when you are ready to apply your rank changes.

Delete a Rank

To delete a rank, click ✖ on the rank on the **Ranks** tab.

Managing Visualization Axes

Lock an Axis

By default, the axes of your visualizations adjust automatically to your data. If you change the data that is displayed by applying filters or ranks or by drilling down, then the data ranges and scaling of your axes will change to fit the data.

For comparison purposes, you might want to lock the axes and retain the same data ranges and scaling. To lock an axis, click the icon beside the axis heading.

To re-enable automatic adjustments for an axis, click the icon beside the axis heading.

Adjust an Axis

For a locked axis, you can adjust the visible data range. You can adjust the range by dragging the scroll bar that appears over the axis tick marks or by right-clicking a measure heading, and then selecting **Set Visible Axis Range**.

Transfer Axis Settings

For a locked axis, you can transfer the axis settings to a compatible visualization. A compatible visualization must have a measure (or frequency) assigned to an axis.

To transfer an axis, right-click the measure heading, and then select **Transfer Axis Settings**. From the Select Compatible Visualizations window, select the visualization to which you want to transfer the axis settings.

Working with Visualization Data Ranges and Color Gradients

Support for Customized Data Ranges and Color Gradients

The following visualization types enable you to customize the data range and color gradient:

- Heat map
- Geo map (with the **Color** role assigned)

- Bubble plot (with continuous data assigned to the **Color** role)
- Network diagram (with the **Node Color** or **Link Color** role assigned)
- Word cloud (with the **Color** role assigned)
- Treemap (with the **Color** role assigned)

Specify a Custom Color Gradient

To specify a custom color gradient:

1 Right-click on the color gradient in the legend, and then select **Edit Color Gradient**.

2 From the Edit Color Gradient window, select a color gradient from the **Color gradient** drop-down list.

3 Click **OK** to apply the new color gradient.

Specify a Custom Data Range

To specify a custom data range:

1 Right-click on the color gradient in the legend, and then select **Configure Color Gradient**.

2 From the Select Color Gradient window, deselect **Automatically adjust color range to data**.

3 In the **Lower** field, specify the lower bound for the data range.

4 If the selected color gradient uses three colors, specify the inflection point of the gradient in the **Inflection** field, or select **Use midpoint** to use the midpoint between the lower and upper values as the inflection point. The inflection point is the point for the middle color in a three-color gradient.

5 In the **Upper** field, specify the upper bound for the data range.

6 Click **OK** to apply the new color gradient.

Share a Color Gradient and Data Range between Visualizations

To share data ranges and color gradients between visualizations, all of the visualizations must support customized data ranges. See “Support for Customized Data Ranges and Color Gradients” on page 153.

To share a data range and color gradient between visualizations:

1 If the data range has not been customized, then right-click on the legend for the color gradient and data range that you want to share, and then select **Lock Data Range**.

2 Right-click on the legend for the color gradient and data range that you want to share, and then select **Transfer Configuration**. The Select compatible visualizations window appears.

3 Select the visualizations that you want to share the color gradient and data range with. The visualizations are sorted based on their similarity to the source visualization. Visualizations that use the same data items and visualizations that are the same type are shown at the top of the list.

4 When you are finished, click **OK** to apply the color gradient and data range to all of the selected visualizations.

Remove Customized or Shared Color Data Ranges

To remove a customized or shared data range from a visualization, right-click on the legend, and select **Remove Custom Data Range**.

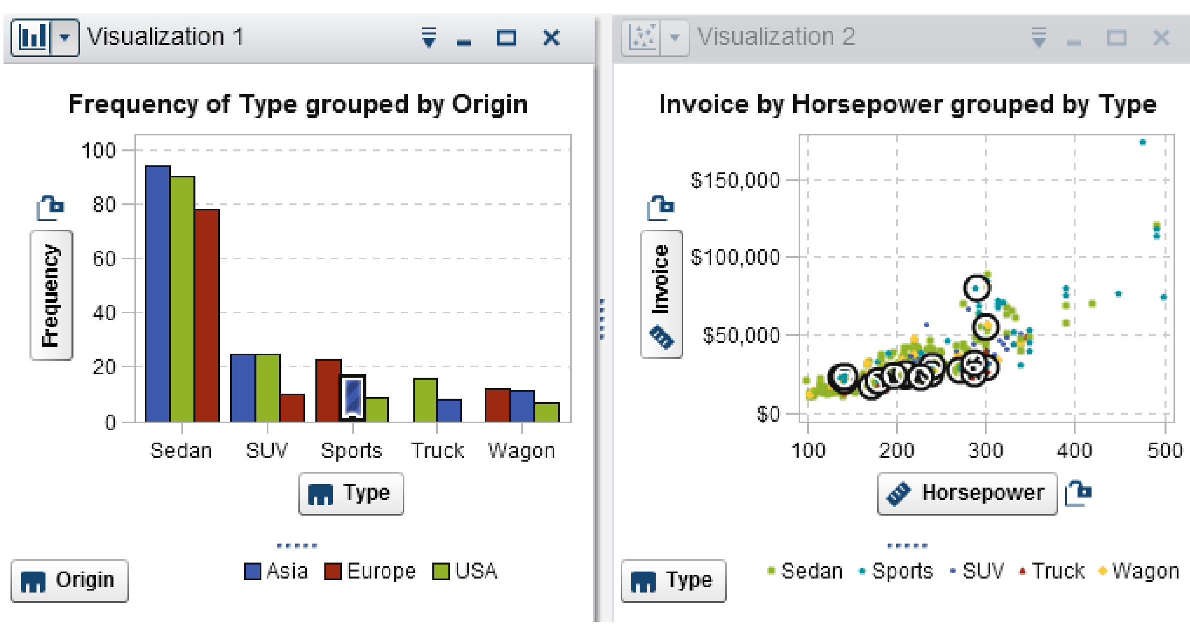
Working with Data Brushing

Overview of Data Brushing

Data brushing is a feature of some visualizations that enables you to select data values in one visualization and highlight the corresponding data values in all of your other visualizations.

For example, you might have a bar chart and a scatter plot in your exploration. If you enable data brushing, and you select a bar in the bar chart, then the markers in the scatter plot that correspond to the selected value in the bar are highlighted.

Figure 25.1 *Data Brushing Example*



The following visualization types support data brushing:

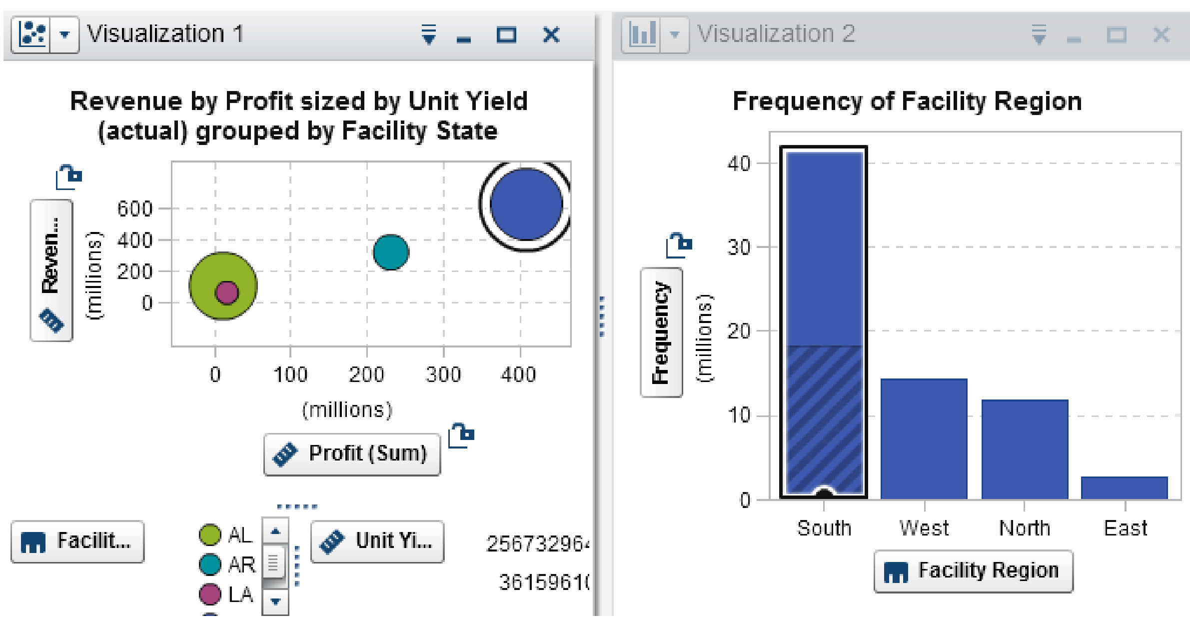
- Bar chart
- Line chart

- Scatter plot
- Bubble plot (if the **Group** role is assigned)
- Network diagram
- Histogram
- Heat map (if categories are assigned to both axes)
- Geo map
- Treemap
- Word cloud

For histograms and for bar charts of frequency that are not grouped or latticed, data brushing can indicate the portion of a bar that corresponds to the brushed data. The bar is highlighted and then partially shaded to indicate the corresponding portion. The value for the shaded portion of the bar is included in the data tip for the bar.

In Figure 25.2 , a selection in a bubble plot highlights a portion of the corresponding bar in a bar chart.

Figure 25.2 Data Brushing with Partial Shading



For bar charts that have a measure assigned or that are grouped or latticed, the entire bar is highlighted.

Enable Data Brushing

To enable data brushing, select **View** ▶ **Data Brushing**.

Select Values in a Visualization

You can select values in your visualization by using any of the following methods:

- Click the data values. This is most useful for bar charts, histograms, and treemaps. To select multiple data values, hold down the Ctrl key while clicking each data value.
- Click and drag to select data values in a rectangular region. This is most useful for scatter plots and heat maps.
- Select data values in the details table for the visualization. Press Shift+click to select multiple adjacent values, or press Ctrl+click to select or deselect individual values.

Working with Automatic Charts

Automatic charts display different types of charts automatically based on the data items that are assigned to the chart.

The type of chart can be any of the following:

Table 25.2 Automatic Chart Types

Data Items	Chart Type
One measure	Histogram
One category	Bar chart
One aggregated measure	Crosstab
One datetime category and any number of other categories or measures	Line chart
One geography and up to two measures	Geo map
One geography and three or more measures	Bar chart
One document collection	Word cloud
Two measures	Scatter plot or heat map
Three or more measures	Scatter plot matrix or correlation matrix
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, click the **Use *chart-type*** link on the **Roles** tab to change the visualization from an automatic chart to the chart type that is currently shown. For example, if your automatic chart shows a bar chart, then click **Use Bar Chart** to change the type to a bar chart and enable advanced data role settings.

Working with Tables

About Tables

A table displays data as text. The data value for each measure or category that is assigned to the table is displayed as a column. 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.

Data Roles for a Table

The basic data role for a table is a column. A column can be any type of data item. You can add any number of columns to a table.

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

By default, the table is sorted by the values of the first column. To change the sorting, click on the heading for the column that you want to sort by. An arrow appears in the column heading to indicate the sorting. If the arrow points upward, then the sort is ascending. If the arrow points downward, then the sort is descending.

Rearrange Columns

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

Resize Columns

To resize a column, click, and then drag and drop the left or right edge of a column heading.

Working with Crosstabs

About Crosstabs

A crosstab displays the intersections of category values and measure values as text. If the 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 of the crosstab contains the frequency of an intersection of category values.

Data Roles for a Crosstab

The basic data roles for a crosstab are columns, rows, and measures. You can assign either a single hierarchy or any number of categories to each column and row role. 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 frequency of each intersection of values is displayed in the cells of the crosstab.

Specify Properties for a Crosstab

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

Name
specifies the name of the visualization.

Show missing labels as blanks
displays missing values as empty cells in the crosstab. By default, missing values are represented by a period (.) character.

Indented
selects the indented layout for the crosstab.

Show column subtotals
adds subtotals to each column for each node on the row axis after the first.

Note: For the indented layout, subtotals are always enabled.

Show column totals
adds totals to each column.

Note: Total values are aggregated based on the aggregations for each measure.

Show row subtotals
adds subtotals to each row for each node on the column axis after the first.

Note: For the indented layout, subtotals are always enabled.

Show row totals
adds totals to each row.

Note: Total values are aggregated based on the aggregations for each measure.

Totals Placement
specifies the location of totals and subtotals. Select **Before** to place the totals and subtotals before the axis headings. Select **After** to place the totals and subtotals after the axis headings.

Note: For the indented layout, totals are always placed before the axis headings.

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 on the heading for the row or column that you want to sort by, and then select **Sort** and select the sorting method.

If you apply sorting to a measure, then the sorting for the rows is applied as a secondary sort. The outer row values are sorted by their subtotals for the measure. To remove measure sorting, select a sort for any category.

Rearrange Rows and Columns

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

Resize Columns

To resize a column, click, and then drag and drop the left or right edge of the column heading.

Create a Hierarchy from a Crosstab

Crosstabs enable you to create hierarchies from the categories on a crosstab axis. To create a hierarchy, right-click on a category heading, and then select **Create Hierarchy**. The categories are replaced with a new hierarchy.

The name of the new hierarchy is generated from the name of the outermost category, with the suffix “Hierarchy.”

Working with Bar Charts

About Bar Charts

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

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 category values 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 these roles:

Group
groups the data based on 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 shown as either 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.

Specify Properties for a Bar Chart

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

Name
specifies the name of the visualization.

Title
specifies the title that appears above the graph.

Note: The **Title** option is disabled if you select **Generate graph title**.

Generate graph title
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.

Frequency
specifies whether the frequency is displayed as a count (**Count**) or as a percentage (**Percent**).

Note: The frequency values are based on the data that is shown in the visualization (after filters and other data selections have been applied).

Note: This option has no effect if a measure is assigned to the visualization.

Grouping style
specifies how grouped data is displayed. If you select **Stack**, then the values of the grouping variable are displayed as segments of each bar. If you select **Cluster**, 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 **Group** 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 labels
shows the data values as text in the visualization.

Note: You can always view a data value as a data tip when you position the cursor over a data value.

Rotate axis labels
displays the category labels at an angle.

Note: The **Rotate axis labels** option has no effect when you select a horizontal bar orientation.

Reference Lines
adds reference lines to the visualization. To create a reference line, click , and then specify the parameters for the line in the New Reference Line window.

The name of each reference line appears below the **Reference Lines** option. You can edit or delete your reference lines by using the icons next to each name.

Sort Data Values

By default, a bar chart is sorted in descending order by the value of the first measure. To change the sorting, right-click on the data item that you want to sort on, and then select **Sort** ▸ ***sort-method***. For a grouped bar chart, the data is sorted by the category values in alphabetical order.

Note: If the visualization contains a rank, then, by default, the data is sorted based on the values of the rank.

Working with Line Charts

About Line Charts

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

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 category values are plotted on the category axis. You can assign many measures, and the measure values are plotted on the response axis. If the line 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 these roles:

Group
groups the data based on the values of the category data item that you assign. A separate line is created for each data 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.

Forecasting
enables forecasting for the line chart. Forecasting estimates future values for your data based on statistical trends.

Note: Forecasting is available only if a date or datetime data item is assigned to the visualization.

Underlying factors
if forecasting is enabled, adds additional measures to the forecast as underlying factors. The forecasting model evaluates the additional measures to determine whether they contribute to the accuracy of the forecast. If the additional measures do not increase the accuracy of the forecast, then they are not used. If the additional measures do increase the accuracy of the forecast, then the forecast line is adjusted, and the confidence bands are narrowed.

The measures that you add as underlying factors can also be used in a scenario analysis.

Note: Underlying factors are available only if your line chart contains a single measure in the **Measures** role.

Specify Properties for a Line Chart

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

Name
specifies the name of the visualization.

Title
specifies the title that appears above the graph.

Note: The **Title** option is disabled if you select **Generate graph title**.

Generate graph title
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.

Frequency
specifies whether the frequency is displayed as a count (**Count**) or as a percentage (**Percent**).

Note: The frequency values are based on the data that is shown in the visualization (after filters and other data selections have been applied).

Note: This option has no effect if a measure is assigned to the visualization.

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 markers
shows markers for the data points in the visualization.

Show data labels
shows the data values as text in the visualization.

Note: You can always view a data value as a data tip when you position the cursor over a data value.

Thin data labels
removes data labels as necessary to avoid overlap. The thinning algorithm retains the labels for local high points and low points on the line chart.

Rotate axis labels
displays the category labels at an angle.

Reference Lines

adds reference lines to the visualization. To create a reference line, click , and then specify the parameters for the line in the New Reference Line window.

The name of each reference line appears below the **Reference Lines** option. You can edit or delete your reference lines by using the icons next to each name.

Duration (if forecasting is enabled)
specifies the number of data intervals to forecast.

Note: This option is available only if forecasting is enabled for the visualization.

Confidence (if forecasting is enabled)
specifies the degree of confidence for the confidence band. The default value is 95%.

Sort Data Values

By default, for category data, a line chart is sorted in descending order by the value of the first measure. To change the sorting, right-click on the data item that you want to sort on, and then select **Sort** and select a sorting method.

Note: If the visualization contains a rank, then, by default, the data is sorted based on the values of the rank.

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

Forecasting

About Forecasting

Forecasting uses the statistical trends in your data source to predict future data values. Forecasting is available only if a date or datetime data item is assigned to the visualization.

In addition to the predicted future data values, the forecast displays a confidence band. For more information, see "Forecasting" on page 227.

Enable Forecasting

To add forecasting to a line chart, select **Forecasting** on the **Roles** tab.

Note: The line chart must contain a date or datetime data item to apply forecasting.

On the **Properties** tab for the line chart, you can adjust the number of data values to predict by using the **Duration** option. You can specify the degree of confidence for the confidence band by using the **Confidence** option.

On the **Roles** tab, you can add additional measures to the forecast as **Underlying factors**. The forecasting model evaluates the additional measures to determine whether they contribute to the accuracy of the forecast. If the additional measures do not increase the accuracy of the forecast, then they are not used. If the additional measures do increase the accuracy of the forecast, then the forecast line is adjusted, and the confidence bands are narrowed.

Note: Underlying factors are available only if your line chart contains a single measure in the **Measures** role.

The measures that you use as underlying factors can also be used in scenario analysis and goal seeking. For more information, see "Working with Scenario Analysis and Goal Seeking" on page 228.

Working with Scatter Plots

About Scatter Plots

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

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

Data Roles for a Scatter Plot

The basic data role for a scatter plot is a measure. You can assign any number of measures. If you assign a single measure to a scatter plot, then the values are plotted along a line.

In addition to measures, you can assign a **Group** variable. The **Group** variable groups the data based on 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.

You can add data items to the **Data tips** role. The values for the data items in the **Data tips** role are displayed in the data tips for the scatter plot.

Specify Properties for a Scatter Plot

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

Name
specifies the name of the visualization.

Title
specifies the title that appears above the graph.

Note: The **Title** option is disabled if you select **Generate graph title**.

Generate graph title
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.

Fit Line
adds a fit line to the scatter plot. For information about the fit types that are available, see “Fit Lines” on page 226.

Note: Fit lines are not available if a grouping variable is assigned to the scatter plot.

Reference Lines
adds reference lines to the visualization. To create a reference line, click , and then specify the parameters for the line in the New Reference Line window.

The name of each reference line appears below the **Reference Lines** option. You can edit or delete your reference lines by using the icons next to each name.

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 visualization. For more information, see “Correlation” on page 225.

Fit Line
: plots a model of the relationship between the variables in the visualization.

  There are many types of fit lines, including linear fit, quadratic fit, cubic fit, and penalized B-spline. For more information, see “Fit Lines” on page 226.

Correlation is applied to your visualization automatically when you add a linear fit line. It is not available with other fit types.

Enable Data Analysis

To add a fit line to your visualization, select the drop-down list from the visualization toolbar, and then select **Fit Line** and select the fit type. For details about the fit types that are available, see “Fit Lines” on page 226.

Working with Bubble Plots

About Bubble Plots

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

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

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.

Bubble size
: specifies the measure that determines the marker size.

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

Group
: groups the data based on 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** role and the **Color** role at 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 measure value is represented by the bubble color.

Note: You cannot assign both the **Group** role and the **Color** role at the same time.

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

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

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.

Specify Properties for a Bubble Plot

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

Name
specifies the name of the visualization.

Title
specifies the title that appears above the graph.

Note: The **Title** option is disabled if you select **Generate graph title**.

Generate graph title
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.

Size scale
specifies the type of scaling that is used to draw the bubbles. Select one of the following values:

Linear
specifies that the bubble sizes are scaled relative to the minimum and maximum values (or frequencies) in your data. A negative value is displayed as a smaller bubble than a positive value.

For the **Linear** scaling type, the difference in bubble sizes might not be proportional to the difference in values.

Magnitude
specifies that the bubble sizes are scaled relative to zero and the greatest absolute value in your data. A negative value is displayed as a hexagon.

For the **Magnitude** scaling type, the difference in bubble sizes is proportional to the difference in absolute values.

Frequency
specifies whether the frequency is displayed as a count (**Count**) or as a percentage (**Percent**).

Note: The frequency values are based on the data that is shown in the visualization (after filters and other data selections have been applied).

Note: This option has no effect if a measure is assigned to the **Bubble size** role.

Color gradient
selects the gradient colors for the visualization.

You can click to select the values that are used to assign the colors. See “Specify a Custom Data Range” on page 154.

Reference Lines
adds reference lines to the visualization. To create a reference line, click , and then specify the parameters for the line in the New Reference Line window.

The name of each reference line appears below the **Reference Lines** option. You can edit or delete your reference lines by using the icons next to each name.

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 25.3 *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.

Task	Action
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.

Note: If you select a bubble to track, the selected bubble is highlighted in the current animation frame.

Working with Network Diagrams

About Network Diagrams

A network diagram displays the relationships between category values as a series of linked nodes.

You can create two types of network diagrams:

Hierarchical
: creates a hierarchical structure by using a hierarchy or a set of categories.

Ungrouped
: creates a structure by using a source data item and a target data item. A node is created for each value of the source data item, and a link is created from each node to the node that corresponds to the value of the target data item.

For example, if your source data item specifies the name of every employee in an organization, and your target data item specifies the manager of each employee, then the network diagram has a node for each employee that is linked to the node for the employee’s manager.

Data Roles for a Network Diagrams

Basic Data Roles for a Hierarchical Network Diagram

The basic data role for a hierarchical network diagram is **Levels**. Specify a hierarchy or any number of categories. The data items in the **Levels** role specify the nodes of the network diagram. If you add categories to the **Levels** role, then the order of the categories determines their hierarchical relationship. You can drag and drop the categories to arrange them.

TIP If your diagram displays multiple duplicate nodes, then an ungrouped diagram might be a better fit for your data.

Basic Data Roles for an Ungrouped Network Diagram

The basic data roles for an ungrouped network diagram are **Source** and **Target**. The **Source** specifies a category that contains all of the node values for the plot. The **Target** specifies a category that creates the links between nodes.

The **Target** category must contain a subset of the values of the **Source** category.

To represent terminal (target-only) values in an ungrouped network diagram, you can add rows to your data where the terminal value is the value for the source data item and the target data item is missing.

For example, in the following table, the final row represents a terminal value:

Employee	Manager	Salary
EMP1	MGR1	40000
EMP2	MGR1	55000
EMP3	MGR1	50000
MGR1		75000

Additional Data Roles for a Network Diagram

In addition to the basic data roles, you can specify the following additional data roles for a network diagram:

Node size
specifies a measure that determines the size of the nodes in the diagram.

Node color
specifies a data item that determines the color of the nodes in the diagram.

Link width
specifies a measure that determines the width of the links in the diagram.

Link color
specifies a data item that determines the color of the links in the diagram.

Data tips
specifies data items whose values are included in the data tips for the diagram. Measure values are aggregated by sum.

Specify Properties for a Network Diagram

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

Name
specifies the name of the visualization.

Title
specifies the title that appears above the graph.

Note: The **Title** option is disabled if you select **Generate graph title**.

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

Show map
displays the network as an overlay on a geographic map.

Note: This option is available only if geographies are assigned to all of the data roles that create nodes. For hierarchical diagrams, all of the **Levels** must be geographies. For ungrouped diagrams, the **Source** and **Target** must be geographies.

Show labels
shows the node values as text inside each node.

Show direction
shows the links in the diagram as arrows.

Map service
if your diagram contains a map, specifies the source for the background map.

Node color
selects the color gradient for the nodes in the diagram. You can click to select the values that are used to assign the colors. See "Specify a Custom Data Range" on page 154.

Note: This option is available only if you assign a measure to the **Node Color** data role.

Link color
selects the color gradient for the links in the diagram. You can click to select the values that are used to assign the colors. See "Specify a Custom Data Range" on page 154.

Note: This option is available only if you assign a measure to the **Link Color** data role.

Additional levels
for hierarchical diagrams only, specifies the number of levels that are displayed beneath the current level.

Node spacing
specifies the amount of space between nodes in the diagram.

Node size
adjusts the size of all of the nodes in the diagram.

Source node
displays the current node selection. To remove the node selection, select .

Predecessors
selects the number of levels of predecessors (parents) of the source node to select. 0 specifies the source node. Note that you can specify a range by making multiple selections. For example, select both 0 and 1 to specify that the source node and the first level of predecessors are selected.

Successors
selects the number of levels of successors (children) of the source node to select. 0 specifies the source node. Note that you can specify a range by making multiple selections. For example, select both 0 and 1 to specify that the source node and the first level of successors are selected.

Arrange Nodes in a Network Diagram

Move Nodes

You can move any node in the diagram by clicking the node and dragging it. You can move multiple nodes in the diagram by selecting the nodes that you want to move and dragging them.

Note: The positions of the nodes in your diagram are saved with your exploration.

Refresh the Node Layout

You can refresh your node layout by clicking . The network diagram creates a new node layout based on your current node layout. This is especially useful after you have moved nodes manually. Refreshing the node layout adjusts the spacing and orientation of your nodes.

Use the Spotlight Tool to View Multiple Data Tips

The spotlight tool enables you to view data tips for all of the nodes that are within a circular area around the cursor.

Click to enable or disable the spotlight tool.

Select Nodes in a Network Diagram

You can select nodes in the diagram by using any of the following methods:

- Click and drag the cursor to create a rectangular selection.
- Hold down the Ctrl key and click the nodes that you want to select.
- Select a series of linked nodes by setting a node as the source node.

  Right-click a node, and then select **Set as Source for Selection**.

  On the **Properties** tab, specify the range of levels of **Predecessors** (parents) and **Successors** (children) of the source node to select. 0 specifies that the source node is selected.

  For example, if you specify a range of 0–1 for **Predecessors** and a range of 0–2 for **Successors**, then the source node, one level of predecessors, and two levels of successors are selected.

Control the View of a Network Diagram

You can control the view of a network diagram by using the following controls:

Zoom	Zoom in and out at the location of the cursor by scrolling the mouse wheel.
Pan (scroll)	Pan (scroll) the diagram by holding down the Shift key and dragging the diagram.

Reposition the view	When you have zoomed in on a diagram and scroll bars are displayed, reposition your zoomed view by holding the Shift and Alt keys and dragging the diagram.
Rotate	Rotate the diagram by holding down the Ctrl key and dragging the diagram.

Create a Hierarchy from a Network Diagram

If your network diagram contains categories in the **Levels** role, then you can create a new hierarchy by using the categories.

On the **Roles** tab, click on the drop-down list for the **Levels** role, and then select **Create Hierarchy**. A new hierarchy is created with the name of the first category in the hierarchy.

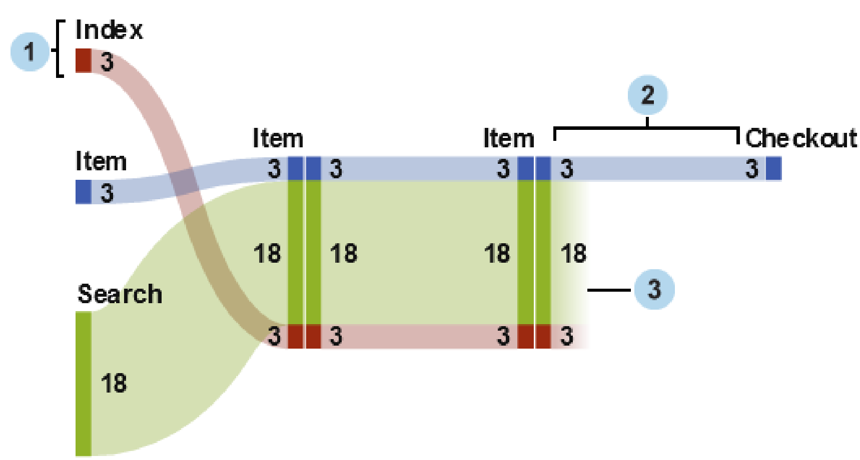
Working with Sankey Diagrams

About Sankey Diagrams

A Sankey diagram enables you to perform path analytics. Path analytics displays flows of data from one event (value) to another as a series of paths.

The following display shows the parts of a Sankey diagram.

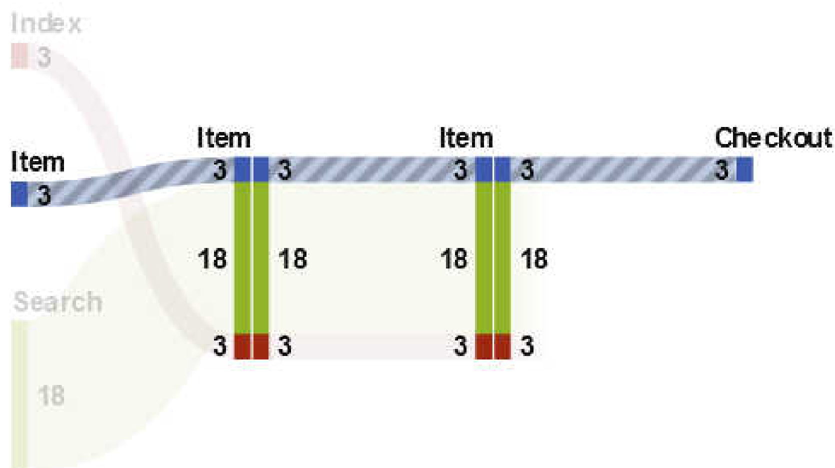
Figure 25.3 *Parts of a Sankey Diagram*



1. Nodes contain the events in each path. The node displays the width of each link that enters and exits the node. The same event can appear at multiple nodes in the diagram.
2. Links between nodes make up the paths in the diagram. The width of each link can represent either the frequency of the path or the value of a weight measure.

3 Drop-off links are links that end at the current node. Drop-off links are displayed only if some links continue onward from the current node.

Figure 25.4 A Path in a Sankey Diagram



A path in a Sankey diagram represents a distinct sequence of events. Each path in the diagram consists of one or more transactions. A transaction is a sequence of events that are associated with a specific transaction identifier value.

For example, if your data contains the activity of visitors to a website, then your transaction identifier might be the unique session identifier for each visitor. The events in your data might be the individual pages on the website that the visitor accessed. In this example, each transaction is the sequence of pages that were accessed by a specific visitor, and each path is a sequence of pages that contains all of the transactions that follow that sequence.

Data Roles for a Sankey Diagram

Here are the basic data roles for a Sankey diagram:

Event
specifies a category whose values identify the events that are represented as nodes in the diagram.

Sequence Order
specifies a datetime data item or a measure whose values identify the order of the events for each transaction.

Transaction Identifier
specifies a data item whose values identify the transactions in the diagram.

Note: The paths in a Sankey diagram exclude any missing values in the data items that are assigned to the **Event**, **Transaction Identifier**, or **Sequence Order** role.

In addition to the basic data roles, you can specify the following role:

Weight
specifies a measure for the weight of each event in a transaction. The weight values for each transaction, for each event in a path are aggregated to determine the path weight.

Note: If the measure that is assigned to the **Weight** role has an aggregation other than SUM or COUNT, then the node width labels in the diagram might be misleading. It is recommended that you disable the **Show nodes** option if your **Weight** measure has an aggregation other than SUM or COUNT.

Specify Properties for a Sankey Diagram

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

Name
specifies the name of the visualization.

Title
specifies the title that appears above the graph.

Note: **Title** is disabled if you select **Generate graph title**.

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

Show labels
displays the event name for each node.

Show nodes
displays link width values at each node.

Show outline
displays a box around each node.

Vertical layout
displays the diagram vertically.

Link color
specifies the method that is used to group and color the paths or links in the diagram.

Select one of the following:

Path
displays each path as a separate group with a distinct color.

Event
groups links together by event. Links that originate from the same event have the same color. For links between the same two events, the links are combined even if their transactions are part of different paths.

Drop off
groups links together based on whether they are drop-off links. A drop-off link represents a link that ends at the current node while other links continue onward from the current node. For links between the same two events, the links are combined even if their transactions are part of different paths.

Link width
specifies whether the width of a link in a path represents the path frequency (**Frequency**) or the aggregated value of the weight measure (**Weight**).

Note: If you do not assign the **Weight** data role, then **Frequency** is always used for the link width.

Note: If the aggregated value of the weight measure for any path is negative, zero, or missing, then **Frequency** is used for the link width.

Minimum path length
specifies the minimum length for a path that is displayed. The path length that you specify indicates the number of nodes in the path.

For example, if the **Minimum path length** is 3 and the **Maximum path length** is 5, then the diagram displays only paths that have lengths of 3, 4, or 5 nodes.

Maximum path length
specifies the maximum length for a path that is displayed. The path length that you specify indicates the number of nodes in the path.

For example, if the **Minimum path length** is 3 and the **Maximum path length** is 5, then the diagram displays only paths that have lengths of 3, 4, or 5 nodes.

Note: A hyphen character (-) specifies that there is no maximum path length.

Minimum frequency
specifies the minimum path frequency for a path that is displayed. For example, if you specify a minimum path frequency of 5, then the diagram displays only paths that have a frequency of 5 or greater.

Maximum frequency
specifies the maximum path frequency for a path that is displayed. For example, if you specify a maximum path frequency of 10, then the diagram displays only paths that have a frequency of 10 or fewer.

Note: A hyphen character (-) specifies that there is no maximum frequency.

Path ranking
specifies a ranking method for the paths in the diagram. If you specified a measure for the **Weight** role, then ranking is based on the aggregated value of the weight measure for each path. If there is no measure for the **Weight** role, then ranking is based on the frequency of each path.

From the drop-down list, select **Top** to select the paths with the highest values. Select **Bottom** to select the paths with the lowest values. In the text field, enter the number of paths to select.

By default, path ranking selects the top 200 paths.

Note: If your ranking method reduces the number of paths that are displayed, then a message appears at the bottom of the diagram.

Compress
combines repeated, consecutive events in each path into a single event. If you specified a measure for the **Weight** role, then the weight value for a compressed event is the average of the weight values for the repeated, consecutive events that were combined.

Note: For all of the options under **Path Analytics**, click ✓ to apply your changes.

Managing the Path Selection for a Sankey Diagram

About Path Selections

You can subset the paths in a Sankey diagram by creating a path selection. A path selection uses a set of conditions to select or exclude paths based on the nodes or events in each path.

Create a New Condition By Using the Add New Condition Window

To create a new condition by using the Add New Condition window, follow these steps:

1 In the visualization toolbar, click the drop-down list, and then select **Add New Condition**. The Add New Condition window appears.

2 From the **Condition** drop-down list, select the condition type.

3 From the **Events** drop-down list, select the events for the condition.

4 From the **Type** drop-down list, select one of the following:

Include Only
: specifies that the path selection includes only paths that contain one or more of the selected events.

Exclude
: specifies that the path selection excludes all paths that contain one or more of the selected events.

5 Click **OK** to apply the new condition to your path selection.

TIP You can create a condition for the path selection when only the **Event** role is assigned to the visualization. For large data sources, you might want to create your path selection before assigning all of the data roles to avoid long loading times.

Create a New Condition from Selected Nodes

To add a new condition from selected nodes in the diagram, follow these steps:

1 Select one or more nodes in the diagram.

Note: To select multiple nodes, press the Ctrl key.

2 Click the drop-down list, and then select either **Include Only** or **Exclude**, and then select the condition type. The condition type can be any of the following:

Paths Containing All the Selected Nodes
: includes or excludes paths that contain all of the selected nodes. This condition type is based on specific nodes in the diagram, rather than the event values.

Paths Containing Any of the Selected Nodes
includes or excludes paths that contain any of the selected nodes. This condition type is based on specific nodes in the diagram, rather than the event values.

Paths Containing the Selected Events on Any Node
includes or excludes paths that contain any of the selected events, on any node. This condition type is based on the event values of the selected nodes.

Paths Starting With the Selected Events
includes or excludes paths that start with any of the selected events.

Paths Ending With the Selected Events
includes or excludes paths that end with any of the selected events.

The new condition appears on the **Path Selection** tab of the details table.

Edit a Condition for a Path Selection

To edit a condition for a path selection, follow these steps:

1 Open the details table for your visualization by clicking the drop-down list from the visualization toolbar, and then selecting **Show Details**.

2 In the details table, select the **Path Selection** tab.

3 In the **Type** column, select the condition type:

Include Only
includes only the paths that satisfy the condition.

Exclude
excludes the paths that satisfy the condition.

Remove Conditions from a Path Selection

To remove a condition from a path selection, perform the following steps:

1 Open the details table for your visualization by clicking the drop-down list from the visualization toolbar, and then selecting **Show Details**.

2 In the details table, select the **Path Selection** tab.

3 To remove specific conditions, select the conditions that you want to remove, right-click one selected condition, and select **Remove Selected Conditions**.

To remove all of the conditions, right-click any condition, and select **Remove All Conditions**.

Explore the Transaction Identifier Values for a Path Selection

You can explore the transaction identifier values for your path selection by creating a new visualization. To create a new visualization from your path selection, perform either of the following steps:

- On the **Path Selection** tab of the details table, right-click a condition, and then select **Create Visualization from All Conditions**. All of the conditions in your path selection are used to filter the data in your new visualization.
- In the Sankey diagram, select one or more nodes, click the drop-down list, select **Create New Visualization**, and select a condition for your path selection. For more information, see "Create a New Condition from Selected Nodes" on page 178.

  The new condition and any existing conditions are all used to filter the data in your new visualization.

The new visualization appears as a bar chart of the data item assigned to the **Transaction Identifier** role. A visualization filter selects all of the transaction identifier values that correspond to your path selection. If your Sankey diagram includes a weight measure, then the weight measure also appears in the bar chart.

Note: You can perform any standard visualization tasks with the new visualization. For example, you can add data items and change the visualization type.

Use the Spotlight Tool to View Multiple Data Tips

The spotlight tool enables you to view data tips for all of the nodes that are within a circular area around the cursor.

Click to enable or disable the spotlight tool.

Display the Overview

For large diagrams, the overview enables you to select the portions of the diagram that are visible.

To display the overview, click the drop-down list from the visualization toolbar, and then select **Show Overview**.

Zoom a Sankey Diagram

You can zoom a Sankey diagram by using either of the following controls:

- in the overview, select the part of the diagram that you want to zoom
- scroll the mouse wheel over the visualization to zoom in or zoom out at the location of the cursor

Pan (Scroll) a Sankey Diagram

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

- In the overview, drag the selection box.
- Hold down the Shift key and drag the diagram.
- When you have zoomed in on a diagram and scroll bars are displayed, reposition your zoomed view by holding down the Shift and Alt keys and dragging the diagram.

Working with Histograms

About Histograms

A histogram displays 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: If you use the default number of bins, then the minimum and maximum values on the histogram axis might not match the actual extent of your data values. If you specify the number of histogram bins, then the histogram axis matches your data values exactly.

Data Roles for a Histogram

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

Specify Properties for a Histogram

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

Name
specifies the name of the visualization.

Title
specifies the title that appears above the graph.

Note: The **Title** option is disabled if you select **Generate graph title**.

Generate graph title
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.

Frequency
specifies whether the frequency is displayed as a count (**Count**) or as a percentage (**Percent**).

Note: The frequency values are based on the data that is shown in the visualization (after filters and other data selections have been applied).

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.

Reference Lines

adds reference lines to the visualization. To create a reference line, click , and then specify the parameters for the line in the New Reference Line window.

The name of each reference line appears below the **Reference Lines** option. You can edit or delete your reference lines by using the icons next to each name.

Working with Box Plots

About Box Plots

A box plot displays the distribution of data values by using a rectangular box and lines called “whiskers.”

Figure 25.5 *Parts of a Box Plot*

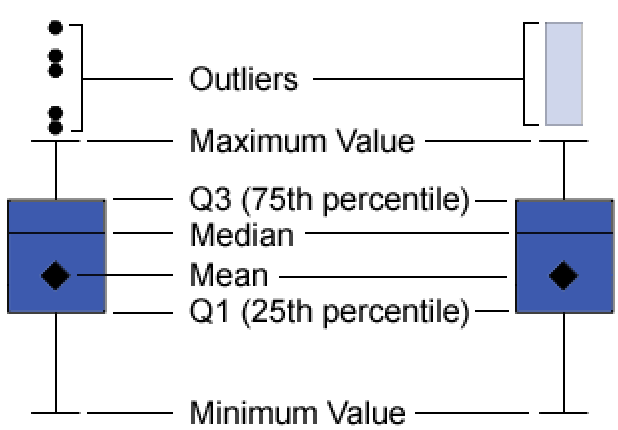


Figure 25.5 on page 182 shows a diagram of a box plot. The bottom and top edges of the box indicate the interquartile range (IQR). That is, the range of values that are between the first and third quartiles (the 25th and 75th percentiles). The marker inside the box indicates the mean value. The line inside the box indicates the median value.

You can enable outliers, which are data points whose distances from the interquartile range are greater than 1.5 times the size of the interquartile range.

The whiskers (lines protruding from the box) indicate the range of values that are outside of the interquartile range. If you do not enable outliers, then the whiskers extend to the maximum and minimum values in the plot. If you enable outliers, then the whiskers indicate the range of values that are outside of the interquartile range, but are close enough not to be considered outliers.

If there are a large number of outliers, then the range of outlier values is represented by a bar. The data tip for the bar displays additional information about the outliers. To explore the outliers, double-click on the outlier bar to view the values as a new histogram visualization.

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 category values 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 these 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.

Specify Properties for a Box Plot

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

Name
specifies the name of the visualization.

Title
specifies the title that appears above the graph.

Note: The **Title** option is disabled if you select **Generate graph title**.

Generate graph title
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.

Outliers
specifies how outliers are displayed. An outlier is a data point whose distance from the interquartile range is greater than 1.5 times the size of the interquartile range.

Select one of the following:

Ignore Outliers
excludes outliers from the plot. If you select this option, then outlier values are not represented in the plot.

Hide Outliers
includes the outliers within the whiskers. If you select this option, then outlier values are not represented differently from the other values in the plot.

Show Outliers
displays outliers separately from the whiskers. If there are a small number of outliers, then each outlier is displayed as a point. If there are a large number of outliers, then the range of outlier values is displayed as a bar.

Show averages
displays the mean value as a marker inside the box.

Rotate axis labels
displays the category labels at an angle.

Note: The **Rotate axis labels** option has no effect if the box plot contains no categories.

Reference Lines
adds reference lines to the visualization. To create a reference line, click , and then specify the parameters for the line in the New Reference Line window.

The name of each reference line appears below the **Reference Lines** option. You can edit or delete your reference lines by using the icons next to each name.

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 a cell's color represents the frequency of each intersection of values. If you assign a measure to the **Color** data role, then a cell's color represents the aggregated measure value for each intersection of values.

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** role, then the cell color indicates frequency.

Specify Properties for a Heat Map

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

Name
specifies the name of the visualization.

Title
specifies the title that appears above the graph.

Note: The **Title** option is disabled if you select **Generate graph title**.

Generate graph title
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** affects only measures.

Show borders
specifies that the borders between cells are visible.

Rotate axis labels
displays the category labels at an angle.

Note: The **Rotate axis labels** option affects only the values on the X axis.

Note: The **Rotate axis labels** option has no effect if the heat map contains no categories.

Color gradient
selects the gradient colors for the visualization.

You can click to select the values that are used to assign the colors. See “Specify a Custom Data Range” on page 154.

Fit line
adds a fit line to the heat map. For information about the fit types that are available, see “Fit Lines” on page 226.

Note: Fit lines are not available if a category is assigned to the heat map.

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 visualization. For more information, see “Correlation” on page 225.

Fit Line
plots a model of the relationship between the variables in the visualization.

There are many types of fit lines, including linear fit, quadratic fit, cubic fit, and penalized B-spline. For more information, see “Fit Lines” on page 226.

Correlation is applied to your visualization automatically when you add a linear fit line. It is not available with other fit types.

Enable Data Analysis

To add a fit line to your visualization, select the ▼ drop-down list from the visualization toolbar, and then select **Fit Line** ▶ ***fit-type***. For information about the fit types that are available, see “Fit Lines” on page 226.

Note: Fit lines are not available if your heat map contains categories or hierarchies.

Working with Geo Maps

About Geo Maps

A geo map overlays your data on a geographic map. You can display your data either as bubbles, as a scatter plot, or as colored regions on the geo map.

To display a geo map, you must define one or more of the categories as geography data items. For more information, see “Defining a Geography Data Item” on page 134.

Data Roles for a Geo Map

The basic data roles for a bubble plot are:

Map style
specifies the type of data overlay for the map. Select one of the following values:

Coordinates
displays your data as a simple scatter plot on the map. Each point is located at the center of a geographic region or at the coordinates of a location.

Bubbles
displays your data as a series of bubbles. Each bubble is located at the center of a geographic region or at the coordinates of a location.

Regions
displays your data as colored regions on the map.

Note: The **Regions** map style is not available for custom geographic roles or for ZIP codes.

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

Geography data items are identified by the 🌐 icon.

Bubble size
for the **Bubbles** map style, specifies the measure that determines the bubble size.

Color
for the **Regions** map style, specifies the measure that determines the regions’ colors.

for the **Bubbles** map style, specifies a measure that determines the bubble color.

Specify Properties for a Geo Map

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

Name
specifies the name of the visualization.

Title
specifies the title that appears above the graph.

Note: The **Title** option is disabled if you select **Generate graph title**.

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

Map service
specifies the source for the background map.

Size scale
specifies the type of scaling that is used to draw the bubbles. Select one of the following values:

Linear
specifies that the bubble sizes are scaled relative to the minimum and maximum values (or frequencies) in your data. A negative value is displayed as a smaller bubble than a positive value.

For the **Linear** scaling type, the difference in bubble sizes might not be proportional to the difference in values.

Magnitude
specifies that the bubble sizes are scaled relative to zero and the greatest absolute value in your data. A negative value is displayed as a circle with a waving edge.

For the **Magnitude** scaling type, the difference in bubble sizes is proportional to the difference in absolute values.

Note: This option is available only for the **Bubbles** map style.

Frequency
specifies whether the frequency values are displayed as the number of values (**Count**) or as the percentage of values (**Percent**).

Show map navigation control
specifies whether the map zoom and pan controls are visible.

Color gradient
selects the gradient colors for the visualization.

You can click to select the values that are used to assign the colors. See “Specify a Custom Data Range” on page 154.

Transparency
specifies the amount of transparency for the data overlay.

Zoom a Geo Map

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

- click the zoom bar to select your zoom level
- press the + and – buttons on the zoom bar
- scroll the mouse wheel to zoom in or zoom out at the location of the cursor

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

Working with Treemaps

About Treemaps

A treemap displays a hierarchy or a category as a set of rectangular tiles. Each tile represents a category value or a hierarchy node. The size of each tile represents either the frequency count or the value of a measure. If you assign a measure to the **Color** role, then the color of each tile represents the value of that measure.

Data Roles for a Treemap

The basic data roles for a treemap are:

Tile
specifies categories or a hierarchy that are used to create the tiles in the treemap. If you specify categories for the **Tile** role, the order of the categories determines the level of each category. Drag and drop the categories to place them in the order that you want.

Size
specifies a measure that determines the size of each tile. If you do not specify the **Size** role, then the tile size is determined by the frequency count.

Note: If any of the aggregated values for the **Size** role results in a negative size value or a value of zero, then an error appears.

Color
specifies a measure that determines the color of the tiles.

Specify Properties for a Treemap

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

Name
specifies the name of the visualization.

Title
specifies the title that appears above the graph.

Note: The **Title** option is disabled if you select **Generate graph title**.

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

Show data labels
shows a text label for each tile in the treemap.

Frequency
specifies whether the frequency values are displayed as the number of values (**Count**) or as the percentage of values (**Percent**).

Arrangement
specifies the layout of the tiles in the treemap. Select one of the following values:

Standard
arranges the tiles into squares, with the largest tiles generally at the bottom left.

Flow
arranges the tiles from largest to smallest, with the largest tile at the top left.

Toggle
arranges the tiles into a single row or column, with the largest tile on the left or at the top.

The orientation of the tiles alternates between hierarchy levels. The top level is arranged as a row, the second level is a column, and so on.

Additional levels
specifies the number of levels that are displayed beneath the current level.

Color gradient
selects the gradient colors for the visualization.

You can click to select the values that are used to assign the colors. See "Specify a Custom Data Range" on page 154.

Create a Hierarchy from a Treemap

If your treemap contains categories in the **Tile** role, then you can create a new hierarchy by using the categories.

On the **Roles** tab, click on the drop-down list for the **Tile** role, and then select **Create Hierarchy**. A new hierarchy is created with the name of the first category in the hierarchy.

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.

A correlation matrix can either compare within a single set of measures or it can compare between two sets of measures.

The correlation values are calculated by using Pearson's product-moment correlation coefficient. Correlation values are identified as weak, moderate, or strong as follows:

Weak
: the absolute value is 0.3 or lower

Moderate
: the absolute value is greater than 0.3 and less than or equal to 0.6

Strong
: the absolute value is greater than 0.6

Data Roles for a Correlation Matrix

The basic data role for a correlation matrix is a measure. You must assign at least two measures.

Note: The maximum number of measures is 60.

The **Show Correlations** option specifies whether the correlation matrix uses a single set of measures (**Within one set of measures**) or two sets of measures (**Between two sets of measures**.)

If you select **Between two sets of measures**, then assign measures to the **X axis** and **Y axis** roles.

Specify Properties for a Correlation Matrix

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

Name
: specifies the name of the visualization.

Title
: specifies the title that appears above the graph.

: **Note:** The **Title** option is disabled if you select **Generate graph title**.

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

Show borders
: specifies that the borders between cells are visible.

Rotate axis labels
displays the axis labels at an angle.

Color gradient
selects the gradient colors for the visualization.

Sort Correlation Values

For a correlation matrix between two sets of measures, you can sort by the correlation values for a measure.

To apply sorting, right-click the **Measures** button on the axis that you want to sort, and then select **Sort**. You can select any of the measures on the selected axis, and you can select whether to sort the correlation values in ascending or descending order.

Explore the Data for a Cell

For any cell (crossing) in the correlation matrix, you can explore the measures for that cell as a heat map.

To explore the measures, right-click the cell that you want to explore, and then select **Explore *measure* by *measure***.

A new heat map visualization displays the two measures from your selected cell.

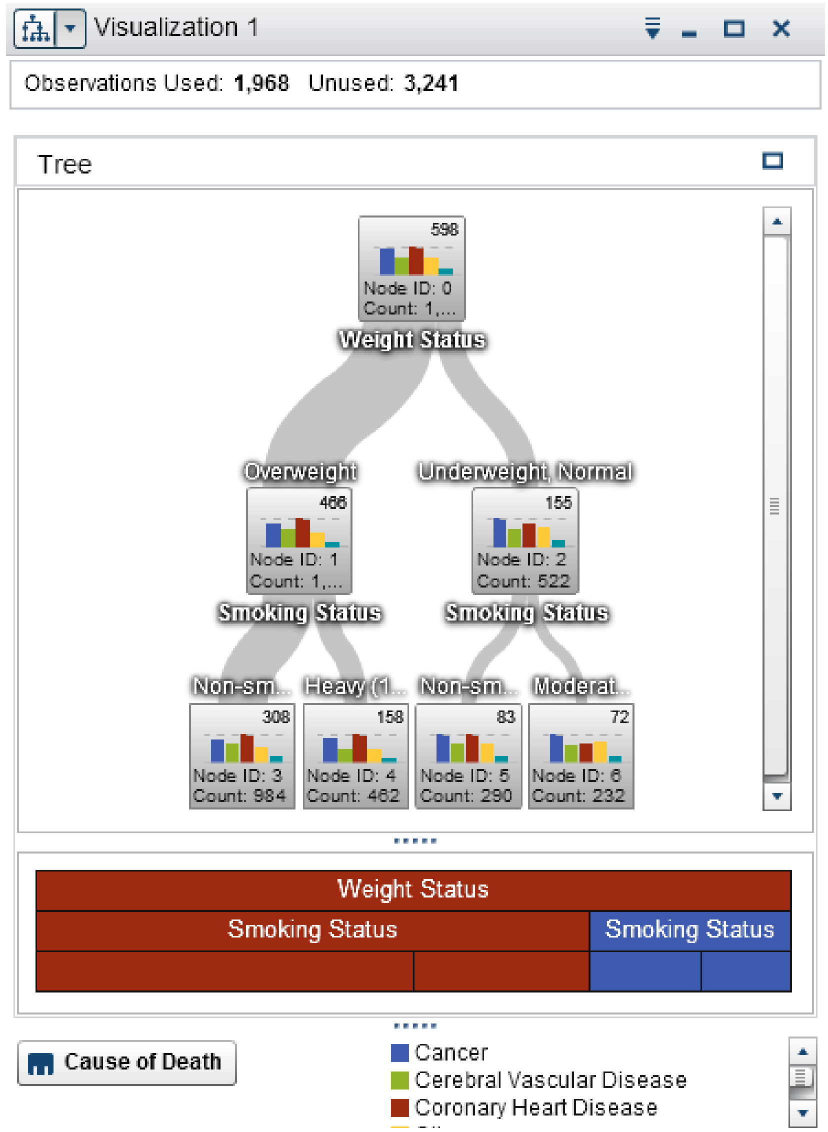
Working with Decision Trees

About Decision Trees

Note: If SAS Visual Statistics is licensed at your site, then the decision tree contains advanced features. See Chapter 37, “Decision Trees,” on page 281.

A decision tree uses the values of one or more predictor data items to predict the values of a response data item. A decision tree displays a series of nodes as a tree, where the top node is the response data item, and each branch of the tree represents a split in the values of a predictor data item. Decision trees are also known as classification and regression trees.

Figure 25.6 Example Decision Tree



Each branch of the tree displays the name of the predictor for the branch at the top of the split. The thickness of the branch indicates the number of values that are associated with each node. The predictor values for each node are displayed above the node.

Each node in the tree displays the data for the node either as a histogram (if the response contains continuous data) or as a bar chart (if the response contains discrete data). The histogram or bar chart in each node displays the values of the response data item that are selected by the splits in the tree. The number at the top right of the node indicates the greatest value or frequency for the bar chart or histogram. At the bottom of each node, the total number of data values (count) for the node is displayed.

Below the decision tree, an icicle plot of the nodes is displayed. The color of the node in the icicle plot indicates the predicted level for that node. When you select a node in either the decision tree or the icicle plot, the corresponding node is selected in the other location.

Decision trees in SAS Visual Analytics use a modified version of the C4.5 algorithm.

The details table for a decision tree contains two additional data columns, Node ID and Parent ID. Node ID specifies a unique value for each node in the tree. Parent ID specifies the ID of the parent node.

Data Roles for a Decision Tree

The basic data roles for a decision tree are:

Response
specifies the response for the decision tree. You can specify any category or measure. The decision tree attempts to predict the values of the response data item. The bar chart or histogram inside each node of the tree displays the frequency of values for the response data item.

Predictors
specifies predictors for the decision tree. You can specify one or more categories or measures as predictors. The values of predictor data items are displayed above the nodes in the tree. The order of the data items in the **Predictors** list does not affect the tree.

Note: If a predictor does not contribute to the predictive accuracy of the tree or the contribution has been pruned, then the predictor is not included in the final tree that is displayed.

Specify Properties for a Decision Tree

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

Name
specifies the name of the visualization.

Include missing
specifies whether missing values are included in the tree.

Frequency
specifies whether the frequency value for each node is displayed as a count (**Count**) or as a percentage (**Percent**).

Note: The frequency values are based on the data that is shown in the visualization (after filters and other data selections have been applied).

Growth strategy
specifies the parameters that are used to create the decision tree. Select one of the following values:

Basic
specifies a simple tree with a maximum of two branches per split and a maximum of six levels. For details, see Table 25.4 on page 194.

Advanced
specifies a complex tree with a maximum of four branches per split and a maximum of six levels. For details, see Table 25.4 on page 194.

Custom
enables you to select the values for each of the parameters.

If you select **Custom** as the value for **Growth strategy**, then the following additional options appear:

Maximum branches
specifies the maximum number of branches for each node split.

Maximum levels
specifies the maximum number of levels in the tree.

Leaf size
specifies the minimum number of values (count) for each node.

Response bins
specifies the number of bins that are used for the response data item.

Note: This option has no effect if the response data item contains discrete data.

Predictor bins
specifies the number of bins that are used for predictor data items.

Note: This option has no effect if the predictor data items contain discrete data.

Rapid growth
enables you to use the information gain ratio and k-means fast search methods for decision tree growth. When disabled, the information gain and greedy search methods are used, which generally produce a larger tree and require more time to create.

Pruning
specifies the level of pruning that is applied to the tree. Pruning removes leaves and branches that contribute the least to the predictive accuracy of the tree. A more **Lenient** pruning value specifies that fewer leaves and branches are removed from the tree. A more **Aggressive** pruning value specifies that more leaves and branches are removed from the tree.

Reuse predictors
specifies that predictors can be used more than once in the tree.

The following parameter values are used for the **Basic** and **Advanced** growth strategies:

Table 25.4 Parameter Values for the Basic and Advanced Growth Strategies

Property	Basic Value	Advanced Value
Maximum branches	2	4
Maximum levels	6	6
Leaf size	1	1
Response bins	10	10
Predictor bins	2	10
Reuse predictors	No	Yes

Explore a Node as a New Visualization

For each node in the tree, you can explore the data values as a new histogram or bar chart visualization. To create a new visualization for a node, right-click on the node, and then select **Create Visualization from Node**. A new visualization appears.

Note: If you create a visualization from a node that displays a histogram, then the new histogram might have visual differences from the histogram that is displayed in the decision tree. For example, the bins and the extent of data on the X axis might be different. However, the two histograms use exactly the same data.

Derive a Leaf ID Data Item from a Decision Tree

You can derive a leaf ID data item to represent the results of a decision tree. The leaf ID data item creates values that correspond to the node IDs in the details table for the decision tree.

You can use the leaf ID data item in a filter to select the values for a decision tree node in other types of visualizations.

To calculate a leaf ID data item from a decision tree:

1 Select the drop-down list from the visualization toolbar, and then select **Derive a Leaf ID Variable**.

2 In the New Calculated Item window, enter a **Name** for the new calculated item.

3 Click **OK** to create the new data item.

Display the Overview

For large decision trees, the overview enables you to select the portions of the tree that are visible.

To display the overview, right-click an empty part of the decision tree, and then select **Show Overview**.

Note: Each leaf node in the overview displays a single bar for the greatest value in that node. To display a bar for each node in the overview, right-click an empty part of the decision tree, and then select **Show Branch Coloring**.

Zoom a Decision Tree

You can zoom a decision tree by using either of the following controls:

- in the overview, select the part of the decision tree that you want to view
- scroll the mouse wheel over the visualization to zoom in or out at the location of the cursor

Note: If you zoom out on the decision tree, then each leaf node displays a single bar for the greatest value in that node. To display a bar for each node in the tree, right-click an empty part of the decision tree, and then select **Show Branch Coloring**.

When you have zoomed in on a decision tree and scroll bars are displayed, you can reposition the decision tree by holding down the Shift and Alt keys and dragging it.

Pan (Scroll) a Decision Tree

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

- in the overview, drag the selection box
- hold down the Shift key and drag the decision tree

Working with Word Clouds

About Word Clouds

A word cloud displays a set of words from a character data item. Depending on the type of word cloud and your data roles, the size of each word in the cloud can indicate the importance (topic term weight) of the word, the frequency of the word, or the value of a measure.

You can create two types of word cloud:

Word clouds that use text analytics

Word clouds that use text analytics analyze each value in a document collection data item as a text document that can contain multiple words. Words that often appear together in the document collection are identified as topics. For the selected topic, the word cloud displays the terms with the greatest topic term weight values. The topic term weight indicates the importance of the term within the topic.

A word cloud that uses text analytics can also display whether the documents in a topic express positive, negative, or neutral sentiment.

The details table for a text analytics word cloud contains additional information about the terms, topics, and documents in the word cloud. For more information, see "Explore Text Analytics Results" on page 200.

To enable text analytics, you must set a unique row identifier and define one or more categories as document collections. See "Define Data Items for Text Analytics" on page 136.

Note: Text analytics can be applied only to English or German text.

Note: Depending on the number of rows in your data source and the length of the values in your document collection, a word cloud with text analytics might require a significant amount of time to display.

Note: Text analytics in SAS Visual Analytics uses a different algorithm from SAS Text Miner. Your results might be different from the results that SAS Text Miner produces.

Word clouds that use category values

Word clouds that use category values analyze each value in a category data item as a single text string. The word cloud can display either the string values that have the highest frequency or the string values that have the greatest value for a measure. The color of each word can indicate the value of a measure.

Note: If you view the word cloud as an automatic chart, then any changes to the **Roles** tab might cause the visualization to reset. It is recommended that you view it as a word cloud.

Data Roles for a Word Cloud

About Data Roles for a Word Cloud

The data roles for a word cloud are dependent on the type of word cloud that you select.

The **Show Word Cloud** option selects whether the word cloud is generated by using text analytics or by using category values.

Data Roles for a Word Cloud That Uses Text Analytics

For a word cloud that uses text analytics, the basic role is a **Document collection**. A document collection is a category data item that contains the words that you will analyze.

Note: To enable text analytics, you must set a unique row identifier and define one or more categories as document collections. See "Define Data Items for Text Analytics" on page 136.

In addition to the basic role, you can specify the following role:

Document details
specifies data items that are displayed as columns in the **Documents** tab of the details table.

Data Roles for a Word Cloud That Uses Category Values

For a word cloud that uses category values, the basic role is **Words**. Specify a category whose values are used in the word cloud.

In addition to the basic role, you can specify these roles:

Size
specifies a measure that determines the size of each word. If you do not specify a measure, then the word size indicates the frequency of each word.

Color
specifies a measure that determines the color of each word.

Specify Properties for a Word Cloud

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

Name
specifies the name of the visualization.

Title
specifies the title that appears above the graph.

Note: The **Title** option is disabled if you select **Generate graph title**.

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

Frequency (for category values only)
specifies whether the frequency is displayed as a count (**Count**) or as a percentage (**Percent**).

Note: The frequency values are based on the data that is shown in the visualization (after filters and other data selections have been applied).

Note: This option has no effect if a measure is assigned to the **Size** role.

Word display limit
specifies the maximum number of words that are displayed in the word cloud.

Font scale
specifies the amount of difference in font sizes between the largest and smallest words in the cloud. The number value specifies the ratio in points of the largest font size to the smallest font size.

For word clouds that use category values, you can specify the following addition option:

Color gradient
selects the gradient colors for the visualization.

You can click to select the values that are used to assign the colors. See "Specify a Custom Data Range" on page 154.

For word clouds that use text analytics, you can specify the following additional basic options:

Analyze document sentiment
enables sentiment analysis for the word cloud.

Sentiment analysis determines whether a document has a positive sentiment, negative sentiment, or neutral sentiment based on the content of the document.

When sentiment analysis is enabled, the number of positive, neutral, and negative documents in the topic is displayed at the top of the word cloud. In addition, sentiment values are displayed on the **Topics** and **Documents** tabs of the details table.

Identify term roles
identifies terms by their parts of speech. In addition, this option identifies groups of nouns as single terms and identifies text entities such as names, addresses, telephone numbers, and so on.

Note: This option is equivalent to the advanced options **Include parts of speech**, **Extract noun groups**, and **Use entity extraction**.

Maximum topics
specifies the maximum number of topics to create. Specify a number from 4 to 20.

For word clouds that use text analytics, you can specify the following additional advanced options:

Analyze document sentiment
enables sentiment analysis for the word cloud.

Sentiment analysis determines whether a document has a positive sentiment, negative sentiment, or neutral sentiment based on the content of the document.

When sentiment analysis is enabled, the number of positive, neutral, and negative documents in the topic is displayed at the top of the word cloud. In addition, sentiment values are displayed on the **Topics** and **Documents** tabs of the details table.

Maximum topics
specifies the maximum number of topics to create. Specify a number from 4 to 20.

Resolution
specifies the resolution that is used to identify topics. A **Low** resolution identifies fewer topics. A **High** resolution identifies more topics.

Cell weight
specifies whether to weight the frequency of each term for every document that it appears in. Selecting **Logarithmic** de-emphasizes terms that appear many times in relatively few documents.

Term weight
specifies a weighting algorithm for the terms in the document collection. The **Entropy** weighting algorithm emphasizes terms that have a low frequency across the document collection.

Document threshold
specifies the minimum number of documents that a term must appear in. Specify a number from 1 to 20. If a term does not appear in the minimum number of documents, then it is not included in the word cloud.

Topic label length
specifies the number of terms that are included in a topic name. Specify a number from 2 to 8. This property does not affect the number of terms that are used to select topics; only the topic names are changed.

Include parts of speech
specifies that terms are classified by parts of speech (for example, a noun, a verb, or an adjective.) The part of speech for each term is displayed in the data tip for the term.

Extract noun groups
specifies whether to identify groups of nouns as terms.

Use entity extraction
specifies whether to identify text entities such as names, addresses, telephone numbers, and so on. If this option is disabled, then text entities are not treated differently from other text.

Stem words
specifies whether all forms of a given word are identified as a single term. For example, if you select **Stem words**, then the words “sell,” “sells,” “selling,” and “sold” are identified as a single term “sell.”

Use stop list (if available)
specifies whether to use a stop list to exclude common words such as “the,” “with,” and “is” when identifying terms. If no stop list is available, then a message appears at the bottom of the word cloud.

Stop list
specifies the stop list that is used, if the **Use stop list** option is enabled.

Explore Text Analytics Results

For a word cloud visualization that uses text analytics, a large amount of additional information is available in the details table. To display the details table, click the drop-down list from the visualization toolbar, and then select **Show Details**.

The details table for a text analytics word cloud contains the following tabs:

Results
: displays all of the terms in the current topic. For each term, the **Topic Term Weight** value indicates the importance of the term in the current topic.

If the **Identify term roles** property or the **Include parts of speech** property is enabled, then the **Role** value identifies the grammatical role of each term.

Note: You can sort any column by clicking the column heading.

Topics
: displays all of the topics in the document collection. If sentiment analysis is enabled, then the number of positive, neutral, and negative documents for each topic is displayed.

Note: You can sort any column by clicking the column heading.

Documents
: displays each of the documents that contains the selected term. For each document, the **Relevance** value indicates how relevant the document is to the current topic.

To view the full text for a document, right-click the document, and then select **View Full Document**.

If sentiment analysis is enabled, then the **Sentiment** value identifies how positive or negative the document is. You can filter the documents to exclude documents with positive, negative, or neutral sentiment.

Note: You can sort any numeric column by clicking the column heading.

Analysis
: provides definitions of the key concepts for text analytics.

Explore Selected Documents as a New Visualization

You can explore a set of selected documents as a new table visualization. To create a new visualization from your selected documents, follow these steps:

1. Select the topic and the term that you want to explore.
2. On the **Documents** tab in the details table, select the documents that you want to explore in a new visualization. To select all of the documents, right-click any document, and then select **Select All**.
3. Right-click any document, and then select **Create Visualization from Selected Documents**.

A new table visualization appears with your selected document values.

26

Working with Filters

About Filters in SAS Visual Analytics Explorer

In SAS Visual Analytics Explorer (the explorer), you can create filters to subset your data.

Three types of filters are available:

Basic filters
: subset the data for your visualizations by using a single data item. Basic filters can be modified by using visual filter controls or by using the Edit Filter window.

  For more information, see "Working with Basic Filters" on page 204.

Advanced filters
: subset the data for your visualizations by using any number of data items. Visual filter controls are not available for advanced filters.

  For more information, see "Working with Advanced Filters" on page 209.

Data source filters
: subset a data source for the entire exploration. Data source filters are applied before all other data processing, so statistics like cardinality values and the percentage of total rows that is displayed in the right pane are based on the filtered data. Data source filters can use any number of data items to select data.

  For information about data source filters, see "Working with Data Source Filters" on page 210.

All of your filters are saved when you save your exploration.

Managing Your Filters

Set the Scope of a Filter (Global or Local)

A filter can be either local (applies only to the current visualization) or global (applies to all of the visualizations in the exploration that use the selected data source).

To set the scope of a filter, click the drop-down list for the filter on the **Filters** tab, and then select **Filter on *data-source*** to make it a global filter or **Filter on Visualization** to make it a local filter.

Note: If you change a global filter's scope to local, then a local filter is created for all of the visualizations in the exploration that use the selected data source.

Note: If local filters exist in your other visualizations for the same data item, then changing a local filter to a global filter replaces those local filters. A confirmation window enables you to continue or cancel.

Collapse or Expand a Filter

To collapse (minimize) or expand (restore) a filter, click the filter name on the **Filters** tab.

To collapse all of the filters in an area, click the drop-down list in the global filters area or in the local filters area on the **Filters** tab, and then select **Collapse All Filters**.

To expand all of the filters in an area, click the drop-down list in the global filters area or in the local filters area on the **Filters** tab, and then select **Expand All Filters**.

Arrange Filters

To arrange the filters in an area of the **Filters** tab, click the drop-down list in the global filters area or in the local filters area and then select **Arrange Filters**.

In the Arrange Filters window, you can change the order of your filters.

Delete a Filter

To delete a filter, click beside the filter on the **Filters** tab, or click the drop-down list for the specific filter, and then select **Delete Filter**.

You can delete all of the filters in an area by clicking the drop-down list in the global filters area or in the local filters area on the **Filters** tab, and then selecting **Delete All Filters**.

Note: If you delete a global filter, then the filter is removed from all of the visualizations that use the selected data source.

Resize the Global Filters and Local Filters Areas

To resize the global filters and local filters areas of the **Filters** tab, drag the resizing tab between the two areas.

Copy a Local Filter to Another Visualization

For a local filter, you can copy your filter to any visualization that uses the same data source.

To copy your filter, click the drop-down list for the filter on the **Filters** tab, and then select **Copy Filter to** ▸ ***visualization-name***.

Copy All Local Filters from the Current Visualization to Another Visualization

For local filters, you can copy all of your filters to any visualization that uses the same data source.

To copy your filters, click the drop-down list for the local filters (**Visualization**) area of the **Filters** tab, and then select **Copy All Filters to** ▸ ***visualization-name***.

Copy a Local Filter to New Visualization

For a local filter, you can copy your filter to new blank visualization.

To copy your filter, click the drop-down list for the filter on the **Filters** tab, and then select **Copy Filter to** ▸ **New Visualization** to copy your filter to a new blank visualization.

Link a Global Filter to Another Data Source

You can create a single filter that affects multiple data sources by linking a global filter from one data source to another.

To create a linked global filter, follow these steps:

1 Create a global, basic filter or select an existing global, basic filter. See "Create a Basic Filter" on page 205.

   Note: You cannot link an advanced filter to another data source.

2 Click the drop-down list for the filter on the **Filters** tab, and then select **Link Filter to** ▸ ***data-source***. The Link Filters window appears.

   Note: If the target data source has a different locale from your current data source, then a message appears. A linked filter between data sources that use different encodings might cause query errors.

3 Specify the **New filter name**, and select the **Target data item**.

   The values for the original data item and the target data item are displayed. For discrete data items, an asterisk marks any values that are identical between the two data items.

4 Click **OK** to create the linked filter. The linked filter appears in the global filters area of each visualization that uses any of the linked data sources.

Note: You can link a linked filter to additional data sources to filter all of the linked data sources at once.

Working with Basic Filters

About Basic Filters

For all visualization types, you can subset your data by using the **Filters** tab. 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 for the selected data source. In the **Filters** tab, the local filters area is labeled **Visualization** and the global filters area is labeled with the name of the selected data source for the current visualization.For more information, see “Set the Scope of a Filter (Global or Local)” on page 202.

The total percentage of values that are selected by your filters is displayed at the bottom of the **Filters** tab. The tooltip for the percentage displays the exact number of data rows. The percentage is calculated after data source filters are applied.

Create a Basic Filter

To create a basic filter:

1. From the **Data** pane, 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 on the data item, and select **Add as Filter on Visualization**, or drag and drop the data item onto the local filters area of the **Filters** tab.

   For a global filter, either right-click on the data item, and select **Add as Filter on *data-source***, or drag and drop the data item onto the global filters area of the **Filters** tab.

   Note: If a local filter exists for a data item in any of your visualizations, then you cannot create a new global filter for that data item. Remove the local filter, or change its scope to global. Similarly, if a global filter exists for a data item, then you cannot create a local filter for that data item.

3. Set the parameters for your filter. You can use either visual filter controls or the Edit Filter window .

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

Create a Basic Filter from a Data Selection

When you select data values in a visualization, you can use the selected values to create a basic filter.

To create a new filter from a data selection:

1. Select one or more data values in a visualization.

2. Right-click on the visualization, and then select one of the following:

   Include Only Selection
   creates a filter that includes the selected values only.

   Exclude Selection
   creates a filter that excludes the selected values.

   The new filter appears on the **Filters** tab.

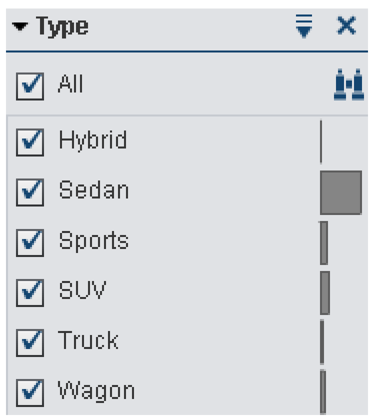
Edit a Filter By Using Visual Filter Controls

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

Figure 26.1 A Visual Filter for a Discrete Data Item



Note: If your data contains a large number of discrete values, then the top 1,000 most frequent values are displayed.

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.

Use the check boxes next to each value to select the values for the filter. You can search for a value in the displayed list by clicking . Enter a text string in the field. The search locates values from the displayed list that begin with your text string.

Select **All** to select or deselect all of the values. If the filter displays the top 1,000 most frequent values, then select **Include unlisted values** to select the unlisted values.

To invert your selection, click the drop-down list, and then select **Invert Selection**.

You can sort the values by frequency or alphabetically by value. Click the drop-down list, and then select **Sort by Frequency** or **Sort by Values**.

To display the selected values at the top of the list, click the drop-down list, and then select **Show Selected at Top**.

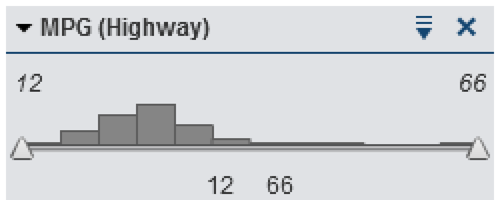
If your data item contains more than 1,000 discrete values, then you can exclude the unlisted values by deselecting **Include unlisted values**.

If your data item contains missing values, then you can exclude missing values by deselecting **Include missing values**.

By default, your filter changes are applied automatically to the active visualization. To apply multiple changes together, deselect **Auto-Update** at the bottom of the **Filters** tab, and then click **Update** when you are ready to apply your filter changes.

Filtering Continuous Data

Figure 26.2 A Visual Filter for a Continuous Data Item



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 data that is selected by the filter.

The visual filter can select data by using different filter conditions. By default, the **Between (Inclusive)** condition is used. To change the condition for your filter, click the drop-down list, select **Condition Type**, and then select the condition. Select one of the following conditions:

Between (Exclusive)
specifies that a matching value must be in the range between the lower bound and the upper bound. Matching values do not include the bounding values.

Between (Inclusive)
specifies that a matching value must be in the range between the lower bound and the upper bound. Matching values can include the bounding values.

Equal To
specifies that a matching value must be equal to the value specified by the filter.

Greater Than
specifies that a matching value must be greater than the value specified by the filter.

Greater Than or Equal To
specifies that a matching value must be greater than or equal to the value specified by the filter.

Less Than
specifies that a matching value must be less than the value specified by the filter.

Less Than or Equal To
specifies that a matching value must be less than or equal to the value specified by the filter.

Not Between (Exclusive)
specifies that a matching value must be outside the range between the lower bound and the upper bound. Matching values do not include the bounding values.

Not Between (Inclusive)
specifies that a matching value must be outside the range between the lower bound and the upper bound. Matching values can include the bounding values.

Not Equal To
specifies that a matching value must not be equal to the value specified by the filter.

You can specify the values for the filter by using sliders or by entering the values explicitly:

- For numeric data, click on a value below the histogram, and then enter a new value in the text field.
- For a continuous datetime data item, click , and then select a date or time.

If your data item contains missing values, then you can exclude the missing values by deselecting **Include missing values**.

To invert your selection, click 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-Update** at the bottom of the **Filters** tab, and then click **Update** when you are ready to apply your filter changes.

Edit a Basic Filter By Using the Edit Filter Window

To edit a basic filter by using the Edit Filter window:

1 Select the **Filters** tab.

2 Click the drop-down list for the filter that you want to edit, and then select **Edit Filter**. The Edit Filter window appears.

3 Edit the expression for the filter. For more information, see "Editing a Filter Expression" on page 211.

Reset a Basic Filter

To restore a basic filter to its initial state, click the drop-down list for the filter, and then select **Reset Filter**.

You can reset all of your filters by clicking the drop-down list in the global filters area or the local filters area of the **Filters** tab, and then selecting **Reset All Filters**.

Working with Advanced Filters

About Advanced Filters

For all visualization types, you can create advanced filters to subset your data by using the **Filters** tab.

Advanced filters enable you to create filters that use more than one data item.

Your advanced filters can be either *local filters* that apply only to the current visualization, or *global filters* that apply to all of the visualizations that use the selected data source.

The total percentage of values that are selected by your filters is displayed at the bottom of the **Filters** tab. The tooltip for the percentage displays the exact number of data rows. The percentage is calculated after data source filters are applied.

Create an Advanced Filter

To create an advanced filter:

1. Click the drop-down list in the global filters area or in the local filters area of the **Filters** tab, and then select **New Advanced Filter**. The Edit Filter window appears.

2. Specify a **Filter name**. The filter name identifies the advanced filter on the **Filters** tab.

3. Build the expression for your filter by dragging and dropping data items, global parameters, conditions, and operators onto the expression.

   For information about the conditions and operators that are available, see Appendix 6, "Conditions for Filters," on page 607 and Appendix 5, "Operators for Data Expressions," on page 583 .

   You can also build your expression as text by using the **Text** tab. See "Editing a Data Expression in Text Mode" on page 577.

4. Click **OK** to apply the filter.

Edit an Advanced Filter

To edit an advanced filter:

1. Select the **Filters** tab.

2. Click the drop-down list for the filter that you want to edit, and then select **Edit Filter**. The Edit Filter window appears.

3. Edit the expression for the filter. For more information, see "Editing a Filter Expression" on page 211.

Working with Data Source Filters

About Data Source Filters

A data source filter enables you to subset a data source for an entire exploration. A data source filter differs from a global advanced filter in the following ways:

- A data source filter updates the cardinality values that appear in the **Data** pane.
- A data source filter updates the total number of rows that is used in the filtered rows percentage at the bottom of the right pane.
- A data source filter is not shown on the **Filters** tab.
- An exploration can contain only one data source filter for each data source.

Create a Data Source Filter from a Global Filter

To convert an existing global filter to a data source filter:

1. On the **Filters** tab, select the global filter that you want to convert.
2. Click the drop-down list, and select **Convert to Data Source Filter**. The global filter disappears from the **Filters** tab.

Note: If your exploration already contains a data source filter for the selected data source, then the global filter is appended to the filter expression using the AND operator.

You can convert all of your global filters together into a single data source filter by clicking the drop-down list in the global filters area of the **Filters** tab, and then selecting **Convert All to Data Source Filter**. The expressions from the global filters are joined with AND operators.

To view the data source filter settings, select **Data** ▸ **Data Source Details** from the main menu. The data source filter settings are displayed in the **Data source filter** field.

Create a New Data Source Filter

To create a data source filter:

1. In the **Data** pane, select the data source that you want to filter.
2. Click and then select **New Data Source Filter**.

   Note: If the current exploration contains an existing data source filter, then select **Edit Data Source Filter** to add conditions to the existing data source filter.

   The Edit Filter window appears.

3 Build the expression for your filter by dragging and dropping data items, global parameters, conditions, and operators onto the expression.

For information about the conditions and operators that are available, see Appendix 6, “Conditions for Filters,” on page 607 and Appendix 5, “Operators for Data Expressions,” on page 583 .

You can also build your expression as text by using the **Text** tab. See “Editing a Data Expression in Text Mode” on page 577.

4 Click **OK** to apply the filter.

Edit a Data Source Filter

To edit a data source filter:

1 In the **Data** pane, select the data source that you want to filter.

2 Click and then select **Edit Data Source Filter**.

The Edit Filter window appears.

3 Edit the expression for the filter. For more information, see “Editing a Filter Expression” on page 211.

Delete a Data Source Filter

To delete a data source filter:

1 In the **Data** pane, select the data source for which you want to remove the filter.

2 Click and then select **Delete Data Source Filter**.

Editing a Filter Expression

Specify the Filter Name

To specify a name for the filter, enter a name in the **Filter name** field. The filter name identifies the advanced filter on the **Filters** tab.

Note: You cannot specify the name for a basic filter or a data source filter.

Edit a Parameter Value

To edit a parameter value for a condition or operator, select the parameter, and enter a new value. Or, right-click the parameter field, and select **Replace with** to select a data item or a global parameter.

Add a Condition

To add a new condition:

1 From the **Data items** list, select the data item on which the condition is based.

2 From the **Conditions** list, select a condition. For a list of the conditions that are available, see Appendix 6, “Conditions for Filters,” on page 607.

3 Drag and drop the condition onto the expression.

4 For any required parameters, select the parameter, and enter a value, or right-click the parameter field, and select **Replace with** to select a data item.

Add an Operator

To add an operator to the expression:

1 From the **Operators** list, select the operator that you want to add. For a list of the operators that are available, see Appendix 5, “Operators for Data Expressions,” on page 583.

2 Drag and drop the operator onto the expression.

3 For any required parameters, select the parameter, and enter a value. Or, right-click the parameter field, and select **Replace with** to select a data item.

Replace an Operator

To replace an operator, drag and drop a new operator onto the existing operator in the expression. You can also right-click an operator in your expression, select **Replace Operator with**, and then select a new operator.

Add a Global Parameter

To add an existing global parameter to your expression, drag and drop the parameter from the **Data Items** list onto your expression.

To create a new global parameter, in the expression editor, right-click the **Parameter** heading in the **Data Items** list, and then select **Create Parameter**.For more information about creating global parameters, see “Working with Global Parameters” on page 123.

Delete Part of an Expression

To delete a part of an expression, highlight the part of the expression that you want to delete, right-click, and select **Delete** or **Clear**.

Subset an Expression

To subset an expression and delete the remainder:

1 Highlight the part of the expression that you want to keep.

2 Right-click, and select **Keep Operand**. The parts of the expression that were not selected are removed.

Edit an Expression as Text

On the **Text** tab of the expression editor, you can edit an expression as text code. See “Editing a Data Expression in Text Mode” on page 577.

Display the Current Expression as Text

To display the current expression as text on the **Display Text** tab, click the drop-down list, and then select **Show display text**.

Manage the Scratch Area

The scratch area of the expression editor enables you to store elements of an expression, and then add them back into the expression as needed.

You can see the contents of the scratch area on the **Scratch** tab.

To move items to the scratch area, right-click part of the expression, and select **Move to Scratch Area**. You can also drag and drop items from your expression onto the **Scratch** tab.

To move an item from the scratch area to your expression, drag and drop the item from the **Scratch** tab onto your expression.

To delete an item in the scratch area, click next to the item.

27

Exporting Content

Export an Exploration as a Report

You can export your exploration as a report, which can be refined in the designer interface or viewed directly on a mobile device or in SAS Visual Analytics Viewer (the viewer).

To export your current exploration as a report:

1 Select **File** ▶ **Export** ▶ **Exploration as Report**.

   Note: If you have only one visualization in your exploration, the Save As window is displayed at this point. Select the location where you want to save the report, and then click **Save** to export it as a 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 the report:

   - visualizations that do not contain any data
   - decision trees
   - network diagrams
   - Sankey diagrams
   - word clouds
   - treemaps that display additional levels
   - line charts with forecasting analysis are not available if the visualization has not been displayed during the current session
   - geo maps that use calculated geography data items

Note: If your scatter plot or heat map contains correlation data, the correlations are not included in the report.

The Export to Report window enables you to perform the following tasks:

Task	Action
Add visualizations to the selection list. (Restore visualizations.)	Select a visualization from the **Available** list, and then click , or click to add all of the available visualizations.
Remove visualizations from the selection list. (Minimize visualizations.)	Select a visualization from the **Selected** list, and then click , or click to remove all of the visualizations.
Change the order of the selected visualizations.	In the **Selected** list, drag and drop the visualizations in the order that you want, or select a visualization and click the up and down arrows to move it.
Search the available visualizations.	In the **Search** field, enter a string to search for. A visualization matches the search if the string appears anywhere in the visualization name. Only the matching visualizations appear in the **Available** list.
Filter the available visualizations.	Click , and then select your filter parameters. You can filter either on the visualization type or on the data items that are used in each visualization. Only the matching visualizations appear in the **Available** list.

When you have finished selecting your visualizations, click **OK**.

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

TIP Report names are limited to 60 characters.

4 Click **OK** to return to the explorer, or select **Open the report now**, and click **OK** to view the report in the designer.

Export an Exploration as a PDF

To export your current 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 that 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: The following types of visualizations are not available to include in the PDF document:

- visualizations that do not contain any data
- crosstabs
- tables that contain more than 1,000 rows

The Export as PDF window enables you to perform the following tasks:

Task	Action
Add visualizations to the selection list. (Restore visualizations.)	Select a visualization from the **Available** list, and then click , or click to add all of the available visualizations.
Remove visualizations from the selection list. (Minimize visualizations.)	Select a visualization from the **Selected** list, and then click , or click to remove all of the visualizations.
Change the order of the selected visualizations.	In the **Selected** list, drag and drop the visualizations in the order that you want, or select a visualization and click the up and down arrows to move it.
Search the available visualizations.	In the **Search** field, enter a string to search for. A visualization matches the search if the string appears anywhere in the visualization name. Only the matching visualizations appear in the **Available** list.
Filter the available visualizations.	Click , and then select your filter parameters. You can filter either on the visualization type or on the data items that are used in each visualization. Only the matching visualizations appear in the **Available** list.

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 your current visualization as an image file:

1. Select the visualization that you want to save as an image.
2. If your visualization contains hierarchies, drill down to the hierarchy levels that you want to show in the image.
3. Select the drop-down list from the visualization toolbar, and then select **Export Image**.
4. If necessary, scroll and zoom 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.

Export Data from a Visualization

You can export the data from a visualization in comma-separated values (.csv) format. You can open .csv files in other software such as SAS Enterprise Miner, SAS Enterprise Guide, or Microsoft Excel.

Note: The explorer uses the global setting for **User locale** to export data.

To export data for a visualization:

1. Select the visualization for which you want to export data.

   Note: For crosstab visualizations, a data summary is exported instead of a full data table.

2. If your visualization contains hierarchies, drill down to the hierarchy levels that you want to export.
3. Select the drop-down list from the visualization toolbar, and then select **Export Data** to open a download window for your browser.

   Note: For a crosstab visualization, select **Export Data Summary** instead of **Export Data**.

4. Specify a filename and select the location where you want to save the file.

   Note: To export data from a table visualization, your browser must allow pop-ups.

E-mail an Exploration as a Link

To e-mail your exploration as a link:

1 Select **File** ▶ **E-mail**. The E-mail window appears.

2 Enter information in the E-mail window.

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

Here are some key points:

- The **From** address is used for notifications if the e-mail cannot be delivered.
- The **Message** field is optional. By default, it includes a link to the current exploration. If you specify text in the **Message** field, then your text appears before the link in the e-mail.

28

Managing Hierarchies

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, you can 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.

Create a New Hierarchy

To create a new hierarchy:

1 Select **Data** ▸ **New Hierarchy**. The New Hierarchy window appears.

2 In the **Name** field, 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 category that you want to remove. Then, click ⬅.

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

Note: You can create a hierarchy from within a crosstab visualization. See “Create a Hierarchy from a Crosstab” on page 160.

Derive a Hierarchy from a Date, Time, or Datetime Data Item

For date, time, and datetime data items, you can derive a hierarchy automatically. To derive a hierarchy, right-click on a date, time, or datetime data item in the **Data** pane, and then select one of the following values:

Create Date Hierarchy
creates a hierarchy with levels for year, quarter, month, and day. Depending on the format of the data item, some of the levels might not be created.

Create Date and Time Hierarchy
creates a hierarchy with levels for year, quarter, month, day, hour, minute, and second. Depending on the format of the data item, some of the levels might not be created.

Create Time Hierarchy
creates a hierarchy with levels for hour, minute, and second. Depending on the format of the data item, some of the levels might not be created.

New calculated items are created for each level in the hierarchy. By default, the new calculated items are hidden in the **Data** pane. If you delete a derived hierarchy, then a window appears. The window enables you to delete the calculated items that are associated with the hierarchy.

Create a Hierarchy from a Visualization

You can create a hierarchy from a treemap, crosstab, or network diagram visualization. See the following topics:

- “Create a Hierarchy from a Treemap” on page 189
- “Create a Hierarchy from a Crosstab” on page 160
- “Create a Hierarchy from a Network Diagram” on page 174

Edit a Hierarchy

To edit an existing hierarchy:

1 From the **Data** pane, right-click on the hierarchy that you want to edit, and then select **Edit**. The Edit Hierarchy window appears.

2 In the **Name** field, 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 click ⬇ to move the category down.

 To remove a category from the hierarchy, select the category 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** pane, right-click on the hierarchy that you want to delete, and then select **Delete**. Click **Yes** to confirm that you want to delete the hierarchy.

29

Performing Data Analysis

Overview of Data Analysis in SAS Visual Analytics Explorer

Types of Data Analysis

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

Correlation
identifies the degree of statistical relationship between measures.

Fit Line
plots a model of the relationship between measures. There are many types of fit lines, including linear fit, quadratic fit, cubic fit, and penalized B-spline.

Forecasting
estimates future values for your data based on statistical trends.

Correlation

Correlation identifies the degree of statistical relationship between measures. The strength of a correlation is described as a number between -1 and 1. A value that is close to -1 implies a strong negative correlation, a value that is close to 0 implies little or no correlation, and a value that is close to 1 implies a strong positive correlation.

To apply correlation to a visualization, add a linear fit line, or select the correlation matrix visualization type.

For a heat map or a simple scatter plot, the correlation is identified by a text label in the visualization legend. Select to view additional details about the correlation, including the exact correlation value.

For a scatter plot matrix, the correlation for each plot is identified by a colored border around the plot. The visualization legend displays a key for the color values. Select to view additional details about the correlation, including the exact correlation values for each plot.

Note: For nonlinear fit types, a scatter plot matrix displays additional plots to show each intersection of variables in two orientations. For example, if a scatter plot matrix plots the variables A, B, and C, then plots are created for both A * B and B * A when a nonlinear fit line is applied.

For a correlation matrix, the correlation for each cell is identified by the color of the cell background. The visualization legend displays a key for the color values. The data tip for each cell displays the correlation value.

Fit Lines

A fit line plots a model of the relationship between measures. You can apply fit lines to scatter plots and heat maps.

You can apply the following types of fit line to your visualization:

Best Fit
: selects the most appropriate model (linear, quadratic, or cubic) for your data. The **Best Fit** method uses backward variable selection to select the highest-order model that is significant. To see the final model that was used, select from the visualization legend.

Linear
: creates a linear fit line from a linear regression algorithm. A linear fit line produces the straight line that best represents the relationship between two measures. For more information about the linear fit line, select from the visualization legend.

  For a linear fit, correlation is automatically added to the visualization. Correlation is not available with other fit types.

Quadratic
: creates a quadratic fit line. A quadratic fit produces a line with a single curve. A quadratic fit line often produces a line with the shape of a parabola. For more information about the quadratic fit line, select from the visualization legend.

Cubic
: creates a cubic fit line. A cubic fit line produces a line with two curves. A cubic fit line often produces a line with an “S” shape. For more information about the cubic fit line, select from the visualization legend.

PSpline
: creates a penalized B-spline. A penalized B-spline is a smoothing spline that fits the data closely. A penalized B-spline can display a complex line with many changes in its curvature. For more information about the penalized B-spline, select from the visualization legend.

Forecasting

Forecasting estimates future values for your data based on statistical trends. Forecasting is available only for line charts that contain date or datetime data items.

A forecast adds a line with predicted values to your visualization and a colored band that represents the confidence interval. For example, a 95% confidence interval is the data range where the forecasting model is 95% confident what the future values will be.

The explorer automatically tests multiple forecasting models against your data, and then selects the best model. To see which forecasting model was used, select ⓘ from the visualization legend.

The forecast model can be any one of the following:

- Damped-trend exponential smoothing
- Linear exponential smoothing
- Seasonal exponential smoothing
- Simple exponential smoothing
- Winters method (additive)
- Winters method (multiplicative)

Note: Forecasting accounts for cyclical patterns by using standard intervals of time (for example, 60 minutes in an hour, 24 hours in a day, and so on). If your data uses nonstandard intervals (for example, 48 30-minute cycles per day), then cyclical patterns are not considered in the forecast.

Add a Fit Line to an Existing Visualization

To add a fit line to a scatter plot or heat map, select the drop-down list from the visualization toolbar, and then select **Fit Line** ▸ ***fit-type***. For information about the fit types that are available, see "Fit Lines" on page 226.

Add Forecasting to an Existing Visualization

To enable forecasting for a line chart, select **Forecasting** on the **Roles** tab.

Note: To apply forecasting, the line chart must contain a date or datetime item.

On the **Properties** tab for the line chart visualization, you can adjust the number of data points to predict by using the **Duration** option. You can also specify the degree of confidence for the confidence band by using the **Confidence** option.

On the **Roles** tab, you can add measures to the forecast as **Underlying factors**. The forecasting model evaluates the additional measures to determine whether they contribute to the accuracy of the forecast. If the additional measures do not increase the accuracy of the forecast, then they are not used. If the additional measures do increase the accuracy of the forecast, then the forecast line is adjusted, and the confidence bands are narrowed.

Note: Underlying factors are available only if your line chart contains a single measure in the **Measures** role.

If your forecast includes underlying factors, then you can apply scenario analysis and goal seeking to the forecast. For more information, see "Working with Scenario Analysis and Goal Seeking" on page 228.

Working with Scenario Analysis and Goal Seeking

About Scenario Analysis and Goal Seeking

Scenario analysis enables you to forecast hypothetical scenarios by specifying the future values for one or more underlying factors that contribute to the forecast.

For example, if you forecast the profit of a company, and material cost is an underlying factor, then you might use scenario analysis to determine how the forecasted profit would change if the material cost increased by 10%.

In addition to scenario analysis, you can perform goal seeking. Goal seeking enables you to specify a target value for your forecast measure, and then determine the values of your underlying factors that would be required to achieve the target value.

For example, if you forecast the profit of a company, and material cost is an underlying factor, then you might use goal seeking to determine what value for material cost would be required to achieve a 10% increase in profit.

Scenario analysis and goal seeking can be used together in the same forecast.

Apply Scenario Analysis and Goal Seeking to a Forecast

To apply scenario analysis and goal seeking:

1. On the **Roles** tab, select **Scenario Analysis**. The Scenario Analysis window appears.

   Note: The **Scenario Analysis** option is available only if one or more underlying factors contribute to the forecast.

2. For the forecast measure and for each of the measures that are assigned to the forecast as underlying factors, a line displays the values of the measure, and a series of points enables you to set the future values for the measure.

   Note: You can also display the forecast values as a table. Click  to display the table mode.

   To perform scenario analysis, set the future values for the underlying factors.

   To perform goal seeking, set the target future values for the forecast measure.

You can set the values by using any of the following methods:

- drag each data point upward or downward on the line chart. To select the accuracy of the points, click the measure heading on the plot axis, and select **Snap Interval**.
- enter the value for each data point. Right-click on a data point, and select **Set Point Value**.
- set all of the values for the measure. Click the measure heading on the plot axis, and select **Set Series Values**. The Change Future Values window enables you to set all of the values to a specific value or to adjust the future values relative to the forecasted values of the measure.

  Note: The **Progressively by** option increments future values by the amount that you specify. For example, if you specify 100, then the first future value is increased by 100, the second future value is increased by 200, the third future value is increased by 300, and so on.

Note: If you change future values for either scenario analysis or goal seeking, you must apply your changes before you can change future values for a different analysis type.

To reset any data point to its original value, right-click on the data point, and then select **Reset Point**.

To reset all of the data points for a measure, click the measure heading on the plot axis, and then select **Reset Entire Series**.

3 If you perform goal seeking and there are multiple underlying factors, then choose which of the underlying factors are optimized to achieve the target value.

4 (Optional) If you perform goal seeking, you can set bounds for each underlying factor to limit the minimum and maximum values that are possible.

To set the minimum value for an underlying factor, click the measure heading, and then select **Bounds** ▸ **Add Lower Bound**.

To set the maximum value for an underlying factor, click the measure heading and then select **Bounds** ▸ **Add Upper Bound**.

5 When you are finished setting the future values for the scenario, click **Apply** to apply the scenario to the forecast.

The forecast is updated to show the results of the scenario. The original forecast is displayed as an additional line that is labeled **(Baseline)**.

Remove Scenario Analysis and Goal Seeking from Your Forecast

To remove the scenario analysis and goal seeking from a forecast, follow these steps:

1. On the **Roles** tab, select **Scenario Analysis**. The Scenario Analysis window appears.
2. For each measure, click the measure heading, and then select **Reset Entire Series**.
3. Click **Apply** to apply your changes.

30

Sharing Comments in the Explorer

Sharing Comments in the Explorer

About Comments in the Explorer

Comments enable you to share feedback with other users. You can create comments for visualizations and for explorations. Exploration comments can be viewed both in the explorer and on the SAS Visual Analytics home page.

View Comments

To display the comments for the current exploration, select **File** ▸ **Exploration Comments** from the main menu.

To display comments for a visualization, select the visualization and then select the **Comments** tab from the right pane.

Comments are grouped by topic. If there are many comments in a topic, then only the most recent comments are displayed. Click **Show All Comments** to view all of the comments for a topic.

To search the comments, enter a search term in the **Search within comments** field.

Create a Comment under an Existing Topic

To create a comment under an existing topic:

1 Enter your comment text in the **Respond to topic** field that is beneath the topic that you want to comment on.

2 (Optional) Click 📎 to attach a file to the comment.

3 When you are finished with the comment, click **Post**. Your comment is saved and shared immediately.

Create a Comment under a New Topic

To create a comment under a new topic:

1 Enter the name of the topic in the **Enter a topic name** field.

2 Enter your comment text in the **Enter a comment** field.

3 (Optional) Click to attach a file to the comment.

4 When you are finished with the comment, click **Post**. Your comment is saved and shared immediately.

Edit a Comment

To edit a comment, select the comment that you want to edit, and then click **Edit**.

Note: To edit other users' comments, you must belong to the **Comments:Administrator** role.

Delete a Comment

To delete a comment, select the comment that you want to delete, and then click **Delete**.

Note: To delete comments, you must belong to the **Comments:Administrator** role.

Part 5

Building Models

31

Overview of SAS Visual Statistics

What Is SAS Visual Statistics?

SAS Visual Statistics is an add-on to SAS Visual Analytics that enables you to develop and test models using the in-memory capabilities of SAS LASR Analytic Server. SAS Visual Analytics Explorer (the explorer) enables you to explore, investigate, and visualize data sources to uncover relevant patterns. SAS Visual Statistics extends these capabilities by creating, testing, and comparing models based on the patterns discovered in the explorer. SAS Visual Statistics can export the score code, before or after performing model comparison, for use with other SAS products and to put the model into production.

Benefits of Using SAS Visual Statistics

SAS Visual Statistics enables you to rapidly create powerful statistical models in an easy-to-use, web-based interface. After you have created two or more competing models for your data, SAS Visual Statistics provides a model comparison tool. The model comparison tool enables you to evaluate the relative performance of two or more models against each other and to choose a champion model. A wide variety of model selection criteria is available. Regardless of whether you choose to perform a model comparison, you are able to export model score code for any model that you create. With exported model score code, you can easily apply your model to new data.

Specify Global Preferences

See “Specifying Your Preferences” on page 9 for details about global preferences for SAS Visual Analytics.

Specify Your Preferences for SAS Visual Statistics

Here are the steps to specify preferences that are specific to SAS Visual Statistics:

1 Select **File** ▶ **Preferences** to open the Preferences window.

2 Select **SAS Visual Analytics Explorer** ▶ **Models**.

3 Specify a value for the **p-value precision** property. This value determines the minimum number of decimal places used when displaying p-values.

4 To return to the default settings, select **Reset to Defaults**.

5 Click **OK** to apply your changes. Your preferences persist between sessions.

32

Getting Started with SAS Visual Statistics

Overview

This is a brief overview of using SAS Visual Statistics to derive a new variable, create two different models, and compare those models. This example uses the Framingham Heart Study data set, located in SASHELP.HEART, to compare the performance of a Linear Regression model and a Generalized Linear Model. The goal is to predict a person's age of death based on a collection of health factors. These factors include gender, weight, height, whether the person is a smoker, blood pressure, and more. The focus of this example is how to use SAS Visual Statistics, not how to build the best model.

This example assumes that you have access to the SASHELP.HEART data set. It is beyond the scope of this example to provide instructions about how to access individual data sets at your location. Your system administrator should be able to provide you access to this data set.

Create the Exploration

This example assumes that you have already signed in to SAS Visual Analytics and you are on the home page.

Here are the steps to create the exploration:

1 From the home page, click **Data Explorer**. This opens the explorer and enables you to open a recent exploration or to create a new exploration.

2 Click **Select a Data Source**, located under **Start a new exploration**, to create a new project. A window appears that enables you to select the data source for this exploration.

3 Select the data source that corresponds to SASHELP.HEART. Click **Open**.

4 By default, the exploration is named **Exploration 1**, which is displayed in the upper left corner of the explorer. Before continuing with the example, rename the exploration by saving it.

5 Click **File** ▸ **Save** from the main menu. This opens the Save As window. In the **SAS Folders** pane, navigate to a location where you have Write permission. In the **Name** field, enter `Heart Study`, and click **Save**.

   Typically, you can save your work in `My Folder`.

By default, an automatic visualization is immediately available for usage. However, in this exploration, a decision tree is created in order to derive a leaf ID variable. The leaf ID variable is then used in the Linear Regression model and Generalized Linear Model.

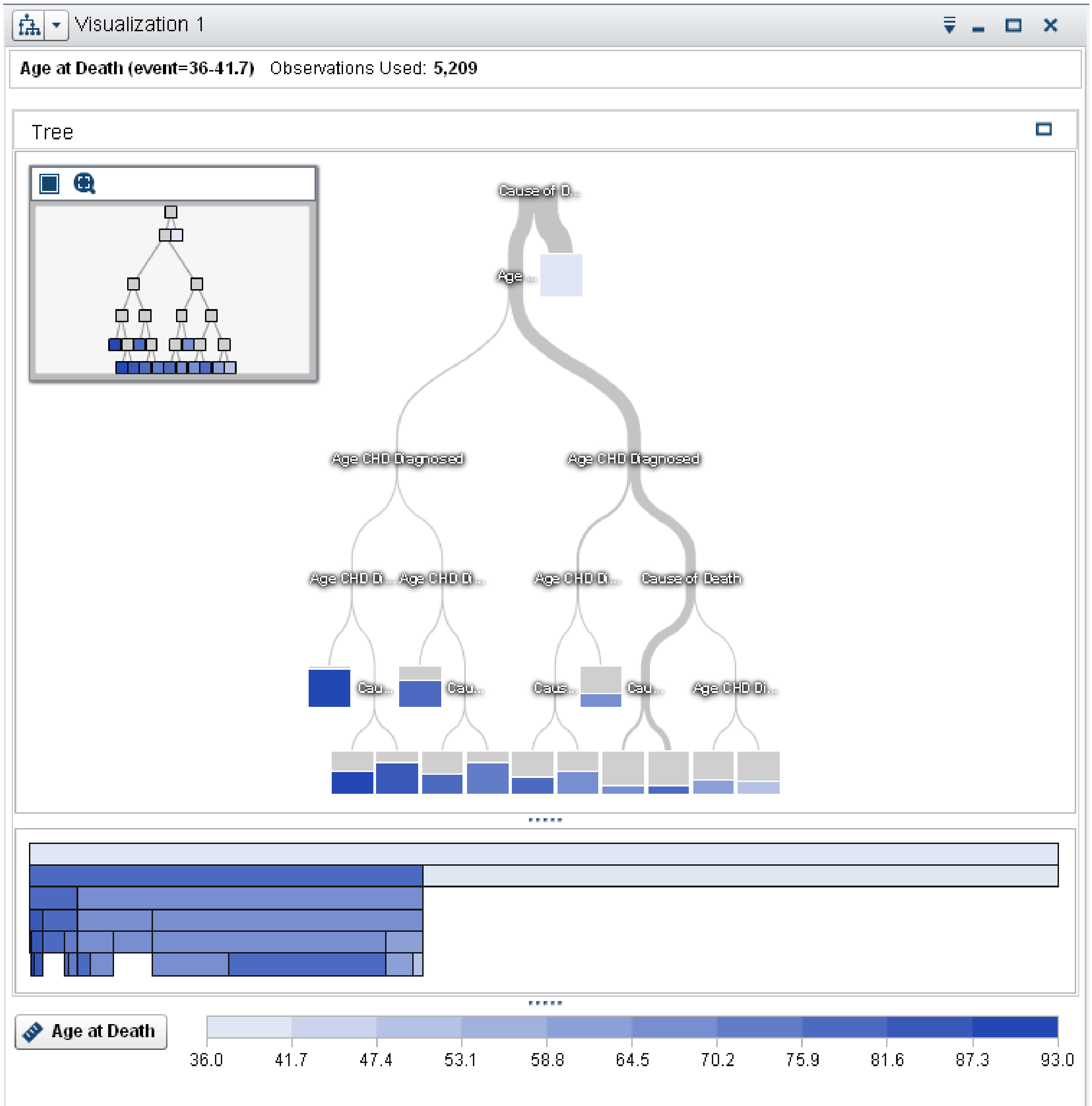
Create a Decision Tree

Here are the steps to create a decision tree:

1 From the toolbar, click to create a decision tree.

2 From the **Data** pane, drag and drop the **Age at Death** variable into the **Response** field in the right pane.

3 In the **Data** pane, select **Diastolic**, **Weight**, **Height**, **Cholesterol**, **Age CHD Diagnosed**, **Sex**, and **Cause of Death**. Drag and drop these items into the model pane. The decision tree automatically updates.



4 Click in the upper right of the model pane.

In the details table, select the **Node Rules** tab. Notice that the only predictors used were **Age CHD Diagnosed** and **Cause of Death**. You can adjust the decision tree properties to include more predictors in the model.

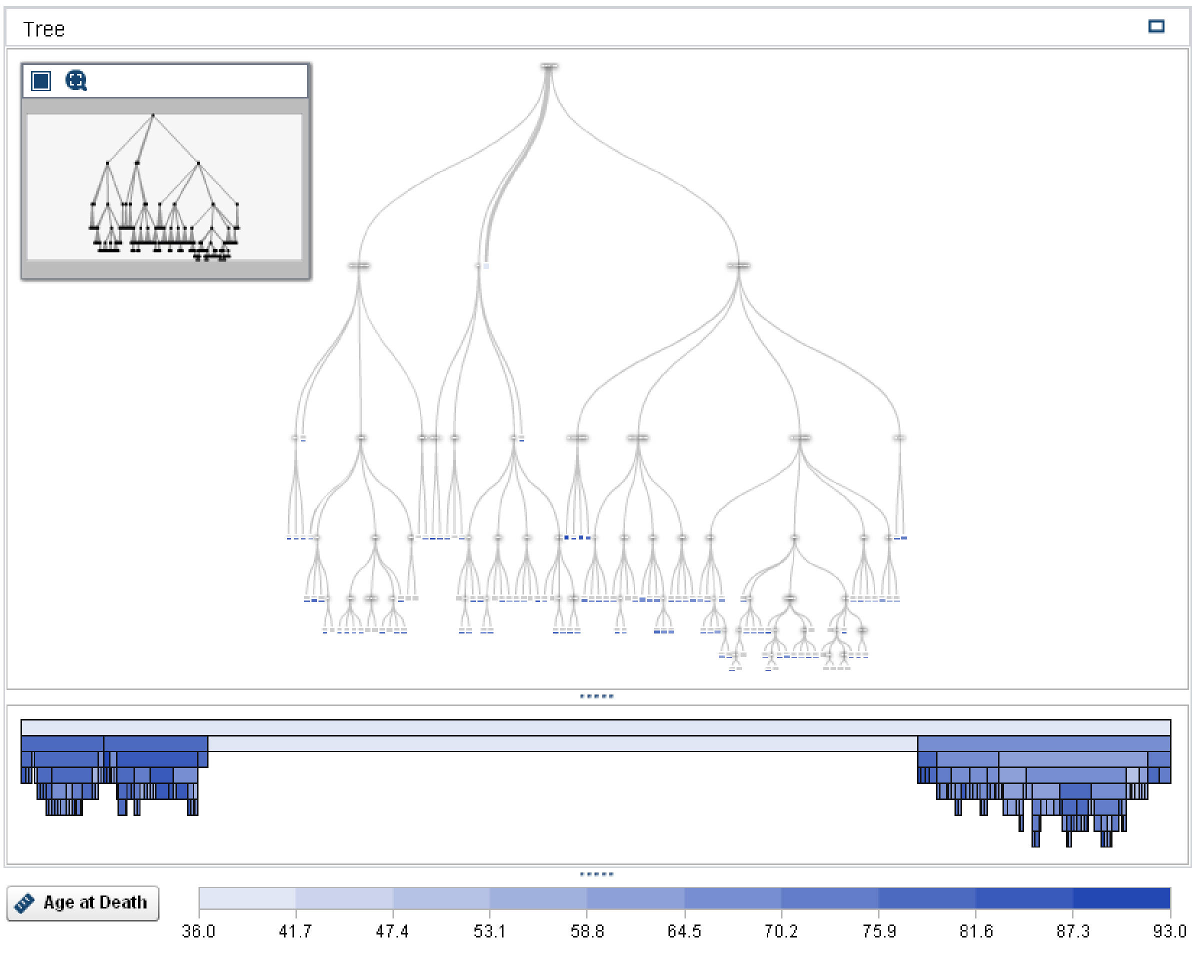
5 Click the **Properties** tab in the right pane. The most obvious property to change is **Reuse predictors**. When you deselect this property, each predictor variable is used in at most one split. However, assume that reusing predictors creates the best split in each node for this example. This might not always be the case for your data.

6 Set the value of **Maximum levels** to `10`. The decision tree now has a maximum depth of 10 levels, instead of the default 6. On the **Node Rules** tab of the details table, every predictor is used at least once.

7 Set the value of **Maximum branches** to `4`. This allows each non-leaf node to split into at most four new nodes.

8 To view the Tree Overview window, click ▣ in the upper right corner of the exploration workspace. In the Tree Overview window, click 🔍 to fit the entire decision tree into the Tree Overview window.

Although each node is difficult to see, your decision tree should resemble the following:



9 In the Tree Overview window, right-click, and select **Derive a Leaf ID Variable**. The default name for this variable is **Leaf ID (1)**.

In the New Calculated Item window, click **OK**. The **Leaf ID (1)** variable appears in the **Data** pane.

10 Save the project.

Create a Linear Regression

Here are the steps to create a linear regression:

1 From the toolbar, click ▾ next to [icon]. Select **Linear Regression** from the drop-down list.

2 Minimize the Decision Tree visualization and Tree Overview window.

3 In this example, the variable of interest is **Age at Death**, which should be the first variable listed in the **Measure** section of the **Data** pane.

Because you want this variable to be the response variable, click, drag, and drop **Age at Death** from the **Data** pane onto the model pane. Notice that **Age at Death** now appears in the **Response** field on the **Roles** tab.

4 Choose the effect variables or interaction terms that you want to include in the analysis. One option is to make every available variable an effect variable and let SAS Visual Statistics perform variable selection. However, this is not always feasible from a computational resources perspective. This example creates an interaction term to use as an effect variable and includes a few other variables as effect variables.

5 Because you suspect that systolic blood pressure and diastolic blood pressure interact with each other, create an interaction term for these variables.

Follow these steps to create an interaction term:

a In the **Data** pane, click , and select **New Interaction Effect**.

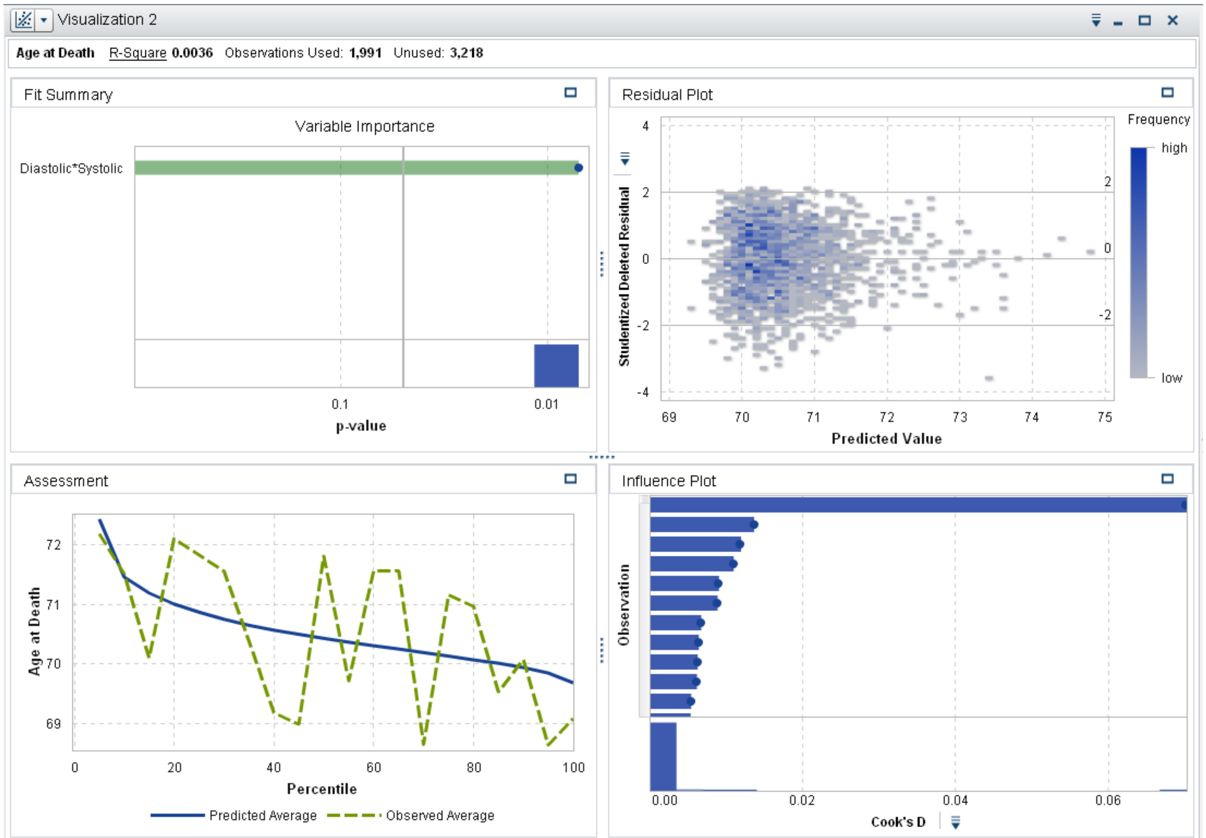
b In the New Interaction Effect window, move **Diastolic** and **Systolic** from the **Available columns** area into the **Effect elements** area.

c Click **Create**.

The interaction term **Diastolic*Systolic** appears in the **Interaction Effects** group of the **Data** pane.

6 Click, drag, and drop **Diastolic*Systolic** onto the model pane. A model is created based on that single effect because the **Auto-update model** option is selected in the right pane.

TIP Each time a change is made to the model, the Linear Regression automatically updates. If you anticipate making many changes or if you are experiencing server performance issues, deselect the **Auto-update model** option. When auto-updates are disabled, you must click **Update** in the right pane to update the model.



7 Add more effects to the model. Hold down the Ctrl key, and select **Blood Pressure Status**, **Cause of Death**, **Leaf ID 1**, **Sex**, **Smoking Status**, **Cholesterol**, **Height**, **Smoking**, and **Weight**. Drag and drop these variables onto the model pane. The Linear Regression updates to include these effects.

8 In the right pane, select the **Properties** tab. In this model, **Informative missingness** and **Use variable selection** are not selected. Disabling **Informative missingness** means that observations with missing values are not included in the analysis. Disabling **Use variable selection** means that all variables are used in the model, regardless of how significant they are to the model. For this model, keep the default properties settings.

The Fit Summary window indicates that **Cause of Death**, **Leaf ID (1)**, and **Height** are the three most important effects in this model.

The Assessment window indicates that the observed average and predicted average are approximately equal for most bins.

9 Save the project.

Create a Generalized Linear Model

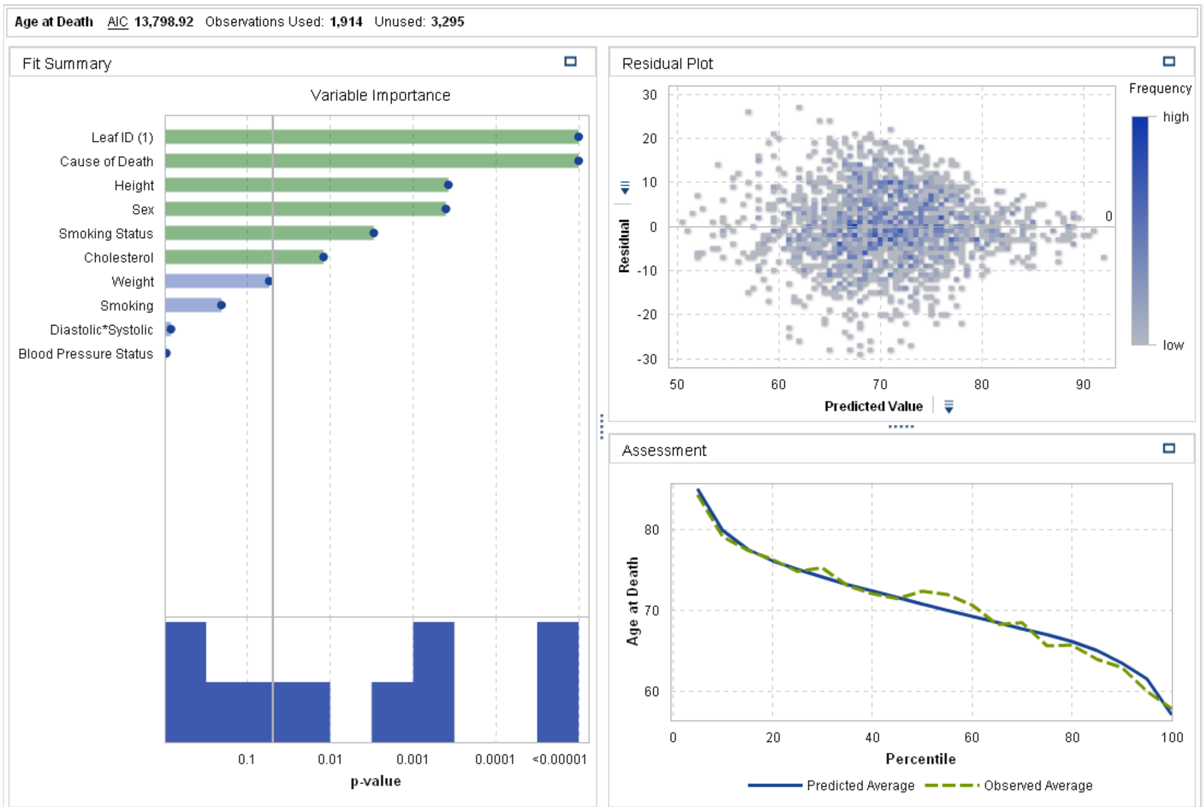
Here are the steps to create a generalized linear model:

1 From the toolbar, click ▼ next to . Select **Generalized Linear Model** from the drop-down list.

2 Minimize the Linear Regression visualization.

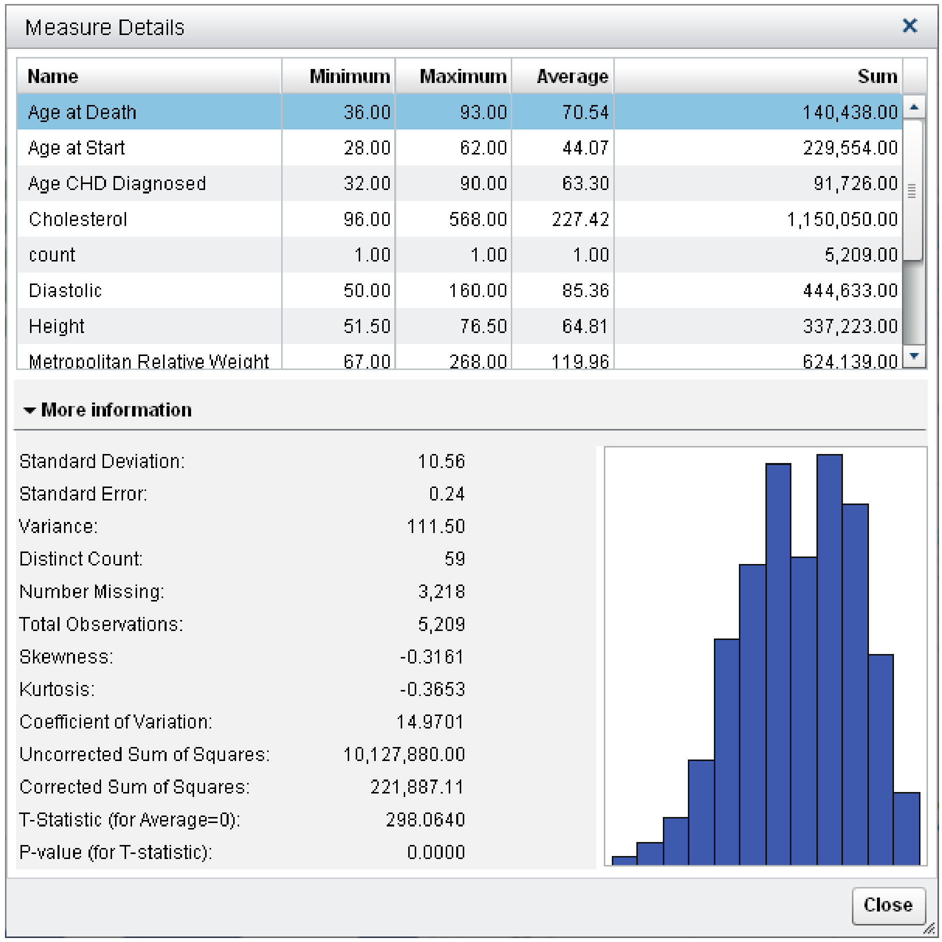
3 From the **Data** pane, drag and drop the **Age at Death** variable into the **Response** field in the right pane.

4 In the **Data** pane, hold down the Ctrl key, and select **Blood Pressure Status**, **Cause of Death**, **Leaf ID (1)**, **Sex**, **Smoking Status**, **Cholesterol**, **Height**, **Smoking**, **Weight**, and **Diastolic*Systolic**. Drag and drop these variables onto the model pane.



5 Click the **Properties** tab in the right pane. The **Distribution** property enables you to specify the distribution of the response variable and to build a model based on that distribution. The default distribution is **Normal**.

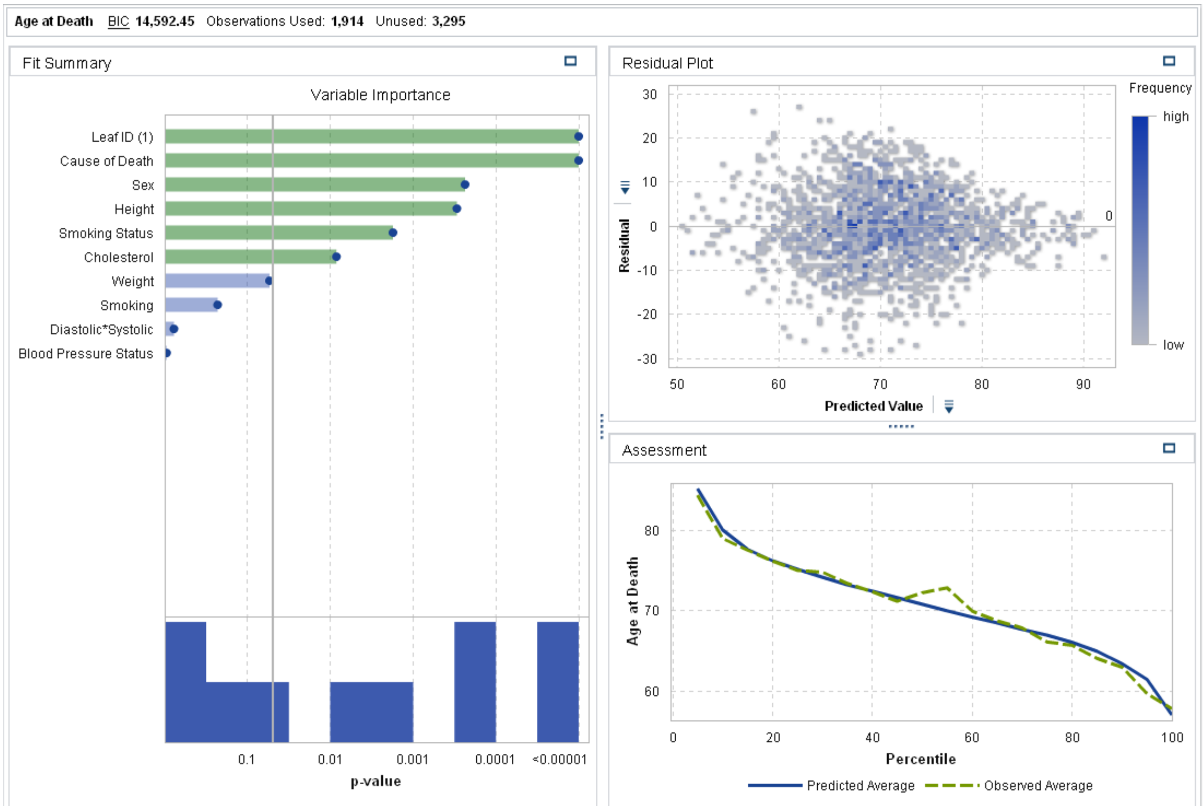
To determine whether the normal distribution applies to the response variable, click in the **Data** pane, and select **Measure Details**. In the Measure Details window, select **Age at Death**.



6 Notice that **Age at Death** is not normally distributed and is slightly skewed left. Click **Close**.

7 Although the distribution is not exactly Poisson, use the Poisson distribution for this example. For the **Distribution** property, select **Poisson**. Next, select **Identity** for **Link function**.

Note: You are encouraged to repeat this example with different distributions and link functions and compare their performance and to familiarize yourself with SAS Visual Statistics.

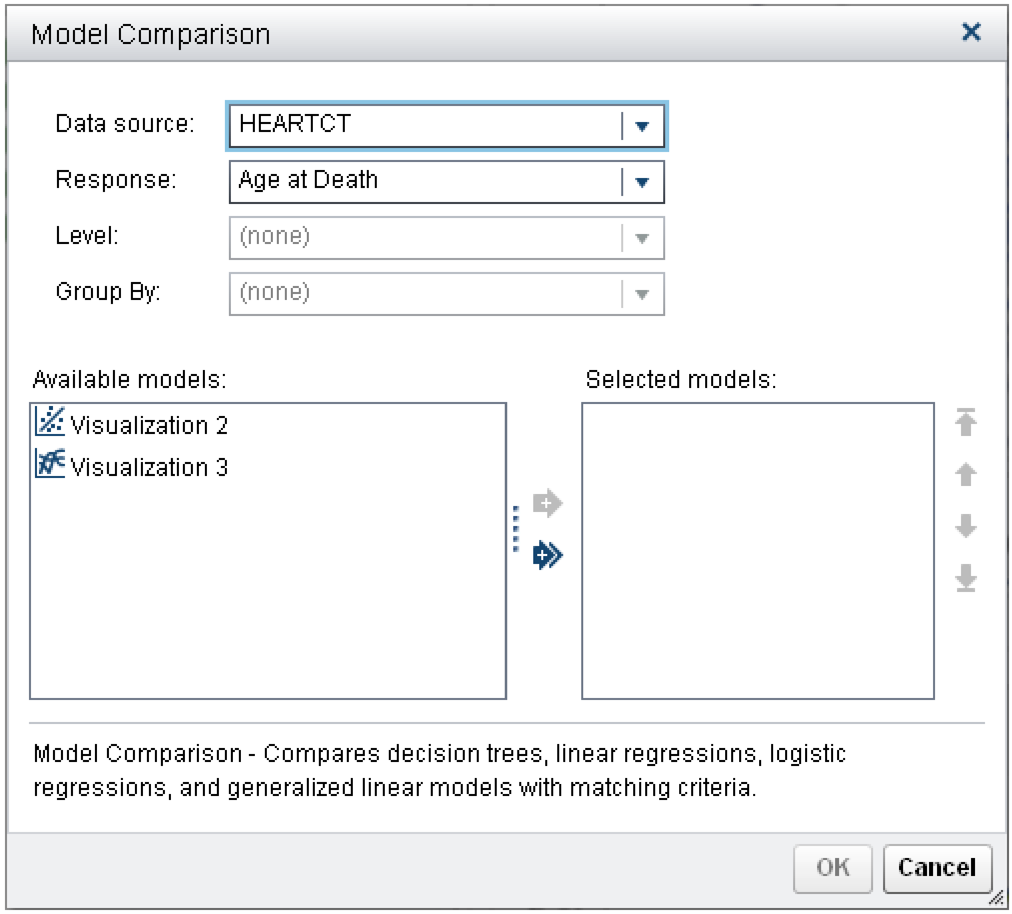


8 Save the project.

Perform a Model Comparison

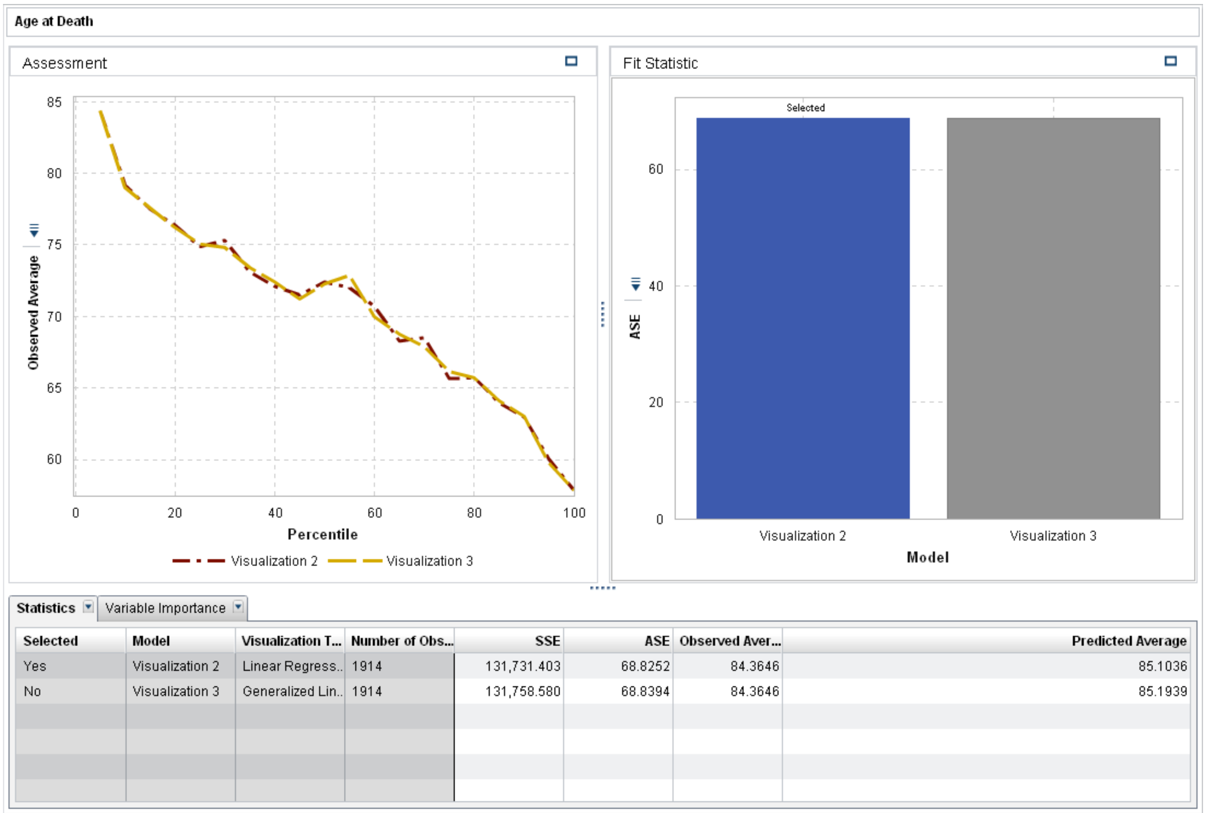
Here are the steps to perform a model comparison:

1 From the toolbar, click to create a new model comparison. The Model Comparison window appears.



The **Response** variable is already set to **Age at Death**, and **Level** and **Group By** are unavailable. With these settings, the available models are **Visualization 2** (the Linear Regression) and **Visualization 3** (the Generalized Linear Model).

2 Click to select both models for comparison. Click **OK**.



3 By default, the fit statistic average squared error, **ASE**, is used to compare the models. The other available fit statistics are **SSE** and **Observed Average**. Because smaller values are preferred, the Linear Regression is chosen as the champion when **ASE** or **SSE** is the criterion. The models are very similar.

When the fit statistic is **Observed Average**, the **Percentile** slider is available. This slider specifies the percentile where the observed average and predicted average are compared. In some percentiles, the Generalized Linear Model might be chosen over the Linear Regression.

If you view the Assessment plot, both the **Observed Average** and **Predicted Average** plots show that the models are relatively similar.

4 Now that you have a champion model, you can export the model score code for that model to score new data.

Here are the steps to export the model score code:

a Open **Visualization 2**, the Linear Regression.

b Click, and select **Export Score Code**.

c In the Export Score Code window, click **Export**.

d In the Save As window, navigate to where you want to save the code, and click **Save**.

33

Modeling Information

Available Models

The following models are available in SAS Visual Statistics:

- Linear Regression on page 257 attempts to predict the value of an interval response as a linear function of one or more effect variables.
- Logistic Regression on page 265 attempts to predict the probability that a binary or ordinal response will acquire the event of interest as a function of one or more effects.
- Generalized Linear Model on page 273 is an extension of a traditional linear model that allows the population mean to depend on a linear predictor through a nonlinear link function.
- Decision Tree on page 281 creates a hierarchical segmentation of the input data based on a series of rules applied to each observation.
- Cluster on page 289 segments the input data into groups that share similar features.

Overview of Variables and Interaction Terms

Variables

Category Variables

Category variables are numeric or nonnumeric variables with discrete levels. The levels of a category variable are considered unordered by SAS Visual Statistics. Examples of category variables include drink size (small, medium, or large), number of cylinders in an engine (2, 4, 6, or 8), or whether a customer has made a purchase (yes or no).

You can create a category variable from a response variable by right-clicking the variable, and selecting **Category**. In this case, each distinct value of the measure variable is turned into a level for the category variable.

Category variables can be used as response variables for classification models, classification effect variables, decision tree predictors, filter variables, and group by variables.

Note: To ensure proper performance and valid modeling results, the maximum number of distinct levels allowed for a category variable is limited based on the model type and variable role.

Measure Variables

Measure variables are continuous numeric variables that can assume an infinite number of possible values between two numbers. Even though some numeric variables are not continuous, such as count variables, these variables can be treated as continuous values for the purpose of modeling. Examples of measure variables include the temperature of a drink, engine displacement amount, or a customer’s total purchase amount.

Summary statistics and a histogram for each measure variable are obtained by right-clicking the variable in the **Data** pane, and selecting **Properties**. Use the **Name** drop-down menu to specify the variable that you want to view.

Measure variables can be used as response variables for continuous models, continuous effect variables, decision tree predictors, offset variables, frequency variables, weight variables, and filter variables.

Interaction Terms

Two variables, A and B, *interact* if the effect of one variable on the model changes as the other variable changes. That is, the effects of variables A and B are not additive in the model.

SAS Visual Statistics enables you to create interactions between two or more input variables, including squared interactions. A squared interaction is the interaction of a variable with itself. You cannot create squared interactions for category variables.

For an example where interaction terms might be useful, consider a situation where you are modeling the fuel mileage (MPG) for several cars. Two of your input variables are engine displacement in liters and engine size (number of cylinders). You expect that as either value increases, fuel mileage will suffer. However, if you suspect that the effects on fuel mileage that are attributable to engine displacement are not constant across engine size, then you should consider creating the interaction term between those variables.

SAS Visual Statistics is not limited to creating just two-way interactions. You can create *n*-way interactions that include an arbitrary number of variables, but not more than the number of available input variables.

The number of distinct levels for an interaction term is the product of the number of levels for each variable in the term. Measure variables are treated as if they contain one level. The number of levels in an interaction term counts against the maximum number of distinct levels allowed in regression models.

Variable Selection

Variable selection is the process of reducing the number of input variables to include just the most significant variables. The Linear Regression and Logistic Regression models provide a property to automatically perform variable selection. When you use this property, SAS Visual Statistics performs backward selection on the input variables to determine the most significant variables. Modeling with just the most significant variables is intended to avoid creating a model that overfits the data. Automated variable selection can actually take longer to run than not performing variable selection.

Missing Values

By default, SAS Visual Statistics handles missing values by dropping all observations that contain a missing value in any assigned role variable. However, the Linear Regression, Logistic Regression, and Generalized Linear Model models provide the **Informative missingness** property. In some cases, the fact that an observation contains a missing value provides relevant modeling information. Selecting this property explicitly models missing values of variables as a separate variable. For measure variables, missing values are imputed with the observed mean, and an indicator variable is created to denote missingness. For category variables, missing values are considered a distinct level.

Group By Variables

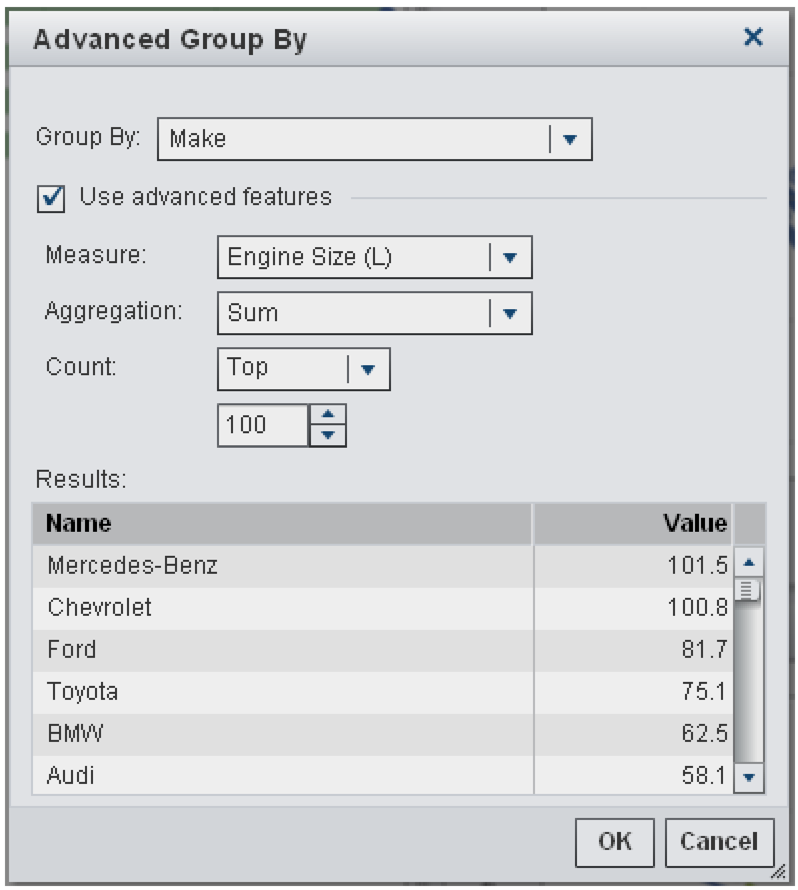
A group by variable enables you to fit a model for each data segment defined by one or more category variables. Each unique combination of levels across all of the group by variables is a specific data segment. For example, if you have one group by variable with three levels, then there are three data segments. But, if you have two group by variables, one with three levels and the other with four levels, then there are at most 12 data segments. A data segment is not created when there are no observations in a combination of classification levels.

SAS Visual Statistics enforces a maximum number of BY groups, except when you use the Advanced Group By feature. By default, the maximum number of BY groups allowed is 1024. Empty data segments count against the maximum number of BY groups allowed in a model.

When you specify two or more group by variables, the results are grouped in the order in which the variables appear in the **Group By** field.

In the Fit Summary window, when you select a specific data segment, the Residual Plot and Influence Plot windows are updated to include only the observations in the specified data segment.

The Advanced Group By window provides more control over variable grouping. To access the Advanced Group By window, click **Advanced** next to **Group By** in the right pane.

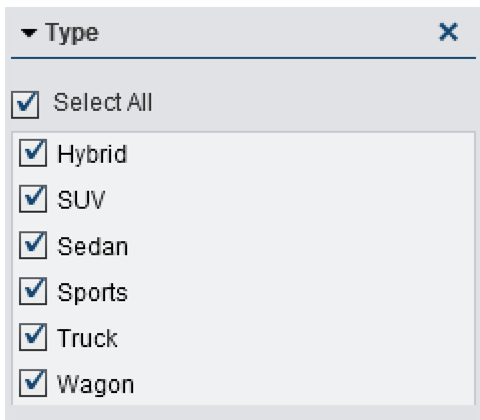


The **Group By** field enables you to select the variable that is used for grouping. Select the **Use advanced features** option to display aggregation statistics for a specified measure variable. Specify the measure variable in the **Measure** field. The **Aggregation** field specifies whether the **Average** or **Sum** is computed. Use the **Count** field to specify whether you want the **Top** or **Bottom** *n* values. The field below **Count** enables you to specify the value of *n*.

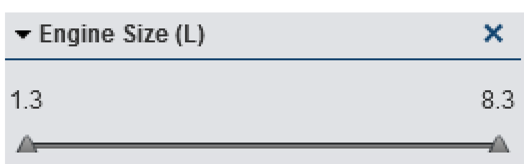
Filter Variables

Filter variables are used to subset the modeling data. You can filter on any variable included in the data, not just on variables used in the model. Filter variables are applied only to the current model.

When you filter on a category variable, you are presented with a list of the levels for that variable. Select only values that you want to include in the model. In the following image, all levels are available.



When you filter on a measure variable, a slider lets you specify a range of values. Use the triangles to specify the lower and upper limits of the filter variable.



Score Code

Model scoring refers to the process of generating predicted values for a data set that might contain the response variable of interest. Score code is exported as a SAS DATA step that can be executed on new data sets in any SAS environment. All variables used by the model in any capacity are included in the score code. This includes interaction terms, group by variables, frequency variables, and weight variables. Score code is not available for interactive decision trees.

To generate model score code, click , and select **Export Score Code**. In the Export Model Score Code window, click **Export**. In the Save As window, navigate to where you want to save the code, and click **Save**.

Score code is saved as a .sas file and can be viewed in any word processing program.

Note: It is possible for your exported score code to have lines of code that exceed the maximum line length of 32768. There are two solutions for this issue. The first solution requires that you edit the exported text file to include a line break on each of the long lines and to insert `/ lrecl=1000000` in the %include statement. The second solution requires that you open the exported text file in a SAS Program Editor and insert a line break on each of the long lines. In the SAS Program Editor there is a limit of 6000 characters per line.

Derive Predicted Values

For all three regression visualizations, SAS Visual Statistics creates two variables that contain prediction information for each observation in the data set. After these variables are created, they can be used in any other visualization, including the other predictive models.

To create the two new variables, complete the following steps:

1 Create a valid linear regression, logistic regression, or generalized linear model visualization.

2 Click ▾ in the upper right corner of the visualization and select **Derive Predicted Values**.

3 In the New Prediction Variables window, enter a name for the **Predicted Values** and either the **Residual Values** or the **Probability Values**. **Residual Values** are available for linear regressions and generalized linear models. **Probability Values** are available for logistic regressions.

4 Click **OK**. The predicted values for the logistic regression appear in the **Category** variables section. All other variables, including the predicted values for the other models, appear in the **Prediction** variables section.

Depending on the chosen visualization, the information contained in each variable is slightly different.

Predicted Values
For linear regressions and generalized linear models, this is a numeric value that is the value generated by the regression model. Or, this is the value that would have been generated by the regression model if the observation was scored by the model.

For logistic regressions, this is the decision generated by the logistic regression based on the calculated probability and **Prediction cutoff** property. All observations are classified into either the event level of interest, not the event level of interest, or missing.

Residual Values
The computed residual for each observation. Available for the linear regression and generalized linear model visualizations.

Probability Values
The computed probability for each observation. Observations with probability values that are greater than or equal to the **Prediction cutoff** property are predicted to be in the event level of interest. Observations with probability values that are less than the **Prediction cutoff** property are considered to be not in the event level or interest. That is, there is no prediction made

regarding each individual measurement level, only between the measurement level of interest and everything else.

34

Linear Regression Model

Overview of the Linear Regression Model

A linear regression attempts to predict the value of a measure response variable as a linear function of one or more effects. The linear regression model uses the least squares method to determine the model. The least squares method creates a line of best fit by minimizing the residual sum of squares for every observation in the input data set. The residual sum of squares is the vertical distance between an observation and the line of best fit. The least squares method requires no assumptions about the distribution of the input data.

The linear regression model requires a measure response variable and at least one effect variable or interaction term.

Linear Regression Model Properties

The following properties are available for the linear regression model:

Name
enables you to specify the name for this model.

Informative missingness
specifies whether the informative missingness algorithm is used. For more information, see Missing Values on page 251.

Use variable selection
specifies whether variable selection is performed. For more information, see Variable Selection on page 251.

Significance level
specifies the significance level that is required in order for variables to be considered for the model. This property is available only when **Use variable selection** is selected.

Assessment

- **Use default number of bins** specifies whether you want to use the default number of bins or to set your own value. By default, measure variables are grouped into 20 bins.
- **Number** specifies the number of bins to use when the **Use default number of bins** property is not selected. You must specify an integer value between 5 and 100.
- **Tolerance** specifies the tolerance value that is used to determine the convergence of the iterative algorithm that estimates the percentiles. Specify a smaller value to increase the algorithmic precision.

Show diagnostic plots
specifies whether the Residual Plot, Assessment, and Influence Plot appear in the model pane.

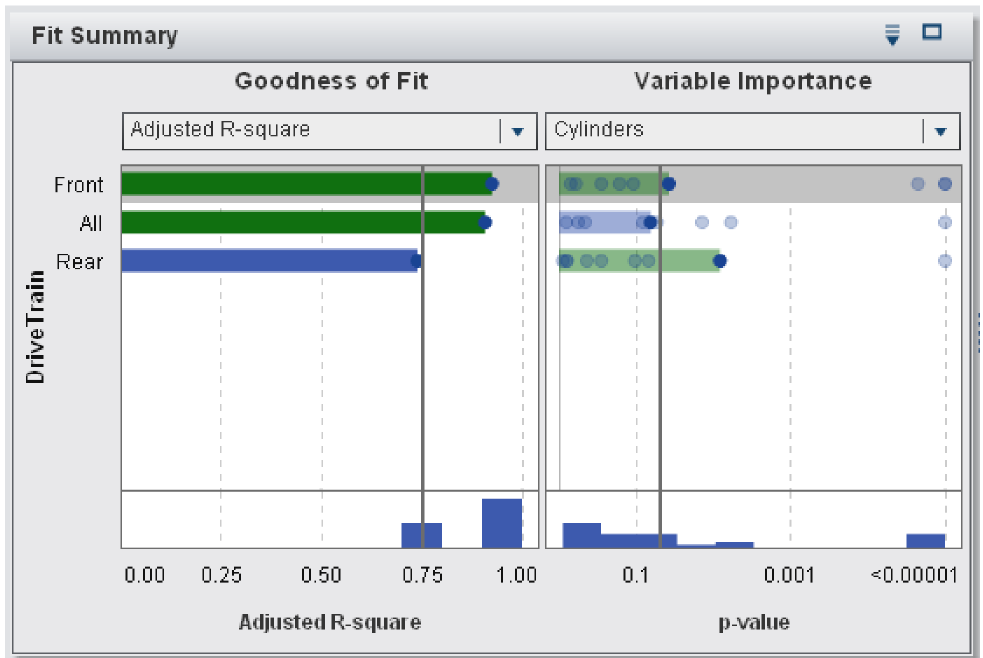
Working with the Fit Summary Window

About the Fit Summary Window

The Fit Summary plots the relative importance of each variable as measured by its *p*-value. The *p*-value is plotted on a log scale and the alpha value (plotted as -log(alpha)), is shown as a vertical line. To adjust the alpha value, click, drag, and drop the vertical line. A histogram of the *p*-values is displayed at the bottom of the window.

Including a Group By Variable

When your analysis includes a group by variable, the Fit Summary displays a **Goodness of Fit** plot and **Variable Importance** plot.



The **Variable Importance** plot enables you to select a single variable to inspect within each level of the group by variable. Use the drop-down menu to select which variable you want to inspect. Each dot in the **Variable Importance** plot represents a model effect. The bars are drawn for the effect selected in the drop-down menu.

The **Goodness of Fit** plot, which is not available when there is no group by variable, displays how well the model predicts the response variable within each level of the group by variable. Use this plot to determine whether your model has a significantly different fit within different levels.

Notice that when you select a group by variable, the **Residual Plot**, **Assessment**, and **Influence Plot** are updated. These plots enable you to further inspect the variable within each level of the group by variable.

Use to specify how the plots are sorted.

Working with the Residual Plot

About Residual Plots

A residual plot shows the relationship between the predicted value of an observation and the residual of an observation. The residual of an observation is the difference between the predicted response value and the actual response value. When using large data sets, the residual plot is displayed as a heat map instead of as an actual plot. In a heat map, the actual observations are binned, and the color of each point indicates the relative number of observations in that bin.

Residual plots have several uses when examining your model. First, obvious patterns in the residual plot indicate that the model might not fit the data. Second, residual plots can detect nonconstant variance in the input data when you plot the residuals against the predicted value. Nonconstant variance is evident when the relative spread of the residual values changes as the predicted values change. Third, in combination with other methods, the residual plot can help identify outliers in your data.

To change the residual that is plotted, click on the vertical axis.

By default, a scatter plot or heat map is displayed. However, a histogram of the residuals is also available. To display a histogram of the residuals, right-click in the **Residual Plot**, and select **Use Histogram**. In the histogram, you can change the residual that is plotted by clicking on the horizontal axis.

From the histogram, it is easy to determine whether the distribution of the residuals is approximately normal or skewed. A non-normal residual histogram can indicate that the model does not fit the data.

Use the Residual Plot

The **Residual Plot** provides several features to investigate your data. These features are available in both the scatter plot and histogram, and the process to activate these features is the same in either plot. The available features enable you to plot the residuals by a category variable, show only the selected observations, exclude the selected observations, and remove all exclusions.

There are two ways to select observations in the **Residual Plot**. If you hold the mouse pointer over an individual point, you can click on that point to select it. Or, when you click and drag the mouse in the **Residual Plot**, a selection rectangle appears. When you release the mouse, all points inside the rectangle are selected.

To plot the residuals by a category variable, right-click in the **Residual Plot**, select **Plot By**, and then select a category variable. A box plot of the residuals appears. The residuals are grouped by the selected category variable.

To show the selected observations, select some data points, right-click in the **Residual Plot**, and select **Show Selected**. A table of the selected observations appears.

To exclude the selected observations, select some data points, right-click in the **Residual Plot**, and select **Exclude Selected**. The selected observations are no longer used to build the model.

To remove all exclusions, right-click in the **Residual Plot**, and select **Remove Exclusions**. All observations are now used to build the model.

Working with the Assessment Window

About the Assessment Window

For a linear regression, the Assessment plots the average predicted and average observed response values against the binned data. Use this plot to determine how well the model fits the data.

Use the Assessment Window

The **Assessment** plot bins the data based on the values specified in the **Assessment** properties. At each bin, you can hold the mouse pointer over one or both of the lines to display a tooltip.

Influence Plot

The Influence Plot displays several measurements that are computed for each observation. A histogram of the displayed measurements is also displayed. When the input data contains a large number of observations, the observations are binned. Use these measurements to help identify outliers and other data points that greatly affect the predicted regression model.

To change the computed measurement that is plotted, click on the horizontal axis.

Fit Statistics

The linear regression model computes several assessment measures to help you evaluate how well the model fits the data. These assessment measures are available at the top of the model pane. Click the currently displayed assessment measure to see all of the available assessment measures.

Adjusted R-square
: The Adjusted R-squared value attempts to account for the addition of more effect variables. Values can range from 0 to 1. Values closer to 1 are preferred.

AIC
: Akaike's Information Criterion. Smaller values indicate better models, and AIC values can become negative. AIC is based on the Kullback-Leibler information measure of discrepancy between the true distribution of the response variable and the distribution specified by the model.

AICC
: Corrected Akaike's Information Criterion. This version of AIC adjusts the value to account for sample size. The result is that extra effects penalize AICC more than AIC. As the sample size increases, AICC and AIC converge.

Average Squared Error
The average squared error (ASE) is the sum of squared errors (SSE) divided by the number of observations. Smaller values are preferred.

F Value for Model
The value of the F test in a one-way ANOVA after the variances are normalized by the degrees of freedom. Larger values are better, but can indicate overfitting.

Mean Square Error
The mean squared error (MSE) is the SSE divided by the degrees of freedom for error. The degrees of freedom for error is the number of cases minus the number of weights in the model. This process yields an unbiased estimate of the population noise variance under the usual assumptions. Smaller values are preferred.

Observations
The number of observations used in the model.

Pr > F
The *p*-value associated with the corresponding F statistic. Smaller values are preferred.

R-Square
The R-squared value is an indicator of how well the model fits the data. R-squared values can range from 0 to 1. Values closer to 1 are preferred.

Root MSE
Square root of the MSE.

SBC
The Schwarz's Bayesian Criterion (SBC), also known as the Bayesian Information Criterion (BIC), is an increasing function of the model's residual sum of squares and the number of effects. Unexplained variations in the response variable and the number of effects increase the value of the SBC. As a result, a lower SBC implies either fewer explanatory variables, better fit, or both. SBC penalizes free parameters more strongly than AIC.

Details Table

When you click at the upper right of the model pane, the details table is displayed at the bottom of the model pane. The details table contains the following information:

Overall ANOVA
The analysis of variance results for the model, error, and corrected total.

Dimensions
An overview of the effect variables used in the model. This tab identifies how many measures and classification effects were chosen for the model, the rank of the cross-product matrix, how many observations were read, and how many observations were used in the model.

Fit Statistics
Lists all of the fit statistics described in the previous section.

Model ANOVA
The analysis of variance results for the model.

Type III Test

Provides details for the Type III test. A Type III test examines the significance of each partial effect with all other effects in the model. For more information, see the chapter “The Four Types of Estimable Functions” in the *SAS/STAT User’s Guide*.

Parameter Estimates

Gives the estimated values for the model parameters.

35

Logistic Regression Model

Overview of the Logistic Regression Model

A logistic regression attempts to predict the value of a binary response variable. A logistic regression analysis models the natural logarithm of the odds ratio as a linear combination of the explanatory variables. This approach enables the logistic regression model to approximate the probability that an individual observation belongs to the level of interest.

The logistic regression model requires a category response variable and at least one effect variable or interaction term. When your category response variable contains more than two levels, SAS Visual Statistics prompts you to select the level of interest. That is, SAS Visual Statistics treats all observations in the level of interest as an event and all other observations as nonevents.

Logistic Regression Model Properties

The following properties are available for the logistic regression model:

Name
enables you to specify the name for this model.

Informative missingness
specifies whether the informative missingness algorithm is used. For more information, see Missing Values on page 251.

Use variable selection
specifies whether variable selection is performed. For more information, see Variable Selection on page 251.

Significance level
specifies the significance level that is required in order for variables to be considered for the model. This property is available only when **Use variable selection** is selected.

Link function
specifies the link function that is used in the logistic regression. Link functions link the response mean to the linear predictor.

The following link functions are available:

- **Logit** (default) specifies the inverse of the cumulative logistic distribution function.

  $$g(M) = \log\left(\frac{M}{1 - M}\right)$$

- **Probit** specifies the inverse of the cumulative standard normal distribution function.

  $$g(M) = \frac{1}{\Phi(M)}$$

Convergence

- **Override function convergence** enables you to manually specify the function convergence value.
- **Value** specifies the function convergence value when **Override function convergence** is selected. When you specify a larger value, the model will converge sooner. This reduces the amount of time spent training the model, but it can create a suboptimal model.
- **Override gradient convergence** enables you to manually specify the gradient convergence value.
- **Value** specifies the gradient convergence value when **Override gradient convergence** is selected. When you specify a larger value, the model will converge sooner. This reduces the amount of time spent training the model, but it can create a suboptimal model.
- **Maximum iterations** specifies the maximum number of iterations performed during model training. If you specify a relatively small value, you reduce the amount of time spent training the model, but it can create a suboptimal model.

Note: When you specify a gradient convergence or function convergence criterion, it is possible for the model to converge based on an internal convergence criterion before your specified criterion is reached. The reason for convergence is on the **Convergence** tab of the details table.

Assessment

- **Use default number of bins** specifies whether you want to use the default number of bins or to set your own value. By default, measure variables are grouped into 20 bins.
- **Number** specifies the number of bins to use when the **Use default number of bins** property is not selected. You must specify an integer value between 5 and 100.
- **Prediction cutoff** specifies the value at which a computed probability is considered an event.
- **Tolerance** specifies the tolerance value that is used to determine the convergence of the iterative algorithm that estimates the percentiles. Specify a smaller value to increase the algorithmic precision.

Show diagnostic plots

specifies whether the Residual Plot, Assessment, and Influence Plot windows appear in the model pane.

Working with the Fit Summary Window

About the Fit Summary Window

The Fit Summary window plots the relative importance of each variable as measured by its *p*-value. The *p*-value is plotted on a log scale and the alpha value (plotted as -log(alpha)), is shown as a vertical line. To adjust the alpha value, click, drag, and drop the vertical line. A histogram of the *p*-values is displayed at the bottom of the window.

Including a Group By Variable

When your analysis includes a group by variable, the Fit Summary window displays a Goodness of Fit plot and Variable Importance plot.

The Variable Importance plot enables you to select a single variable to inspect within each level of the group by variable. Use the drop-down menu to select the variable that you want to inspect. Each dot in the Variable Importance plot represents a model effect. The bars are for the model effect chosen in the drop-down menu.

The Goodness of Fit plot, which is not available when there is no group by variable, displays how well the model predicts the response variable within each level of the group by variable. Use this plot to determine whether your model has a significantly different fit within different levels.

Notice that when you select a group by variable level, the Residual Plot, Assessment, and Influence Plot windows are updated. This enables you to further investigate each level of the group by variable.

Use to specify how the plot is sorted.

Working with the Residual Plot

About Residual Plots

A residual plot shows the relationship between the predicted value of an observation and the residual of an observation. The residual of an observation is the difference between the predicted response value and the actual response value. When using large data sets, the residual plot is displayed as a heat map instead of as an actual plot. In a heat map, the actual observations are binned, and the color of each point indicates the relative number of observations in that bin.

Residual plots have several uses when examining your model. First, obvious patterns in the residual plot indicate that the model might not fit the data. Second, residual plots can detect nonconstant variance in the input data when you plot the residuals against the predicted values. Nonconstant variance is evident when the relative spread of the residual values changes as the predicted values change. Third, in combination with other methods, the residual plot can help identify outliers in your data.

To change the residual that is plotted, click on the vertical axis.

By default, a scatter plot or heat map is displayed. However, a histogram is also available. To display a histogram of the residuals, right-click in the **Residual Plot**, and select **Use Histogram**. In the histogram, you can change the residual that is plotted by clicking on the horizontal axis.

From the histogram, it is easy to determine whether the distribution of the residuals is approximately normal or skewed. A non-normal residual histogram can indicate that the model does not fit the data.

Use the Residual Plot

The Residual Plot provides several features to investigate your data. These features are available in both the scatter plot and histogram. The process to activate these features is the same in either plot. The available features are plot the residuals by a category variable, show only the selected observations, exclude selected observations, and remove all exclusions.

There are two ways to select observations in the Residual Plot. If you hold the mouse over an individual point, you can click on that point to select it. Alternatively, when you click and drag the mouse in the Residual Plot, a selection rectangle appears. When you release the mouse, all points inside the rectangle are selected.

To plot the residuals by a category variable, right-click in the Residual Plot, select **Plot By**, and then select a category variable. A box plot of the residuals appears. The residuals are grouped by the selected category variable.

To show selected observations, select some data points, right-click in the Residual Plot, and select **Show Selected**. A table of the selected observations appears.

To exclude selected observations, select some data points, right-click in the Residual Plot, and select **Exclude Selected**. The selected observations are no longer used to build the model.

To remove all exclusions, right-click anywhere in the Residual Plot, and select **Remove Exclusions**. All observations are now used to build the model.

Working with the Assessment Window

About the Assessment Window

For a logistic regression, the Assessment window plots lift, ROC, and misclassification rates. Use the Assessment window to determine how well the model fits the data.

Lift is the ratio of the percent of captured responses within each percentile bin to the average percent of responses for the model. Similarly, *cumulative lift* is calculated by using all of the data up to and including the current percentile bin.

A receiver operating characteristic (ROC) chart displays the ability of a model to avoid false positive and false negative classifications. A false positive classification means that an observation has been identified as an event when it is actually a nonevent (also referred to as a Type **I** error). A false negative classification means that an observation has been identified as a nonevent when it is actually an event (also referred to as a Type **II** error).

The misclassification plot displays how many observations were correctly and incorrectly classified for each value of the response variable. When the response variable is not binary, the logistic regression model considers all levels that are not events as equal. A significant number of misclassifications could indicate that your model does not fit the data.

Use the Assessment Window

Lift

The default lift chart displays the cumulative lift of the model. To view the noncumulative lift, click on the vertical axis, and select **Lift**.

For comparison, the lift chart plots a best model based on complete knowledge of the input data.

ROC

The *specificity* of a model is the true negative rate. To derive the false positive rate, subtract the specificity from 1. The false positive rate, labeled **1 – Specificity**, is the X axis of the ROC chart. The *sensitivity* of a model is the true positive rate. This is the Y axis of the ROC chart. Therefore, the ROC chart plots how the true positive rate changes as the false positive rate changes.

A good ROC chart has a very steep initial slope and levels off quickly. That is, for each misclassification of an observation, significantly more observations are correctly classified. For a perfect model, one with no false positives and no false negatives, the ROC chart would start at (0,0), continue vertically to (0,1), and then horizontally to (1,1). In this instance, the model would correctly classify every observation before a single misclassification could occur.

The ROC chart includes two lines to help you interpret the ROC chart. The first line is a baseline model that has a slope of 1. This line mimics a model that correctly classifies observations at the same rate it incorrectly classifies them. An ideal ROC chart maximizes the distance between the baseline model and the ROC chart. A model that classifies more observations incorrectly than correctly would fall below the baseline model. The second line is a vertical line at the false positive rate where the difference between the Kolmogorov-Smirnov values for the ROC chart and baseline models is maximized.

Influence Plot

About the Influence Plot

The Influence Plot displays several measurements that are computed for each observation. When the input data contains a large number of observations, the observations are binned. By default, the **Likelihood Displacement** value is plotted on the X axis. Use these values to help identify outliers and other points that greatly affect the predicted regression model.

Use the Influence Plot

By default, a horizontal bar chart of a selected measurement and a histogram of its values are displayed. To view just the histogram, right-click in the Influence Plot, and select **Use Histogram**.

In either view, you can change the measurement that is plotted by clicking on the horizontal axis.

Fit Statistics

The logistic regression model computes several assessment measures to help you evaluate how well the model fits the data. These assessment measures are available at the top of the model pane. Click the currently displayed assessment measure to see all of the available assessment measures.

-2 Log Likelihood
: The likelihood function estimates the probability of an observed sample given all possible parameter values. The log likelihood is simply the logarithm of the likelihood function. The likelihood function value is -2 times the log likelihood. Smaller values are preferred.

AIC
: Akaike's Information Criterion. Smaller values indicate better models, and AIC values can become negative. AIC is based on the Kullback-Leibler information measure of discrepancy between the true distribution of the response variable and the distribution specified by the model.

AICC

Corrected Akaike’s Information Criterion. This version of AIC adjusts the value to account for sample size. The result is that extra effects penalize AICC more than AIC. As the sample size increases, AICC and AIC converge.

BIC

The Bayesian Information Criterion (BIC), also known as Schwarz’s Bayesian Criterion (SBC), is an increasing function of the model's residual sum of squares and the number of effects. Unexplained variations in the response variable and the number of effects increase the value of the BIC. As a result, a lower BIC implies either fewer explanatory variables, better fit, or both. BIC penalizes free parameters more strongly than AIC.

R-Square

The R-squared value is an indicator of how well the model fits the data. R-squared values can range from 0 to 1. Values closer to 1 are preferred.

Max-rescaled R-Square

The observed R-squared value divided by the maximum attainable R-squared value. This value is useful when there are multiple independent category variables. Values can range from 0 to 1. Values closer to 1 are preferred.

Details Table

When you click at the top of the model pane, the details panel is displayed at the bottom of the model pane. The details table contains the following information:

Dimensions

An overview of the effect variables used in the model. This tab identifies how many measures and classification effects were chosen for the model, the rank of the cross-product matrix, how many observations were read, and how many observations were used in the model.

Iteration History

The function and gradient convergence results. This tab shows at which iteration the function and gradient converged.

Convergence

Provides the reason for convergence.

Fit Statistics

Lists all of the fit statistics described in the previous section.

Type III Test

Provides details for the Type III test. A Type III test examines the significance of each partial effect with all other effects in the model. For more information, see the chapter “The Four Types of Estimable Functions,” in the *SAS/STAT User’s Guide*.

Parameter Estimates

Gives the estimated values for the model parameters.

Response Profile

Displays the event and nonevent counts.

36

Generalized Linear Model

Overview of the Generalized Linear Model

A generalized linear model is an extension of a traditional linear model that allows the population mean to depend on a linear predictor through a nonlinear link function. A generalized linear model requires that you specify a distribution and a link function. The distribution should match the distribution of the response variable. The link function is used to relate the response variable to the effect variables.

The generalized linear model requires a measure response variable and at least one effect variable or interaction term. The distribution imposes range requirements on the measure response variable. These requirements are provided in the following table:

Distribution	Range Requirements
Beta	Values must be between 0 and 1, exclusive
Binary	Two distinct values
Exponential	Nonnegative real values

Distribution	Range Requirements
Gamma	Nonnegative real values
Geometric	Positive integers
Inverse Gaussian	Positive real values
Negative Binomial	Nonnegative integers
Normal	Real values
Poisson	Nonnegative integers

Generalized Linear Model Properties

The following properties are available for the Generalized Linear Model:

Name
enables you to specify the name for this model.

Informative missingness
specifies whether the informative missingness algorithm is used. For more information, see Missing Values on page 251.

Distribution
specifies the distribution used to model the response variable.

Link function
specifies the link function used to relate the linear model to the distribution of the response variable. Available link functions are different for each distribution and are shown in the following table:

Distribution	Available Link Functions
Beta	Logit, Probit, Log-log, C-log-log
Binary	Logit, Probit, Log-log, C-log-log
Exponential	Log, Identity
Gamma	Log, Identity, Recip
Geometric	Log, Identity
Inverse Gaussian	Power(-2), Log, Identity
Negative Binomial	Log, Identity
Normal	Log, Identity
Poisson	Log, Identity

Convergence

- **Override function convergence** enables you to manually specify the function convergence value.
- **Value** specifies the function convergence value when **Override function convergence** is selected. When you specify a larger value, the model will converge sooner. This reduces the amount of time spent training the model, but it can create a suboptimal model.
- **Override gradient convergence** enables you to manually specify the gradient convergence value.
- **Value** specifies the gradient convergence value when **Override gradient convergence** is selected. When you specify a larger value, the model will converge sooner. This reduces the amount of time spent training the model, but it can create a suboptimal model.
- **Maximum iterations** specifies the maximum number of iterations performed during model training. If you specify a relatively small value, you reduce the amount of time spent training the model, but it can create a suboptimal model.

Note: When you specify a gradient convergence or function convergence criterion, it is possible for the model to converge based on an internal convergence criterion before your specified criterion is reached. The reason for convergence is on the **Convergence** tab of the details table.

Assessment

- **Use default number of bins** specifies whether you want to use the default number of bins or to set your own value. By default, measure variables are grouped into 20 bins.
- **Number** specifies the number of bins to use when the **Use default number of bins** property is not selected. You must specify an integer value between 5 and 100.
- **Tolerance** specifies the tolerance value that is used to determine the convergence of the iterative algorithm that estimates the percentiles. Specify a smaller value to increase the algorithmic precision.

Show diagnostic plots

specifies whether the Residual Plot and Assessment windows appear in the model pane.

Working with the Fit Summary Window

About the Fit Summary Window

The Fit Summary window plots the relative importance of each variable as measured by its *p*-value. The *p*-value is plotted on a log scale and the alpha value (plotted as -log(alpha)), is shown as a vertical line. To adjust the alpha value, click, drag, and drop the vertical line. A histogram of the *p*-values is displayed at the bottom of the window.

Including a Group By Variable

When your analysis includes a group by variable, the Fit Summary window displays a Goodness of Fit plot and Variable Importance plot.

The Variable Importance plot enables you to select a single variable to inspect within each level of the group by variable. Use the drop-down menu to select the variable that you want to inspect. Each dot in the Variable Importance plot represents a model effect. The bars are for the model effect chosen in the drop-down menu.

The Goodness of Fit plot, which is not available when there is no group by variable, displays how well the model predicts the response variable within each level of the group by variable. Use this plot to determine whether your model has a significantly different fit within different levels.

Notice that when you select a group by variable level, the Residual Plot, Assessment, and Influence Plot windows are updated. This enables you to further investigate each level of the group by variable.

Use to specify how the plot is sorted.

Working with the Residual Plot

About Residual Plots

A residual plot shows the relationship between the predicted value of an observation and the residual of an observation. The residual of an observation is the difference between the predicted response value and the actual response value. When using large data sets, the residual plot is displayed as a heat map instead of as an actual plot. In a heat map, the actual observations are binned, and the color of each point indicates the relative number of observations in that bin.

Residual plots have several uses when examining your model. First, obvious patterns in the residual plot indicate that the model might not fit the data. Second, residual plots can detect nonconstant variance in the input data when you plot the residuals against the predicted values. Nonconstant variance is evident when the relative spread of the residual values changes as the predicted values change. Third, in combination with other methods, the residual plot can help identify outliers in your data.

To change the residual that is plotted, click on the vertical axis.

By default, a scatter plot or heat map is displayed. However, a histogram is also available. To display a histogram of the residuals, right-click in the **Residual Plot**, and select **Use Histogram**. In the histogram, you can change the residual that is plotted by clicking on the horizontal axis.

From the histogram, it is easy to determine whether the distribution of the residuals is approximately normal or skewed. A non-normal residual histogram can indicate that the model does not fit the data.

Use the Residual Plot

The Residual Plot provides several features to investigate your data. These features are available in both the scatter plot and histogram. The process to activate these features is the same in either plot. The available features are plot the residuals by a category variable, show only the selected observations, exclude selected observations, and remove all exclusions.

There are two ways to select observations in the Residual Plot. If you hold the mouse over an individual point, you can click on that point to select it. Alternatively, when you click and drag the mouse in the Residual Plot, a selection rectangle appears. When you release the mouse, all points inside the rectangle are selected.

To plot the residuals by a category variable, right-click in the Residual Plot, select **Plot By**, and then select a category variable. A box plot of the residuals appears. The residuals are grouped by the selected category variable.

To show selected observations, select some data points, right-click in the Residual Plot, and select **Show Selected**. A table of the selected observations appears.

To exclude selected observations, select some data points, right-click in the Residual Plot, and select **Exclude Selected**. The selected observations are no longer used to build the model.

To remove all exclusions, right-click anywhere in the Residual Plot, and select **Remove Exclusions**. All observations are now used to build in the model.

Working with the Assessment Window

About the Assessment Window

For a Generalized Linear Model, the Assessment window plots the average predicted and average observed response values against the binned data. Use this plot to reveal any strong biases in your model. Large differences in the average predicted and average observed values can indicate a bias.

Use the Assessment Window

The Assessment window bins the data based on the values specified in the **Assessment** properties. At each bin, you can hold the mouse over one or both of the lines to display a tooltip.

Fit Statistics

The Generalized Linear Model computes several assessment measures to help you evaluate how well your model fits the data. These assessment measures are available at the top of the model pane. Click the currently displayed assessment measure to see all available assessment measures. The available assessment measures are the following:

-2 Log Likelihood
The likelihood function estimates the probability of an observed sample given all possible parameter values. The log likelihood is simply the logarithm of the likelihood function. This value is -2 times the log likelihood. Smaller values are preferred.

AIC
Akaike's Information Criterion. Smaller values indicate better models. AIC values should be compared only when two models have an approximately equal number of observations. AIC values can become negative. AIC is based on the Kullback-Leibler information measure of discrepancy between the true distribution of the response variable and the distribution specified by the model.

AICC
Corrected Akaike's Information Criterion. This version of AIC adjusts the value to account for a relatively small sample size. The result is that extra effects penalize AICC more than AIC. As the sample size increases, AICC and AIC converge.

BIC
The Bayesian Information Criterion (BIC), also known as Schwarz's Bayesian Criterion (SBC), is an increasing function of the model's residual sum of squares and the number of effects. Unexplained variations in the response variable and the number of effects increase the value of the BIC. As a result, a lower BIC implies either fewer explanatory variables, better fit, or both. BIC penalizes free parameters more strongly than AIC.

Details Table

When you click at the top of the model pane, the details panel is displayed at the bottom of the model pane. The details table contains the following information:

Dimensions
An overview of the effect variables used in the model. This tab identifies how many measures and classification effects were chosen for the model, the rank of the cross-product matrix, how many observations were read, and how many observations were used in the model.

Iteration History
The function and gradient iteration results. This tab shows the value of the objective (likelihood) function, its change in value, and its maximum gradient.

Convergence
Provides the reason for convergence.

Fit Statistics
Lists all of the fit statistics described in the previous section.

Type III Test
Provides details for the Type III test. A Type III test examines the significance of each partial effect with all other effects in the model. For more information,

see the chapter “The Four Types of Estimable Functions,” in the *SAS/STAT User’s Guide*.

Parameter Estimates
Gives the estimated values for the model parameters.

37

Decision Trees

Overview of the Decision Tree

A decision tree creates a hierarchical segmentation of the input data based on a series of rules applied to each observation. Each rule assigns an observation to a segment based on the value of one predictor. Rules are applied sequentially, which results in a hierarchy of segments within segments. The hierarchy is called a tree, and each segment is called a *node*. The original segment contains the entire data set and is called the *root node*. A node and all of its successors form a *branch*. The final nodes are called *leaves*. For each leaf, a decision is made about the response variable and applied to all observations in that leaf. The exact decision depends on the response variable.

The decision tree requires a measure response variable or category response variable and at least one predictor. A predictor can be a category or measure variable, but not an interaction term.

The decision tree enables you to manually train and prune nodes by entering interactive mode. In interactive mode, you are unable to modify the response variable, growth properties are locked, and you cannot export model score code. Certain modifications to predictors are allowed, such as converting a measure to a category. When you are in interactive mode and modify a predictor, the decision tree remains in interactive mode, but attempts to rebuild the splits and prunes using the same rules.

To enter interactive mode, you can either start making changes to the decision tree in the Tree window or you can click **Use Interactive Mode** on the **Roles** tab in the right pane. To leave interactive mode, click **Use Non-Interactive Mode** on the **Roles** tab.

Note: When you leave interactive mode, you lose all of your changes.

Decision Tree Properties

The following properties are available for the decision tree:

Name
enables you to specify the name for this model.

Maximum branches
specifies the maximum number of branches allowed when splitting a node.

Maximum levels
specifies the maximum depth of the decision tree.

Leaf size
specifies the minimum number of observations allowed in a leaf node.

Response bins
specifies the number of bins used to categorize a measure response variable.

Predictor bins
specifies the number of bins used to categorize a predictor that is a measure variable.

Pruning
specifies the aggressiveness of the tree pruning algorithm. A more aggressive algorithm creates a smaller decision tree. Larger values are more aggressive.

Rapid growth
enables you to use the information gain ratio and k-means fast search methods for decision tree growth. When disabled, the information gain and greedy search methods are used, which generally produce a larger tree and require more time to create.

Include missing
enables you to include observations with missing values. For category variables, a missing value is assigned to its own level. For measure variables, a missing value is assigned to the smallest available machine value (negative infinity).

Reuse predictors
allows more than one split in the same branch based on a predictor.

Frequency
specifies whether nodes report how many observations they contain or what percentage of the observations they contain.

Assessment

- **Use default number of bins** specifies whether you want to use the default number of bins or to set your own value. By default, measure variables are grouped into 20 bins.

- **Number** specifies the number of bins to use when the **Use default number of bins** property is not selected. You must specify an integer value between 5 and 100.
- **Prediction cutoff** specifies the value at which a computed probability is considered an event.
- **Tolerance** specifies the tolerance value that is used to determine the convergence of the iterative algorithm that estimates the percentiles. Specify a smaller value to increase the algorithmic precision.

Show diagnostic plots
specifies whether the Leaf Statistics and Assessment windows appear in the model pane.

Show tree overview
displays the tree overview. The tree overview enables quick navigation of large decision trees. When you zoom in to view a specific area of the decision tree, the tree overview shows the entire decision tree and highlights the area that you are viewing. You can click and drag the highlighted area to change the display of the decision tree. Click  in the upper left corner of the tree overview to view the entire decision tree. Click  in the upper left corner of the tree overview to minimize the tree overview.

Information Gain and Gain Ratio Calculations

When the **Rapid growth** property is enabled, node splits are determined in part by information gain ratio instead of information gain. The information gain and information gain ratio calculations and their benefits and drawbacks are explained in this section. In these explanations, an attribute is considered any specific measurement level of a classification variable or bin of a measure variable.

The information gain method chooses a split based on which attribute provides the greatest information gain. The gain is measured in bits. Although this method provides good results, it favors splitting on variables that have a large number of attributes. The information gain ratio method incorporates the value of a split to determine what proportion of the information gain is actually valuable for that split. The split with the greatest information gain ratio is chosen.

The information gain calculation starts by determining the information of the training data. The information in a response value, *r*, is calculated in the following expression:

$$-\log_2\left(\frac{freq(r, T)}{|T|}\right)$$

T represents the training data and |*T*| is the number of observations. To determine the expected information of the training data, sum this expression for every possible response value:

$$I\left(T\right) = -\sum_{i=1}^{n} \frac{freq(r_i, T)}{|T|} \times \log_2\left(\frac{freq(r_i, T)}{|T|}\right)$$

Here, n is the total number of response values. This value is also referred to as the *entropy* of the training data.

Next, consider a split S on a variable X with m possible attributes. The expected information provided by that split is calculated by the following equation:

$$I_S\left(T\right) = \sum_{j=1}^{m} \frac{|T_j|}{|T|} \times I\left(T_j\right)$$

In this equation, T_j represents the observations that contain the j^{th} attribute.

The information gain of split S is calculated by the following equation:

$$G(S) = I(T) - I_S(T)$$

Information gain ratio attempts to correct the information gain calculation by introducing a split information value. The split information is calculated by the following equation:

$$SI\left(S\right) = -\sum_{j=1}^{m} \frac{|T_j|}{|T|} \times \log_2\left(\frac{|T_j|}{|T|}\right)$$

As its name suggests, the information gain ratio is the ratio of the information gain to the split information:

$$GR\left(S\right) = \frac{G(S)}{SI(S)}$$

Pruning

The **Pruning** property of the decision tree visualization determines how aggressively your decision tree is pruned. The growth algorithm creates a decision tree based on the properties that you specify. The pruning algorithm considers each node to be a root node of its own subtree, starting from the bottom. If the misclassification rate of the subtree is significantly better than the misclassification rate of the root node, then the subtree is kept. If the misclassification rate of the subtree is similar to the misclassification rate of the root node, then the subtree is pruned. In general, smaller decision trees are preferred.

If the **Pruning** property slider is closer to **Lenient**, then the difference in the misclassification rates must be relatively small. If the **Pruning** property is closer to **Aggressive**, then the difference in the misclassification rates must be relatively large. That is, a lenient pruning algorithm allows the decision tree to grow much deeper than an aggressive pruning algorithm.

Variables that are not used in any split can still affect the decision tree, typically due to one of two reasons. It is possible for a variable to be used in a split, but the subtree that contained that split might have been pruned. Alternatively, the variable might include missing values, but the **Include missing** property is disabled.

Note: If a predictor does not contribute to the predictive accuracy of the decision tree or the contribution is too small, then it is not included in the final, displayed decision tree.

Working with the Tree Window

The Tree window contains the decision tree, tree overview, and icicle plot.

> **TIP** To navigate the decision tree, you can use the mouse and keyboard. Hold down the **Shift** key and click anywhere in the Tree window to move the decision tree within the window. Use your mouse's scroll wheel to zoom in and out of the decision tree. Scroll up to zoom in, and scroll down to zoom out. The zoom is centered on the position of your cursor.

The color of the node in the icicle plot indicates the predicted level for that node. When you select a node in either the decision tree or the icicle plot, the corresponding node is selected in the other location. When you select a leaf node, that node is selected in the Leaf Statistics window. A legend is available at the bottom of the model pane.

When the response variable is a measure variable, a gradient is used to denote the predicted bin. Darker colors represent larger values.

Right-click outside of a node in the Tree window to open a pop-up menu. The first item in this menu is **Derive a Leaf ID Variable**. When you click this item, SAS Visual Statistics creates a category variable that contains the leaf ID for each observation. You can use this variable as an effect in other models.

Right-click inside a node to open a different pop-up menu. The available menu options depend on whether you clicked a leaf node.

For leaf nodes, you can select from the following menu options:

Split
: opens the Split Decision Tree window. Use this window to select the variable that is used to split the node. Click **OK** to split the node based on the selected variable. Click **Cancel** to not split the node. Variables are sorted in descending order by their log worth.

  Some variables are not available for a split if the value of the split is too small or the split would violate the **Leaf size** property.

Split Best
: splits the node based on the variable with the best information gain ratio when **Rapid growth** is enabled. In addition, splits the node based on the variable with the best information gain when **Rapid growth** is disabled.

Train
: opens the Train Decision Tree window. Use this window to train more than one level beyond the leaf node. First, select every variable that you want to be available for training. Only those variables selected in the Train Decision Tree window are available for training. Specify the maximum depth of training in the **Maximum depth of subtree** property. Click **OK** to train the decision tree.

For other nodes, select **Prune** to remove all nodes that follow the selected node. This turns the selected node into a leaf node. After pruning a node, you can select **Restore** to undo the prune.

Working with the Leaf Statistics Window

The Leaf Statistics window plots the percentage of each observation in each leaf node. The most common level in a node is the predicted value assigned to that node. Leaf nodes that contain approximately equal amounts of more than one level might benefit from additional training.

When you select a column in the Leaf Statistics window, the corresponding leaf is selected in the Tree window.

Working with the Assessment Window

About the Assessment Window

For a decision tree, the Assessment window plots lift, ROC, and misclassification rates. Use the Assessment window to determine how well the model fits the data.

Lift is the ratio of the percent of captured responses within each percentile bin to the average percent of responses for the model. Similarly, *cumulative lift* is calculated by using all of the data up to and including the current percentile bin.

A receiver operating characteristic (ROC) chart displays the ability of a model to avoid false positive and false negative classifications. A false positive classification means that an observation has been identified as an event when it is actually a nonevent (also referred to as a Type **I** error). A false negative classification means that an observation has been identified as a nonevent when it is actually an event (also referred to as a Type **II** error).

The misclassification plot displays how many observations were correctly and incorrectly classified for each value of the response variable. When the response variable is not binary, the logistic regression model considers all levels that are not events as equal. A significant number of misclassifications could indicate that your model does not fit the data.

Use the Assessment Window

Lift

The default lift chart displays the cumulative lift of the model. To view the noncumulative lift, click on the vertical axis, and select **Lift**.

For comparison, the lift chart plots a best model based on complete knowledge of the input data.

ROC

The *specificity* of a model is the true negative rate. To derive the false positive rate, subtract the specificity from 1. The false positive rate, labeled **1 – Specificity**, is the X axis of the ROC chart. The *sensitivity* of a model is the true positive rate. This is the Y axis of the ROC chart. Therefore, the ROC chart plots how the true positive rate changes as the false positive rate changes.

A good ROC chart has a very steep initial slope and levels off quickly. That is, for each misclassification of an observation, significantly more observations are correctly classified. For a perfect model, one with no false positives and no false negatives, the ROC chart would start at (0,0), continue vertically to (0,1), and then horizontally to (1,1). In this instance, the model would correctly classify every observation before a single misclassification could occur.

The ROC chart includes two lines to help you interpret the ROC chart. The first line is a baseline model that has a slope of 1. This line mimics a model that correctly classifies observations at the same rate it incorrectly classifies them. An ideal ROC chart maximizes the distance between the baseline model and the ROC chart. A model that classifies more observations incorrectly than correctly would fall below the baseline model. The second line is a vertical line at the false positive rate where the difference between the Kolmogorov-Smirnov values for the ROC chart and baseline models is maximized.

Misclassification

The misclassification plot displays how many observations were correctly and incorrectly classified. A significant number of misclassifications might indicate that the model does not fit the data.

When the ratio of events to non-events in your data is relatively large, the misclassification plot might show a large number of true positives and false positives. In this case, your model predicts most observations as events and is correct more often than not.

Assessment

When the number of **Response bins** is set to more than 10, the Assessment window plots the predicted average and observed average values. Use this plot to determine how well the model fits the data.

The Assessment window bins the data based on the values specified in the **Assessment** properties. At each bin, you can hold the mouse over one or both of the lines to display a tooltip.

Details Table

When you click at the top of the model pane, the details panel is displayed at the bottom of the model pane. The details table contains the following information:

Node Statistics
: provides summary statistics for each node in the decision tree. Available statistics include **Depth**, **Parent ID**, **N Children**, **Type**, **Observations**, **% Observations**, **N Missing**, **Gain**, **Predicted Value**, **Split**, and the number and percentage of observations in each bin.

Node Rules
: provides the sorting rule used for each node in the decision tree. Every available variable is listed as a column in the table. If a rule was applied for a

variable in a node or any of its parent nodes, then it is listed in the table. Otherwise, the entry is blank.

38

Clustering

Overview of the Cluster Tool

Clustering is a method of data segmentation that puts observations into groups that are suggested by the data. The observations in each cluster tend to be similar in some measurable way, and observations in different clusters tend to be dissimilar. Observations are assigned to exactly one cluster. From the clustering analysis, you can generate a cluster ID variable to use in other explorations.

The cluster tool requires at least two measure variables as input. You cannot specify an interaction term or category variable.

Cluster Properties

The following properties are available for the cluster tool:

Name
enables you to specify the name for this model.

Cluster Matrix

- **Number of clusters** specifies the number of clusters that are generated.
- **Seed** specifies the seed value of the random number generator that is used during initial cluster assignments.
- **Initial assignment** specifies the method that is used to create the initial cluster assignments. The available methods are:

- **Forgy** specifies that *k* data points are selected at random to use as the centroids of the *k* clusters.
- **Random** assigns observations to a cluster at random.

- **Visible roles** determines how many effects are shown in the Cluster Matrix. Valid values are integers between 2 and 6, inclusive.

  When you specify a value *n*, the first *n* effects listed in the **Variables** table on the **Roles** tab are displayed. To change the effect pairs that are plotted in the Cluster Matrix, you can remove an effect from the analysis, and then immediately add it back in. The clustering results remain unchanged because you are using the same input data. However, the **Variables** table adds new effects to the bottom of the list.

- **Variable standardization** transforms the effect variables so that they have a mean of zero and a standard deviation of 1. This property is enabled by default and affects the results displayed in the details table. The Cluster Matrix window and the Parallel Coordinates window display the original variables.

Parallel Coordinates

- **Number of bins** specifies the number of bins used when generating the parallel coordinate polyline plots.
- **Maximum polylines** specifies the maximum number of polylines generated by the parallel coordinate algorithm.
- **Visible roles** determines how many effects are shown in the Parallel Coordinates plot. Valid values are integers between 2 and 10, inclusive.

Show ellipses
enables you to display the cluster projection ellipses in the Cluster Matrix.

Show centroids
enables you to display the centroids in the Cluster Matrix.

Working with the Cluster Matrix Window

About the Cluster Matrix Window

The Cluster Matrix displays a two-dimensional projection of each cluster onto a specified number of effect pairs. These projections are useful for spotting cluster similarities and differences within the plotted effect pairs.

Each cluster is assigned a unique color. Although each cluster is unique in *n*-space, the two-dimensional projections will overlap. It is important to note that every observation can belong to exactly one cluster. However, because the Cluster Matrix displays a projection in just two dimensions, multiple clusters can overlap an observation.

When a heat map is not used, individual observations are color-coded to indicate cluster membership.

Use the Cluster Matrix Window

To view a larger plot for an effect pair, right-click inside that plot, and click **Explore**. In the Explore window, it is easier to view and select observations.

To view a box plot for a variable that segments the observations by cluster, right-click inside a plot that contains the variable of interest, and select **Plot *variable_name* by Cluster ID**. Each variable in the selected plot has a menu item. The box plot is used to determine how similar the clusters are for the selected variable.

To derive a cluster ID variable that can be used in other explorations, right-click inside a plot and select **Derive a Cluster ID Variable**. When you select this item, SAS Visual Statistics creates a category variable that contains the cluster ID for each observation. Cluster ID -1 is reserved for observations with missing values.

Working with the Parallel Coordinates Plot

About the Parallel Coordinates Plot

The Parallel Coordinates plot shows patterns in the data and clusters. In this plot, the cluster ID is on the far left, and each variable is a column with its binned range of values displayed vertically. Color-coded polylines are drawn from each cluster and show which range of values the cluster contains for every variable.

Use the Parallel Coordinates Plot

You can use the Parallel Coordinates plot to make several inferences. Adjust the plot to explore the data based on cluster membership, a specified range for one or more variables, or both.

To view just the polylines for a single cluster, select that cluster ID on the far left. Notice that the polylines for all other clusters are grayed out. This enables you to focus on a single cluster. To view multiple clusters, hold the **Control** key and select the clusters that you want to view.

To view the Parallel Coordinates plot by a variable, click the button that corresponds to that variable. This action changes the color gradient of the polylines so that larger values are darker than smaller values. To undo the variable selection, click anywhere to the left of **Cluster ID**.

To restrict the number of bins displayed for a variable, click the top or bottom of a variable’s range and drag the selection rectangle to the desired range. To move the selection rectangle, click inside the rectangle, and drag the rectangle upward or downward. You can repeat this process for other variables. To undo the selection for a variable, click the name of that variable at the top of the Parallel Coordinates plot.

Combining these two features, you can restrict the display to specific clusters and variable ranges that interest you.

Details Table

When you click at the top of the model pane, the details table is displayed at the bottom of the model pane. The details table contains the following information:

- **Cluster Summary** provides summary statistics for each cluster. Available statistics include **Observations**, **RMS of STD**, **Within-Cluster SS**, **Min centroid-to-observation**, **Max centroid-to-observation**, **Nearest Cluster**, and **Centroid Distance**.

39

Model Comparison

Overview of Model Comparison

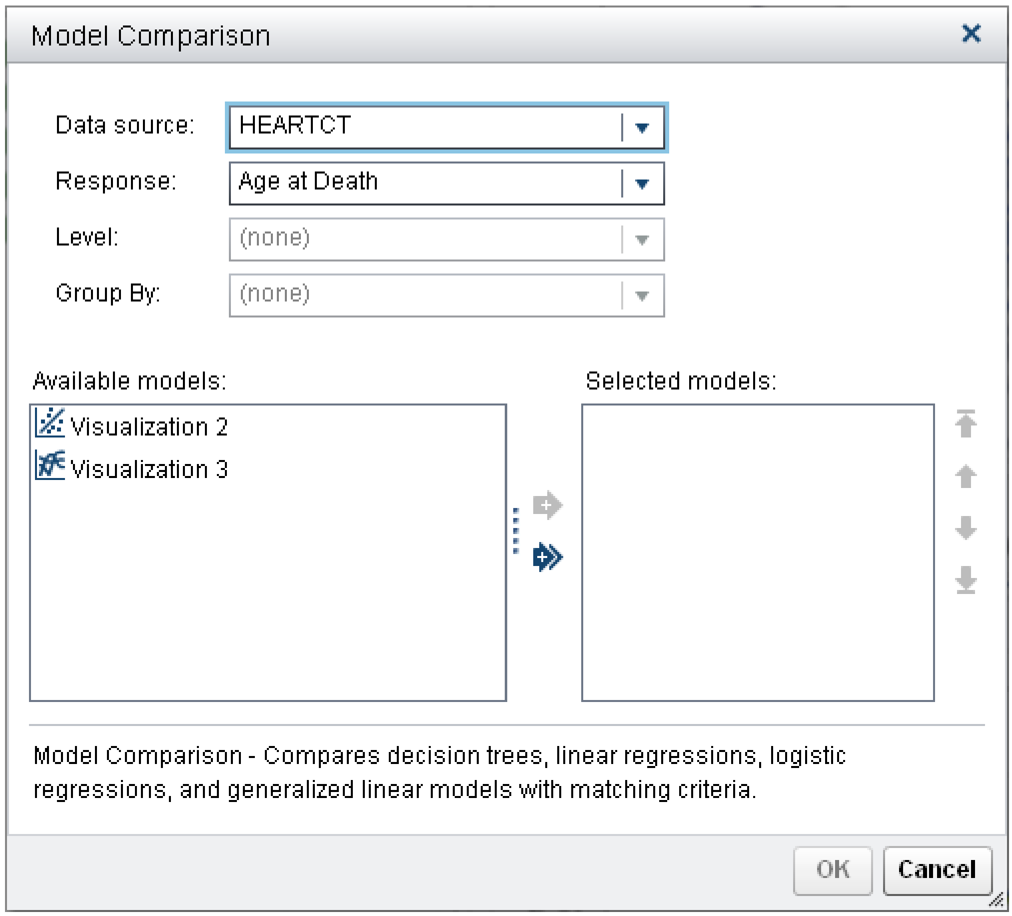
The model comparison tool enables you to compare the performance of competing models using various benchmarking criteria. The comparison criteria available depends on the models and response variable used in your analysis. A model comparison requires that at least one model is trained before you can perform a comparison.

Before performing a model comparison, ensure that all models are initialized and updated. If the **Auto-update model** property is disabled for a model, you must manually update it before you can compare it to another model. A model is not considered initialized until it has been trained.

When you change a model after a comparison has been created, changes are not carried over to the model comparison.

Model Comparison Usage

When you click in the toolbar, the Model Comparison window appears.



The Model Comparison window enables you to specify the response variable of interest, the level of interest, a group by variable, and the models for comparison. You must specify a response variable and at least two models.

Note: You are able to compare two or more models only when the response variable, level of interest, and group by variable are identical.

Model Comparison Properties

The following properties are available for model comparison:

Name
enables you to specify the name for this comparison.

Fit statistic
specifies the comparison criterion that is plotted in the Fit Statistic window and used to determine the champion model. The fit statistics available depend on the models being compared.

For the error sum of squares (SSE) fit statistic, the linear regression model and logistic regression model use the weighted SSE. The generalized linear model uses the unweighted SSE.

Prediction Cutoff
specifies the cutoff probability that determines whether an observation is a modeled event.

Percentile
when available, specifies the percentile at which the specified fit statistic is plotted.

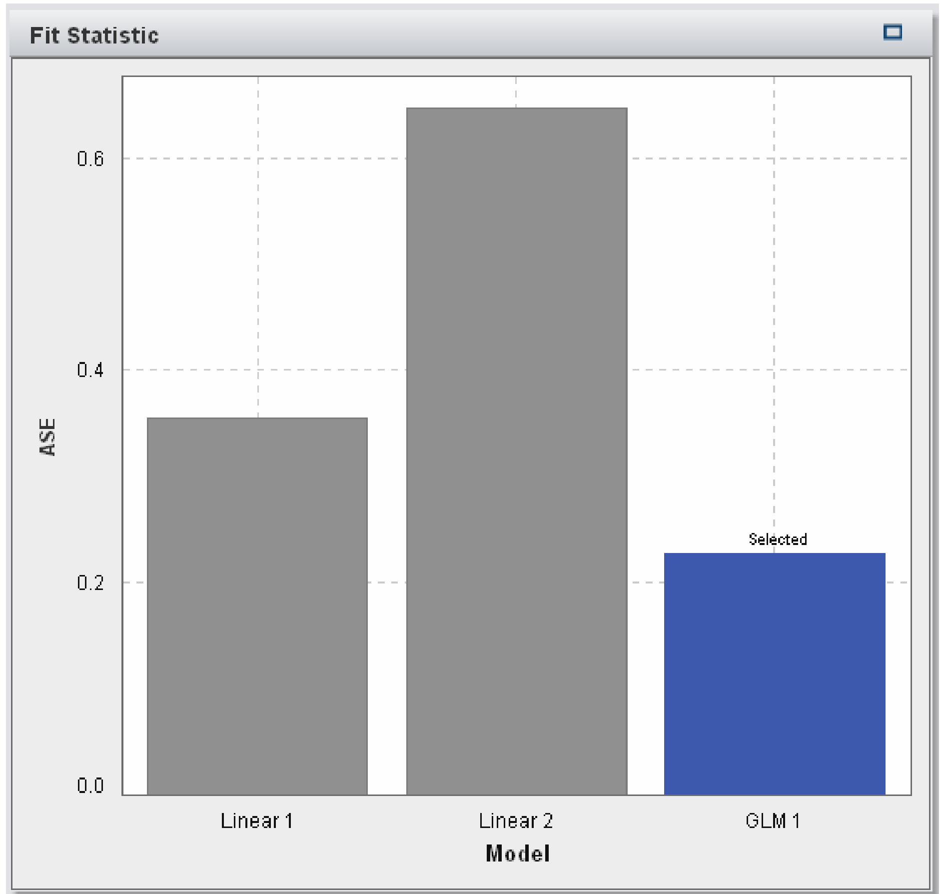
Model Comparison Results Windows

Assessment

The assessment plots available depend on the models being compared. For classification models, the plots displayed are Lift, ROC, and Misclassification. For numerical models, the plots displayed are observed response value and predicted response value.

Fit Statistic

The Fit Statistic plot displays the criterion specified in the **Fit statistic** property. In the following image, the observed average value is plotted for a linear regression and a GLM model. The champion model is indicated in the plot. It is displayed different from the other models.



Details Table

When you click at the top of the model pane, the details panel is displayed at the bottom of the model pane. The details table contains the following information:

Statistics
Provides summary statistics for each model in the comparison. The value in the **Selected** column, either **Yes** or **No**, indicates which model the model comparison tool prefers based on the criterion specified in the **Fit statistic** property. However, the statistics listed in the details table can differ from those listed in the **Fit statistic** property.

Variable Importance
Indicates which variables had the greatest impact on each of the models in the comparison.

Part 6

Designing Reports

40

Overview of SAS Visual Analytics Designer

About the SAS Visual Analytics Designer

The SAS Visual Analytics Designer (the designer) enables users to easily create reports or dashboards that can be saved and viewed on either a mobile device or in the viewer. The designer is part of the SAS Visual Analytics product that enables a user with either the **Visual Analytics: Analysis** role or the **Visual Analytics: 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 designers 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, images, stored processes, and controls to reports. All of this is accomplished by using the designer, which runs in a web browser. You do not need to understand a programming language to create reports.

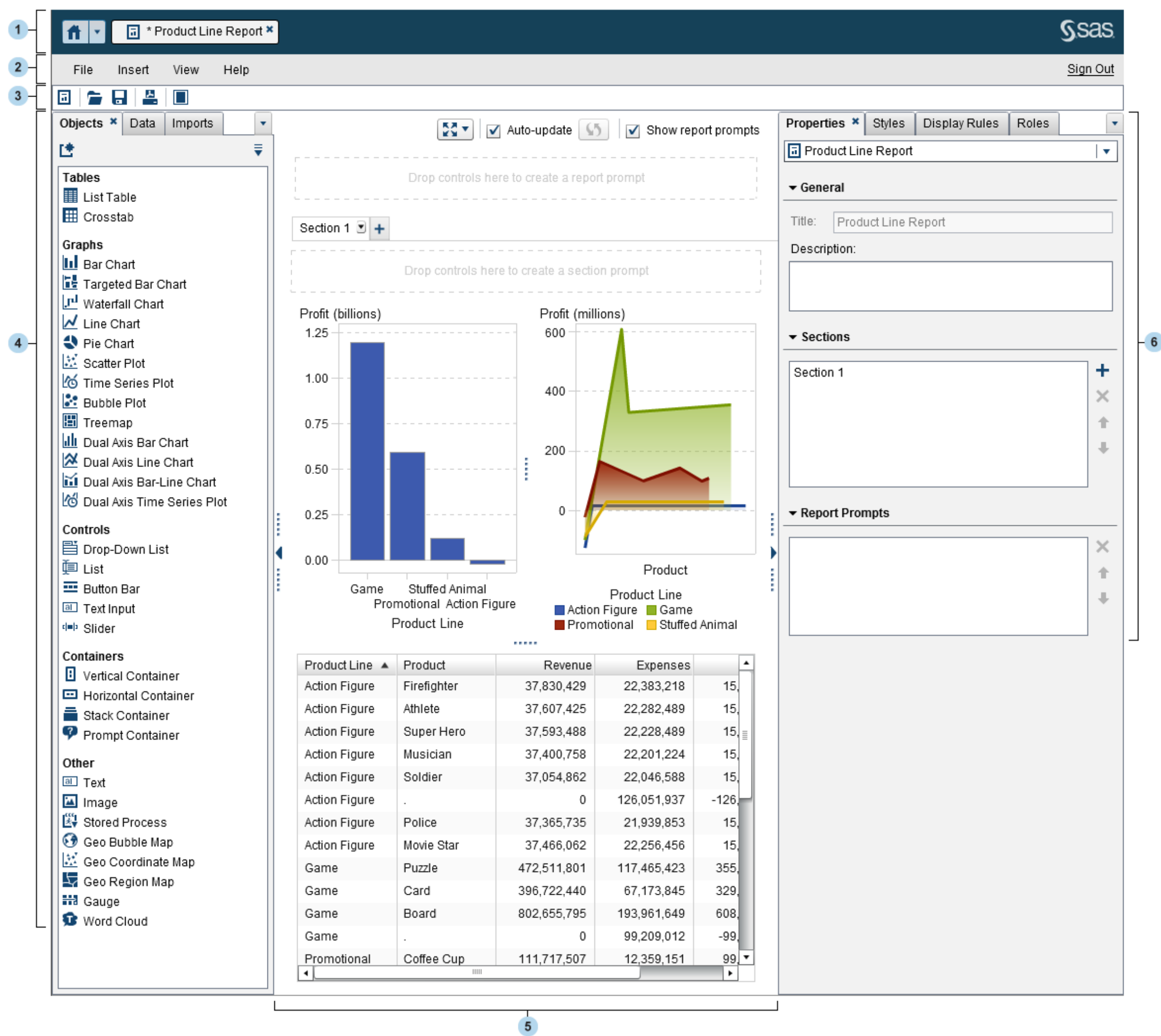
Report designers can easily create reports and dashboards based on data sources that have been provided by a system administrator. They can update reports that were created from visual explorations. Report designers can create reports by importing report objects or visuals from other reports. They can also define interactions (either filtering or brushing) for report objects, and then include SAS analytical results in a single report.

Your First Look at the Designer

When you are designing reports in SAS Visual Analytics, you see the designer. In general, data-related tasks are initiated from the leftmost pane, and presentation-related tasks are initiated from the rightmost pane.

Here are the features of the designer:

Figure 40.1 *The Designer*



Product Line	Product	Revenue	Expenses	
Action Figure	Firefighter	37,830,429	22,383,218	15,
Action Figure	Athlete	37,607,425	22,282,489	15,
Action Figure	Super Hero	37,593,488	22,228,489	15,
Action Figure	Musician	37,400,758	22,201,224	15,
Action Figure	Soldier	37,054,862	22,046,588	15,
Action Figure	.	0	126,051,937	-126,
Action Figure	Police	37,365,735	21,939,853	15,
Action Figure	Movie Star	37,466,062	22,256,456	15,
Game	Puzzle	472,511,801	117,465,423	355,
Game	Card	396,722,440	67,173,845	329,
Game	Board	802,655,795	193,961,649	608,
Game	.	0	99,209,012	-99,
Promotional	Coffee Cup	111,717,507	12,359,151	99,

1 The application bar enables you to return to the home page and to access other parts of SAS Visual Analytics and other SAS applications that integrate with the home page. You can access your recently created or viewed reports, explorations, stored processes, data queries, or other objects in your recent history. Buttons are displayed for each open application.

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, adding interactions, and launching the viewer without returning to the home page. You can also sign out of SAS Visual Analytics.

3 The toolbar contains icons that enable you to manage and print 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 tabs in the left pane enable you to work with new report objects, data, imported report objects, and shared rules.

5 The canvas is the workspace for building a report. The appearance of the canvas is affected by the report theme.

6 The tabs in the right pane enable you to work with details about the report and report objects.

Understanding the Tabs in the Designer

Tabs in the Left Pane

The **Objects**, **Data**, **Imports**, and **Shared Rules** tabs are in the left pane. The **Objects**, **Data**, and **Imports** tabs are displayed by default. You can choose to display icons instead of names on the tabs. To display the icons on the tabs, click after the last tab name, and select **Show icons only**. From this menu, you can also choose which tabs are displayed.

The following table lists the available tabs:

Tab	Description
Objects	This tab provides a list of the tables, graphs, gauges, controls, containers, other objects, and custom objects that can be used in a report or dashboard. **Note:** You can use the Show or Hide Objects window to specify which report objects are displayed on the **Objects** tab. For more information, see "Show or Hide Report Objects in the Objects Tab" on page 321.

Tab	Description
Data	This tab enables you to select a data source (or data sources) and the data items for your report. You can add, refresh, import, or remove a data source using the icons above the list of data items. Using the menu, you can change the data source, define a hierarchy, define a calculated item, define an aggregated measure, show or hide data items, and filter on a data item. You can check the details for the measures in the data set. You can use this tab to add parameters to a filter, a calculated item, a display rule, or a rank. The data item table on the **Data** tab provides information about a selected item, including the name, classification, format, aggregation, sort options, and parameters. These data item properties can be modified, which impacts all of the report objects that use the data item. When a report has multiple data sources, the **Data** tab provides information about the data source and data items for the selected report object. When you select a report object that has a different data source, the **Data** tab updates automatically to provide information about the selected data source.
Imports	This tab provides a list of reports and report objects that have been created in the designer or exported from the explorer. This enables you to create reports from multiple data sources and create full-featured reports. You can choose objects or sections in these reports to include in either a new report or an existing report.
Shared Rules	This tab enables you to create a new display rule for a gauge, which is used by other gauges to designate intervals and colors for ranges. You can edit or delete an existing shared display rule. These display rules are shared across multiple gauges and can be created at any time.

Tabs in the Right Pane

The **Properties**, **Styles**, **Display Rules**, and **Roles** tabs are displayed by default. You can choose to display icons instead of tab names on the tabs. To display the icons on the tabs, click after the last tab name, and select **Show icons only**. From this menu, you can also choose which tabs are displayed.

The following table lists the available tabs:

Tab	Description
Properties	This tab lists the properties for the currently selected report, section, info window, or report object. ■ When a report is selected, the report's title and description are listed. ■ When a section is selected, the section's name and layout are listed, as well as any section prompts. When an info window is selected, the name and layout, any section prompts, and the size of the window are listed. For more information, see "Overview of Report Sections and Info Windows" on page 471. ■ When a report object is selected, the object's name, title, and description are listed. Information that is specific to the report object type, such as the axes and legend, is listed. ■ If a report object is not selected on the canvas, then the properties are for the section, info window, or report.
Styles	This tab enables you to specify the data styling, frame styling, text styling, and data colors for a selected report object. You can specify the report theme for the report.
Display Rules	This tab enables you to populate intervals, add intervals, or add color-mapped values to an object that is currently selected on the canvas. You can edit or delete an existing display rule for the selected table, graph, gauge, or control.
Roles	This tab enables you to add or update data role assignments in a selected report object that has a data source.
Alerts	This tab enables you to add, update, or delete alerts for report objects.
Comments	This tab enables you to add comments to the report after you save it.
Filters	This tab enables you to add a filter (or filters) to the selected report object.
Interactions	This tab enables you to create interactions and links.
Ranks	This tab enables you to add rankings to report objects. For example, you might want to see the top 10 products that sold last year.

About the Canvas in the Designer

The canvas is the workspace for building a report.

You can change the report view by clicking the icons at the top of the canvas. You can access the report views by clicking beside the icon. For more information about report views, see “Choose a Report View” on page 309.

Above the canvas, there is an **Auto-update** check box. You can clear this check box so that you can design your report without waiting for a query to finish. When you are finished designing your report, select the **Auto-update** check box to run and refresh all of the queries.

There is an area at the top of the canvas on which you can drag and drop filter controls and then categories to create report-level prompts. If you do not want to see this area, clear the **Show report prompts** check box. For more information, see “Use a Control to Create a Report Prompt” on page 334.

A report can have multiple sections or info windows, which can be accessed by using the tabs at the top of the canvas.

There is an area below the section tabs on which you can drag and drop filter controls and then categories to create section prompts. For more information, see “Use a Control to Create a Section Prompt” on page 335.

About Report Themes

Report themes are available in the designer. You can select a report theme for each report using the **Styles** tab. Alternatively, you can set a preference for a default report theme. For more information, see “Specifying Your Preferences for the Designer” on page 305.

The designer provides the following report themes: **SAS Snow**, **SAS Light**, **SAS Dark**, or **SAS High Contrast**. Starting in the 7.2 release, SAS Snow is the default report theme for all new reports. If you want existing reports to use the SAS Snow report theme, you can change the report theme using the **Styles** tab.

Your site might also have custom report themes. Custom report themes are automatically created when a coordinating custom application theme is created in SAS Theme Designer for Flex. For more information, see “About Application Themes in SAS Visual Analytics” on page 9.

You can use the **Styles** tab in the designer to override the default report background color, prompt background color, font, and font color for the report theme. You can use the **Reset Theme** button on the **Styles** tab to reset the default report style when you have overridden the background color, font, or font color.

The **Customize Theme** button is displayed on the **Styles** tab for users who have the Customize Themes capability (through a direct or indirect role membership). For more information about application themes, see *SAS Theme Designer for Flex: User’s Guide*.

Specifying Your Preferences for the Designer

Specify Global Settings

If you receive distributed reports, it is recommended that you specify the **User locale** setting. The report distribution feature does not have access to the browser locale, so it depends on the locale specified for the **User locale** in the modern home page settings. For more information, see "Specify Global Settings Using SAS Home" on page 10.

TIP If you change the **User locale** setting or preference, then you must sign out and sign in to SAS Visual Analytics for the change to take effect.

For information about global preferences for the classic home page, see "Specifying Your Preferences for the Classic SAS Visual Analytics Home Page" on page 637.

Specify General Preferences for the Designer

To specify general preferences for how alert notifications are sent from both the designer and the viewer:

1 Select **File** ▸ **Preferences** to open the Preferences window.

2 Click **General** in the left pane.

3 Specify your **Alert Notifications** preference. The options are **Send e-mail messages**, **Send SMS text messages**, or **Use the system default for alert notifications**.

   Note: If your preference for alert notifications is set to e-mail when you add an alert, you will always receive notifications for that alert via e-mail. If you decide later that your preference for alert notifications is a text message, then after you change your preference to **Send SMS text messages**, you will have to delete your existing alerts and create new ones.

   The **Use the system default for alert notifications** option specifies that you want to use the system default that is set by your system administrator in the SAS Preferences Manager. For more information, see the SAS Preferences Manager topic in *SAS Intelligence Platform: Middle-Tier Administration Guide*. When you select either **Send e-mail messages** or **Send SMS text messages**, you override the system default.

   TIP If you select the **Send SMS text messages** preference, and do not receive alerts via text message, contact your system administrator. Your mobile number must be set up correctly as an **SMS** type in SAS Management Console.

To specify general preferences for SAS Visual Analytics, see "Specify Settings Using SAS Home" on page 10.

Specify Preferences for the Designer

To specify preferences that are specific to the designer:

1 Select **File** ▸ **Preferences** to open the Preferences window.

2 Click **SAS Visual Analytics Designer** in the left pane.

3 Specify your **Report** preferences.

- Specify your default view for new or existing reports. Select **Full screen**, **Tablet**, or **Wide-screen tablet**.
- Select your preferred report theme. The designer provides the following report themes: **Application theme**, **SAS Snow**, **SAS Light**, **SAS Dark**, or **SAS High Contrast**. Your site might also have custom report themes.

  TIP **Application theme** specifies that reports match the selected application theme. **Application theme** is available only as a preference. It is not available as a report theme selection on the **Styles** tab. For more information, see "About Application Themes in SAS Visual Analytics" on page 9.

4 Specify your **Geo Map** preferences.

- Select your default **Map provider mode**. Your choices are **OpenStreetMap** and **Esri**.
- If you select **Esri** as the map provider, you can specify the **Esri map service** in the Esri Map Service Selector window. Your choices in the Esri Map Service Selector window depend on the Esri server.

  Note: The **Esri** option is available only if your system administrator has set the `va.SASGeomapEsriURL` configuration property in SAS Management Console.

5 Specify your **Data Tab** preference. If you are trying to optimize performance, then select the **Bypass cardinality checks** check box.

Note: If you select the **Bypass cardinality checks** check box, then the distinct counts of data items are not displayed on the **Data** tab.

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

41

Creating and Working with Reports

About Reports

You can drag and drop tables, graphs, gauges, and controls to create a well-designed report in SAS Visual Analytics Designer (the designer). You can add text, images, and other controls to reports. A report can be divided into one or more sections. (Sections are like pages.) Each section can have a different layout and contain different report objects.

When you design a report, keep in mind that it might look slightly different on a mobile device or in the viewer. For example, the layout of the tiles in the treemap is dependent on the size of the display area. This means that the same treemap might appear slightly different in the designer than it does in the viewer or on a mobile device.

Reports that are imported from SAS Visual Analytics Explorer (the explorer) might look slightly different in the designer.

The designer provides a report view that enables you to change the size of the canvas that you use for laying out reports. For more information, see "Choose a Report View" on page 309.

Two report layouts are available in the designer: **Precision** and **Tiled**. For more information, see "Choose a Report Layout" on page 309.

Create a New Report

There is no single process for creating a new report in the designer. For example, some users select their data source (or data sources) before they add the report objects while other users add report objects to the canvas and then select their data source (or data sources). Some users choose to update the properties and styles for their reports while other users choose not to. The following list of steps is one way that you can create a new report in the designer.

To create a new report:

1. Select one or more data sources with the associated data items.
2. (Optional) Select the layout (**Precision** or **Tiled**) for the first section of the report using the **Properties** tab for Section 1.
3. Add report objects by dragging and dropping them onto the canvas or by double-clicking the report object on the **Objects** tab. You can tab over a report object and press Enter to add the report object to the canvas.
4. Add data items to report objects by dragging and dropping one or more data items onto the report object or by right-clicking on the report object and using the pop-up menu.
5. (Optional) Modify the report.
   - Update the properties for the report and the report objects.
   - Update the styles for the report objects.
   - Modify data role assignments.
   - Create or modify advanced data items (for example, hierarchies or geographic data items).
   - Create or modify display rules and add alerts for report objects.
   - Add filters to the report.
   - Add interactions between the report objects in a section.
   - Add a rank to a report object.
   - Import report objects from the explorer.
   - Add a new section to the report.
   - Add comments to the report.

   Note: If you add a graph to your report, and there is not enough space for both the graph and the legend to display, then the legend will be dropped.
6. (Optional) View the report in SAS Visual Analytics Viewer (the viewer) by selecting **File** ▸ **View Report**. Once you have viewed the report, you can return to the designer by selecting **File** ▸ **Edit Report**.
7. Save the report. The default location for the first save is `My Folder`. After that, the default save location is the last accessed folder.

TIP Report names are limited to 60 characters.

8 (Optional) View or change the report properties using the **File** menu.

9 (Optional) E-mail the report.

10 (Optional) Print the report.

You can create a new report based on an existing report or on existing report objects. For more information, see "Repurpose an Existing Report" on page 314.

Choose a Report View

There are three report views available in the designer: **Full Screen**, **Tablet**, or **Wide-screen Tablet**. The default report view is **Full Screen**. You can access the report views from the **View** menu or by clicking beside the icon. You can choose , , or .

The report layout is not automatically adjusted for optimal viewing on a device. However, the report does not necessarily display poorly. Before making the report available to users, see how the report is displayed in each view.

Note: Reports can appear to be distorted when they are viewed using different screen resolutions. It is recommended that you design a report using the smallest resolution that you think a user would use to view the report.

Choose a Report Layout

The following report layouts are available in the designer:

Precision
: enables you to place, align, and size report objects. The precision layout allows report objects to overlap, and it lets users control the depth order of these overlapping objects by sending objects backward or bringing them forward. (For example, you might want your company logo to display behind a bar chart and a pie chart in your report.)

  This layout option is keyboard accessible.

  Note: Stored processes are not allowed in the precision layout.

Tiled
: enables you to place report objects directly next to other report objects. The report objects cannot overlap. All report objects in the section are sized to fit in one screen. If you adjust the size of an object, then the other objects automatically resize to ensure that all objects continue to fill the entire screen.

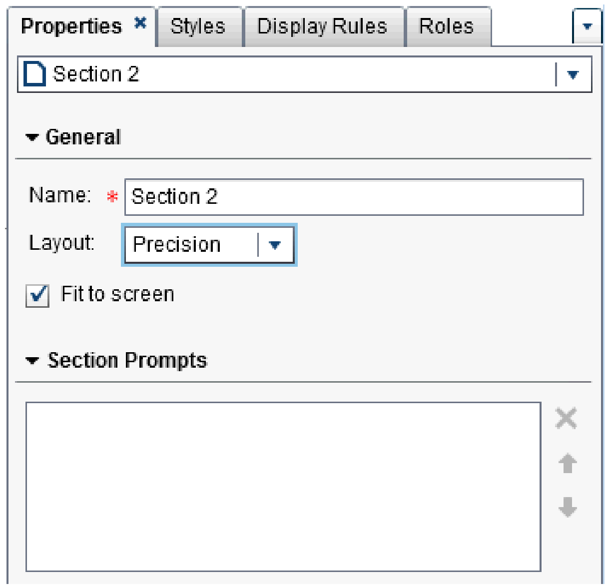
Note: You can switch from the tiled layout to the precision layout after you have added report objects to your report. However, this action can change the size and the position of the report objects. It can also change whether a report object is allowed in the layout. For example, stored processes are not allowed in the precision layout.

To choose your report layout:

1 Select a section tab.

2 In the right pane, click the **Properties** tab.

3 For **Layout**, select either **Precision** or **Tiled**. The **Tiled** layout is the default.



4 (Optional) If you select **Precision**, then you can specify **Fit to screen**. The **Fit to screen** option prevents objects from being sized too wide or too tall, which can cause the report viewing area for the section to scroll.

Note: The **Fit to screen** option is only for designing reports. Your **Fit to screen** selection is not saved with the report, so it does not affect the web viewer or mobile devices.

View Report Properties

You can use the Report Properties window to get an overview of the report. With the exception of the **Name** field, the information in the Report Properties window is different from the information that appears on the **Properties** tab in the right pane of the designer. The Report Properties window provides information such as who created the report and when it was last modified.

All of the information in the Report Properties window is generated after a report is saved. You can add or change only the **Description** and **Keywords** for the report in the Report Properties window.

TIP Starting in the 7.2 release, text in the **Location** field in the Report Properties window is now selectable, which makes it easier for you to share the location of your report with other users.

To view a report’s properties:

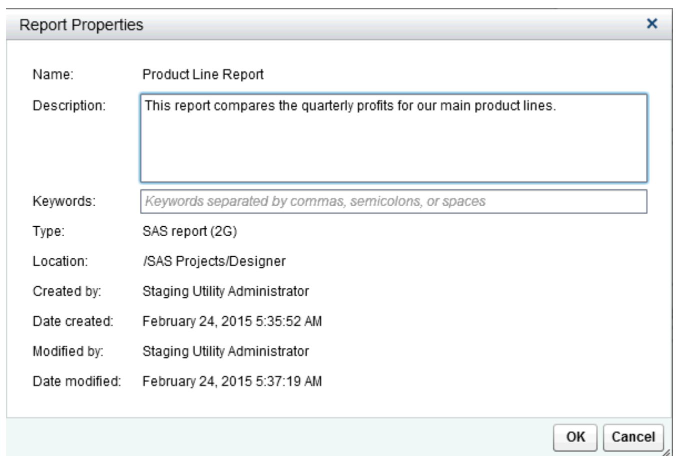
1 Select **File** ▶ **Report Properties**. The Report Properties window is displayed.

2 (Optional) Update the **Description** or **Keywords** for the report.

> **TIP** There is a limit of 200 characters in the **Description** field.

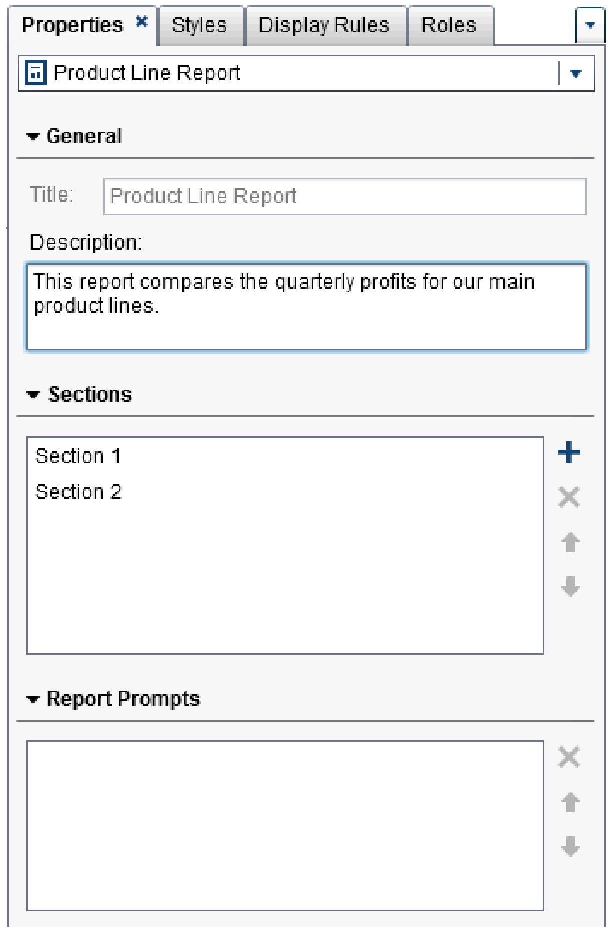
> **TIP** There is a limit of 60 characters for each keyword in the **Keywords** field.

Here is an example of the Report Properties window:



By comparison, the **Properties** tab in the right pane of the designer can be used to update the properties for individual report objects or report sections. The available properties depend on the selected report object. To move to the properties of different report objects or report sections, click ▼ after the report name, and then select another report object or report section from the list.

This example shows a report title and description and lists two sections.



For information about updating report object properties, see "About Report Objects" on page 318.

Change a Report's Name or Title

A report's name and title are the same in the designer. The title is displayed on the **Properties** tab and the name is displayed in the Report Properties window.

To change a report's name and title:

1 Select **File** ▶ **Save As**. The Save As window is displayed.

2 Enter a new **Name**, and click **Save**. The new report name is displayed as the title on the **Properties** tab and as the name in the Report Properties window.

Import a Report or Report Objects

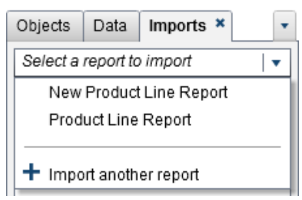
You can import an entire existing report, any section from an existing report, or any report object from an existing report. You can immediately save the imported report with a new name, or you can customize it, and then save it.

Note: Reports that you export from an exploration in the explorer might look slightly different in the designer.

To import a report into the designer:

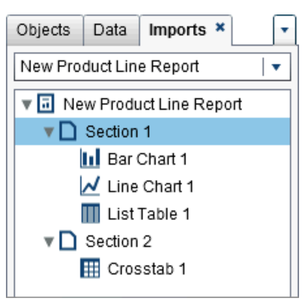
1. Click **Imports** in the left pane. If the **Imports** tab is not visible, select **View** ▶ **Imports**.

2. Select a report name from **Select a report to import**, or click **Import another report** to display the Open window, where you can select a report. The report is displayed on the **Imports** tab.

   Note: Report objects that cannot be imported from the explorer are not displayed on the **Imports** tab.

3. In the tree view on the **Imports** tab, a list of the report name, the sections, and the report objects is displayed. Drag the report, a single report object, multiple report objects, or an entire section and drop it onto the canvas.

4. Select **File** ▶ **Save As**, or click , which displays the Save As window. Enter a **Name**. Report names cannot use these characters: / \

5. Modify the report objects, properties, styles, roles, filters, display rules, ranks, and interactions as needed.

6. Select **File** ▶ **Save**, or click .

Repurpose an Existing Report

If there is an existing saved report that closely matches the report that you want to create, then you can open the existing report that contains the report objects that you want to reuse. (For example, you might want to reuse calculated data items or hierarchies from an existing report in your current report.)

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

1. Open the existing report by doing one of the following:
   - On the home page, double-click on the report that you want to open.
   - In the designer, do one of the following:
     - Select **File** ▸ **Recent**, and then select the report name.
     - Click ▾ beside the recent content drop-down list at the top of the window, and then select a report name.
     - Click on the menu bar to navigate to a report, and then select it.
     - Select **File** ▸ **Open** to display the Open window and select a report.
2. Select **File** ▸ **Save As**, or click , which displays the Save As window. Enter a **Name**. Report names cannot use these characters: / \
3. Modify the report as needed.
   - For existing objects, modify the properties, styles, roles, filters, display rules, ranks, and interactions.
   - To import objects from other reports, click **Imports** in the left pane. If the **Imports** tab is not visible, select **View** ▸ **Imports**. From the **Imports** tab, drag a single report object, multiple report objects, or an entire section and drop it onto the canvas.
4. Select **File** ▸ **Save**, or click .

Open a Report

Reports in the designer are saved when you use either the Save window or the Save As window. A saved report contains at least one section. Typically, a section uses data items from a data source to perform queries. The section displays the results with one or more report objects (for example, a table, a graph, a gauge, a control, and so on). For more information about sections, see "Overview of Report Sections and Info Windows" on page 471.

A section is not required to contain any report objects. For example, you might have a report that you use only as a template to create other reports if you want all reports to have a similar appearance. A template-like report might contain data sources, calculated items, global data filters, and shared display rules, but it might not contain report objects.

To open a report:

1 Select **File** ▸ **Open** or click . The Open window is displayed.

2 (Optional) Search for a report by clicking . The Search window is displayed. Enter your search criteria, and then click **Search**. When the results are displayed, select the report name. Click **OK** to open the report.

> **TIP** Searching does not include report content.

> **TIP** If you search for a single word, then the search assumes a wildcard character before and after the word. For example, if you perform a search with *low* in the **Name** field, then the search results include report names such as *Low Activity*, *Regions with Lowered Sales*, and *Monthly Travel Allowance*.

> **TIP** The search is not case sensitive. For example, if you search for *profit*, then the search results include report names such as *Sports Equipment Profits* and *Company profits last year*.

3 Select a report name, and then click **Open**. The report is displayed in the canvas.

Alternatively, you can open a report using the object inspector on the SAS Visual Analytics home page. For more information, see "Discover Details Using the Object Inspector on the Classic Home Page" on page 633.

Refresh a Report

You can reopen the current report and save any unsaved changes by selecting **File** ▸ **Refresh Report**. When you are prompted to save changes, choose **Save** or **Don't Save**. Unsaved changes can include changes to the underlying data sources or changes that another user might have saved to the report.

This is especially useful if you have closed the current report without saving your changes.

Note: You are not prompted to save the report if there are no unsaved changes.

Delete a Report

To delete a report:

1 Select **File** ▸ **Open**. The Open window is displayed.

2 Select the report that you want to delete, and then click .

You can also delete a report from the home page. Click , and then click **Manage** (next to **My Content** or **Other Content**). For more information, see "Manage Content on the Classic Home Page" on page 628.

42

Using Report Objects

About Report Objects

After selecting your data source and data items, add one or more report objects to display the results. The SAS Visual Analytics Designer (the designer) provides report objects for all of your reports. (You can also select data after you add report objects to the canvas.) Report objects in the designer are grouped into types on the **Objects** tab in the left pane and in the **Insert** menu: **Tables**, **Graphs**, **Gauges**, **Controls**, **Containers**, **Other**, and **Custom**. **Custom** is displayed only if one or more custom graphs are saved in your `My Folder` location or if custom graphs are selected in the Show or Hide Objects window. For more information, see "Show or Hide Report Objects in the Objects Tab" on page 321.

For a definition and a picture of each report object, see Appendix 2, "Gallery of Report Objects," on page 553.

The following table lists the available table objects:

Icon	Table Type
	List Table
	Crosstab

The following table lists the default graph objects:

Icon	Graph Type
	Bar Chart
	Targeted Bar Chart
	Waterfall Chart
	Line Chart
	Pie Chart

Icon	Graph Type
	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

Note: You can display the additional graph objects using the Show or Hide Objects window. For more information, see “Show or Hide Report Objects in the Objects Tab” on page 321. The additional graph objects are displayed in the gallery in the graph builder.

The following table lists the available controls:

Icon	Control Type
	Drop-down List
	List
	Button Bar
	Text Input
	Slider

The following table lists the available containers:

Icon	Container Type
	Vertical Container
	Horizontal Container
	Stack Container

Icon	Container Type
	Prompt Container

The following table lists the other report objects:

Icon	Other Type
	Text
	Image
	Stored Process
	Geo Bubble Map
	Geo Coordinate Map
	Geo Region Map
	Gauge
	Word Cloud

Report objects of the **Custom** type are created using the graph builder. They are identified by the icon. For more information, see "Using Custom Graphs to Display Results" on page 348.

In the designer, you have access to report objects from SAS Visual Analytics Explorer (the explorer). You can open a histogram, heat map, box plot, or correlation matrix in a report that has been exported from the explorer. However, you cannot create new histograms, heat maps, box plots, or correlation matrices in the designer.

Insert a Report Object into a Report

To insert a report object into a report, choose one of the following methods:

- Drag the report object from the **Objects** tab in the left pane and drop it onto the canvas.
- Double-click the report object on the **Objects** tab in the left pane. The report object is automatically placed in the canvas. If you want the report object to appear in a different location, then drag and drop it in a new location.
- Select the report object on the **Objects** tab in the left pane, and press Enter. The report object is automatically placed in the canvas. If you want the report

object to appear in a different location, then drag and drop it in a new location.

> **TIP** Use the Show or Hide Objects window to specify which report objects are displayed on the **Objects** tab. For more information, see "Show or Hide Report Objects in the Objects Tab" on page 321.

- Use the **Insert** menu to select the report object that you want to insert. The report object is automatically placed in the canvas. If you want the report object to appear in a different location, then drag and drop it in a new location.

Additional steps are required for some report objects.

- If you insert a container, then you can drag and drop other report objects onto the container.
- Inserting images requires additional steps. For more information, see "Insert an Image into a Report" on page 340.
- Inserting stored processes requires additional steps. For more information, see "Use a Stored Process" on page 341.

> **TIP** Use the **Clear Selection** or **Select All** option on the report object's pop-up menu to clear the selected data or to select data in the report object.

Show or Hide Report Objects in the Objects Tab

You can customize which report objects you want to see on the **Objects** tab. Once you hide a report object, it remains hidden until you choose to show it again.

> **TIP** To quickly hide a report object on the **Objects** tab, right-click the object, and select **Hide Object**.

Note: Custom graphs appear under the **Custom** heading on the **Objects** tab if they are saved to your `My Folder` location or if custom graphs are selected in the Show or Hide Objects window.

To show or hide report objects:

1. On the **Objects** tab, click , and then select **Show or Hide Objects**. The Show or Hide Objects window is displayed. Report objects that are available in the designer by default are selected. The list of report objects includes additional graph objects, which are not selected by default. Graph objects are also displayed in the graph builder gallery.
2. Select the report objects that you want to show on the **Objects** tab. If there are report objects that you do not want to see on the **Objects** tab, then clear one or more check boxes for one or more report objects.

3 (Optional) To find a custom graph object that is not listed, click **Select Custom**. The Choose an Item window is displayed. Select the custom graph object, and click **OK** to return to the Show or Hide Objects window.

4 Click **OK**. The **Objects** tab is updated.

Using Tables to Display Results

For a definition and a picture of each table type, see “Tables” on page 554.

About List Tables

By default, a list table contains aggregated data with one row for each distinct combination of category values. However, if the **Show detail data** check box has been selected, then all of the data is not aggregated.

TIP To rearrange the columns in a list table, drag and drop the column headings.

You can add sparklines to a column (if the data source contains a date data item) when aggregated data is displayed in the list table. For more information, see “Add Sparklines to a List Table” on page 325.

Starting in the 6.4 release, list tables are sorted in ascending order by the first data item that you add. Only new list tables have a default sort selection. The sorting of data items in list tables in existing reports does not change. To sort the list table by a column, click on the column heading. An arrow appears in the column heading to indicate the sorting.

Here are some key points about list tables:

- If you sort a list table, then the list table displays only the first 5,000 sorted rows. For more information, see “Sort Data in a List Table” on page 395.
- You cannot select the totals in a list table.
- List tables that show detail data cannot be the source of an interaction or a link.

About Crosstabs

You can show subtotals and totals by selecting the appropriate check box (or check boxes) on the **Properties** tab for the crosstab. You can show a percentage of total or percentage of subtotal in a crosstab. For more information about percentage of total or percentage of subtotal, see “Create Derived Items for Measures” on page 369.

You cannot add a sparkline to a crosstab.

You should consider placing lower cardinality (fewer distinct values) categories on the columns and higher cardinality (more distinct values) categories on the rows. Crosstabs can help you improve readability especially when there are several category data items to include in your table.

By default, frequency is displayed only when there are no measures in the crosstab. If you add a category data item first, then the Frequency column is automatically added. When you add a measure data item, the Frequency column is automatically replaced by the measure that you added. If you add a measure data item first, then the Frequency column is added only if you manually add it.

Here are some key points about crosstabs:

- You can create a brush interaction for totals and subtotals in a crosstab.
- A crosstab does not show data if the query is too large.

Specify Table Properties

To specify the properties for list tables and crosstabs:

1. If it is not already selected, select the table in the canvas that you want to update.
2. In the right pane, click the **Properties** tab.
3. Update the general properties for the table. You can update the **Name**, **Title**, **Format** (for the title's font style), and **Description**.
4. Update the object-specific properties for the table. The available properties depend on the selected table type.

   Here are some details about the properties for list tables:

   - By default, the **Enable selection in the viewers** property is selected for list tables. This means that users who use the web viewer or a mobile device can select the list table, and click ⓘ to see the list table name and any incoming filter information.
   - You can use the **Data Options** properties to enable sorting, show detail data, combine excluded rows into “all other,” and to show totals. These properties are available when you use summarized data. They apply to both ranks and post-aggregate filters.

     Note: By default, a list table contains aggregated data with one row for each distinct combination of category values. However, if the **Show detail data** check box has been selected, then all of the data is not aggregated. List tables that show detail data cannot be the source of an interaction or a link.

     For list tables, you can select the **Combine excluded rows into “All Other”** property, under the **Data Options** heading, to summarize all excluded rows. For more information, see “Use the Combine Excluded Rows (or Cells) into “All Other” Properties” on page 327.

     > **TIP** Select the **Show labels for totals** property to turn off or turn on the aggregation labels for totals.

   Here are some details about the properties for crosstabs:

   - By default, the **Enable selection in the viewers** property is selected for crosstabs. This means that users who use the web viewer or a mobile device can select the crosstab, and click ⓘ to see the crosstab name and any incoming filter information.

- You can specify that you want to display missing labels as blanks instead of the string `(missing)` being displayed.
- You can specify that you want an indented layout and totals and subtotals for columns, rows, or both.
- You can specify the placement of totals and subtotals.

Specify Table Styles

To specify styles for tables:

1. If it is not already selected, select the table in the canvas that you want to update.
2. In the right pane, click the **Styles** tab.
3. Update the styles for the table. The available styles depend on the selected table type. For example, you can specify **Border and Fill**, **Cells**, **Column Headings**, and **Totals** for list tables and crosstabs.

   Here are some details about the styles for list tables:

   - Select **Wrap text** to wrap text in the cells of a list table.
   - Select **Enable alternating background color** to specify that the color of every other row appears darker in the list table. Use the **Alternating background color** to specify the color of every other row in the list table. Use the **Selection color** style to specify the color of the selected row in the list table.

Your custom colors are saved between SAS Visual Analytics sessions. Your custom colors are displayed in the color palette. Here is an example:

Figure 42.1 A Color Palette in the Designer

Add Sparklines to a List Table

A *sparkline* is a small line graph without axes or labels that presents a single trend over time. A sparkline is about the size of one or two words, so it fits in a single cell and repeats for each row in a column. A sparkline does not have axes or labels. They are frequently used to present stock trends or production rates over time. A sparkline is intended to be both succinct and noteworthy.

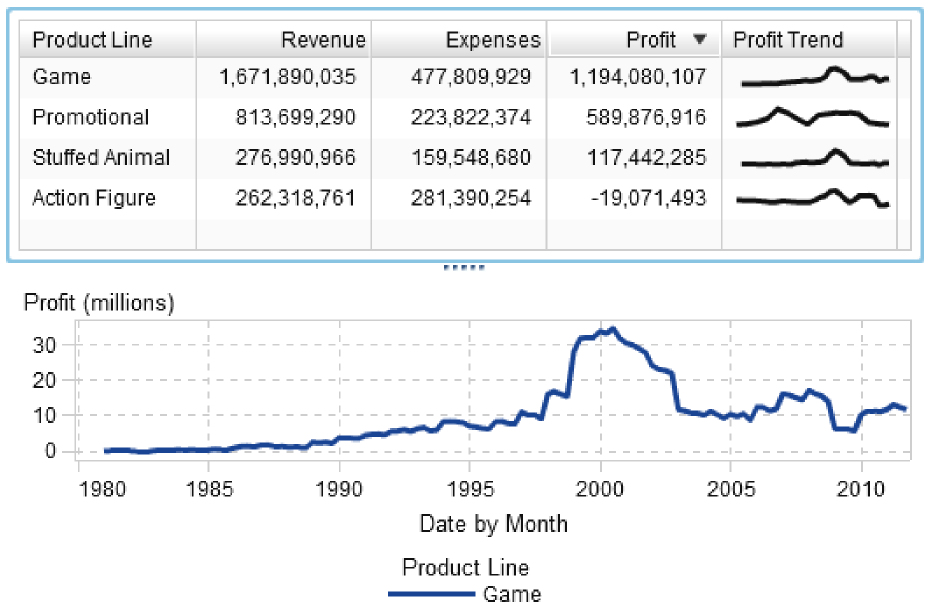
In the designer, you can add sparklines to a column in a list table. The data source for the list table must include a date, datetime, or time data item before you can add a sparkline.

A sparkline in the designer can have up to 40 bins. (A *bin* is a way to group continuous values into a smaller number of intervals.) When the sparkline data is binned, it is grouped by a boundary. A boundary is a minute, hour, day, month, quarter, or year. For example, if there are two years' worth of data, then the data will be grouped by month, and the sparkline will have 24 bins (or one for each month). If there is one month (30 days) of data, the sparkline will have 30 bins. If you have two months of data, the sparkline will have two bins because the data is binned by a month boundary. Having a lot of data for the list table does not guarantee that the sparkline provides more detail because the designer prioritizes bin boundaries over maximum number of bins.

The data tip values on the sparkline show the high and low values of the sparkline and the last value in the sparkline. The data tip values are affected by the boundary at which the data is binned and by the aggregation of the data item. For example, suppose that the aggregation is Sum. If the data is binned by day, then the minimum and maximum values for a given day are displayed in the data tip. However, if the data changes and it is binned by month, then the minimum and maximum values for the sum of all days in the month are displayed in the data tip.

Note: The designer does not let users control the binning of the sparkline.

The data for each sparkline is displayed as a miniature time series plot. Here is an example of a report that contains a simple list table with a sparkline and a time series plot, which is filtered to represent the data shown in the sparkline:



Product Line	Revenue	Expenses	Profit ▼	Profit Trend
Game	1,671,890,035	477,809,929	1,194,080,107	
Promotional	813,699,290	223,822,374	589,876,916	
Stuffed Animal	276,990,966	159,548,680	117,442,285	
Action Figure	262,318,761	281,390,254	-19,071,493	

In the example, both the time series plot and the sparkline are using *Profit* for the measure. The time series plot has more detail than the sparkline because in the time series plot, the data is grouped at a more granular level. The sparkline shows the same overall line as the time series plot, but has less detail.

The sparkline does not display the trend at the most granular date, datetime, or time level. Instead, the sparkline summarizes the trend depending on the unit of time that is used in the list table. For example, the sparkline might be summarized for the month, quarter, or year, depending on the data. The report designer cannot change the level of summary in the sparkline.

To add a sparkline:

1 If it is not already selected, select the list table in the canvas that you want to update.

2 Right-click the list table, and then select **Add Sparkline**. The Add Sparkline window is displayed.

3 Enter a **Column label**.

4 For **Time Axis**, select a date, datetime, or time data item in the current data source.

5 Select a **Measure (line)**.

6 (Optional) Select the **Set baseline** check box. Enter a **Value**, and select a **Fill type**. Your choices are **Gradient** or **Solid**.

The **Set baseline** option draws a horizontal line through the graph at the point on the Y axis where the baseline value resides. Everything above or below the baseline is filled in with either a solid or gradient color.

7 Click **OK**. The sparkline is added to the last column in the list table. You can move the sparkline to another location in the table.

To edit a sparkline, right-click in the sparkline column in the list table, and then select **Edit Sparkline**. The Edit Sparkline window is displayed. Update the information, and then click **OK** to save your changes.

To delete a sparkline, right-click in the sparkline column in the list table, and then select **Remove Sparkline**.

Use the Combine Excluded Rows (or Cells) into “All Other” Properties

You can use the **Combine excluded rows into “All Other”** property for list tables or the **Combine excluded cells into “All Other”** property for some graphs. These properties combine content across categories (for example, the content that is formed by combining the report object’s visible categories). These properties are frequently used with filters and ranks.

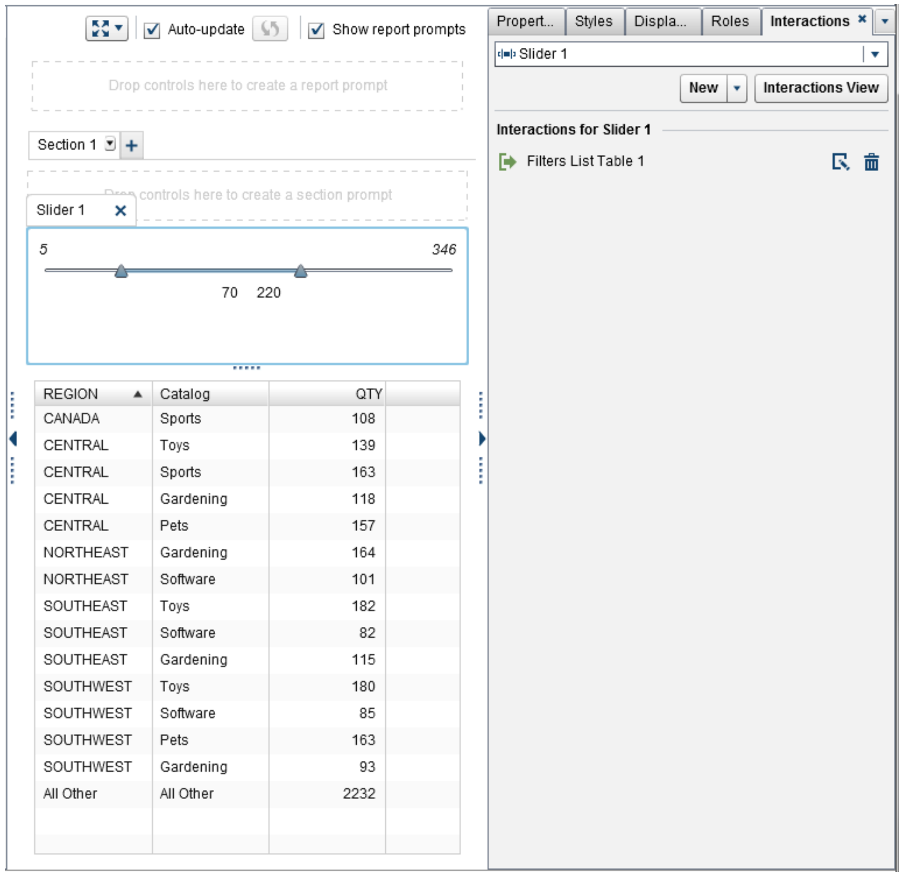
The effect of these properties can be influenced by certain types of interactions. For example, if you have a list table with a *Sales* measure data item, the list table might have the **Combine excluded rows into “All Other”** property selected and be targeted by a slider showing a range of *Sales* figures. The category combinations that are dropped based on the range selected in the slider control are grouped into a category named "All Other."

Here are some key points about the **Combine excluded rows into “All Other”** and the **Combine excluded cells into “All Other”** properties:

- The **Combine excluded rows into “All Other”** property is not available for crosstabs.
- The **Combine excluded cells into “All Other”** property is not available for time series plots, bubble plots, scatter plots, step plots, needle plots, geo maps, or word clouds.
- The **Combine excluded cells into “All Other”** property is not available for custom graphs if they depend on multiple data definitions. The property might be available for some custom graphs. For example, suppose that you have a custom graph with a bar chart and a line chart side by side, and the category and measure data role assignments for each chart are shared between the charts. Since one query produces the results for both charts, the **Combine excluded cells into “All Other”** property is available.

There is a similar, but distinct per-category “All Other” option that is provided for ranking in the designer. For more information about the All Other concept for ranking, see “Add a New Rank” on page 459.

The following list table shows how the content of combining visible categories is factored into “All Other”. The list table has two categories and a measure. The **Combine excluded rows into “All Other”** property is selected on the **Properties** tab. The list table is filtered by aggregate quantity values using the adjacent slider control.



Here are some key points about the **Combine excluded rows into “All Other”** property and the **Combine excluded cells into “All Other”** property:

- The property cannot be set when the report object is displaying detail data.
- The property cannot be set when the report object includes a rank with the per-category **All Other** option selected.
- The property cannot be set for pie charts when the **Create “Other” slice for minimal values** property is selected.
- The property cannot be set when a hierarchy is assigned to the report object.
- The property cannot be set for line charts that are imported from the explorer.

Using Graphs to Display Results

About Graphs

For a definition and a picture of each graph type, see "Graphs, Charts, and Plots" on page 555.

Some graphs are available by default in the designer. (These graphs are listed under the **Graphs** heading on the **Objects** tab.) There are additional graph objects, which are displayed in the graph builder gallery. For more information about the gallery, see "About the Graph Template Gallery" on page 496. You can also display the additional graph objects using the **Show or Hide Objects** window in the designer. For more information, see "Show or Hide Report Objects in the Objects Tab" on page 321.

You can also create custom graphs. For more information, see "Using Custom Graphs to Display Results" on page 348.

Specify Graph Properties

To specify the properties for graphs:

1. If it is not already selected, select the graph in the canvas that you want to update.
2. In the right pane, click the **Properties** tab.
3. Update the general properties for the graph. You can update the **Name**, **Title**, **Format** (for the title's font style), and **Description**.
4. Update the specific properties for the graph. The available properties depend on the selected graph type.

   Here are details about the properties for graphs:

   - By default, the **Enable selection in the viewers** property is selected for graphs. This means that users who use the web viewer or a mobile device can select the graph, and click to see the graph name and any incoming filter information.
   - For select graphs only, you can use the **Combine excluded cells into "All Other"** property under **Data Options**. This property applies to both ranks and post-aggregate filters. For more information, see "Use the Combine Excluded Rows (or Cells) into "All Other" Properties" on page 327.
   - To add a new horizontal or vertical reference line to all graph types except the pie chart and the treemap, select **Create new reference line**. You can specify a **Label**, an **Axis**, a **Value**, and the **Style** of the new line.

     Note: Properties for the axes or reference lines are not displayed for a custom graph that has merged or common axes.
   - To rotate a graph's X-axis category labels by 45 degrees, select the **Rotate value label** property. This property is available for the bar chart, targeted bar chart, waterfall chart, line chart, dual axis bar chart, dual axis

line chart, the dual axis bar-line chart, and custom graphs with a single X axis that has discrete values.

Note: You cannot use the **Rotate value label** property if you assign a data item to a lattice row or lattice column data role.

- For bar charts that are assigned multiple measure data items (in the **Measures** role) or if you group a data item, you can create a 100% stacked bar chart using the **Grouping style** and **Grouping scale** properties. Select **Cluster** or **Stack** for the **Grouping style**. Select **Normalize groups to 100%** for the **Grouping scale**.

  Note: The **Grouping scale** property is not available for the dual axis bar chart, the dual axis bar-line chart, or any custom graph with at least one bar chart and a second Y axis.

  Note: Negative values are ignored in the 100% stacked bar chart.

  Note: If you select **Normalize groups to 100%** for a bar chart, then the **Set baseline** property is not available.

  Note: The **Callout** and **Outside** options for the **Data label location** are not supported for grouped pie charts. Selecting either of these options has no effect on the pie chart.

- Use the **Set baseline** property to place a baseline on the response axis using a value that you specify in the text field. When this property is not selected, the graph defaults to placing the baseline at 0. The **Set baseline** property is available for bar charts, targeted bar charts, waterfall charts, line charts, dual axis bar charts, dual axis line charts, dual axis bar-line charts, needle plots, and butterfly charts.

  Note: The **Set baseline** property is not available for bar charts if you select **Normalize groups to 100%** for the **Grouping scale** property.

- To fill a line chart or a time series plot with color, select a **Grouping style**. The style choices are **Overlay Unfilled**, **Overlay Filled**, and **Stack Filled**. The default grouping style is **Overlay Unfilled**.

- To change the percentage for the “other” slice in a pie chart, change the amount listed in the **Minimum percentage for “Other”** field. The default is 4%.

- To display the measure label at the top of a pie chart, select the **Show label** check box.

- To change the transparency of the markers in a scatter plot, needle plot, step plot, or dot plot; or for the bubbles in a bubble plot or a bubble change plot; or for the bars in a schedule chart, move the slider for **Transparency** to any value between **0%** and **100%**. The default value for transparency is **0%**.

- For the treemap, the **Layout** property determines how the tiles are arranged. The **Show level indicator** check box determines whether the label of the data item that is in the **Tile** role is displayed above the treemap.

- Use the **Binning interval** for a time series plot or a dual axis time series plot when you have an uneven grouping of time data. The options are:

  Automatic

  Determines the best bin size for your data. This option is the default.

Fixed count
Enables you to specify an interval between 10 and 500 in the **Fixed bin count** field.

Use format
Uses the format of the date data item as the interval.

- To remove the arrowheads for the vector plot or the stock high-low plot, clear the **Show arrowheads** property. To change the transparency of the vector plot or the stock high-low plot, move the slider for **Transparency** to any value between **0%** and **100%**. The default value for transparency is **0%**.

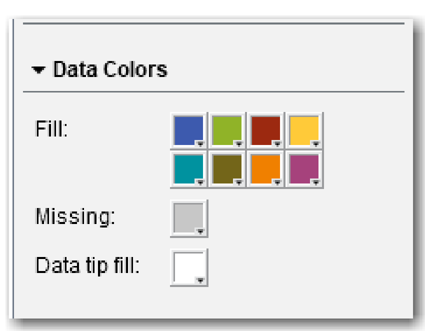
Specify Graph Styles

To specify styles for graphs:

1. If it is not already selected, select the graph in the 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, you can specify **Border and Fill**, **Data Styling**, **Frame Styling**, **Text Styling**, and **Data Colors** for all graphs.

   Here are some details about the styles for graphs:

   - By default, the background of a graph is set to white. Use the **Wall background** option (under **Frame Styling**) to specify a different color.
   - Move the slider beside the color palette to set the transparency for the **Wall background**, the **Legend background**, or the **Header background** for graphs, except for treemaps. The default value for transparency is **0%**.
   - A three-color **Gradient** data color style is available for bubble plots, treemaps, some bar charts, some waterfall charts, and word clouds.
   - To modify the colors that are used in a graph, use the styles that are available under **Data Colors**.

     To change a color, click its tile to open the color palette. Using the palette, select a new color. The new color is automatically applied to the graph, and the tile changes to the new color.

Your custom colors are saved between SAS Visual Analytics sessions. Your custom colors are displayed in the color palette. For an example of the color palette, see Figure 42.1 on page 325.

Using Controls to Display Results

For a definition and a picture of each control type, see “Controls” on page 570.

About Controls

A control is a report object that filters or narrows the scope of the data that you are currently viewing. A control enables you to group your data by a selected category, and then select which group you want to view. When you drag and drop a data item onto a control, the control creates a group based on that data item. For example, you might have a data item called *Cars* that contains all of the models that a manufacturer produces. When you drag and drop the *Cars* data item onto a drop-down list, the control groups the car models, and then you can select a car model to use as a filter. Controls can be used in a report with interactions.

Report prompts are controls that are placed in the special area at the top of the canvas. A report prompt automatically filters all of the other report objects as long as the report object uses the same data source as the report prompt control. For more information, see “Use a Control to Create a Report Prompt” on page 334.

Section prompts are the controls that are placed in the special row area below the section tabs at the top of the canvas. A section prompt automatically filters all of the other report objects in the same section, as long as the report object uses the same data source as the section prompt control. For more information, see “Use a Control to Create a Section Prompt” on page 335.

You can place any control in the main area of the canvas below the section prompt row. You must define explicit interactions (using either the **Interactions** tab or the Interactions view) between these controls (as the source report objects) and one or more target report objects. For more information about interactions, see “Overview of Report Interactions” on page 433.

The text input control, button bar control, drop-down list control, and slider control (single-point only) support parameters. For more information, see Chapter 49, “Working with Parameters in Reports,” on page 463.

Here are some key points about filtering using controls:

- Filters use the AND operator.
- Filters are applied as separate steps.
- The filter results are impacted by the type of data used in the control.

Specify Control Properties

To specify the properties for a control:

1. If it is not already selected, select the control in the canvas that you want to update.
2. In the right pane, click the **Properties** tab.

3 Update the general properties for the control. You can update the **Name**, **Title**, **Format** (for the title's font style), and **Description**.

4 Update the specific properties for the control. The available properties depend on the selected control.

Here are some details about the properties for controls:

- By default, the **Enable selection in the viewers** property is not selected for controls. This means that users who use the web viewer or a mobile device cannot select the control, and click to see the control name and any incoming filter information. However, users can still modify values for the control.
- For drop-down lists, lists, and button bars, select the **Required** property if you want to require users to make a selection in the control. If you select the **Required** property for a list, at least one check box must always be selected.
- By default, the **Allow multiple selection** property is selected for lists. If you clear the **Allow multiple selection** check box, radio button are displayed instead of check boxes and the **Required** property is applied automatically.
- For button bars and sliders, the **Horizontal** property is selected by default.
- For sliders, select the **Interact on the data in view** property if you want to have the control interactively filter the post-aggregated data.
- For sliders, the **Set value to dynamic minimum** and **Set value to dynamic maximum** properties automatically adjust the slider to the minimum or maximum values in the current data query. These properties are available for only dates and measures.
- For sliders, select the **Set fixed range** property so that you can specify the **Minimum** and **Maximum** properties for the slider end points.

  Note: You cannot have a filter or rank when the **Set fixed range** property is selected for a slider.

Specify Control Styles

To specify styles for controls:

1 If it is not already selected, select the control in the canvas that you want to update.

2 In the right pane, click the **Styles** tab.

3 Update the styles for the control. The available styles depend on the selected control type. For example, for a drop-down list, you can specify **Border and Fill**, **Drop-down Styling**, and **Text Styling**.

Your custom colors are saved between SAS Visual Analytics sessions. Your custom colors are displayed in the color palette. For an example of the color palette, see Figure 42.1 on page 325.

Use a Control to Create a Report Prompt

If you use a control to create a report prompt, then the user can select a value to filter the data in the report. For some control types, when the **Required** property is not selected, the user might need to press Ctrl+click to clear the value in the filter.

You can use a report prompt to cascade filters to section prompts.

Starting in the 7.4 release, you can create cascading (or dependent) report prompts. Cascading report prompts enable you to create filter (but not brush) interactions between objects in the report prompt bar. For example, suppose you have a report that has filters from *Region* to *State* and *State* to *City*. When you change the value for *Region*, *City* should also be filtered. For more information, see “About Report Filters” on page 419.

To use a control to create a report prompt:

1 Drag the control icon from the **Objects** tab in the left pane and drop it onto the area above the tabs on the canvas. (Look for the hint text that says, “Drop controls here to create a report prompt.”) The control appears above the tabs on the canvas.

   Note: You can also use a prompt container to create a report prompt.

2 Drag and drop a category, measure, or parameter onto the control. For example, if you drag and drop a drop-down list control, then you can assign a category like *Facility City* or *Facility State*. Then, the drop-down list is populated with the cities or states that are used in that category.

   You can also use the **Roles** tab in the right pane to specify the **Category** and **Frequency** roles for the report prompt.

3 (Optional) Update the general properties for the report prompt. You can update the **Name** and **Description**.

4 (Optional) Update the specific properties for the report prompt. The available properties depend on the selected control.

   Here are some details about the properties for controls:

   - By default, the **Enable selection in the viewers** property is not selected for controls. This means that users who use the web viewer or a mobile device cannot select the control, and click to see the control name and any incoming filter information. However, users can still modify values for the control.

   - For drop-down lists and button bars, select the **Required** property if you want to require a user to make a selection in the control. If you select the **Required** property for a drop-down list, at least one check box must always be selected.

   - For sliders, for **Value**, the **Range** property is selected by default. Select the **Interact on the data in view** property to have the control filter only the aggregated data that is currently displayed in the report. If you clear this property, then the detail data is filtered.

     Note: If the **Interact on the data in view** property is not selected, then a slider does not filter crosstabs or time series plots.

- For button bars and sliders, the **Horizontal** property is selected by default.

If the report prompt uses one data source and the report objects on the canvas use another data source, you can change the data source mappings by right-clicking the control, and then selecting **Edit Data Source Mapping**. For more information, see “Map Data Sources” on page 444.

Reorder Report Prompts

You can reorder the display of report prompts.

To change the order in which report prompts are displayed:

1. In the right pane, click the **Properties** tab.
2. Select the report name in the drop-down list.
3. In the **Report Prompts** area, select a report prompt. Click or to reorder the report prompts.

Use a Control to Create a Section Prompt

If you use a control to create a section prompt, then the user can select a value to filter other report objects in the same section, as long as the report object uses the same data source as the section prompt control.

Here are some key points about section prompts:

- The drop-down list, button bar, and text input controls are the only controls that can be used as section prompts.
- If you use a control to create a section prompt, then the user can select a value to filter the data. For some control types, when the **Required** property is not selected, the user might need to press Ctrl+click to clear the value in the filter.
- Section prompts can be affected by report prompts.
- Section prompts are not allowed in info windows.

Starting in the 7.4 release, you can create cascading (or dependent) section prompts. Cascading section prompts enable you to create filter interactions between objects in the section prompt bar. Cascading section prompts do not allow data brushing.

To use a control to create a section prompt:

1. Drag the control icon from the **Objects** tab in the left pane and drop it onto the area above the report objects and below the tabs on the canvas. (Look for the hint text that says, “Drop controls here to create a section prompt.”) The control appears below the tabs on the canvas.

   Note: You can also use a prompt container to create a section prompt.

2. Drag and drop a category, a measure, or a parameter onto the control. For example, if you drag and drop a drop-down list control, then you can assign a category like *Facility City* or *Facility State*. Then the drop-down list is populated with the cities or states that are used in that category.

You can also use the **Roles** tab in the right pane, and then specify the **Category** and **Frequency** roles.

TIP If you move a section prompt from one section to another section in a report, you must edit the data source mapping for an interaction to work. Right-click the control, and select **Edit Data Source Mapping**. For more information, see "Map Data Sources" on page 444.

If the section prompt uses one data source and the report objects on the canvas use another data source, you can change the data source mappings by right-clicking the control, and then selecting **Edit Data Source Mapping**.

Reorder Section Prompts

You can reorder the display of section prompts.

To change the order in which section prompts are displayed:

1 In the right pane, click the **Properties** tab.

2 Select the section name in the drop-down list.

3 In the **Section Prompts** area, select a section prompt. Click or to reorder the section prompts.

Using Container Object Types in Reports

About Container Objects

You can use containers to group other report objects. There are these types of containers:

- Vertical and horizontal containers.
- Stack containers. The report objects are displayed as if they are in a slide deck. Only one report object is displayed at a time. The stack container has a control bar instead of a scroll bar that lets you move between report objects. You cannot nest stack containers. However, you can place them side by side. When you use precision layout, you can resize the stack container.
- Prompt containers. These containers group prompt controls. Prompt containers are affected by report prompts, but not by other section prompts. They are not affected by interactions. Report objects inside prompt containers are filtered by the same rules as other objects.

  Prompt containers can be added to the report prompt area and section prompt area on the canvas. This enables you to add control types (for example, list controls) that are not otherwise allowed in those prompt areas.

  TIP If a prompt container is open and there are unsaved changes, then the button bar changes. An **Apply** button is displayed so that you can apply changes.

Note: Stored processes cannot be added to any type of container.

Specify Container Properties

To specify the properties for a container:

1. If it is not already selected, select the container in the canvas that you want to update.

2. In the right pane, click the **Properties** tab.

3. Update the general properties for the container. You can update the **Name**, **Title**, **Format** (for the title's font style), and **Description**.

4. Update the object-specific properties for the container. Here are some details about the properties for containers:

   - By default, the **Enable selection in the viewers** property is selected for containers. This means that users who use the web viewer or a mobile device can select the container, and click to see the container name and any incoming filter information.

   - Update the order in which the report objects appear inside the container.

   - For stack containers, the **Navigation control location** property changes the location of the control bar, and the **Navigation button type** property changes the appearance of the control bar. You can reorder the display of the report objects in the stack container using the **Objects** list.

   - For prompt containers, you can select the **Layout** and **Button text**. The **Automatically apply values** property is selected by default. If you clear the **Automatically apply values** check box, then you cannot work with any interactions to or from the prompt container until you apply or cancel changes.

Specify Container Styles

To specify styles for containers:

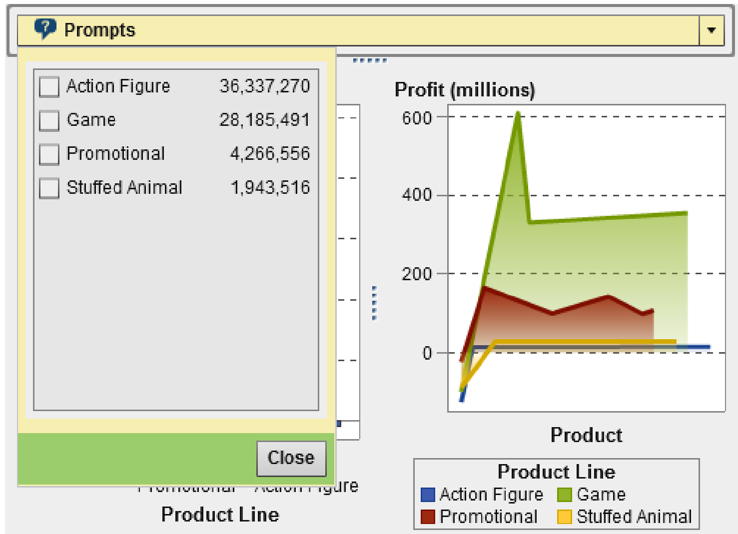
1. If it is not already selected, select the container in the canvas that you want to update.

2. In the right pane, click the **Styles** tab.

3. Update the styles for the container. The available styles depend on the selected container type. **Border and Fill** is available for all containers. Prompt containers also have **Drop-down Styling** and **Text Styling** options.

   The **Button bar color** style option for the prompt container enables you to change the background around the **Close** button.

Here is an example of a prompt container that has the **Background color** and **Button bar color** styles set.



Your custom colors are saved between SAS Visual Analytics sessions. Your custom colors are displayed in the color palette. For an example of the color palette, see Figure 42.1 on page 325.

Using Other Object Types in Reports

Use a Text Object

About Text Objects

Text objects can display both static text and dynamic text. You can use text objects to convey messages (such as confidentiality statements), to annotate other objects, and to display key values.

Dynamic text supports display rules. The displayed values are affected by interactions, filters, and ranks. For more information, see "Display Dynamic Text in a Text Object" on page 339.

Text objects can contain links in static text. See "Create a Link from a Text Object" on page 454.

Specify Text Object Properties

To specify the properties for a text object:

1 If it is not already selected, select the text object in the canvas that you want to update.

2 In the right pane, click the **Properties** tab.

3 Update the general properties for the text. You can update the **Name**, **Title**, **Format** (for the title's font style), and **Description**.

4 (Optional) Update the object-specific properties for the text object. By default, the **Enable selection in the viewers** property is not selected for text objects. This means that users who view the report in the web viewer or on a mobile device cannot select the text in a report.

Specify Text Object Styles

You can use the floating toolbar to change the font, font size, text color, and text background color. You can specify whether the text is bold, italic, or underlined, and whether it is left-aligned, centered, or right-aligned. You can also use the floating toolbar to create a link from a text object. For more information, see "Create a Link from a Text Object" on page 454.

TIP You can use the pop-up menu to cut, copy, and delete text. However, you have to use the keyboard shortcut (Ctrl+V) to paste text.

Note: You cannot change text styles using the **Styles** tab.

Display Dynamic Text in a Text Object

Text objects support several types of dynamic text:

- Measure values
- Parameter values
- A timestamp for the most recent update of the current data source
- A description of the current interactive filters (from prompts and from interactions with other objects)

To add a measure or a parameter to the text object, either use the **Roles** tab or drag the data item from the **Data** tab and drop it onto the text object. The parameter can be of any type, but parameters with multiple values (for example, character list parameters) are limited to 25 characters.

Note: To display the frequency value in a text object, you must drag Frequency or Frequency Percent from the **Data** tab and drop it onto the text object.

To add the timestamp for the most recent update of the current data source, click ⊕ in the floating toolbar, and then select **Table Modified Time**.

To add a filter description, click ⊕ in the floating toolbar, and then select **Interactive Filters**.

Note: If you replace or remove the data source that is associated with the text object, then all dynamic text is removed.

Use an Image

About Images

You can use images to include your corporate logo or other graphics in your reports. 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.

Insert an Image into a Report

To insert an image into a report:

1 Choose one of the following:

- Drag the from the **Objects** tab in the left pane and drop it onto the 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 canvas. If you want the image to appear in a different location, then drag and drop it in a new location.

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 Image Properties

To specify the properties for an image:

1 If it is not already selected, select the image in the canvas that you want to update.

2 In the right pane, click the **Properties** tab.

3 Update the general properties for the image. You can update the **Name**, **Title**, and **Description**.

4 Update the properties specific for the image. Your choices are **Location**, **Scale type**, and **Tooltip text**.

   By default, the **Enable selection in the viewers** property is not selected for images. This means that users who use the web viewer or a mobile device cannot select the image, and click to see the image name.

Specify Image Styles

No styles are available for images.

Use a Stored Process

About Stored Processes

A *stored process* is a SAS program that is stored on a server and that can be executed as requested by client applications such as SAS Visual Analytics. The embedded SAS code can contain instructions for displaying report elements that include queries, prompted filters, titles, images, and statistical analyses.

You can add one or more stored processes to a report. You can edit the prompts for a stored process using the **Properties** tab in the designer. When you set the value for a prompt in the designer, that prompt value becomes the default value for the viewer.

Any prompts in a stored process must follow the guidelines for entering prompt values. This is important if the stored process contains date values. See “Entering Prompt Values in the SAS Stored Process Web Application” in the *SAS Stored Processes: Developer’s Guide*.

You can also add a *stored process report*, which contains the predefined output from a stored process.

There are prerequisites for printing stored process output. For more information, see “Printing Reports” on page 477.

There are limitations to where stored processes can be used in the designer:

- A stored process cannot be used in the precision layout.
- A stored process cannot be the source or target of an interaction in a report.
- A stored process cannot be added to a container.

Note: Users who view stored processes using the SAS Visual Analytics Apps (previously called SAS Mobile BI) cannot be prompted. Instead, the stored process runs using the prompt values that were added when the report was created.

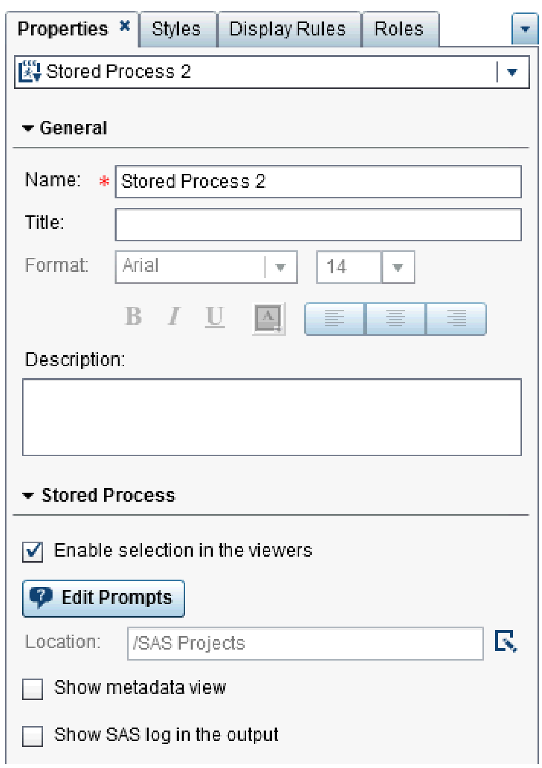
Note: The stored process output in a report is rendered as HTML regardless of the requested output type.

For information about creating a stored process and registering it in metadata, see *SAS Stored Processes: Developer's Guide*.

Specify Stored Process Properties

To specify the properties for a stored process:

1. If it is not already selected, select the stored process in the canvas that you want to update.
2. In the right pane, click the **Properties** tab.
3. Update the general properties for the stored process. You can update the **Name**, **Title**, **Format** (for the title's font style), and **Description**.
4. Update the properties specific to the stored process. Your choices are **Show metadata view** and **Show SAS log in the output**. If the stored process contains prompts, the **Edit Prompts** button is displayed.

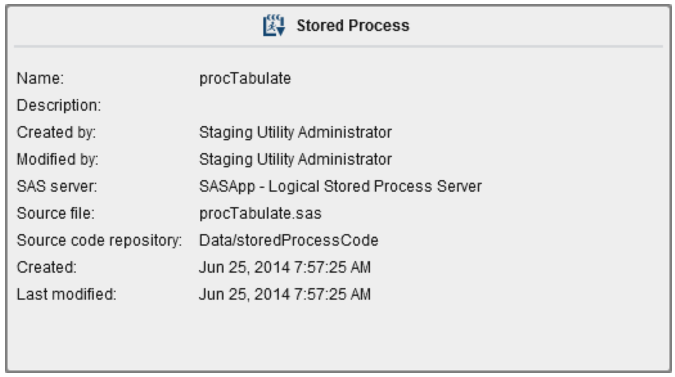


By default, the **Enable selection in the viewers** property is selected for stored processes. This means that users who use the web viewer or a mobile device can select the stored process, and click ⓘ to see the stored process name.

Click **Edit Prompts** to open a new window where you can edit the parameters for the stored process. Click **OK** to save your changes.

Selecting the **Show metadata view** check box can make it easier to work with the stored process while you are designing a report. In the metadata view, you can see information about when the stored process was created and last modified. You can also see the name of the stored process file.

Here is an example of the metadata view for a stored process:



Selecting the **Show SAS log in the output** option means that both the log output and the stored process output are displayed in the report. This can assist you in debugging problems.

Specify Stored Process Styles

No styles are available for stored processes.

Use a Geo Map

About Geo Maps

Many types of data have a spatial aspect, including demographics, marketing surveys, and customer addresses. For example, if a user needs to evaluate population data for U.S. Census tracts, a report designer could display the information in a table. However, it would be easier and more effective for the person using the report to see the information in the context of the geography of the tracts. When evaluating information that has a spatial component, users might find it easier to recognize relationships and trends in the data if they see the information in a spatial context.

A geo map overlays your data on a geographic map. You can add a geo map only if the report uses data items from a data source that is enabled for geographic mapping.

You can create a filter or brush interaction between a geo map and another report object in your report or dashboard. When you click on a specific region or city in the geo map, the other report object filters or highlights to show the same location.

The following geo maps are available in the designer:

Geo Bubble Map

A bubble plot that is overlaid on a geographic map. For a complete definition and picture of a geo bubble map, see "Other Report Objects" on page 571.

A geo bubble map requires a geography variable with the role type of geography.

Geo Coordinate Map

A simple scatter plot that is overlaid on a geographic map. For a complete definition and picture of a geo coordinate map, see "Geo Coordinate Maps" on page 571.

A geo coordinate map requires a geography variable with the role type of geography.

Geo Region Map

A two-dimensional map that uses color combinations to represent different regions on the map. For a complete definition and picture of a geo region map, see "Geo Region Maps" on page 572.

A geo region map requires a geography variable with the role type of geography. However, it does not support custom geography data items or ZIP code data.

Here are some key points about geo maps:

- Calculated data items and grouped category data items can be changed into geographic data items and used in geo maps.
- Certain geo maps that are exported from the explorer (for example, geo maps that use custom roles or data sets that have centroids) cannot be fully modified in the designer.
- When a **Color** data role and a display rule are applied to a geo bubble map or geo region map, the **Color** data role takes precedence over the display rule. The geo coordinate map does not have a **Color** data role. For more information about data roles, see "Working with Data Role Assignments" on page 386.

Specify Geo Map Properties

To specify the properties for a geo bubble map, geo coordinate map, or geo region map:

1. If it is not already selected, select the geo bubble map, geo coordinate map, or geo region map in the canvas that you want to update.
2. In the right pane, click the **Properties** tab.
3. Update the general properties for the geo bubble map, geo coordinate map, or geo region map. You can update the **Name**, **Title**, **Format** (for the title's font style), and **Description**.
4. Update the properties that are specific to the geo bubble map, geo coordinate map, or geo region map. You can update **Show map navigation control**, **Transparency**, and **Show legend**.

   Here are some details about the properties for geo maps:

   - Use the **Map Service** property to change the map provider for a geo bubble map, geo coordinate map, or geo region map after you have placed it on the canvas. This property changes the map type from OpenStreetMap to any available Esri service if an Esri service is available.
   - For a geo bubble map, you select the **Transparency** for the bubble plot. For a geo coordinate map, you select the **Transparency** for the scatter plot. For a geo region map, you select the **Transparency** for the regions.

- There are **Legend** properties for geo coordinate maps. However, geo coordinate maps cannot have a legend. Any changes that you make to the **Legend** properties will be discarded.

Specify Styles for Geo Maps

To specify styles for a geo bubble map, geo coordinate map, or geo region map:

1. If it is not already selected, select the geo bubble map, geo coordinate map, or geo region map in the canvas that you want to update.
2. In the right pane, click the **Styles** tab.
3. Update the styles for the geo bubble map, geo coordinate map, or geo region map. You can specify **Border and Fill**, **Data Styling**, **Frame Styling**, **Text Styling**, and **Data Colors** for the geo bubble map, geo coordinate map, or geo region map.

   Note: Performance for geo maps is negatively impacted when you change the **Data skin** style. The recommendation is to leave it as **None**.

Your custom colors are saved between SAS Visual Analytics sessions. Your custom colors are displayed in the color palette. For an example of the color palette, see Figure 42.1 on page 325.

Use a Gauge

About Gauges

A gauge is a dashboard indicator (also known as a KPI) that 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. They often appear in business intelligence dashboards.

Qualitative ranges are required for all gauges in the designer. You can populate the range intervals manually, or you can have them generated for you based on the range of the actual data.

Gauges in the designer support high cardinality.

For a definition and a picture of each gauge type, see “Gauges” on page 573.

Specify Gauge Properties

Starting in the 7.1 release, there is one gauge object on the **Objects** tab. When the gauge is on the canvas, you can use **Type** on the **Properties** tab to specify which type of gauge (bullet, dial, slider, speedometer, or thermometer) you want displayed in a report.

To specify properties for a gauge:

1. If it is not already selected, select the gauge in the canvas.
2. In the right pane, click the **Properties** tab.
3. Update the general properties for the gauge. You can update the **Name**, **Title**, **Format** (for the title's font style), and **Description**.

4 Update the specific properties for the gauge. The available properties depend on the selected gauge type.

Here are some details about the properties for gauges:

- By default, the **Enable selection in the viewers** property is selected for gauges. This means that users who use the web viewer or a mobile device can select the gauge and click to see the gauge name and any incoming filter information.
- (Optional) Clear **Show value label**, **Show range labels**, or both properties.
- Select a **Type** for the gauge. You can choose **Bullet**, **Dial**, **Slider**, **Speedometer**, or **Thermometer**.
- For the bullet, slider, and thermometer gauges, you can specify the **Direction** for displaying the gauge. The default for the bullet and slider gauges is **Horizontal**. The default for the thermometer gauge is **Vertical**.

Specify Gauge Styles

To specify styles for gauges:

1 If it is not already selected, select the gauge in the 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, you can specify **Border and Fill**, **Data Styling**, **Frame Styling**, **Text Styling**, and **Data Colors** for gauges.

Note: The **Header value** enables you to set only the text color.

Your custom colors are saved between SAS Visual Analytics sessions. Your custom colors are displayed in the color palette. For an example of the color palette, see Figure 42.1 on page 325.

Use a Word Cloud Object

About Word Cloud Objects

A word cloud displays a set of words from a character data item. Depending on the type of word cloud and your data roles, the size of each word in the cloud can indicate the relevance of the word to a topic, the frequency of the word in a category, or the value of a measure.

Every word cloud has a descending sort based on its size.

TIP There is a limit of 100 rows for a word cloud. To reduce the amount of data, add a rank for the word cloud, and use a number less than 100. For more information, see “Add a New Rank” on page 459.

Word clouds support display rules and interactions.

Specify Word Cloud Properties

To specify the properties for a word cloud:

1. If it is not already selected, select the word cloud in the canvas that you want to update.
2. In the right pane, click the **Properties** tab.
3. Update the general properties for the word cloud. You can change the **Name**, **Title**, **Format** (for the title's font style), and **Description**.
4. Update the object-specific properties for the word cloud. Here are some details about the properties for word clouds:
   - By default, the **Enable selection in the viewers** property is selected for word clouds. This means that users who use the web viewer or a mobile device can select the word cloud, and click to see the word cloud name and any incoming filter information.
   - Use the **Arrangement** property to specify how the words are displayed in the word cloud. **Cloud** is the default. Use **Rows** to display the words in rows as if you were reading text in a book.
   - Use the **Font scale** to control the difference in the size of the smallest and largest words. If you select `1`, there is a one-to-one ratio, so the smallest and largest words are displayed as the same size.
   - By default, the word cloud has a limit of 100 terms. After the results of a query come back, only the top 100 terms, based on a descending sort of the size measure, will be displayed. Use the **Word display limit** property to display fewer terms. The slider can be set to a number between five and 100.

Specify Word Cloud Styles

To specify styles for word clouds:

1. If it is not already selected, select the word cloud in the canvas that you want to update.
2. In the right pane, click the **Styles** tab.
3. Update the styles for the word cloud. For example, you can specify **Border and Fill**, **Frame Styling**, **Text Styling**, and **Data Colors**.

   > **TIP** A three-color **Gradient** data color style is available for word clouds that have at least two measures.

Your custom colors are saved between SAS Visual Analytics sessions. Your custom colors are displayed in the color palette. For an example of the color palette, see Figure 42.1 on page 325.

Using Custom Graphs to Display Results

About Custom Graphs

The graph builder enables you to create custom graph objects. To access the graph builder, on the **Objects** tab, click . The Create Custom Graph window is displayed. For more information, see Chapter 52, "Creating and Using Custom Graph Objects," on page 495.

When a custom graph is saved in the `My Folder` location, it is displayed under the **Custom** heading on the **Objects** tab. Then, you can insert the custom graph into a report and add data to it. For information about adding a custom graph to the **Objects** tab, see "Show or Hide Report Objects in the Objects Tab" on page 321.

Note: If a data role is not available for a report object in the designer, then the data role is not available in the graph builder. For example, a bar chart in the designer does not have a **Color** data role. Therefore, a bar chart in the graph builder does not have a **Color** data role either.

For a definition and a picture of each custom graph type, see "Graphs, Charts, and Plots" on page 555.

Specify Custom Graph Properties

To specify the properties for custom graphs:

1. If it is not already selected, select the custom graph in the canvas that you want to update.
2. In the right pane, click the **Properties** tab.
3. Update the general properties for the graph. You can update the **Name**, **Title**, **Format** (for the title's font style), and **Description**.
4. Update the specific properties for the graph. The available properties depend on the selected custom graph type.

   By default, the **Enable selection in the viewers** property is selected for custom graphs. This means that users who use the web viewer or a mobile device can select the custom graph and click to see the custom graph name and any incoming filter information.

   Note: No properties are available for the schedule chart.

Specify Custom Graph Styles

To specify styles for custom graphs:

1. If it is not already selected, select the custom graph in the 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 custom graph type. For example, you can specify **Border and Fill**, **Data Styling**, **Frame Styling**, **Text Styling**, and **Data Colors** for custom graphs.

   By default, the background of a graph is set to white. Use the **Wall background** option (under **Frame Styling**) to specify a different color.

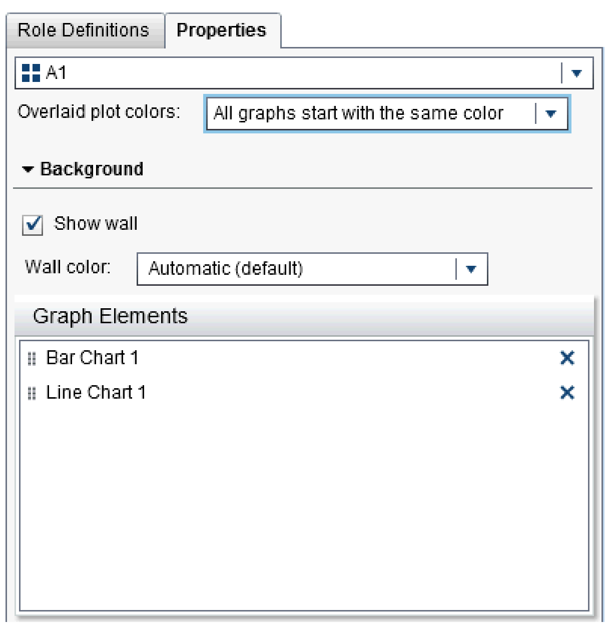
   Note: A three-color **Gradient** data color style is available for custom graphs that have a **Color** role.

Your custom colors are saved between SAS Visual Analytics sessions. Your custom colors are displayed in the color palette. For an example of the color palette, see Figure 42.1 on page 325.

Assign Colors to Overlays in Custom Graph Elements

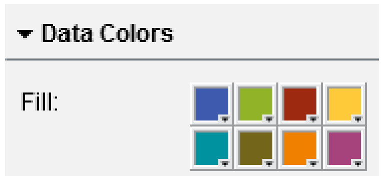
Each time a new graph element is overlaid on top of an existing graph element, the data colors that are assigned to each subsequent graph element are the colors after all of the colors have been assigned to the first graph and its elements. If 12 colors are assigned to the first graph element (for example, a bar chart), then the graph element that is overlaid (for example, a line chart) is assigned color 13.

To have the overlaid graph element start with the first color, change the **Overlaid plot colors** property for the cell in the graph builder to **All graphs start with the same color**.



Setting the **Overlaid plot colors** property affects the **Fill** colors on the **Styles** tab when you open the custom graph in the designer.

Figure 42.2 Data Colors on the Styles Tab in the Designer



If a custom graph contains more than eight groupings, the **Fill** colors are repeated. Here is the pattern for colors:

Colors	Pattern
1-8	The colors are specified in the **Fill** color palette.
9-16	Repeat colors 1 through 8 in a lighter hue.
17-24	Repeat colors 1 through 8 in a darker hue.
More than 24 colors	Repeat the color patterns above.

Duplicate a Report Object

Duplicating a report object in the designer enables you to use a copy of the same object in the same section or another section of your report.

Note: If you duplicate a control with a parameter, the parameter is not copied from the original control because the parameter can get its value from only one control.

To duplicate a report object:

1. On the canvas, right-click the report object that you want to duplicate.
2. Select **Duplicate <ReportObject>**, where <ReportObject> is the name of the report object in the report. (For example, List Table 1, Bar Chart 1, and so on.) The duplicated report object is placed on the canvas with a name based on the original name. For example, if the original report object name is List Table 1, then the duplicate report object is displayed as List Table 1 (1). If you choose to duplicate the same report object again, then it is displayed as List Table 1 (2).
3. (Optional) Move the duplicate report object to another section. Right-click the report object that you want to move. Select **Move <ReportObject> to ▸ <SectionName>**, where <ReportObject> is the name of the report object and <SectionName> is the name of the section.

4 (Optional) If you want the report object to appear in a different location, then drag and drop it in a new location.

If you move an object to a section with precision layout, then you must manually move the object to its proper location. All objects are put in the top left corner by default.

Working with Alerts for Report Objects

Overview of Alerts

You can create alerts for a report object so that subscribers are notified via e-mail or a text message when the alert condition is met. You can specify how frequently the system checks to see whether the alert condition is been met.

You can add alerts to report objects using the **Alerts** tab. You can also add alerts when you create a display rule. For more information about display rules, see "Adding Table-Level Display Rules" on page 403.

Note: You can specify a preference for receiving alert notifications via e-mail or a text message in both the designer and the viewer. For more information, see "Specify General Preferences for the Designer" on page 305.

Add an Alert

To add an alert to a report object:

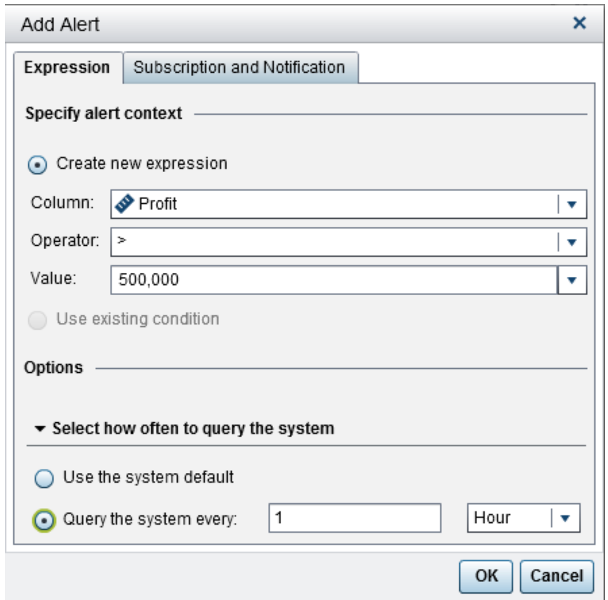
1 If it is not already selected, select the report object in the canvas to which you want to add an alert.

2 Do one of the following:

- Right-click the report object, and select **Add Alert**. The Add Alert window is displayed.
- In the right pane, click the **Alerts** tab. Click . The Add Alert window is displayed.

3 On the **Expression** tab, specify the criteria for the alert. You can create a new expression or use an existing expression.

4 (Optional) Specify how often you want the system to check for the criteria. You can use the system default, which is set by your administrator, or you can limit the check to a minute or hourly increment.



5 On the **Subscription and Notification** tab, add or remove subscribers for the alert. You can specify how frequently alert notifications are sent. For example, suppose that you specify that you want alert notifications sent every five days. Then, it will be at least five days before you receive an alert notification.

Note: Only users who have e-mail addresses stored in metadata are displayed in the **Manage subscribers** list on the **Subscription and Notification** tab.

Note: The timing of when alert notifications are sent can vary depending on when the alert condition is met.

6 Click **OK**. The details for the alert are displayed at the bottom of the **Alerts** tab.

Note: An alert notification has a blank subject. Some cellular carriers convert text messages to e-mail messages. When a subject is not specified in an e-mail, these carriers try to add a subject. Some carriers add the alert condition as the subject. Other carriers cannot add the subject, so the e-mail message has a blank subject.

Edit an Alert

To edit an alert:

1 In the right pane, click the **Alerts** tab.

2 Select the alert that you want to edit, and click . The Edit Alert window is displayed.

3 Update the alert criteria, and then click **OK** to save your changes.

Delete an Alert

Alerts are not automatically deleted when a report is deleted. You can delete an alert on the **Alerts** tab. Select the alert that you want to delete, and click 🗑. Then, click **Delete** in the confirmation message that is displayed.

43

Working with Data in SAS Visual Analytics Designer

Overview of Data Sources and Data Items

Each data source that is available in SAS Visual Analytics Designer (the designer) includes one or more data items that you can use in reports. 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.

For information about exporting data, see “Exporting Content from the Designer” on page 480.

Working with Data Sources in Reports

About Data Sources

Many data sources that are available in the designer are prepared by a data administrator or analyst so that you can easily define a report. Data administrators load tables into memory using SAS Visual Analytics Administrator (the administrator). Analysts can use SAS Visual Data Builder (the builder) to design queries that load tables into memory too.

The Add Data Source window can be used to add or import data sources. If you have the Import and Load Data capability, then you can import data from a file into the designer. Supported files are SAS data sets, Microsoft Excel spreadsheets, and delimited text files (such as CSV files). If you can import data sources, the Add Data Source window has **Import Data** in the right pane.

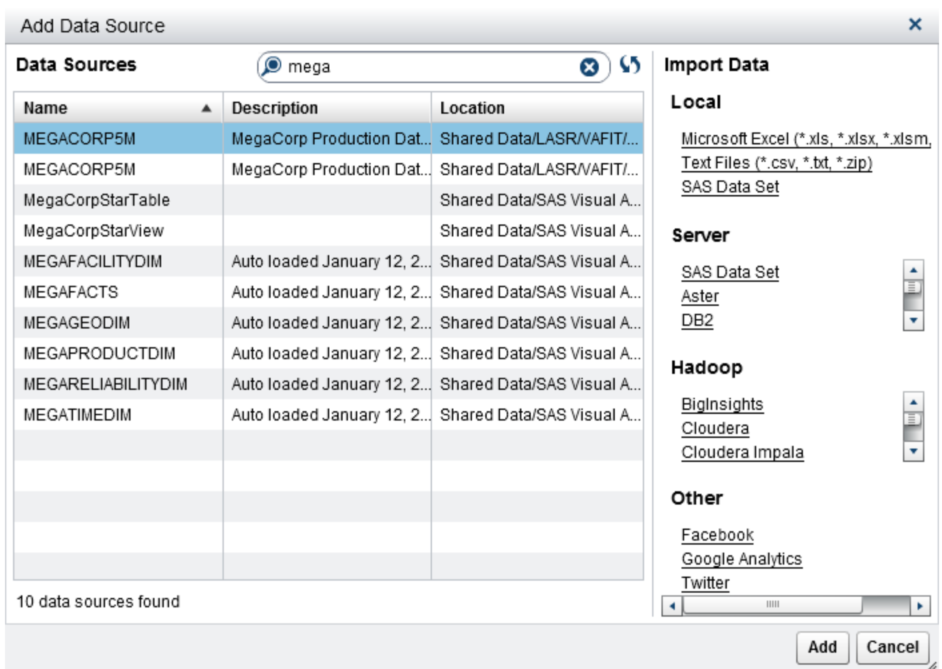
All data sources contain data items, which can refer to calculations or columns in physical data (tables). Reports can include query results from more than one data source.

Import a Data Source for a Report

If you have the Import and Load Data capability, then you can import a data source into the designer using either the Add Data Source window or the Change Data Source window. When you import data, the data source is automatically added to the open report. For more information about importing, see Chapter 4, “Importing Local Data Files,” on page 19 or Chapter 5, “Importing Data from Servers,” on page 25.

To import a data source for a report using the Add Data Source window:

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



2 In the **Import Data** pane, select a data source that you want to import.

> **TIP** If you select a delimited text file, then you can specify additional options. For example, you can specify the delimiter, whether the first row contains column heading names, and where the data rows begin.

> **TIP** If you select a spreadsheet, then you can specify additional options. For example, you can specify which worksheets to import, whether the first row contains column heading names, and where the data rows begin.

3 Click **Add**.

Add a Data Source to a Report

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

To add a data source to a report:

1 On the **Data** tab, click ▼ beside the **Select a data source** text.

The Add Data Source window is displayed.

2 In the Add Data Source window, select one or more data sources.

TIP Use the **Search** field to narrow the list of data sources that are displayed in the Add Data Source window. The search searches the **Name** and **Description** fields. It is a “begins with” search rather than a “contains” search. If you receive a message that a data source is not available, contact your data administrator.

3 Click **Add**. The list of available data items is displayed on the **Data** tab.

4 (Optional) To add additional data sources, click the add data source icon on the **Data** tab, which displays the Add Data Source window. Select the data sources 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 sources.

When you add multiple data sources, the last data source that you selected is displayed on the **Data** tab. If one of the data sources that you selected is not available, the last available data source that you selected is displayed on the **Data** tab.

Note: When you open a saved report that has multiple data sources, the designer displays the same data source that was displayed in the **Data** tab when the report was saved.

Refresh a Data Source for a Report

You can refresh the columns in a data source in the designer at any time. Be aware that refreshing a data source means that all live report objects that are connected to that data source will have their queries rerun.

Note: Refreshing a data source adds any new columns that have been added to the table metadata in the SAS LASR Analytic Server. The default formats and names of existing columns will be updated the next time you open the report.

Note: Data is refreshed from the table that is currently loaded into the SAS LASR Analytic Server.

To refresh a data source for a report, on the **Data** tab, select the data source, and then click .

To refresh the list of data sources when you are using the Add Data Source window, click beside the search field. This updates the list of all possible data sources. Individual data sources are not refreshed.

When you refresh a data source for a report, columns that have been deleted from the table metadata in the SAS LASR Analytic Server are automatically removed if they do not impact any objects in the report. If deleted columns do impact objects in the report, then the Repair Report window is displayed so that you can repair the objects that are impacted by the deleted columns. For more information, see Appendix 8, "Troubleshooting in SAS Visual Analytics Designer," on page 615.

Remove a Data Source from a Report

You can remove all references to a data source from a report in the designer. Be aware that removing a data source means that all related data items are also removed from the report objects within the report.

Note: Other reports that use the same data source are not affected when you remove a data source from a report.

To remove a data source from a report:

1 On the **Data** tab, select the data source, and then click .

2 Click **Delete** in the confirmation message that is displayed.

Change a Data Source in a Report

Note: The currency format will not adjust the locale if you change the data source and the locale associated with the second data source is different from the first data source.

To change a data source in a report:

1 On the **Data** tab, click , and then select **Change Data Source**. The Change Data Source window is displayed.

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

3 Click **Change**.

If a data item with the same name does not exist in the replacement data source, then it is automatically removed from the replacement data source if there are no report objects that use the data item. If there are report objects that use the data item, then those report objects will not work. The Repair Report window is displayed so that you can repair the objects that use the data item. For more information about repairing reports, see Appendix 8, "Troubleshooting in SAS Visual Analytics Designer," on page 615.

Data items in the replacement data source that have names that do not exist in the original data source are added to the **Data** tab automatically.

Note: For data item names, the case is ignored when data sources are compared by the designer.

View Measure Details

To view the details about all of the measures in a data source:

1. On the **Data** tab, click ▾, and then select **Measure Details**. The Measure Details window is displayed.
2. Click **Close**.

Show or Hide Data Items on the Data Tab

You can specify which data items you want to see for the data source on the **Data** tab. Your selections for which data items are shown or hidden are stored with the report. For example, suppose that you hide data items in one report, and then you open a second report that uses the same data source. The data items in the second report are not hidden unless you specifically hide them in that report, too.

Note: Hiding data items on the **Data** tab does not hide them in the entire report. This feature is not a way to implement column-level security.

To show or hide data items:

1. On the **Data** tab, click ▾, and then select **Show or Hide Items**. The Show or Hide Data Items window is displayed.
2. Select the data items that you want to appear on the **Data** tab. If there are data items that you do not want to see on the **Data** tab, then clear the check box (or check boxes) for that data item (or data items).
3. Click **OK**. The **Data** tab is updated.

Alternatively, you can select a data item that you want to hide on the **Data** tab. Right-click the data item, and select **Hide Data Item**.

TIP Starting in the 7.2 release, you can hide all of the data items that are not used in the current report. On the **Data** tab, click ▾, and then select **Only Show Used Items**.

You can also use a data source filter to restrict the data that is displayed in a report. For more information, see "Use a Data Source Filter in a Report" on page 427.

Sort Data Items on the Data Tab

To sort data items on the **Data** tab, click , and then select one of the following:

- **Sort Items ▸ Ascending By Name**
- **Sort Items ▸ Descending By Name**

The data items are sorted on the **Data** tab within each grouping. The default sort is **Ascending By Name**.

For information about sorting data values in report objects, see “Sorting Data in Reports” on page 395.

Group Data Items on the Data Tab

To group data items on the **Data** tab, click , and then select one of the following:

- **Group Items ▸ By First Letter**
- **Group Items ▸ By Data Type**
- **Group Items ▸ By Classification**
- **Group Items ▸ By Format**
- **Group Items ▸ By Aggregation**

The data items are grouped on the **Data** tab. The default is grouping **By Role**.

Working with Hierarchies in a Report

About Hierarchies

Creating hierarchies enables you to add drill-down functionality to your reports. 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 datetime columns with *Year* as the top level, *Month* as the next level, and *Day* as the bottom level.

You can also have a geographic hierarchy. For example, you might have a hierarchy with *Region* as the top level, *State* as the next level, and *City* as the bottom level.

You can have a maximum of two hierarchies for a report object.

Keep the following considerations in mind:

- List tables, controls, and gauges do not support hierarchies.
- Data item auto-assignment does not support hierarchies.
- Crosstabs can have either a hierarchy or categories, but not a combination of both, on each row or column.
- Time series plot report objects allow only datetime data items in a hierarchy.

- Geo bubble maps, geo coordinate maps, and geo region maps allow only geographic data items in a hierarchy.

You can also have a date hierarchy.

Create a New Hierarchy for a Report

To create a new hierarchy:

1 On the **Data** tab, click , and then select **New Hierarchy**. The New Hierarchy window is displayed.

2 Enter a **Name**.

3 Select at least two categories, and drag them to the **Hierarchy** list.

4 (Optional) Use the up and down arrows to arrange the data items in the **Hierarchy** list.

5 Click **OK** to save the new hierarchy. The icon identifies the new hierarchy in the list of data items on the **Data** tab.

TIP You can create a date hierarchy by right-clicking a date or datetime data item on the **Data** tab, and selecting **Create Date Hierarchy**. If you use a date data item, it must have a format that specifies the year or the **Create Date Hierarchy** option is not available.

Crosstab report objects enable you to create hierarchies from the categories on a crosstab axis. To create a hierarchy, right-click a category heading, and then select **Create Hierarchy**. The categories are replaced with a new hierarchy. The name of the new hierarchy is generated from the name of the outermost category with the suffix `Hierarchy`.

Edit a Hierarchy for a Report

To edit a hierarchy:

1 Right-click the hierarchy name on the **Data** tab, and select **Edit Hierarchy**. The Edit Hierarchy window is displayed.

2 (Optional) Edit the **Name**.

3 Add and remove categories. There must be at least two categories.

   Note: If a geographic hierarchy is used in a geo bubble map, geo coordinate map, or geo region map, only geographic data items are displayed when you edit the hierarchy.

4 Click **OK** to save the updated hierarchy.

If the hierarchy that you edit is already used in a report object and is drilled or expanded, it returns to the top level after it is edited.

Delete a Hierarchy for a Report

To delete a hierarchy:

1. Right-click the hierarchy name on the **Data** tab, and select **Delete Hierarchy**.
2. Click **Delete** in the confirmation message that is displayed. The hierarchy is removed from the list of data items, as well as from any report objects, filters, or ranks that were using it.

Working with Data Items in a Report

About 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.

The designer can display data items using an existing user-defined format that has already been specified externally for a data column in a data source. However, you cannot specify a new or different user-defined format for a data item in the designer.

You can create a custom sort so that data items in a table or graph can sort to the top or to the bottom. Category data items, calculated items that are categories, and custom categories also support custom sorts. For more information, see "Using a Custom Sort" on page 397.

For information about filtering data items, see "Use a Data Source Filter in a Report" on page 427.

Table 43.1 *Data Items That Are Available in the Designer*

Data Item	Icon	Description
Aggregated Measure or Time Period Calculation		A data item that represents special predefined operations, like distinct count, percentage of totals, percentage of subtotals, or frequency percent. Or, users can define their own aggregated measure calculations. Aggregate measures can be used in only some report objects. They cannot be used in filters, controls, spark lines, or time series graphs. Percentage of subtotal items (including row total, row subtotal, column total, and column subtotal) can be used only in crosstabs. Some aggregated measure calculations cannot be used in a detail rank.

Data Item	Icon	Description
Calculated	, , or 	A data item that is calculated from existing data items by using an expression. For example, you could create a calculated data item called *Profit*, which is created by using this expression: **`[Revenue] – [Cost]`**, where *Revenue* and *Cost* are measures in a data source. Calculated dates and times are treated as categories with distinct values being governed by the date or time format that you have chosen. Numeric calculated items can be treated as measures (with an aggregation type such as Sum, which is applied to each distinct category combination). Or, you can change numeric calculated items into category data items with distinct values being governed by the number of decimal places in the numeric format.
Category		A data item whose distinct values are used to group and aggregate measures. There are five types of categories: alphanumeric, date, datetime, time, and numeric. 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. Examples of alphanumeric categories include data items such as *Product ID*, *Country*, *Employee Number*, and *Employee Name*. Alphanumeric categories sort lexically. Date, datetime, time, and numeric categories are sorted by their underlying numeric values. Category data items can also be numeric. A category data item sorts differently than an alphanumeric data item. Numeric category data items sort by number. **Note:** If you change a measure to a category, then it uses this category icon. The  icon indicates a user-defined format category data item. User-defined format categories can be based on underlying numeric or character data.
Date and Time		A category data item whose distinct values are used to group and aggregate measures. There are three types of date categories: date, datetime, and time. Examples of date, datetime, and time categories are *Order Year*, *Date and Time of Sale*, and *Customer Wait Time*.

Data Item	Icon	Description
Frequency		A measure data item whose value represents the number of times an observation occurs in the selected data source. SAS Visual Analytics automatically adds this data item to the **Data** tab under the **Measure** heading when you select a data source. You cannot change the classification for the frequency data item. The frequency data item is automatically displayed in a crosstab when no measures are assigned. It is also automatically assigned to objects that require a measure when you have not specified one (for example, a bar chart).
Frequency Percent		A measure data item whose value is based on the percentage of occurrence in the selected data source. SAS Visual Analytics automatically adds this data item to the **Data** tab under the **Aggregated Measure** heading when you select a data source. You cannot change the classification for the frequency percent data item.
Geography		A category data item whose values are mapped to geographical locations or regions. Geography data items can be used in reports to show your data on a geographic map. For example, a geography data item can identify geographic information that is specific to your organization (for example, sales regions, warehouse locations, oil platforms, and so on). For more information, see “Working with Geography Data Items” on page 379. The geography icon can also indicate that all data items in a hierarchy are based on geography.
Hierarchy		A data item whose values are arranged with more general information at the top and more specific information at the bottom. The first level in the hierarchy is the root level. For example, you might have a *Date* hierarchy, which includes the *Year* (the root level), the *Quarter*, and then the *Month*. You can also have geographic hierarchies.
Measure		A data item whose values can be used in computations. These values are numeric. Examples of measures include *Sales Revenue*, *Units Sold*, and *Salary*. The designer assigns a default aggregation method to every measure. Almost all measures are assigned Sum. You can change the aggregation method.

Note: Report objects that are imported from SAS Visual Analytics Explorer (the explorer) use either live or on-demand data. For on-demand data, you can update the properties and styles for these report objects in the designer, but you cannot change the data assigned to them.

Assign Data Items

To assign 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 . Click **Delete** in the confirmation message that is displayed.

2 Select an existing report object that uses the same data source name or add a new report object to the section.

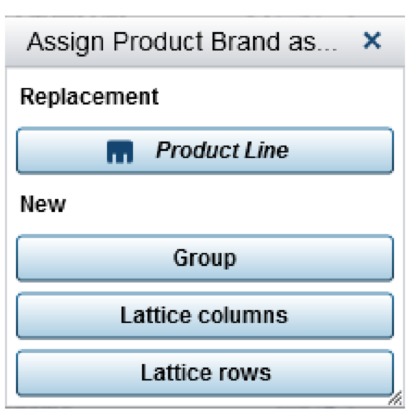
3 Drag and drop a data item onto the canvas. The data item is automatically assigned a data role. For more information, see “Automatic Data Item Assignments” on page 367.

   > **TIP** You can use the Ctrl key to select multiple data items, and then drag and drop them onto the canvas.

   Alternatively, you can right-click on a data item, and select **Add Data Item to <ReportObject>**, where `<ReportObject>` is the name of the report object in the report. (For example, `List Table 1`, `Bar Chart 1`, and so on.)

   Note: You cannot double-click the icon for a data item and assign it to the report.

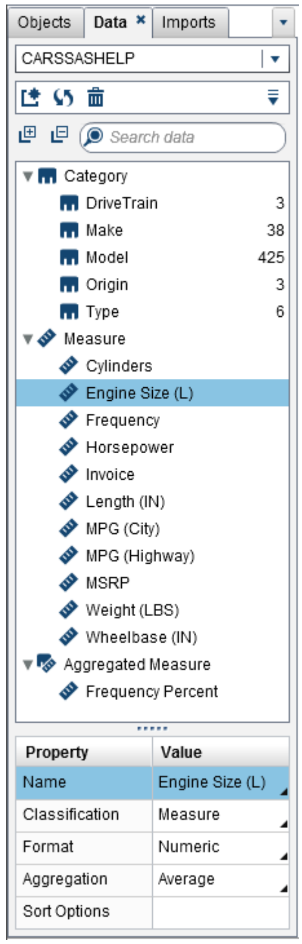
4 If the selected data item can replace a current data item in the report object or if the selected data item is valid for multiple data roles, then select the data assignment from the Assign <DataItemName> as window. A data assignment that is marked with a red asterisk (*) is required before the query can be run.



   For more information about aggregations, see Appendix 4, “Aggregations for Measures,” on page 581.

5 (Optional) To see more information about a data item, select it in the list. The **Name**, **Classification**, **Format**, and **Aggregation** are displayed in the data

item table below the list of data items. For a category data item with a user-defined format that has an underlying numeric value, you can specify **Sort Options**.



Automatic Data Item Assignments

When you drag and drop data items on a report object, the designer automatically assigns them a data role. For a single data item, the data item is assigned to an empty and required data role before you are prompted to replace an already assigned data item. For multiple data items, the data items are assigned to all empty and required data roles that accept the data items. There is a special case for a measure data role that allows multiple data items. In this special case, the designer automatically assigns all of the data items to the **Measures** data role.

Note: Hierarchy data items cannot be used in the automatic assignment of multiple data items. Also, the animation role and data tip role are not automatically assigned.

For more information about data roles, see “Working with Data Role Assignments” on page 386.

Duplicate Data Items

Duplicating measure data items in the designer enables you to see the aggregations of a data item side by side in a table. For a list of the available aggregation types, see Appendix 4, “Aggregations for Measures,” on page 581. You can duplicate a numeric measure if you want to use it as a category to group other values in some tables or graphs. If you save a report with duplicate data items, then those data items are available when you edit the report the next time.

You can duplicate a calculated data item to make a variation of a calculation. For example, you might make similar calculations involving miles per gallon for a vehicle, but you create one calculation using `MPG (City)` and another using `MPG (Highway)`. You can duplicate any data item if you want to use it with more than one format in your report. For example, you might change `Month` to `Year` for a date data item.

To duplicate a data item:

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

   All of the properties of the original data item are copied to the duplicate data item. The duplicate data item appears in the list of data items on the **Data** tab. 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)`.

2 (Optional) Rename the duplicate data item.

3 (Optional) Change the format or aggregation for the duplicate data item.

4 (Optional) Edit the calculation for a calculated data item or aggregated measure.

5 (Optional) Change the sort options for a category data item with a user-defined format that is based on an underlying numeric value.

6 (Optional) Change the classification for the data item. For example, a numeric data item that has been duplicated and is not yet assigned to a report object can be a category or a measure.

Rename Data Items

You can rename data items in the data source using the **Data** tab.

To rename data items:

1 On the **Data** tab, right-click on a data item, and then select **Rename Data Item**. The Rename Data Item window is displayed.

2 Enter a new name. The name cannot be used by another data item in the same data source.

3 Click **OK**.

Alternatively, you can use the data item table at the bottom of the **Data** tab to rename a data item. For the **Name** property, enter a new name for **Value**.

Search for Data Items

If your data source contains many data items, you can search for particular data items using the **Data** tab.

To search for data items:

1 Enter the name of a data item in the search field on the **Data** tab. The field is located above the list of data items.

2 (Optional) Click to collapse the list of data item groupings or click to expand the list of data item groupings.

3 Click to clear the search term and display all of the data items in the data source.

Create a Distinct Count for a Category Data Item

A distinct count query is useful in many ways. For example, you might want to know the number of distinct products that were purchased during a specific time period. Or, you might want to know which products have the most customers or which products have the most customers in a particular geographic region. You can create a distinct count for category data items only.

Note: If your category contains missing values, then distinct count is increased by one.

To create a distinct count aggregated measure data item:

1 On the **Data** tab in the left pane, right-click the category data item that you want to use for the distinct count.

2 Select **Create Distinct Count**.

The distinct count data item appears in the list of data items with a name that is derived from the original name. For example, if the original data item name is `Date`, then the distinct count data item is displayed as `Date (Distinct Count)`. The icon identifies the new distinct count data item on the **Data** tab.

Create Derived Items for Measures

You can create derived data items that are aggregated measures in the designer. The aggregated measure does not contain data values in itself, but when it is used in a report object, it displays the value for the measure and formula type on which it is based. An example is a percentage of total.

Here are some key points about derived items:

- Derived data items cannot be used in filters or controls.

- Period calculations cannot be derived from measures with certain aggregations.

To create a derived item from a report object:

1. In the report object, right-click on the measure data item that you want to use for the derived item.

2. Select **Create and Add**, and then select one of the following:

   Difference from Previous Period
   Displays the difference between the value for the current time period and the value for the previous time period. For example, you might derive the difference between sales for the current month and sales for the previous month.

   Note: This derived item is not available if your data source does not contain a date data item that includes the year.

   Difference from Previous Parallel Period
   Displays the difference between the value for the current time period and the value for the previous parallel time period within a larger time interval. For example, you might derive the difference between sales for the current month and sales for the same month of the previous year.

   Note: This derived item is not available if your data source does not contain a date data item that includes the year.

   Percent Difference from Previous Period
   Displays the percentage difference between the value for the current time period and the value for the previous time period. For example, you might derive the percentage difference between sales for the current month and sales for the previous month.

   Note: This derived item is not available if your data source does not contain a date data item that includes the year.

   Percent Difference from Previous Parallel Period
   Displays the percentage difference between the value for the current time period and the value for the previous parallel time period within a larger time interval. For example, you might derive the percentage difference between sales for the current month and sales for the same month of the previous year.

   Note: This derived item is not available if your data source does not contain a date data item that includes the year.

   Percent of Subtotals
   Displays the percentage of the subtotal value for the measure on which it is based. You can create a percentage of subtotal only when the source data item has an aggregation of Sum or Count.

   Note: The **Percent of Subtotals** derived item is available only for crosstabs.

   Note: The **Percent of Subtotals** derived item is relative to the subset of data that is selected by your filters and ranks.

   Percent of Total
   Displays the percentage of the total value for the measure on which it is based. You can create a percentage of total only when the source data item has an aggregation of Sum or Count. For example, you might create the percentage of the total value for a measure that contains revenue

values. If you create a bar chart of the aggregated measure and a category that contains product lines, then the bar chart shows the percentage of total revenue for each product line.

Note: The percentage of the total value is relative to the subset of data that is selected by your filters and ranks.

Period to Date
Displays the aggregated value for the current time period and all of the previous time periods within a larger time interval. For example, you might derive the year-to-date total for each month.

Note: This derived item is not available if your data source does not contain a date data item that includes the year.

Year to Date
Displays the aggregated value for the current time period and all of the previous time periods within the year. For example, you might derive the year-to-date total for each month.

The year-to-date calculation subsets the data for each year using today's date (where today is evaluated each time you view the report). To use all data for every period, edit the expression for the derived item.

Note: This derived item is not available if your data source does not contain a date data item that includes the year.

Year to Date Growth
Displays the percentage difference between the year-to-date value for the current time period and the year-to-date value for the same time period of the previous year. For example, you might derive the difference in year-to-date sales between the current month and the same month of the previous year.

The year-to-date calculation subsets the data for each year using today's date (where today is evaluated each time you view the report). To use all data for every period, use a **Period to Date** item or edit the expression for the derived item.

For the month that contains today's date, the data for an earlier year is subset to the same corresponding date.

Note: This derived item is not available if your data source does not contain a date data item that includes the year.

Year over Year Growth
Displays the percentage difference between the current time period and an equivalent time period from the previous year. For example, you might derive the difference in sales between the current month and the same month of the previous year.

The year-over-year growth calculation subsets the data for each year using today's date (where today is evaluated each time you view the report). To display a percentage of growth using full periods, use **Percent Difference from Previous Parallel Period** or edit the generated formula.

For the month that contains today's date, the data for an earlier year is subset to the same corresponding date.

Note: This derived item is not available if your data source does not contain a date data item that includes the year.

Alternatively, you can use the **Data** tab to create derived data items.

Create a Percentage of Total Using the Data Tab

A measure is required to have a Sum or Count aggregation before you can create a percentage of total.

To create a percentage of total from a measure data item using the **Data** tab:

1 On the **Data** tab in the left pane, right-click on the measure data item that you want to use for the percentage of total.

2 Select **Create** ▸ **Percent of Total**.

   The percentage of total measure data item appears in the list of aggregated data items with a name that is derived from the original name. For example, if the original measure data item name is `Revenue`, then the percentage of total measure data item is displayed as `Revenue (Percent of Total)`. The icon identifies the new percentage of total measure data item on the **Data** tab.

Create a Custom Category

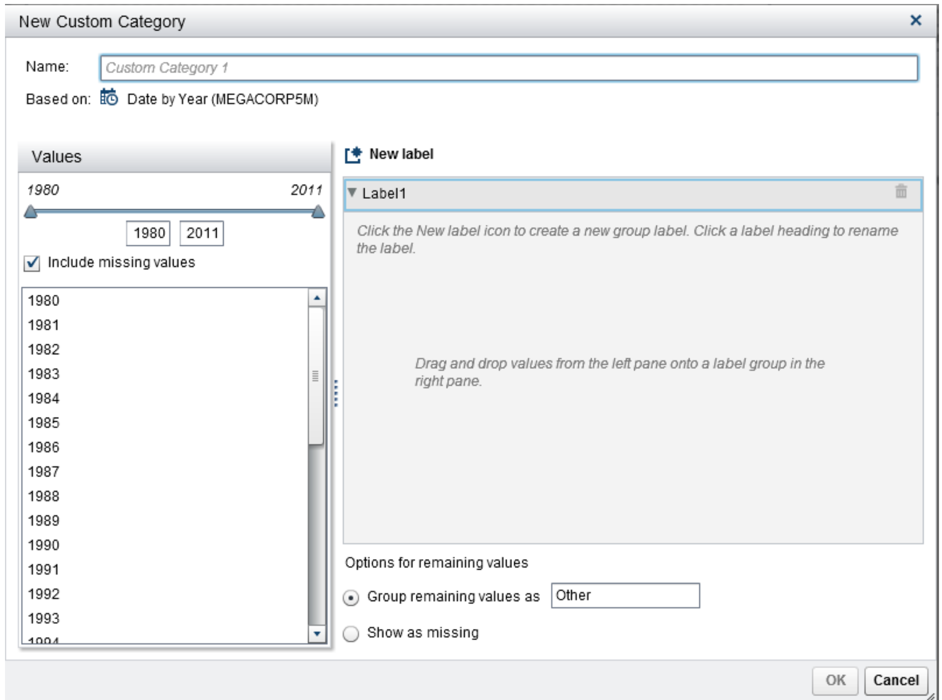
You can create a custom category based on either a category or measure data item. A custom category data item is always a category data item with an alphanumeric value.

When you create a custom category from a measure, you can use intervals, ranges, or specific values to group the data.

TIP The labels for your custom categories must use characters that are compatible with the locale of the data source. If the data source uses Unicode, then your labels can contain characters from any locale.

To create a custom category:

1 On the **Data** tab, right-click on the category or measure data item that you want to use to create the new custom category. Select **New Custom Category**. The New Custom Category window is displayed.



Note: In the New Custom Category window, **Based on** identifies the type and the name of the data item that the custom category is based on.

2 Select the **Values** that you want to include. Drag and drop the values onto the right pane under **New label**.

TIP To rename a custom group label, right-click the label name, and select **Edit**.

3 (Optional) Specify the **Options for remaining values**. You can specify a name when you select **Group remaining values as**. The default label for **Group remaining values as** is `Other`. Alternatively, you can specify **Show as missing** or **Show as is**.

Note: The **Show as is** option is available only for data items that are based on string categories. It is not available for numeric or date values.

4 Specify a **Name** for the new custom category. The default name is `CustomCategory1`.

5 Click **OK**. The new custom category is displayed on the **Data** tab.

Modify 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. The name cannot be used by another data item in the same data source. Your change is saved automatically.

Alternatively, you can right-click the data item, and select **Rename Data Item**.

Modify a Data Item's Classification

You can modify a data item's classification. For example, you might want to modify a measure data item to be a category data item. A data item's classification cannot be modified if that data item is in use in the report or if the data item can have one classification only. However, duplicating the data item allows the new data item to have a different classification.

Here are some key points about modifying a data item's classification:

- You can modify a category data item to be a measure data item only if the data item started as a numeric measure. For example, if you change a measure to a category in the report, then the designer allows you to change it back to a measure. In addition, you can change a category to a measure if it was originally in the data source as a numeric column. In this case, your data administrator converted the numeric column to a character string by applying a user-defined format. As a result, you can use the designer to change the category's format to one of the standard numeric formats, and then you can change it to a measure.
- You cannot change the classification for the frequency data item or the frequency percent data item.
- You cannot change the classification for an aggregated measure.
- You cannot convert calculated data items into geography data items.

> **TIP** For geography data items, use the pop-up menu instead of the drop-down menu to change the classification. The pop-up menu enables you to specify additional classification information for the geography data item.

To modify a data item's classification:

1. Select a data item on the **Data** tab.
2. In the data item table, select the existing classification for the data item.
3. Click to open the drop-down menu. Select **Measure** or **Category**. Your change is saved automatically.

Modify the Format of a Numeric Measure Data Item or a Date, Datetime, or Time Data Item

You can modify the format of a numeric measure data item or a date, datetime, or time data item. You can also modify the format of a data item with a user-defined format as long as the user-defined format is based on an underlying numeric value. For more information about user-defined formats, see "Modify User-Defined Formats" on page 375.

Here are some key points about modifying a data item's format:

- You cannot modify the format of a data item that is being used in a filter, as part of a calculated or aggregated measure, or in a custom category.
- For line charts with multiple measures, formats are removed from all of the measure labels if the formats are different.

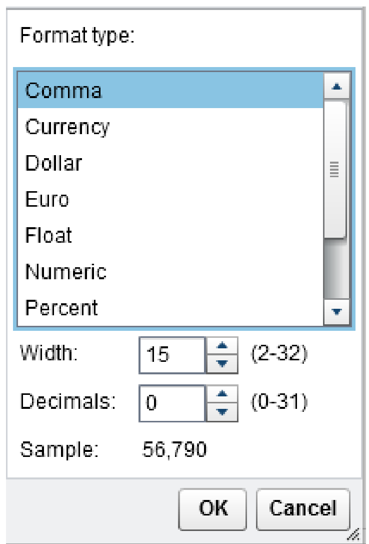
To change the format:

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

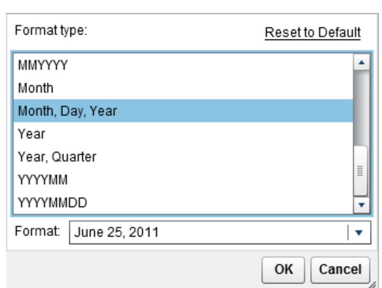
2 In the data item table, select the existing format. A list is displayed with the **Format type**, **Width**, and **Decimals** (for numeric data items). Make your selections. A sample of your selection is displayed under the list.

Note: There are different format variations available for some format types for date, datetime, and time data items. Select the format variation based on the sample value displayed in the **Format** drop-down list.

Here is an example of the list for a numeric data item:



Here is an example of the list for a date data item:



Note: The **Reset to Default** option is displayed only if the format has been changed from the default.

Note: The **Reset to Default** option is available for user-defined format data items after they have been modified to a standard numeric format as long as the data item is still a category data item.

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

Modify User-Defined Formats

In the designer, user-defined formats that are defined in the SAS LASR Analytic Server are applied to the results. You can change the format for an underlying numeric data item, but you cannot change the format for an underlying character-based data item. If you change the format for an underlying numeric data item, you can restore the user-defined format by selecting **Reset to Default**.

The **Format** property of the data item displays the name for a user-defined format.

The icon identifies a category data item with an active user-defined format on the **Data** tab.

Modify How a Measure Is Aggregated

You can change the aggregation method for a measure in a data source using the **Data** tab or a report object 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:

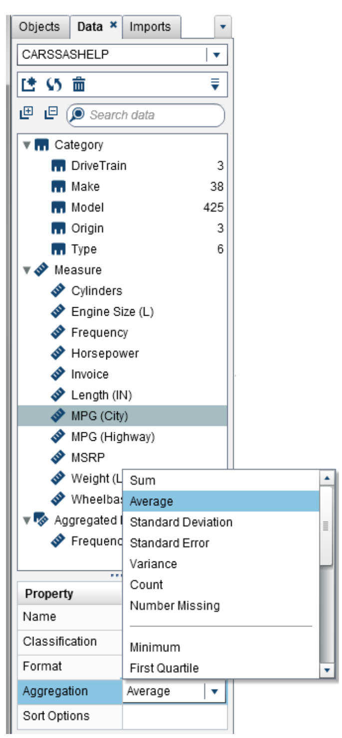
1 Select a measure data item on the **Data** tab.

2 In the data item table, select the existing aggregation, and then click . A drop-down list is displayed with aggregations. For a list of the available aggregation types, see Appendix 4, "Aggregations for Measures," on page 581.

   When you select an aggregation, your change is saved automatically.

   Note: All of the report objects in the report that use this data item are affected by this change unless you have selected a local aggregation override.

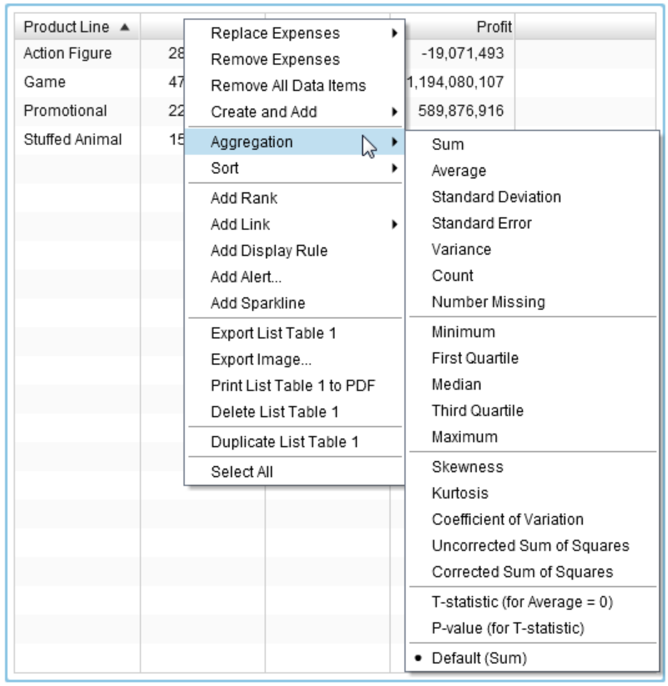
   Note: Depending on the aggregation, formats might be overridden when they are used in report objects. For example, skewness becomes a floating point number with four decimals.



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

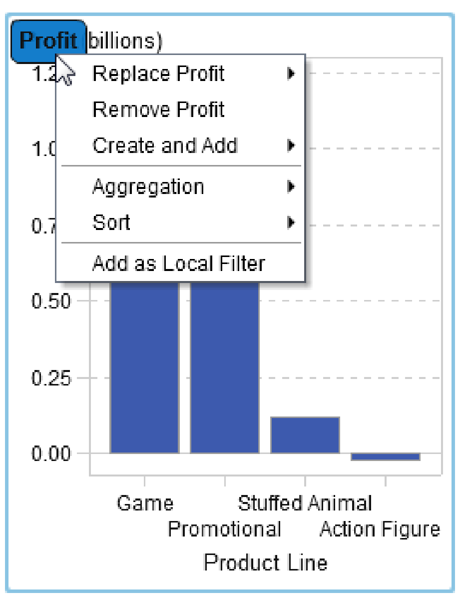
1 Choose one of the following:

- For a list table, right-click on a measure header in the report object.



- For a crosstab, right-click on the header row with the measure. Depending on the aggregation, formats might be overridden when they are used in a crosstab.

- For a graph, right-click on the measure name hotspot. Or, you can right-click on a measure data item name on the **Roles** tab.



2 Select **Aggregation ▸ <aggregation-name>**, where `aggregation-name` is one of the available aggregations. For a description of the available

aggregation types, see Appendix 4, “Aggregations for Measures,” on page 581.

Your change is saved automatically.

Note: This is a local override for only this report object. It does not affect the default aggregation for this data item in other report objects.

Modify the Sort Options for a Category Data Item

Using the data item table on the **Data** tab, you can change the **Sort Options** for a category data item with a user-defined format that is based on an underlying numeric value. The sort options are the following:

Formatted
: Uses the formatted character output of the user-defined format and sorts lexically. (For example, the names of the months in the year would sort as April, August, December, February, and so on.) Formatted is the default.

Unformatted
: Uses the underlying numeric value and sorts numerically. (For example, if the underlying numeric value for January is 1, for February is 2, for March is 3, and for April is 4, then the months of the year would sort as January, February, March, April, and so on.)

Delete Data Items

You can delete data items that you have created in the designer (for example, calculated data items or duplicated data items) so that they no longer appear on the **Data** tab in the left pane.

You cannot delete a data item that is inside a hierarchy if it reduces the hierarchy to a single level. You cannot delete a data item if it is the last or only reference to a column in the original data source. You cannot delete a data item if it is used in a calculated data item, aggregated measure, a geography data item, or a custom category.

> **TIP** You can hide a data item that you cannot delete by right-clicking on the data item name in the **Data** tab, and then selecting **Hide Data Item**. For more information, see “Show or Hide Data Items on the Data Tab” on page 360.

To delete a data item:

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

2 Select **Delete Data Item**.

3 Click **Delete** in the confirmation message that is displayed. The data item is removed from the list of data items, as well as from any report objects, filters, or ranks that were using it.

Working with Geography Data Items

About Geographic Data Items

A geography data item can be useful if your data contains values that are mapped to geographical locations or regions. For example, a geography data item can identify geographic information that is specific to your organization (for example, sales regions, warehouse locations, oil platforms, and so on).

If you change a numeric measure to a geography data item, then it automatically becomes a category data item.

Note: Custom data items cannot be changed into geography data items.

Use a Predefined Geography Data Item

You can use the geography data items that are available in SAS Visual Analytics or you can create custom geography data items. To access existing geography data items, select the **Data** tab in the left pane. Right-click the data item that you want to use for the geography data item, and then select **Geography**. Then, select one of the following items:

- **Country or Region Names**
- **Country or Region ISO 2-Letter Codes**
- **Country or Region ISO Numeric Codes**
- **Country or Region SAS Map ID Values**
- **Subdivision (State, Region, Province) Names**
- **Subdivision (State, Region, Province) SAS Map ID Values**
- **US State Names**
- **US State Abbreviations**
- **US ZIP Codes**

Note: For predefined geographic roles, the values of your geography data items must match the lookup values that are used by SAS Visual Analytics. To view the lookup values, see http://support.sas.com/va72geo.

TIP To get geographic maps to work with any data set, add a column with the predefined lookup values to your data set.

Create a Custom Geography Data Item

A custom geography data item consists of three values:

- Latitude
- Longitude
- A category (other than the latitude or the longitude)

For example, suppose you have data that includes airport location identifiers, latitude, and longitude. You can convert the airport location identifiers to custom geography data items.

To create a custom geography data item:

1 On the **Data** tab in the left pane, right-click the data item that you want to use for the custom geography data item. Select **Geography**, and then select **Custom**. The Geography window is displayed.

> **TIP** Make sure that the data item that you select is a category data item other than the latitude or the longitude.

2 Select a measure for the **Latitude**. You can also enter the first letter of the name of the latitude column to quickly search for it in the drop-down menu.

3 Select a measure for the **Longitude**. You can also enter the first letter of the name of the longitude column to quickly search for it in the drop-down menu.

4 Select a **Coordinate space** (coordinate system) that is used to project the longitude and latitude coordinate values. Your choices are **World Geodetic System (WGS84)**, **Web Mercator**, and **British National Grid (OSGB36)**. The default is **World Geodetic System (WGS84)**.

   Note: The coordinate space should match the projection that your data is in. The designer supports the World Geodetic System (WGS84), Web Mercator (EPSG:3857), and British National Grid (EPSG:27700) for data coordinate space.

5 Click **OK**. The icon identifies the new geography data item on the **Data** tab.

Working with Calculated Items in Reports

About Calculated Data Items

The designer enables you to calculate new data items from your existing data items by using an expression. For example, you might want to calculate a company's profits by subtracting expenses from revenues.

In addition to performing mathematical calculations on numeric values, you can use calculated data items to create date and time values. For example, if your data contains separate categories for *month*, *day*, and *year*, then you can calculate a date value from those categories.

Here are some key points about calculated data items:

- All calculations are performed on unaggregated data. The calculation expression is evaluated for each row in the data source before aggregations are performed. To perform calculations on aggregated data, see "Add a New Aggregated Measure to a Report" on page 384.
- Calculated data items can accept parameters. For more information, see Chapter 49, "Working with Parameters in Reports," on page 463.
- A hierarchy can contain calculated data items as long as they are categories.

- Starting in the 7.2 release, calculated data items can be changed into geographic data items and used in geo maps.

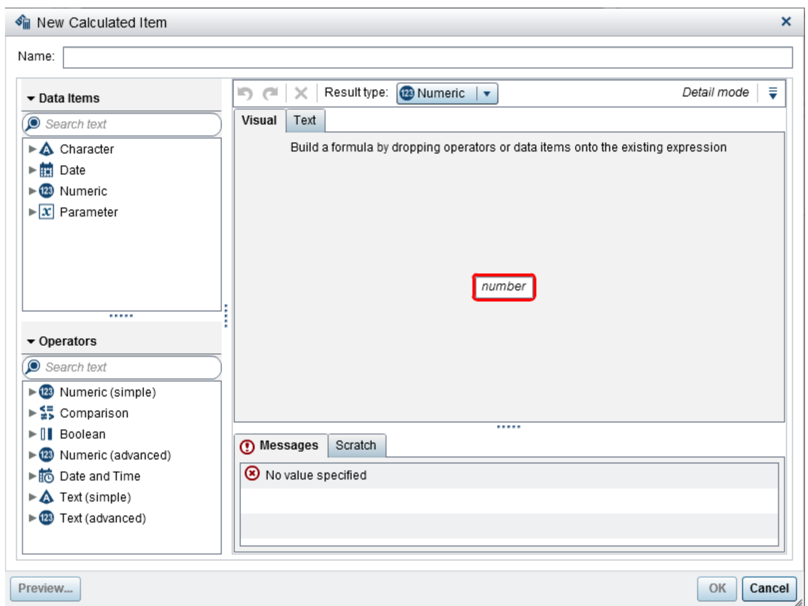
Using the designer, you can work with calculated data items or aggregated measures.

For information about deleting calculated data items, see "Delete Data Items" on page 378.

Add a New Calculated Data Item to a Report

To add a new calculated data item:

1 On the **Data** tab, click , and then select **New Calculated Item**. The New Calculated Item window is displayed.



2 Enter a **Name**.

3 Select a **Result type** from the drop-down list. **Numeric** is the default data type.

The following table lists the available result types:

Icon	Result Type
	Character
	Date
	Datetime
	Numeric
	Time

Calculated data items in the designer always default to the following formats, which are based on the data type:

- Date: DATE9
- Datetime: DATETIME10
- Time: TIME8
- Numeric: COMMA12.2

After you create the new calculated data item, you can change its format using the data item table on the **Data** tab unless it is a character calculated data item. You cannot specify a format for a character calculated data item.

Sum is the default aggregation for new numeric calculated data items in the designer. You can change the aggregation for numeric calculated data items using the data item table on the **Data** tab.

4 (Optional) Click beside **Detail mode** to clear the **Show all drop zones** option. You can also select **Show display text**, which adds a new tab between the **Messages** and **Scratch** tabs.

5 Use the **Visual** tab to build the expression for your calculated data item by dragging **Data Items** and **Operators** onto the expression in the right pane. For each rectangular field in the expression, you can insert a data item, an operator, or a specific value.

When you drag and drop data items or operators onto your expression, the precise location of the cursor determines where and how the new element is added to the expression. As you drag the new element over the expression, a preview appears that displays how the expression changes if you drop the element at that location.

For example, if your current expression is `( Profit / Revenue )`, and you drag and drop the **x - y (subtract)** operator inside the open parenthesis symbol, then the expression changes to `( [number] - ( Profit / Revenue ))`. If you drag and drop the operator over the division symbol, then the expression changes to `( Profit - Revenue )`, and so on.

Alternatively, you can use the **Text** tab to enter the expression.

You can use the **Scratch** tab to build temporary expressions.

There are a large number of operator types available to perform mathematical functions, process datetime values, handle text, and evaluate logical processing such as IF clauses. For more information, see Appendix 5, "Operators for Data Expressions," on page 583.

6 (Optional) Click **Preview** to preview the results of the calculation. The Preview Results window is displayed. Click **Close** to return to the New Calculated Item window.

7 Click **OK**. The new calculated data item appears on the **Data** tab. The , , or icon identifies the new calculated data item on the **Data** tab.

Note: The icon is displayed if you change a calculated numeric measure to a category data item. It is also displayed for a character calculated data item (which is always a category).

Preview the Expression for a Calculated Data Item

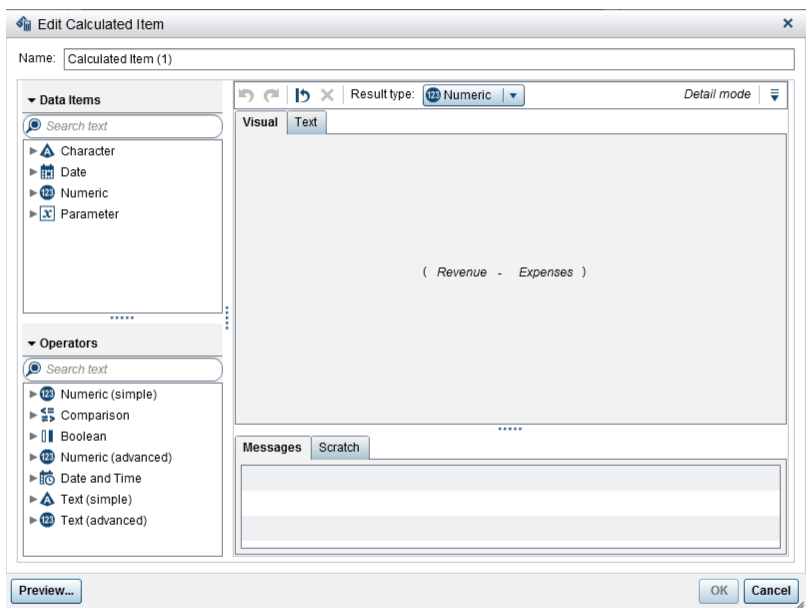
For calculated data items only, you can preview the results of your expression by clicking **Preview**.

You can preview the results of a subset of your expression by right-clicking a part of your expression, and then selecting **Preview Subexpression Results**.

Edit a Calculated Data Item

To edit a calculated data item:

1 Right-click on a calculated data item on the **Data** tab, and select **Edit Calculated Item**. The Edit Calculated Item window is displayed.



2 Modify the **Data Items** and **Operators** for the calculated data item as needed. For information about the operators that are available, see Appendix 5, “Operators for Data Expressions,” on page 583.

Note: If the calculated data item has not been used in a report, then you can modify the **Result type**.

3 Click **OK**.

You can duplicate, rename, hide, and delete calculated data items using the same steps as any other data item. If a calculated data item is used inside another calculated data item, then it cannot be removed.

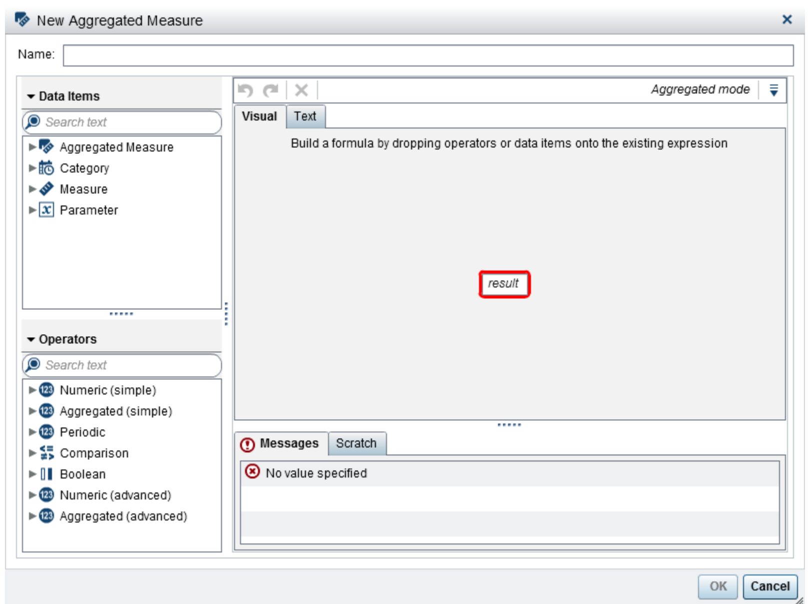
TIP You can cut and paste from the **Text** area between different reports as well as between the explorer and the designer to transfer calculations or to e-mail them to others. For more information, see “Editing a Data Expression in Text Mode” on page 577.

Add a New Aggregated Measure to a Report

Aggregated measures enable you to calculate new data items by using aggregated values. For example, you might want to calculate a company's profit margin by region by taking the aggregated sum of the profit for all of the stores in a region group and dividing it by the aggregated sum of the revenue for all of the stores in that same region group. Aggregations are evaluated as part of a calculated expression.

To add an aggregated measure:

1 On the **Data** tab, click ▼, and then select **New Aggregated Measure**. The New Aggregated Measure window is displayed.



2 Enter a **Name**.

3 (Optional) Click ▼ to **Show all drop zones**. You can also choose to **Show scratch area** to build temporary expressions.

4 Use the **Visual** tab to build the expression for your aggregated measure by dragging and dropping **Data Items** and **Operators** onto the expression in the right pane. For each field in the expression, you can insert a data item, an operator, or a specific value.

When you drag and drop data items and operators onto the expression, the precise location of the cursor determines where and how the data item or operator is added to the expression. As you drag the new element over the expression, a preview appears, which displays how the expression would change if you drop the element at the current location.

Alternatively, you can use the **Text** tab to enter the expression.

There are a large number of operator types available to perform mathematical functions and evaluate logical processing such as IF clauses. For more information, see Appendix 5, “Operators for Data Expressions,” on page 583.

5 For each data item in your expression, select an aggregation type. By default, **Sum** is used for measures and **Distinct** is used for categories. To select a new aggregation type, drag and drop an aggregated operator from the **Operators** list onto the aggregation type in the expression. See Appendix 5, “Operators for Data Expressions,” on page 583 for a list of the aggregated operators that are available.

6 For each aggregation in your expression, select the aggregation context. A drop-down list beside each aggregation enables you to select one of the following context values:

ByGroup
: calculates the aggregation for each subset of the data item that is used in a visualization. For example, in a bar chart, an aggregated measure with the **ByGroup** context calculates a separate aggregated value for each bar in the chart.

ForAll
: calculates the aggregation for the entire data item (after filtering). For example, in a bar chart, an aggregated measure with the **ForAll** context uses the same aggregated value (calculated for the entire data item) for each bar in the chart.

By using the **ForAll** and **ByGroup** contexts together, you can create measures that compare the local value to the global value. For example, you might calculate the difference from mean by using an expression such as the following:

```
Avg ByGroup(X) - Avg ForAll(X)
```

For more information, see “Periodic Operators” on page 591.

7 Click **OK**. The new aggregated measure appears on the **Data** tab. The  icon identifies the new aggregated measure.

Edit an Aggregated Measure

To edit an aggregated measure:

1 Right-click on an aggregated measure on the **Data** tab, and select **Edit Aggregated Measure**. The Edit Aggregated Measure window is displayed.

2 Modify the **Data Items** and **Operators** for the aggregated measure as needed. For information about the operators that are available, see Appendix 5, “Operators for Data Expressions,” on page 583.

3 Click **OK**.

You can duplicate, rename, hide, and delete aggregated measures using the same steps as any other data item.

> **TIP** You can cut and paste from the **Text** area between different reports, as well as between the explorer and the designer, to transfer calculations or to e-mail them to others. For more information, see “Editing a Data Expression in Text Mode” on page 577.

Working with Data Role Assignments

About Data Roles

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 roles.

A *data role* is a designation that describes how a particular data item is to be used in a report object. In the designer, each report object has data roles, some are required, and others are optional. For example, the data roles for a bar chart are **Category**, **Measures**, **Group**, **Lattice columns**, **Lattice rows**, **Data tip values**, and **Animation**. For the bar chart, the category and measures data roles are required.

Not all report objects have the same data roles. For example, the **Color** data role is available for only scatter plots, bubble plots, treemaps, geo bubble maps, geo region maps, and word clouds.

Note: You cannot add Frequency or Frequency Percent data items for dynamic text using the **Roles** tab.

Modify Data Role Assignments in Report Objects

Note: You cannot change data role assignments for any report object that is imported from the explorer and uses on-demand data.

To modify data role assignments:

1. Select a report object in the canvas that has one or more data items assigned.

2. Click the **Roles** tab in the right pane.

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

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

   Different data roles are available for different report objects.

   In addition to the basic data roles, you can assign additional data roles. Here is a list of all the data roles available for each report object:

Report Object	Basic Data Roles	Additional Data Roles
Tables		
List table	**Columns**	None
Crosstab	■ **Columns** ■ **Rows**	**Measures**

Report Object	Basic Data Roles	Additional Data Roles
Graphs		
Bar chart	■ **Category** ■ **Measures**	■ **Group** ■ **Lattice columns** ■ **Lattice rows** ■ **Data tip values** ■ **Animation**
Targeted bar chart	■ **Category** ■ **Measure** ■ **Target**	■ **Lattice columns** ■ **Lattice rows** ■ **Data tip values** ■ **Animation**
Waterfall chart	■ **Category** ■ **Measure**	■ **Lattice columns** ■ **Lattice rows** ■ **Data tip values**
Line chart*	■ **Category** ■ **Measures**	■ **Group** ■ **Lattice columns** ■ **Lattice rows** ■ **Data tip values** ■ **Animation**
Pie chart	■ **Category** ■ **Measures**	■ **Group** ■ **Lattice columns** ■ **Lattice rows** ■ **Data tip values** ■ **Animation**
Scatter plot	■ **X axis** ■ **Y axis**	■ **Color** ■ **Lattice columns** ■ **Lattice rows** ■ **Data tip values**
Time series plot	■ **Time axis** ■ **Measure (line)**	■ **Group** ■ **Data tip values**
Bubble plot	■ **X axis** ■ **Y axis** ■ **Size**	■ **Group** ■ **Color** ■ **Lattice columns** ■ **Lattice rows** ■ **Data tip values** ■ **Animation**
Treemap	■ **Tile** ■ **Size**	■ **Color** ■ **Data tip values**

Report Object	Basic Data Roles	Additional Data Roles
Dual axis bar chart	■ **Category** ■ **Measure (bar)** ■ **Measure (bar 2)**	■ **Lattice columns** ■ **Lattice rows** ■ **Data tip values** ■ **Animation**
Dual axis line chart	■ **Category** ■ **Measure (line)** ■ **Measure (line 2)**	■ **Lattice columns** ■ **Lattice rows** ■ **Data tip values** ■ **Animation**
Dual axis bar-line chart	■ **Category** ■ **Measure (bar)** ■ **Measure (line)**	■ **Lattice columns** ■ **Lattice rows** ■ **Data tip values** ■ **Animation**
Dual axis time series plot	■ **Time axis** ■ **Measure (line)** ■ **Measure (line 2)**	**Data tip values**
Schedule chart**	■ **Task** ■ **Start** ■ **Finish**	■ **Group** ■ **Label** ■ **Data tip values** ■ **Lattice columns** ■ **Lattice rows**
Vector plot**	■ **X axis** ■ **Y axis** ■ **X Origin** ■ **Y Origin**	■ **Color** ■ **Group** ■ **Data tip values** ■ **Lattice columns** ■ **Lattice rows**
Numeric series plot**	■ **X axis** ■ **Y axis**	■ **Group** ■ **Label** ■ **Data tip values** ■ **Lattice columns** ■ **Lattice rows**
Needle plot**	■ **X axis** ■ **Y axis**	■ **Group** ■ **Data tip values** ■ **Lattice columns** ■ **Lattice rows**

Report Object	Basic Data Roles	Additional Data Roles
Step plot**	■ **X axis** ■ **Y axis**	■ **Group** ■ **Label** ■ **Data tip values** ■ **Lattice columns** ■ **Lattice rows**
Dot plot**	■ **X axis** ■ **Y axis**	■ **Data tip values** ■ **Lattice columns** ■ **Lattice rows**
Butterfly chart**	■ **Category** ■ **Measure (bar)** ■ **Measure (bar 2)**	**Data tip values**
Stock high-low plot**	■ **Time axis** ■ **Close** ■ **High** ■ **Low**	None
Stock volume and volatility plot**	■ **Time axis** ■ **Bollinger Upper** ■ **Bollinger Lower** ■ **Moving Average** ■ **High** ■ **Low** ■ **Close** ■ **Volume**	None
Bubble change plot**	■ **Group** ■ **Y end** ■ **X end** ■ **Y start** ■ **X start** ■ **Size start** ■ **Size end**	None
Comparative time series plot**	■ **Time axis** ■ **Measure (time series)** ■ **Measure (time series 2)**	**Data tip values**
Controls		

Report Object	Basic Data Roles	Additional Data Roles
Drop-down list	■ **Category** ■ **Frequency** ■ **Parameter**	None
List	■ **Category** ■ **Frequency**	None
Button bar	■ **Category** ■ **Frequency** ■ **Parameter**	None
Text input	■ **Category** ■ **Frequency** ■ **Parameter**	None
Slider	■ **Measure/Date** ■ **Parameter**	None
Containers		
Vertical container Horizontal container Stack container Prompt container	None	None
Other		
Text Image Stored process	None	None
Geo bubble map	**Geography**	■ **Size** ■ **Color** ■ **Data tip values** ■ **Animation**
Geo coordinate map	**Geography**	**Data tip values**
Geo region map	**Geography**	■ **Color** ■ **Data tip values**
Gauge	**Measure**	■ **Target** ■ **Group** ■ **Data tip values**

Report Object	Basic Data Roles	Additional Data Roles
Word cloud	■ **Word** ■ **Size** ■ **Color**	None

* For line charts with multiple measures, formats are removed from all of the measure labels if the formats are different.

** This report object is not displayed by default in the designer.

Here are definitions for some of the additional data roles:

Animation
animates the date category that you assign. You can specify a **Loop** for the animation, and you can specify the **Speed** of the animation. For more information, see "Add Animation to Charts, Bubble Plots, and Geo Bubble Maps" on page 391.

Note: For a bubble plot only, you must assign the **Group** role to enable the **Animation** role. If you remove the **Group** role after assigning animation, the animation control is disabled.

Data tip values
enables you to add more measures to your data tips.

Note: For pie charts, the Other slice does not show the additional information.

Group
groups the data based on the category data item that you assign.

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.

Add Animation to Charts, Bubble Plots, and Geo Bubble Maps

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.

Starting in the 7.1 release, you can select bubbles, bars, pie slices, and lines in an animated graph. The interaction filter is based on the category value (or values) and the selected frame in the animation. When a bubble, bar, pie slice, or line is selected, its row lists all of the category and date values of the selection. For example, the resulting filter that is applied to any downstream interactions could be filtered by `car="Toyota"` and `year="2014"`.

If you select a bubble for one frame in the animation and click , then any downstream interactions play with the animation. This means that as the animation is running, the interaction targets are filtered with each frame in the animation.

TIP To improve animation performance, use a color for the graph's reference line instead of a pattern.

TIP When you add animation to a pie chart, it is recommended that you sort the Category in ascending or descending order. To sort, right-click on a category name in the pie chart, and select **Sort**. Then, select either **Ascending** or **Descending**.

To add animation:

1 Select an existing chart, bubble plot, or geo bubble map, or create a new one.

2 Add the required roles:

- For a bubble plot, on the **Role** tab, assign a data item to the **Group** data role.

  Note: If you remove the **Group** data role after assigning animation, the animation control is disabled.

- For a bar chart or geo bubble map, on the **Role** tab, assign the other required roles.

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

When an animated bubble plot is displayed, a set of animation controls appears at the bottom of the report object.

Table 43.2 Animation Controls

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.

Note: If you select a bubble to track, the selected bubble is highlighted in the current animation frame.

Add Data Roles for Lattice Columns or Lattice Rows in a Graph

A lattice is a multi-cell graph in which you create each cell independently. Each cell can contain different types of plots. In the designer, you can create a multi-cell graph by using data roles to add lattice columns, lattice rows, or both. For a list of graphs that have lattice column or lattice row data roles, see "Modify Data Role Assignments in Report Objects" on page 386.

Note: There is a limit of 15 unique values for lattice column and lattice row data roles.

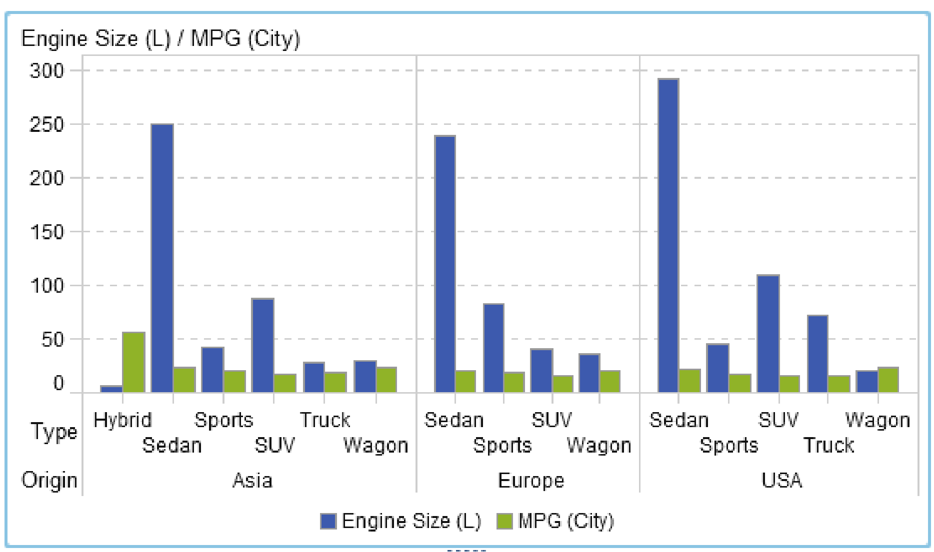
To add a lattice column or lattice row data role:

1 Select a report object in the canvas that has one or more data items assigned and that allows a lattice data role.

2 Click the **Roles** tab in the right pane.

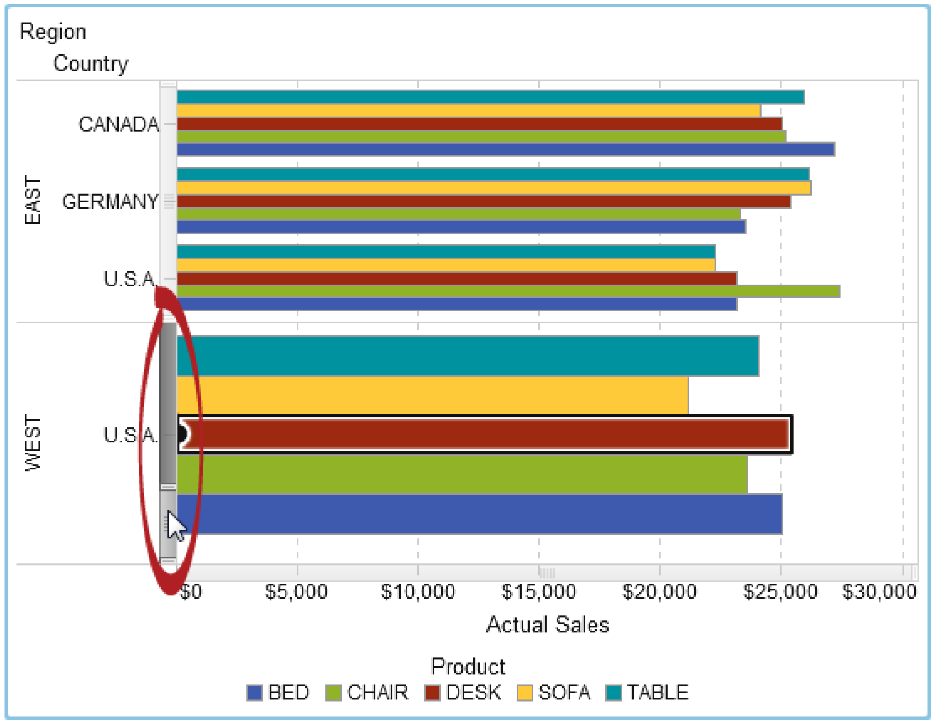
3 Click ▾ beside the **Lattice columns** or **Lattice rows** role that you want to edit.

The following example shows a bar chart with a lattice column specified for the data role:



TIP Scroll bars are generated by the graph if you have a lot of data. These scroll bars work differently than other scroll bars in the designer. They initially fill the maximum area so that you can see all of the bars in the lattice row or lattice column. To zoom in and see specific bars, you have to drag the top or the bottom of the scroll bar to adjust the height of the scrolling bar. Labels are added as space becomes available.

The following example shows a horizontal bar chart with lattice columns. It has a scroll bar activated so that you are zoomed in on the bars for the West region.

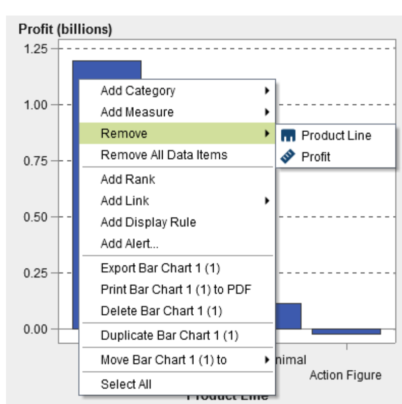


Remove Data Role Assignments from Report Objects

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

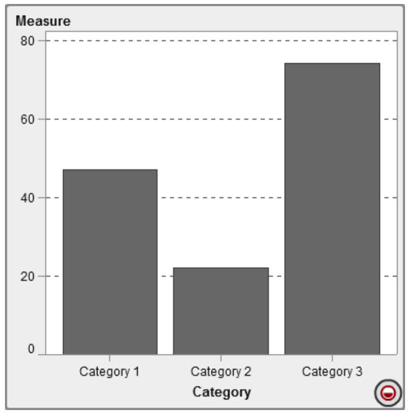
1 Right-click the report object in the canvas. A menu is displayed.

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



2 Select **Remove All Data Items**. All data items are removed from the assigned data roles as well as the data-dependent features like filters, ranks, conditions, and so on. The report object turns gray, and 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 data roles:



Sorting Data in Reports

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 in the designer 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.

You can create a custom sort so that data items in a table or graph can sort to the top or to the bottom. For more information, see "Using a Custom Sort" on page 397.

Ranking can help reduce the amount of visible data and is often used in combination with sorting. For more information, see "Add a New Rank" on page 459.

Note: The designer uses the locale of the SAS LASR Analytic Server to sort data items. For example, if you want to sort data items in the Swedish language, then the table needs to be loaded into a SAS LASR Analytic Server that is initialized for the Swedish locale (`sv_SE`). Contact your system administrator for additional assistance.

Sort Data in a List Table

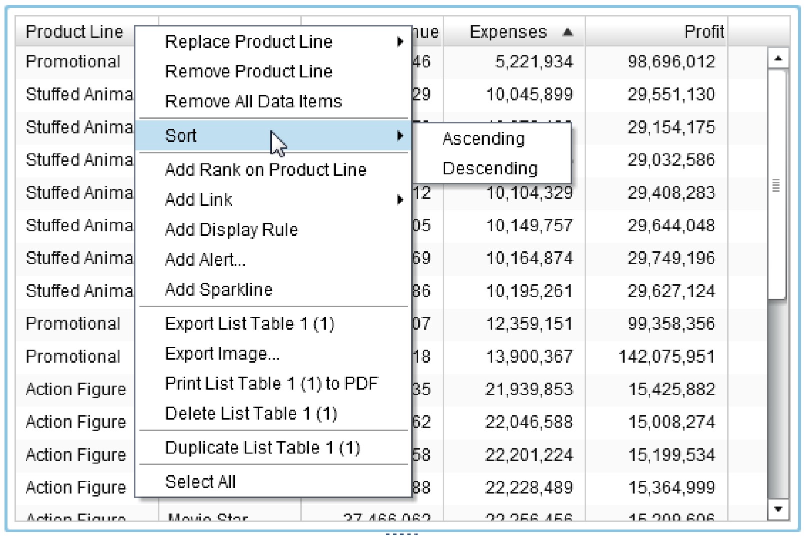
List tables are automatically sorted in ascending order by the first column added to the table. To sort by a different column, right-click on that column heading, and select **Sort**. Then, select either **Ascending** or **Descending**.

Here are some key points about sorting data in a list table:

- If you replace the data item in the first column, the list table is not automatically sorted by the replacement, even though the replacement is the first column in the table. You must manually sort the replacement column.

- Click a column heading to sort the column or toggle an existing sort.
- Use the Ctrl key to select and sort multiple columns.
- You can also press the spacebar to sort a single column in a table.

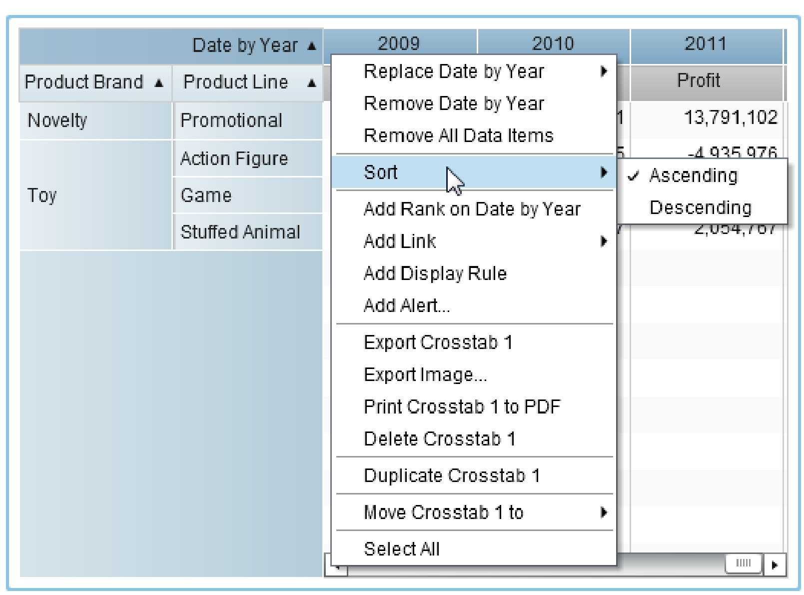
Figure 43.1 Sort Menu Items for List Tables



Sort Data in a Crosstab

To sort by category or measure values in a crosstab, right-click on a column or row heading, and select **Sort**. Then, select either **Ascending** or **Descending**.

Figure 43.2 Sort Menu Items for a Crosstab



You can also sort categories, rather than measures, by clicking the category name.

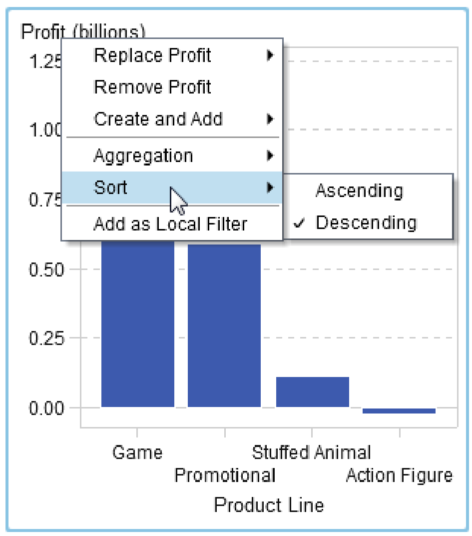
Sort Data in a Graph

To sort by values in a graph, right-click on a measure name or category name, and select **Sort**. Then, select either **Ascending** or **Descending**.

Here are some key points about sorting data in a graph:

- Sorting in a pie chart is based on the measure. If you do not use a measure, then sorting is based on the frequency (and automatically applied).
- Only the Category role in a schedule chart can be sorted.

Figure 43.3 Sort Menu Items for a Graph



Using a Custom Sort

Add a Custom Sort

You can create a custom sort so that data items in a table or graph can sort to the top or to the bottom. Category data items, calculated items that are categories, and custom categories support custom sorts. You can select up to 25 data items for a custom sort. For example, suppose that your company has manufacturing facilities in multiple cities across the country. You can use a custom sort to make sure that certain cities are always sorted before or after other cities in a list table.

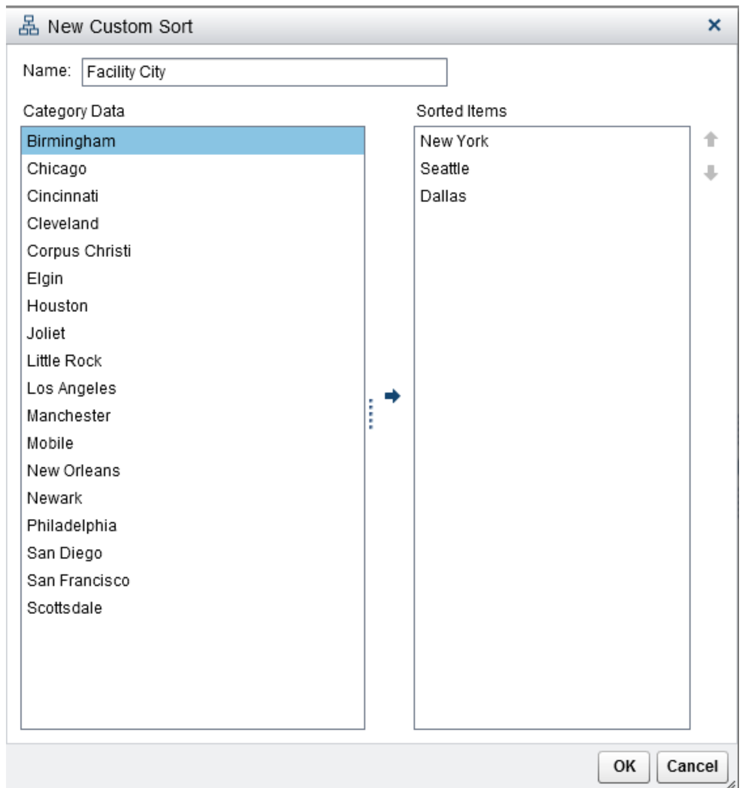
Here are some key points for custom sorting:

- For graphs, a custom sort works only if the data item containing the custom sort is on the X axis, Y axis, or the outermost lattice role.
- For graphs, a custom sort does not work if the data item is assigned to one of the legend data roles (either **Color** or **Group**). The same restriction applies when sorting on the raw values of a data item that has a user-defined format.
- Because user-defined formats have their own sort order, a custom sort works only on the X axis (the Category role).
- A custom sort takes precedence over user-defined formats.

To create a custom sort:

1 On the **Data** tab, select a category data item. Right-click, and select **New Custom Sort**. The New Custom Sort window is displayed.

2 Select the category data items that you want to sort. Click ➡ to move the data items to the **Sorted Items** list. You can also drag and drop data items in the New Custom Sort window. Use the Ctrl key to select multiple data items.



3 (Optional) Use the up and down arrows to arrange the data items in the **Sorted Items** list.

4 Click **OK**.

Edit a Custom Sort

To edit a custom sort:

1 On the **Data** tab, select a category data item, calculated item that is a category, or a custom category. Right-click, and select **Edit Custom Sort**. The Edit Custom Sort window is displayed.

2 Modify the **Sorted Items** list.

3 Click **OK**.

Delete a Custom Sort

You can delete a custom sort on the **Data** tab. Select the category data item with a custom sort. Right-click, and select **Delete Custom Sort**. Click **Delete** in the confirmation message that is displayed.

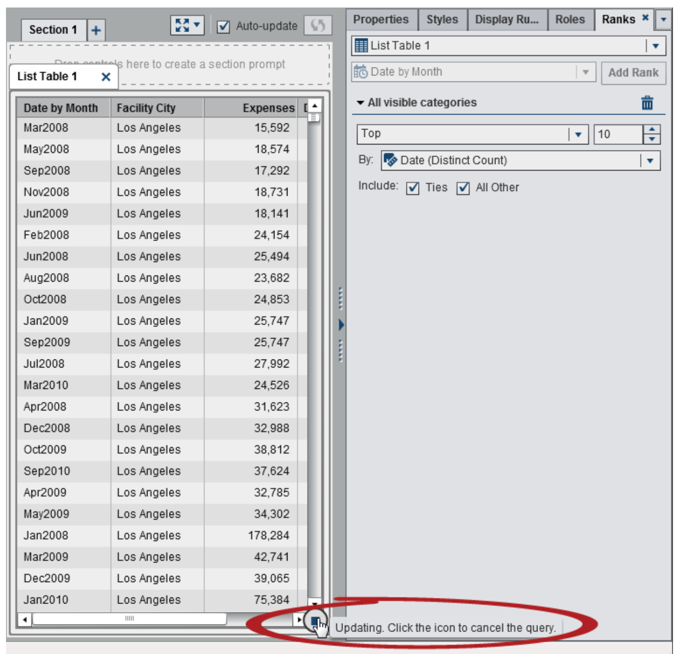
When you delete a custom sort, the data item is sorted on the data.

Cancel a Slow-Running Query for a Report Object

To cancel slow-running queries for a report object, position your mouse over the progress indicator for the report object, and click the icon.

Here is an example of the progress indicator for a list table with a slow-running query:

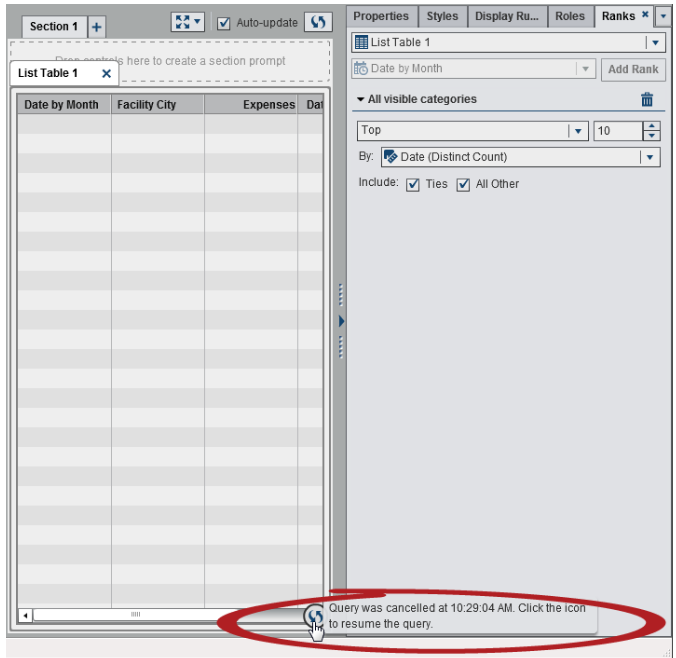
Figure 43.4 *The Progress Indicator for List Table with a Slow-Running Query*



After you cancel the query, the progress indicator is replaced by the refresh icon. To resume a query that you canceled, click .

Here is an example of a list table after a query was canceled:

Figure 43.5 *The Progress Indicator for List Table with Query That Has Been Stopped*



44

Working with Display Rules for Reports

Overview of Display Rules

Display rules include all types of highlighting of report objects. They provide a flexible structure to specify conditions. There are several types of display rules. Display rules enable conditions to be shared across objects, but not all display rules apply across all report object types.

SAS Visual Analytics Designer (the designer) provides the following interfaces for display rules:

- The **Display Rules** tab in the right pane enables you to populate intervals, add intervals, or add color-mapped values for the report object that is currently selected in the canvas. You can use this pane to specify both report-level or object-level display rules, depending on what you have selected in the canvas.
- The **Shared Rules** tab in the left pane enables you to create a new display rule for a gauge, which is used by other gauges to designate intervals and colors for ranges. You can edit or delete an existing shared display rule. These rules are shared across multiple gauges and can be created at any time.

TIP The **Shared Rules** tab is not displayed by default. To display the tab, click ▼ after the last tab name, and select **Shared Rules**.

Display rules can accept parameters. For more information, see Chapter 49, “Working with Parameters in Reports,” on page 463.

For list tables and graphs, you can create display rules that reference a measure that is not included in the currently displayed report object. This feature is not available for crosstabs.

Here are some key points about color-mapped display rules:

- Color-mapped display rules are not allowed if there are multiple overlays.
- Color-mapped values for a report object can be applied only to category data items.
- Color-mapped values cannot be applied to date or datetime data items.
- Color-mapped display rules are not applied if there are two measures.
- If a color-mapped display rule is applied to one measure and you add a second measure, the display rule is removed.

The following report objects that are imported from SAS Visual Analytics Explorer (the explorer) or created in SAS Visual Analytics Graph Builder (the graph builder) do not support display rules:

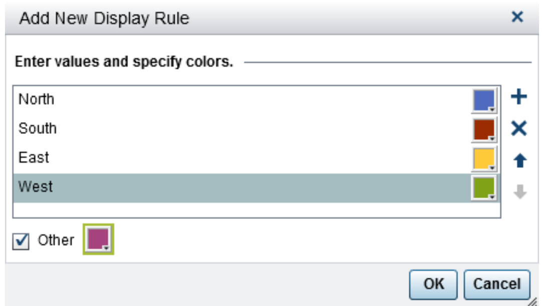
- heat maps (which are imported from the explorer)
- band plots (which are created in the graph builder)
- step plots (which are created in the graph builder)

Add Report-Level Display Rules

To add a report-level display rule:

1. Without any report objects or the section selected in the canvas, click the **Display Rules** tab in the right pane. Click **New**. The Add New Display Rule window is displayed.
2. Enter a value for the display rule in the field.
3. Select a color for the display rule.
4. (Optional) Repeat the steps for entering a value and selecting a color.
5. (Optional) Select the **Other** check box. Then, select a color so that any of the other categories that do not have a color will have the one that you just selected.

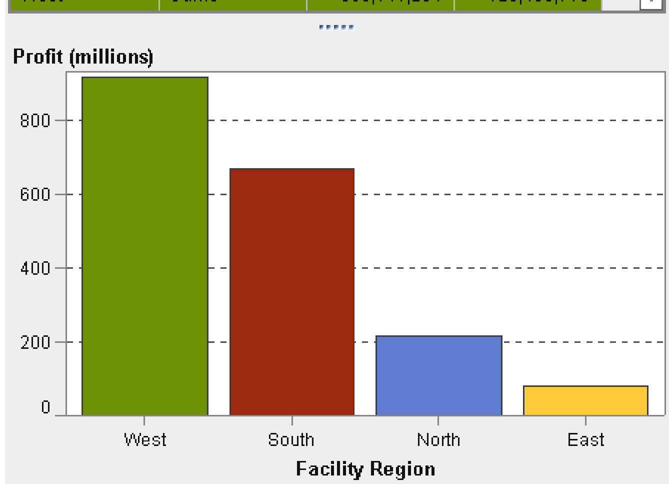
In the following example, values and colors have been specified for each region in a list table. And, a color has been selected for the **Other** check box.



6 Click **OK**. The report objects in the report update with the new display rule. And, the display rule appears on the **Display Rules** tab in the right pane.

The following table illustrates the report-level display rules that were defined in the Add New Display Rule window above.

Facility Region	Product Line	Profit	Expenses
East	Action Figure	-10,515,368	29,014,382
East	Game	5,328,719	1,900,170
East	Promotional	81,456,122	45,807,236
East	Stuffed Animal	1,544,003	8,284,070
North	Action Figure	-21,211,851	60,883,053
North	Game	220,653,512	110,853,384
North	Stuffed Animal	16,039,542	56,849,706
South	Action Figure	9,202,487	165,466,767
South	Game	607,956,582	241,557,659
South	Stuffed Animal	51,908,425	39,659,003
West	Action Figure	3,453,239	26,026,052
West	Game	360,141,294	123,498,715



Adding Table-Level Display Rules

You can add three different types of display rules to tables. You cannot create a display rule to highlight dates in a table.

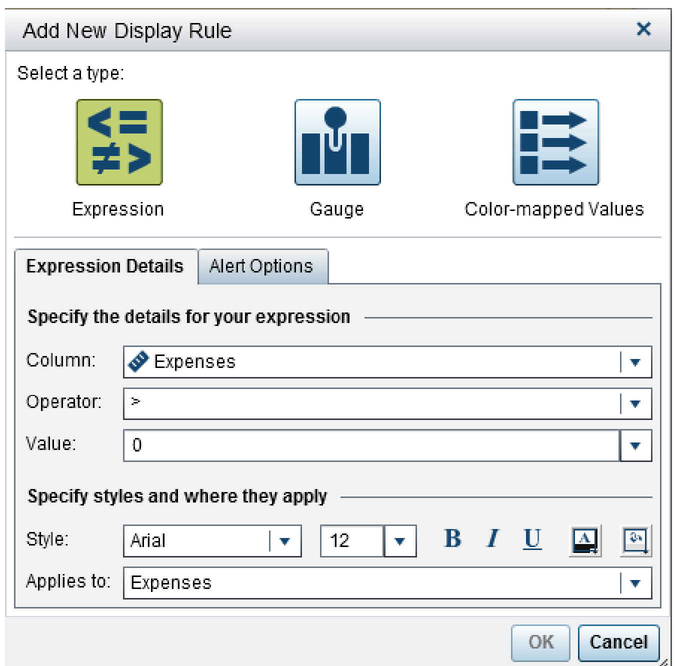
Add Display Rules to a List Table Using an Expression

To specify a new display rule for a table using an expression:

1 If it is not already selected, select the table in the canvas that you want to update.

2 In the right pane, click the **Display Rules** tab. Click **New**. The Add New Display Rule window is displayed.

3 Click **Expression**. The Add New Display Rule window expands to show the details for the expression.



4 On the **Expression Details** tab, select the **Column** or any measure value.

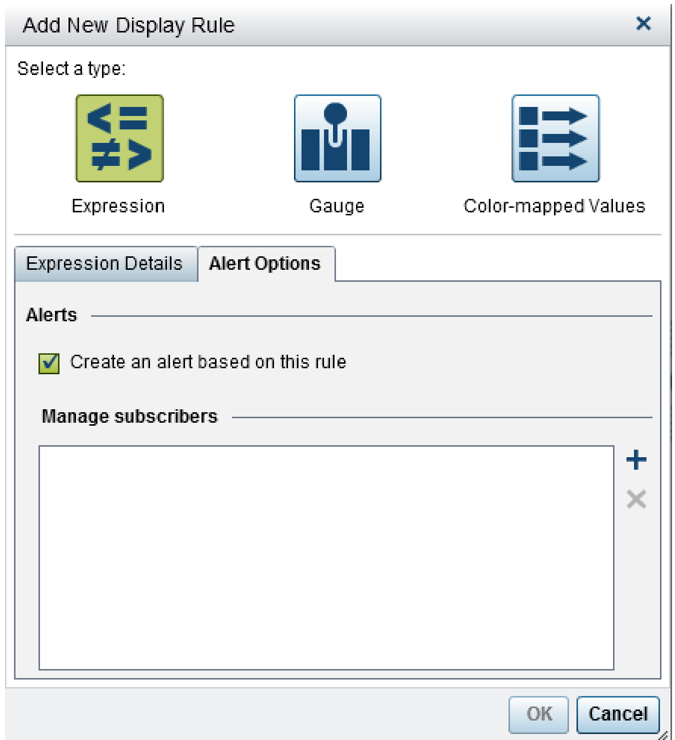
5 Select the **Operator**. You can select **=**, **< >**, **BetweenInclusive**, **<**, **<=**, **>**, **>=**, **Missing**, or **NotMissing**. The default is **>**.

6 Enter or select a **Value**.

7 Modify the style, size, and color of the font. There is no default style for list table display rules. You must specify a style to enable the **OK** button in the Add New Display Rule window.

8 Select the row or column in the **Applies to** drop-down list.

9 (Optional) On the **Alert Options** tab, select the **Create an alert based on this rule** check box. For more information about alerts, see “Working with Alerts for Report Objects” on page 351.



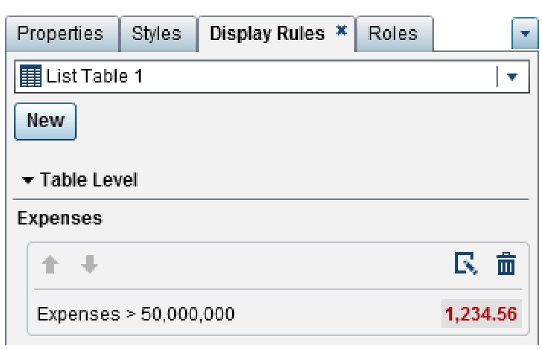
a Click beside the **Manage subscribers** list to add users who you want notified when the alert condition is met. The Add Recipients window is displayed.

b Select one or more users. Click **OK** to return to the Add New Display Rule window.

Note: A user's e-mail address must be stored in metadata to be displayed on the **View All** tab in the Add Recipients window.

10 Click **OK**. The table updates with the new display rule. The display rule appears on the **Display Rules** tab in the right pane.

Here is an example of a display rule that uses an expression:



11 (Optional) Click on the **Display Rules** tab to edit the new display rule.

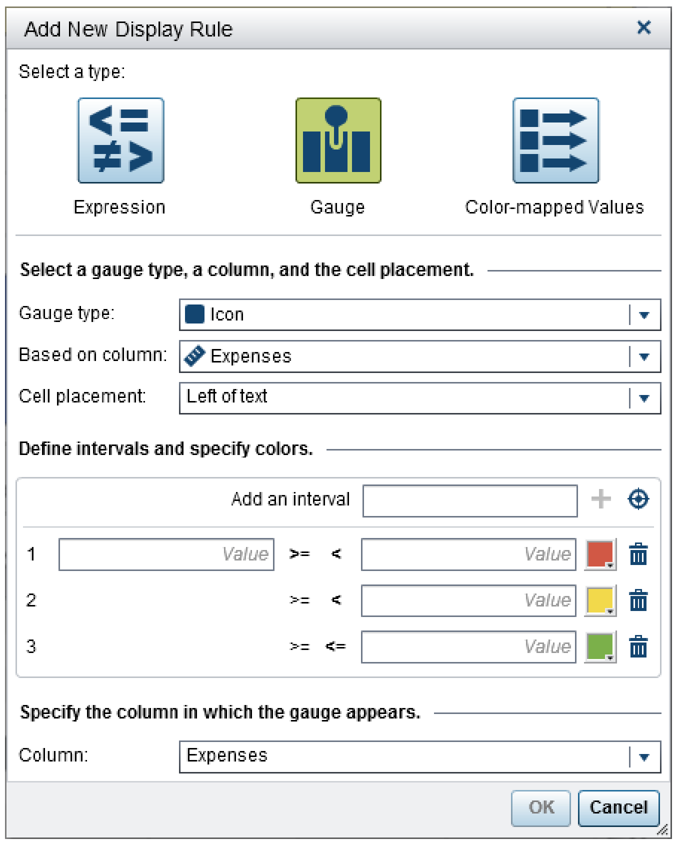
Add Display Rules to a List Table Using a Gauge

To specify a new display rule for a table using a gauge:

1 If it is not already selected, select the table in the canvas that you want to update.

2 In the right pane, click the **Display Rules** tab. Click **New**. The Add New Display Rule window is displayed.

3 Click **Gauge**. The Add New Display Rule window expands to show the details for the gauge.



4 Select a **Gauge type** from the drop-down list. Your choices are a bullet, an icon, a slider, or a thermometer. **Icon** is the default.

5 Using the **Based on column** drop-down list, specify which column the rule should be based on in the report.

6 Specify where the gauge should appear in the column for the **Cell placement**. Your choices are **Left of text**, **Right of text**, or **Replace text**.

7 Define the intervals and colors for the rule:

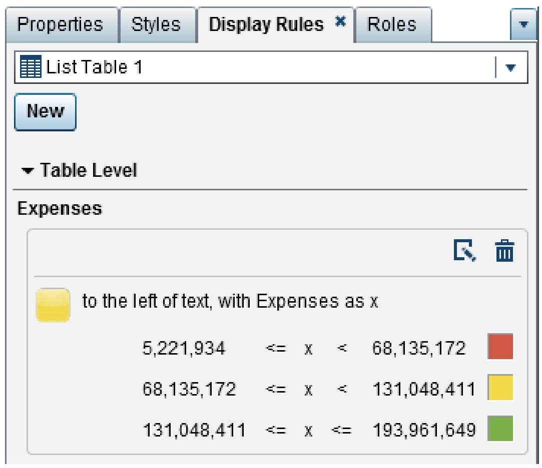
- Enter the individual values for the intervals and then select a color. You can click the operator between the intervals to change it.
- Click to automatically populate the intervals. The Populate Intervals window is displayed.

You can specify **Number of intervals**, **Lower bounds**, and **Upper bounds**. Click **OK**.

8 Specify the **Column** in which you want the gauge to be displayed

9 Click **OK**. The table updates with the new display rule. The display rule appears on the **Display Rules** tab in the right pane.

Here is an example of an automatically populated display rule that uses an icon:



Here is a list table with the display rules applied:

Product Li...	Product	Revenue	Expenses	Profit
Action Figure	Firefighter	37,830,429	22,383,218	15,447,211
Action Figure	Athlete	37,607,425	22,282,489	15,324,936
Action Figure	Super Hero	37,593,488	22,228,489	15,364,999
Action Figure	Musician	37,400,758	22,201,224	15,199,534
Action Figure	Soldier	37,054,862	22,046,588	15,008,274
Action Figure	.	0	126,051,9...	-126,051,937
Action Figure	Police	37,365,735	21,939,853	15,425,882
Action Figure	Movie Star	37,466,062	22,256,456	15,209,606
Game	Puzzle	472,511,801	117,465,4...	355,046,378
Game	Card	396,722,440	67,173,845	329,548,595
Game	Board	802,655,795	193,961,6...	608,694,146
Game	.	0	99,209,012	-99,209,012
Promotional	Coffee Cup	111,717,507	12,359,151	99,358,356
Promotional	Backpack	278,485,882	113,686,2...	164,799,628

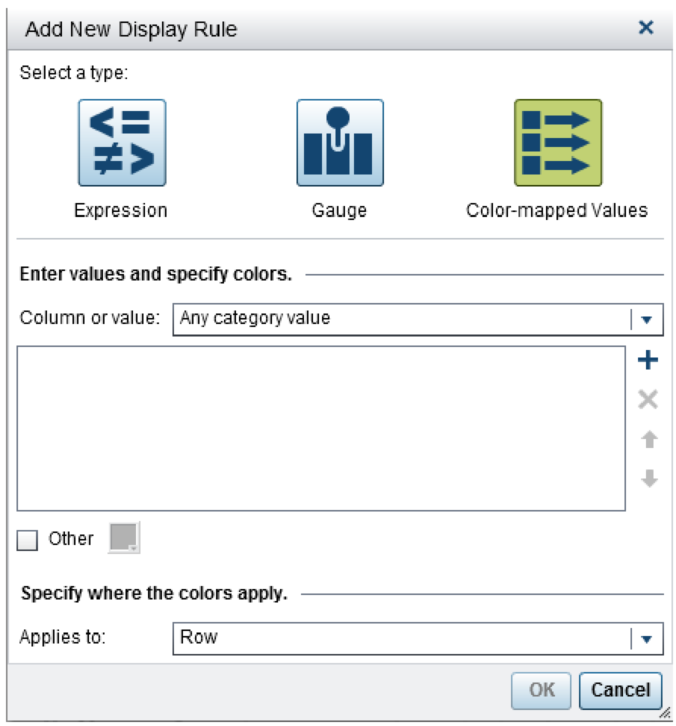
Note: If a value falls outside the bounds of the display rule intervals, no icon is displayed in the list table

10 (Optional) Click on the **Display Rules** tab to edit the new display rule.

Add Display Rules to a List Table Using Color-Mapped Values

To specify a new display rule for a list table using color-mapped values:

1. If it is not already selected, select the list table in the canvas that you want to update.
2. In the right pane, click the **Display Rules** tab. Click **New**. The Add New Display Rule window is displayed.
3. Click **Color-mapped Values**. The Add New Display Rule window expands to show the details for the color-mapped values.



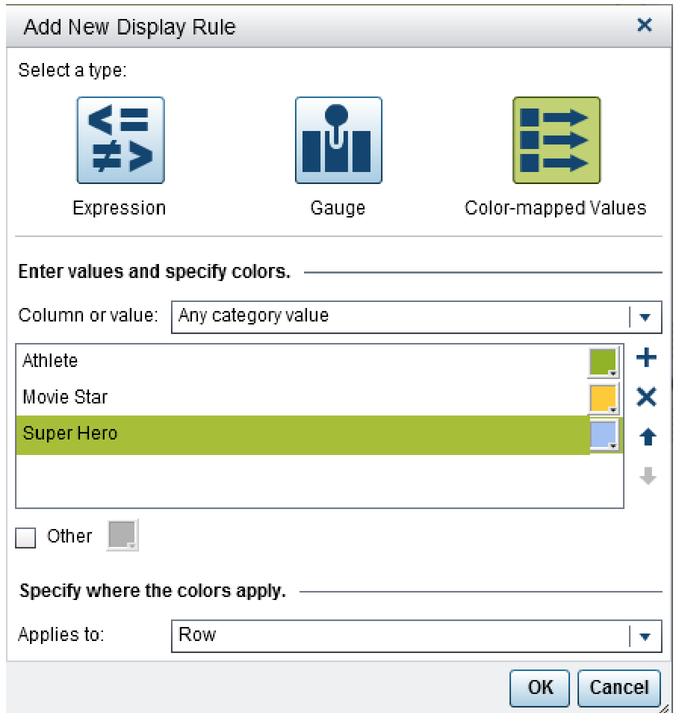
4. Select the **Column or value** to which you want to apply the display rule.

   Note: Color-mapped values for a report object can be applied only to category data items.

   Note: Color-mapped values cannot be applied to date or datetime data items.

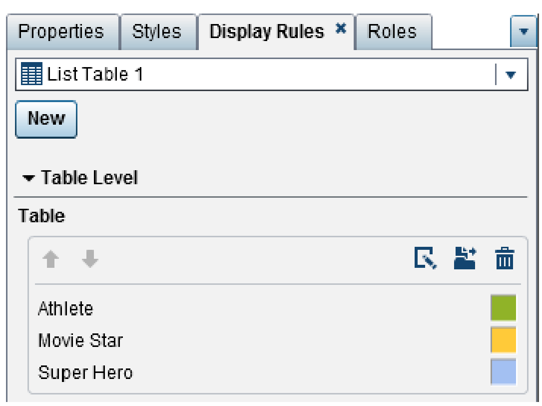
5. Click in the box to enter a value for the display rule.
6. Select a color for the display rule.
7. (Optional) Repeat the steps for entering a value and selecting a color.
8. (Optional) Select the **Other** check box. Then, select a color.
9. Using the list, specify where you want to apply the colors.

In the following example, values and colors have been specified for products in a list table.



10 Click **OK**. The table updates with the display rules.

The display rule appears on the **Display Rules** tab in the right pane.



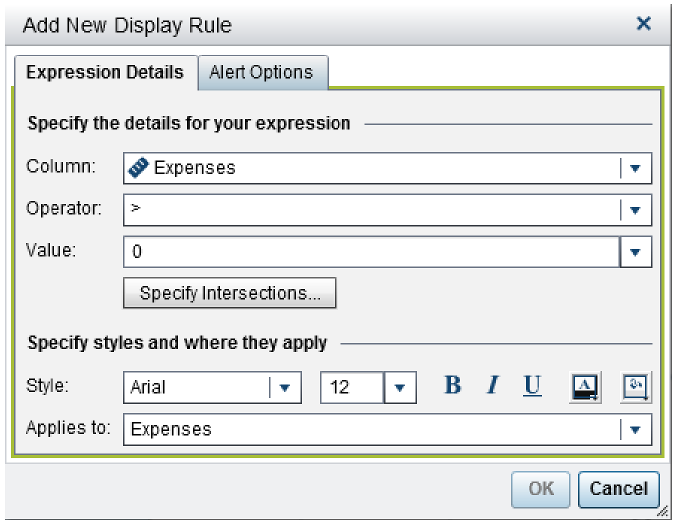
11 (Optional) Click on the **Display Rules** tab to change the table-level display rule to a report-level display rule.

Add Display Rules to a Crosstab

To specify a new display rule for a crosstab using an expression:

1 If it is not already selected, select the crosstab in the canvas that you want to update.

2 In the right pane, click the **Display Rules** tab. Click **New**. The Add New Display Rule window is displayed.



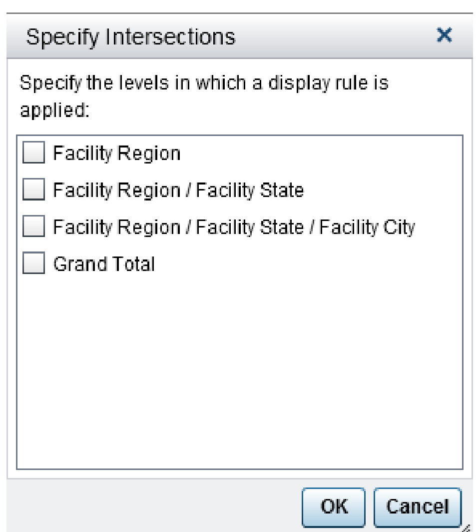
3 On the **Expression Details** tab, select the **Column** or any measure value.

4 Select the **Operator**. You can select **=**, **< >**, **BetweenInclusive**, **<**, **<=**, **>**, **>=**, **Missing**, or **NotMissing**. The default is **>**.

5 Enter or select a **Value**.

6 If your crosstab contains a hierarchy, then you can specify the hierarchy levels in which the display rule is applied.

a Click **Specify Intersections**. The **Specify Intersections** window is displayed.



b Select one or more hierarchy levels, the grand total, or all of the levels. Click **OK** to return to the Add New Display Rule window. The intersections that you selected are displayed above the **Edit Intersections** button.

7 Modify the style, size, and color of the font.

8 Select the row or column in the **Applies to** drop-down list. For crosstabs, the **Applies to** drop-down list displays only the measures that have been added to the crosstab. There are no row or column options.

9 (Optional) On the **Alert Options** tab, select the **Create an alert based on this rule** check box. For more information about alerts, see "Working with Alerts for Report Objects" on page 351.

a Click + beside the **Manage subscribers** list to add users who you want notified when the alert condition is met. The Add Recipients window is displayed.

b Select one or more users. Click **OK** to return to the Add New Display Rule window.

Note: A user's e-mail address must be stored in metadata to be displayed on the **View All** tab in the Add Recipients window.

10 Click **OK**. The crosstab updates with the new display rule. The display rule appears on the **Display Rules** tab in the right pane.

11 (Optional) Click on the **Display Rules** tab to edit the new display rule.

Adding Graph-Level Display Rules

You can add display rules to graphs. Note that you cannot create a display rule to highlight dates in a graph.

Add Display Rules to a Graph Using an Expression

Here are some key points about adding an expression-based display rules to a graph:

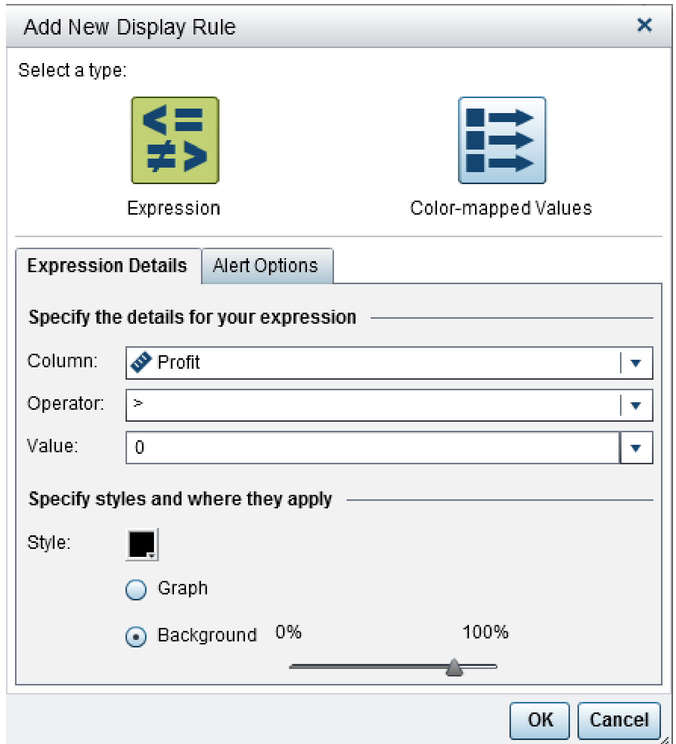
- Display rules can be added to a waterfall chart only if you clear the **Color by response sign** check box on the **Properties** tab in the right pane.
- You can create an expression-based display rule for a 100% stacked bar chart, which uses the **Grouping style** and **Grouping scale** properties. However, the display rule cannot be based on a percentage.

To specify a new expression-based display rule for a graph:

1 If it is not already selected, select the graph in the canvas that you want to update.

2 In the right pane, click the **Display Rules** tab. Click **New**. The Add New Display Rule window is displayed.

3 Click **Expression**. The Add New Display Rule window expands to show the details for the expression.



4 On the **Expression Details** tab, select the **Column** or any measure value.

5 Select the **Operator**. You can select **=**, **< >**, **BetweenInclusive**, **<**, **<=**, **>**, **>=**, **Missing**, or **NotMissing**. The default is **>**.

6 Enter or select a **Value**.

7 If your graph contains a hierarchy, then you can specify the hierarchy levels in which the display rule is applied.

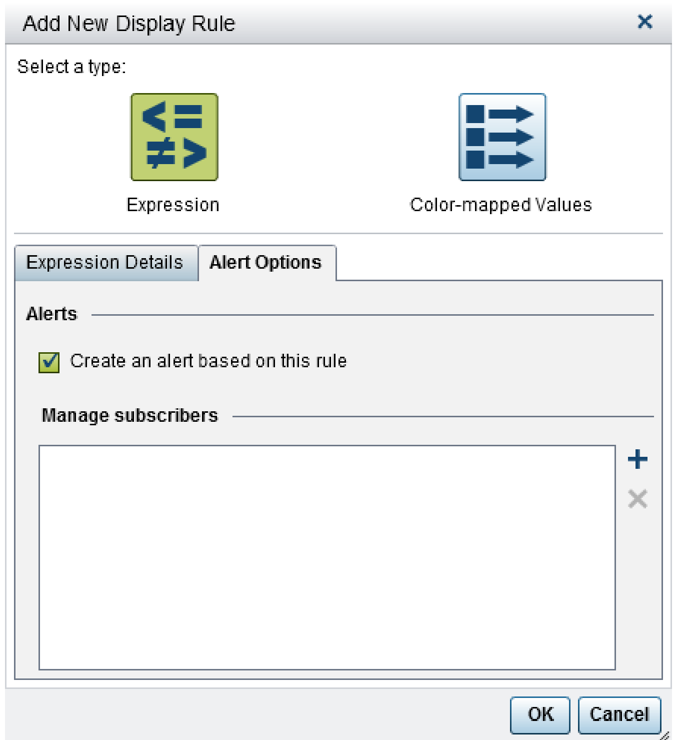
   a Click **Specify Intersections**. The **Specify Intersections** window is displayed.

   b Select one or more hierarchy levels. Click **OK** to return to the Add New Display Rule window. The intersections that you selected are displayed above the **Edit Intersections** button.

8 Modify the **Style**, which includes color and graph or background.

   Note: **Background** is available only for bar charts, waterfall charts, line charts, scatter plots, time series plots, and bubble plots. It is also available for needle plots, which are created in the SAS Visual Analytics Graph Builder (the graph builder).

9 (Optional) On the **Alert Options** tab, select the **Create an alert based on this rule** check box. For more information about alerts, see “Working with Alerts for Report Objects” on page 351.



Use the **Manage subscribers** list to add users who you want notified when the alert condition is met.

10 Click **OK**. The graph updates with the new display rule. The display rule appears on the **Display Rules** tab in the right pane.

Here is an example of a display rule that uses an expression:

11 (Optional) Click on the **Display Rules** tab to edit the new display rule.

Note: A warning badge is displayed on a graph when the color is overloaded. This happens when multiple measures are assigned, a color or group role is assigned, or multiple overlays are present (which have cycling colors).

Add Display Rules to a Graph Using Color-Mapped Values

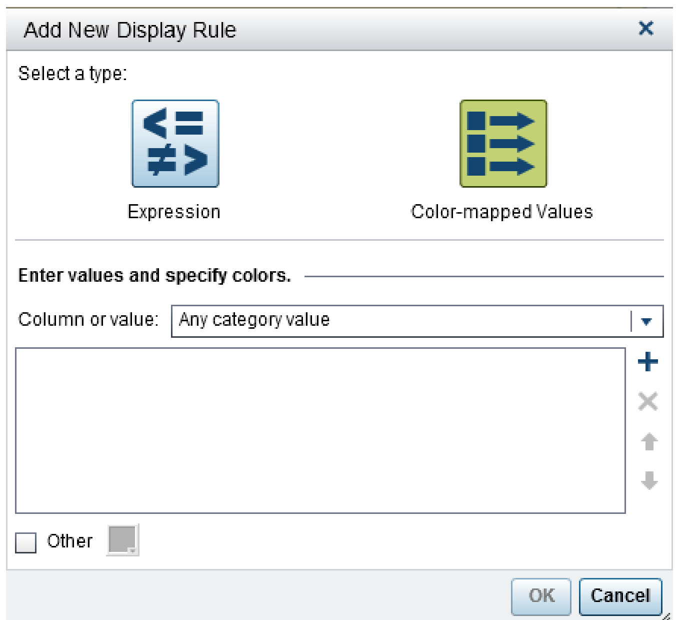
Note: Display rules can be added to a waterfall chart only if you clear the **Color by response sign** check box on the **Properties** tab in the right pane.

To specify a new display rule for a graph using color-mapped values:

1 If it is not already selected, select the graph in the canvas that you want to update.

2 In the right pane, click the **Display Rules** tab. Click **New**. The Add New Display Rule window is displayed.

3 Click **Color-mapped Values**. The Add New Display Rule window expands to show the details for the color-mapped values.



4 Select the **Column or value** to which you want to apply the display rule.

Note: Color-mapped values for a report object can be applied only to category data items.

Note: Color-mapped values cannot be applied to date or datetime data items.

5 Click in the box to enter a value for the display rule.

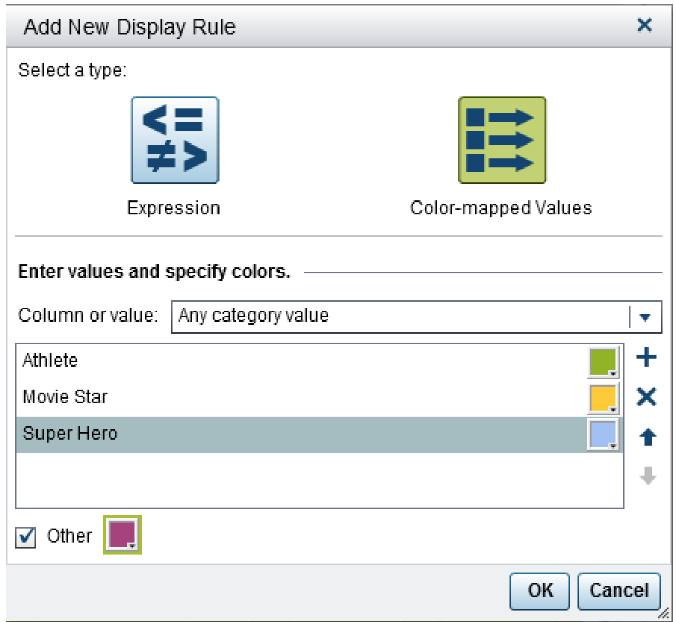
6 Select a color for the display rule.

7 (Optional) Repeat the steps for entering a value and selecting a color.

8 (Optional) Select the **Other** check box. Then, select a color.

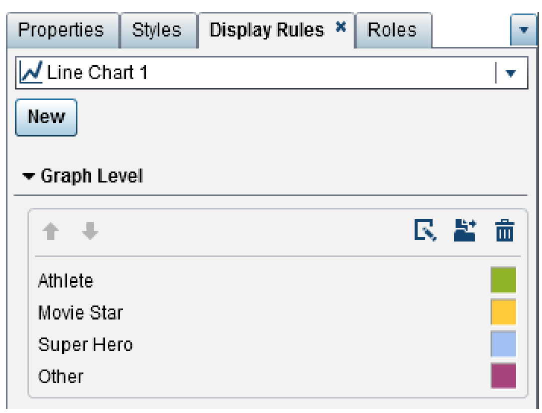
9 Using the list, specify where you want to apply the colors.

In the following example, values and colors have been specified for products in a graph.



10 Click **OK**. The graph updates with the display rules.

The display rule appears on the **Display Rules** tab in the right pane.



11 (Optional) Click on the **Display Rules** tab to change the graph-level display rule to a report-level display rule.

Adding Gauge-Level Display Rules

A display rule is used by a gauge to designate intervals and colors for ranges. You can add a shared display rule for a gauge, which is used by other gauges to designate intervals and colors for ranges.

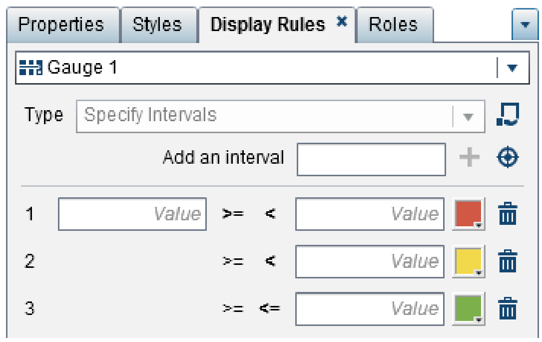
Add Display Rules to a Gauge

To specify a new display rule for a gauge:

1 If it is not already selected, select the gauge in the canvas that you want to update.

2 In the right pane, click the **Display Rules** tab.

Here is an example of the **Display Rules** tab for a gauge:



3 Specify (or modify) the display rules for the gauge. You can populate intervals, edit the display rule, and specify the **Type**.

To automatically populate the intervals:

a Click beside the **Add an interval** field to display the Populate Intervals window.

b (Optional) Review or change the values for the **Number of intervals**, **Lower bounds**, and **Upper bounds** fields.

c Click **OK**. The gauge updates and the new display rule is displayed on the **Display Rules** tab.

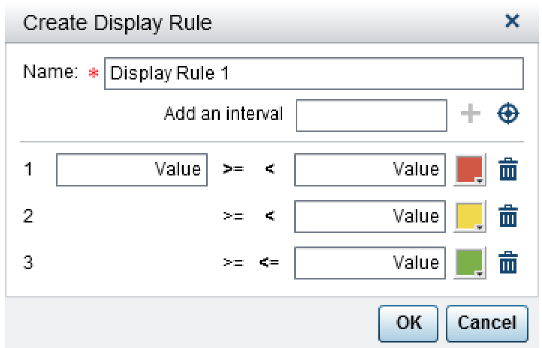
To edit the display rule for a gauge, use the fields on the **Display Rules** tab. You can update the conditions for the values, the intervals, and the color for the range. Select a color to open the color palette. To add new intervals to the display rule, enter a number, and click + beside the **Add an interval** field.

Add a Shared Display Rule for Gauges

There are two ways to create a shared display rule for a gauge in the designer. You can use either the **Shared Rules** tab in the left pane or the **Display Rules** tab in the right pane.

To create a shared display rule for a gauge using the **Shared Rules** tab:

1 Click on the **Shared Rules** tab. The Create Display Rule window is displayed.



2 Enter a **Name**.

3 Add an interval or automatically populate the intervals.

To automatically populate the intervals:

a Click beside the **Add an interval** field to display the Populate Intervals window.

b (Optional) Review or change the values for the **Number of intervals**, **Lower bounds**, and **Upper bounds** fields.

c Click **OK** to return to the Create Display Rule window.

4 Click **OK**. The new shared display rule is added to the **Shared Rules** tab.

To create a shared display rule for a gauge using the **Display Rules** tab:

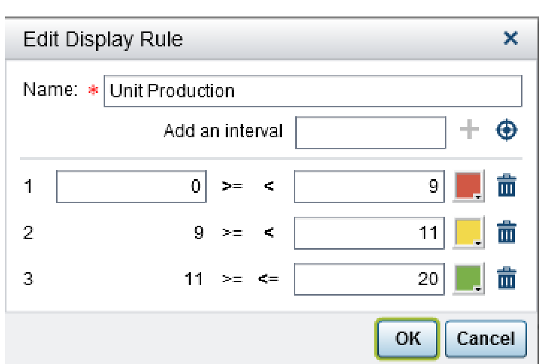
1 Click on the **Display Rules** tab. The Shared Rule Name window is displayed.

2 Enter a **Name**.

3 Click **OK**. The new shared display rule is displayed below the **Type** field on the **Display Rules** tab. The new shared display rule is also added to the **Shared Rules** tab.

To edit a shared display rule for a gauge:

1 Select the shared display rule on the **Shared Rules** tab, and then click . The Edit Display Rule window is displayed.



2 Update the conditions for the values, the intervals, and the color for the range. Select a color to open the color palette. To add new intervals to the

shared display rule, enter a number, and click + beside the **Add an interval** field.

3 Click **OK**.

45

Working with Report Filters

About Report Filters

In SAS Visual Analytics Designer (the designer), you can create filters to subset your data. These types of filters are available:

- Detail Report Filters

  Basic filters
  : subset the data for individual report objects in your reports by using a single data item. A basic filter is constrained to use only the data item that was selected when the basic filter was created. Basic filters can be created and modified by using the pop-up menu for a report object or by using the Edit Filter window. For more information, see "Use a Basic Report Filter" on page 420.

  Advanced filters
  : subset the data for individual report objects in your reports by using any number of data items and operators (for example, OR and AND) in the same expression. Advanced filters can be created and modified by using the Advanced Filter window. For more information, see "Use an Advanced Report Filter" on page 424.

  Data source filters
  : subset the data for the entire report. The data source filters that you create are applied to every report object in the report that uses that data source. For more information, see "Use a Data Source Filter in a Report" on page 427.

- Post-aggregate Report Filters

For more information, see "Using Post-Aggregate Report Filters" on page 429.

Here are some key points about report filters:

- If you change a report filter from a detail filter to a post-aggregate filter, then all of the selections for the filter will be lost. The same is true if you change a post-aggregate filter to a detail filter. There is a warning message for both cases.
- One data item cannot have both a detail data item filter and a post-aggregate data item filter.
- Clear the **Auto-update** check box above the report canvas until you are ready to apply your filter changes.

All of your filters are saved when you save your report.

Filters can accept parameters. For more information, see Chapter 49, "Working with Parameters in Reports," on page 463.

When a report object imported from SAS Visual Analytics Explorer (the explorer) has a local filter, that filter is displayed on the **Filters** tab. You cannot edit or delete these filters on imported report objects that use on-demand data.

Note: Report linking is a type of filter. The report section that is the target of the link is filtered by the values selected in the linked report object. For more information, see "Overview of Report Links" on page 449.

Using Detail Report Filters

Use a Basic Report Filter

About Basic Report Filters

For certain report objects, you can subset your data by using the **Filters** tab in the right pane of the designer. You can base your filters on any data item in the current data source for this report object, regardless of whether the data item is assigned to a report object in the current report.

Note: If you have multiple basic filters, the designer assumes that there is an AND operator between the filters.

Create a Basic Report Filter

To create a basic report filter:

1. If it is not already selected, select the report object in the canvas that you want to filter. The report object must have at least one data item assigned.

   TIP Clear the **Auto-update** check box above the report canvas until you are ready to apply your filter changes.

2. In the right pane, click the **Filters** tab.

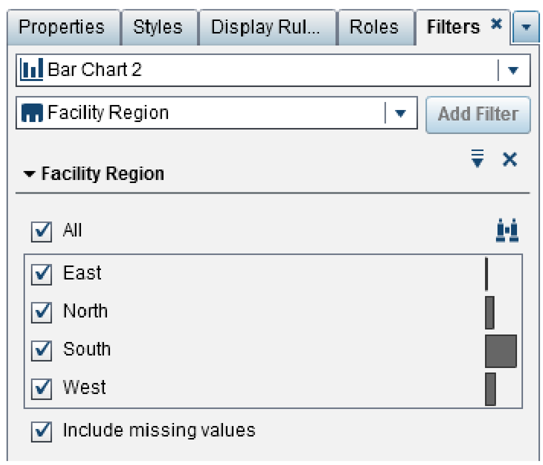
3 Click beside the **Add Filter** button, and select a data item from the list. Then, click **Add Filter**. The filter appears on the tab.

4 Select the data values for the filter.

> **TIP** Use the arrow to the left of the filter name on the **Filters** tab to expand or collapse the filter details when you are working with multiple filters.

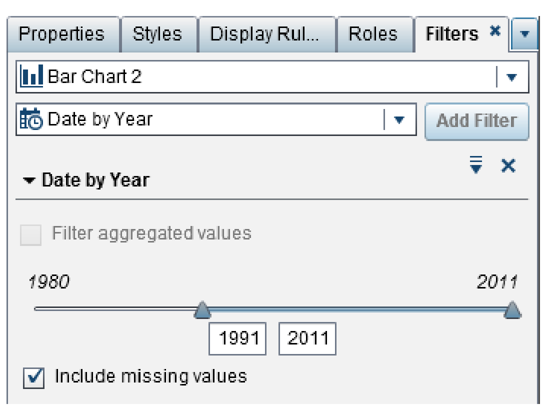
- For a basic filter that uses discrete values, a check box is displayed for each distinct value that uses the current format applied to the data item. To the right of each value, a bar indicates the frequency. Select the data values that you want to filter or clear the selections for the data values that you do not want to filter. Select **All** to select all of the values.

Here is an example of a basic filter for discrete values:



- For a filter that uses continuous values, a slider shows you the maximum values and the minimum values that exist for the data item using the current data item format. Use the slider to select a range of target values.

Here is an example of a basic filter for continuous values:



5 (Optional) If your data contains missing values, and you want to exclude those missing values from your report, then clear the **Include missing values** check box.

6 Click (to the left of the delete icon) for options. The available options depend on whether you are filtering characters, dates, or numerics. You can sort the values or frequencies in the filter.

The following options are available:

Option	Availability
Edit Filter	This option is always available.
Delete Filter	This option is always available.
Filter Detail Values	This option is available only for detail values.
Filter Aggregated Values	This option is available only for aggregated values.
Filter using continuous values	This option is available only for measure data items.
Filter using discrete values	This option is available for measure and category data items.
Include Missing Values	This option is available when you filter continuous values. It works the same as selecting the **Include missing values** check box.
Exclude Missing Values	This option is available when you filter continuous values. It works the same as clearing the **Include missing values** check box.
Select All	This option is available when you filter discrete values. It selects all of the current discrete values that are listed. This option does not affect the **Include missing values** setting.
Clear All	This option is available when you filter discrete values. It clears the selections of all of the current discrete values that are listed. This option does not affect the **Include missing values** setting.
Invert Selection	This option is available when you filter discrete values. It changes all of the discrete value check boxes so that if they are selected, then they are cleared. If they are cleared, then they are selected. This option does not affect the **Include missing values** setting.
Sort by Values	This option enables you to sort by the values in the filter.
Sort by Frequency	This option enables you to sort by the frequencies in the filter.
Show Selected at Top	This option is available for basic filters that use discrete values.

7 (Optional) If you cleared the **Auto-update** check box above the report canvas while you worked on your filter, then select it when you are ready to apply your filter changes.

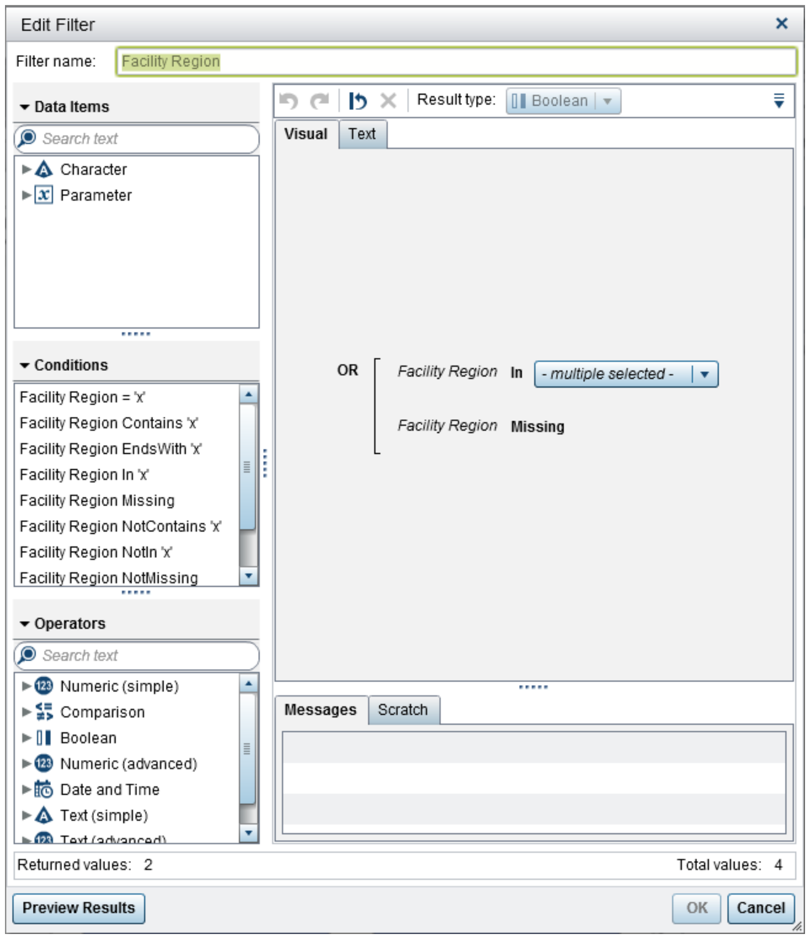
Edit a Basic Report Filter

To edit a basic report filter:

1 If it is not already selected, select the report object in the canvas that you want to filter. The report object must have at least one data item assigned.

2 In the right pane, click the **Filters** tab.

3 Click beside the filter name. Then, select **Edit Filter**. The Edit Filter window is displayed.



4 Edit the expression for your filter.

- You can drag and drop conditions and operators onto the expression on the **Visual** tab in the right pane.
- You can enter the expression on the **Text** tab in the right pane.
- You can create an expression using both the **Visual** and **Text** tabs in the right pane.

For information, see Appendix 6, "Conditions for Filters," on page 607.

Note: The AND and OR operators can accept more than two conditions. To add a condition to the operator, drag and drop a condition onto the operator name in the right pane. For example, to add a third condition to an AND operator, drag and drop the new condition onto AND in the expression.

TIP Right-click the AND or OR operator in the expression, and then select **Add** ▶ **New Condition**.

5 Click **OK** to apply the filter.

Delete a Basic Report Filter

To delete a basic report filter, click ✕ beside the filter on the **Filters** tab.

Use an Advanced Report Filter

About Advanced Report Filters

For most report objects, you can create advanced filters to subset your data by using the **Filters** tab in the right pane of the designer.

Advanced filters enable you to create filters that use more than one data item.

Here are some key points about advanced filters:

- Advanced filters that are created in the explorer might contain expressions that you cannot create in the designer.
- Advanced global filters that are created in the explorer are supported by the designer, but you cannot create advanced global filters in the designer.

Create an Advanced Report Filter

To create an advanced report filter:

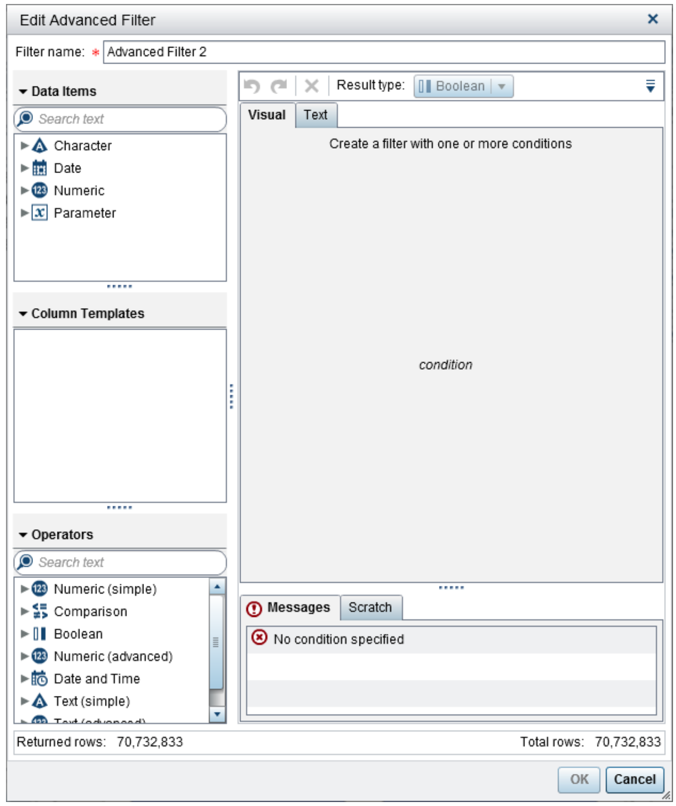
1 If it is not already selected, select the report object in the canvas that you want to filter. The report object must have at least one data item assigned.

TIP Clear the **Auto-update** check box above the report canvas until you are ready to apply your filter changes.

2 In the right pane, click the **Filters** tab.

3 Click ▼ to select **Advanced** instead of a specific data item.

4 Click **Add Filter**. The Edit Advanced Filter window is displayed.



5 Specify a **Filter name**. The filter name identifies the advanced filter on the **Filters** tab.

6 Create a condition for the filter. You can use the **Visual** tab, **Text** tab, or a combination of both tabs.

To create a condition using the **Visual** tab:

a From the **Data Items** list, select the data item on which the condition is based.

b From the **Column Templates** list, select a condition. For information, see Appendix 6, “Conditions for Filters,” on page 607.

c Drag and drop the condition onto the expression.

d For any required parameters, select the parameter, and enter a value, or right-click on the parameter field, and select **Replace with** to select a data item.

7 (Optional) Add operators to the expression. You can use the **Visual** tab, **Text** tab, or a combination of both tabs.

To add operators using the **Visual** tab:

a From the **Operators** list, select an operator to join the conditions in your expression. For more information, see Appendix 6, “Conditions for Filters,” on page 607.

b Drag and drop the operator onto the expression.

   c. From the **Column Templates** list, select another condition. Then, drag and drop the additional condition onto the expression. Complete any required parameters.

   > **TIP** Right-click the AND or OR operator in the expression, and then select **Add** ► **New Condition**.

8. Click **OK** to apply the filter.

9. (Optional) If you cleared the **Auto-update** check box above the report canvas while you worked on your filter, then select it when you are ready to apply your filter changes.

Edit an Advanced Report Filter

To edit an advanced report filter:

1. If it is not already selected, select the report object in the canvas that you want to filter. The report object must have at least one data item assigned.

2. In the right pane, click the **Filters** tab.

3. Click beside the advanced filter name. Then, select **Edit Filter**. The Edit Advanced Filter window is displayed.

4. Edit or add a condition for the filter. You can use the **Visual** tab, **Text** tab, or a combination of both tabs.

   To add conditions using the **Visual** tab:

   a. From the **Data Items** list, select the data item on which the condition is based.

   b. From the **Column Templates** list, select a condition. For more information, see Appendix 6, “Conditions for Filters,” on page 607.

   c. Drag and drop the condition onto the expression.

   d. For any required parameters, select the parameter, and enter a value, or right-click the parameter field, and select **Replace with** to select a data item.

5. (Optional) Replace a condition by dragging and dropping a new condition onto the existing condition in the expression on the **Visual** tab. Or, remove a condition using the **Text** tab.

6. (Optional) Add an operator to the expression. You can use the **Visual** tab, **Text** tab, or a combination of both tabs.

   To add an operator using the **Visual** tab:

   a. From the **Operators** list, select an operator to join the conditions in your expression. For more information, see Appendix 6, “Conditions for Filters,” on page 607.

   b. Drag and drop the operator onto the expression.

   c. (Optional) Add a condition.

> **TIP** Right-click the AND or OR operator in the expression, and then select **Add** ▸ **New Condition**.

7 (Optional) Delete part of an expression by highlighting the part of the expression that you want to delete, and then selecting **Delete** or **Clear**.

8 Click **OK** to apply the advanced filter.

Delete an Advanced Report Filter

To delete an advanced filter, click ✕ on the filter on the **Filters** tab.

Use a Data Source Filter in a Report

About Data Source Filters

Data source filters are used to restrict the data that is displayed in a report. The data source filters that you create in the designer are applied to every report object in the report that uses the data source. A report that has multiple data sources can contain multiple data source filters.

You can use either continuous values or discrete values to create a data source filter. Continuous value filters can be used only for measures. Discrete value filters can be used for any character; numeric; or date, datetime, or data item; as long as the total number of distinct values does not exceed a maximum number.

Data source filters are not displayed on the **Filters** or the **Data** tabs. A data source filter updates the cardinality values that appear on the **Data** tab.

There is a limit of one data source filter per data source. However, if you want to filter on more than one data source, you can create a combination filter.

Create a Data Source Filter

1 On the **Data** tab, click , and then select **New Data Source Filter**. The New Data Source Filter window is displayed.

2 Add a condition for the filter. You can use the **Visual** tab, **Text** tab, or a combination of both tabs.

   To add conditions using the **Visual** tab:

   a From the **Data Items** list, select the data item on which the condition is based.

   b From the **Column Templates** list, select a condition. For more information, see Appendix 6, “Conditions for Filters,” on page 607.

   c Drag and drop the condition onto the expression.

   d For any required parameters, select the parameter, and enter a value, or right-click the parameter field, and select **Replace with** to select a data item.

3 (Optional) Replace a condition by dragging and dropping a new condition onto the existing condition in the expression on the **Visual** tab. Or remove a condition using the **Text** tab.

4 (Optional) Add an operator to the expression. You can use the **Visual** tab, **Text** tab, or a combination of both tabs.

To add an operator using the **Visual** tab:

a From the **Operators** list, select an operator to join the conditions in your expression. For more information, see Appendix 6, “Conditions for Filters,” on page 607.

b Drag and drop the operator onto the expression.

c (Optional) Add a condition.

> **TIP** Right-click the AND or OR operator in the expression, and then select **Add** ▸ **New Condition**.

5 (Optional) Delete part of an expression by highlighting the part of the expression that you want to delete, and then selecting **Delete** or **Clear**.

6 Click **OK** to apply the filter.

Edit a Data Source Filter

1 If a data source is not already added, add a data source.

2 On the **Data** tab, click , and then select **Edit Data Source Filter**. The Edit Data Source Filter window is displayed.

3 Add or modify the condition for the filter. You can use the **Visual** tab, **Text** tab, or a combination of both tabs.

To edit conditions using the **Visual** tab:

a From the **Data Items** list, select the data item on which the condition is based.

b From the **Column Templates** list, select a condition. For more information, see Appendix 6, “Conditions for Filters,” on page 607.

c Drag and drop the condition onto the expression.

d For any required parameters, select the parameter, and enter a value, or right-click the parameter field, and select **Replace with** to select a data item.

4 Edit the operator in the expression. You can use the **Visual** tab, **Text** tab, or a combination of both tabs.

5 Click **OK** to apply the filter.

Delete a Data Source Filter

To delete a data source filter:

1 On the **Data** tab, click , and then select **Delete Filter**.

2 Click **Delete** in the confirmation message that is displayed.

Using Post-Aggregate Report Filters

About Post-Aggregate Report Filters

Post-aggregate filters subset the data for individual report objects in your reports by using the aggregated values, not the summarized values. You can use the **Filters** tab in the right pane of the designer to filter data in a report object using an aggregated value instead of a detail value. Post-aggregate filters are available only for measure data items.

When a report object has both ranks and post-aggregate filters applied, the ranks are applied before the post-aggregate filters.

Here are some key points about post-aggregate filters:

- Post-aggregate filters are not available for report objects that use detail data.
- Crosstabs, time series plots, and dual axis time series plots do not support post-aggregate filters.

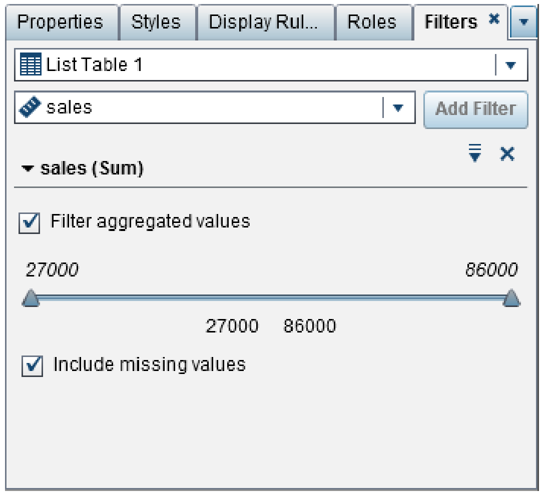
Create a Post-Aggregate Report Filter

To create a post-aggregate filter:

1. If it is not already selected, select the report object in the canvas that you want to filter. The report object must have at least one data item assigned.

   > **TIP** Clear the **Auto-update** check box above the report canvas until you are ready to apply your filter changes.

2. In the right pane, click the **Filters** tab.
3. Click ▾ beside the **Add Filter** button, and select a measure data item from the list. Then, click **Add Filter**. The filter appears on the **Filters** tab.
4. Select the data values for the post-aggregate filter. If the data item allows aggregate values, the **Filter aggregated values** check box is displayed. A slider shows you the maximum and minimum data values that exist for the data item using the current data item format. Use the slider to select a range of target values.



Note: The post-aggregate filter tracks the aggregation associated with the data item. If you change the **Aggregation** using the data item table on the **Data** tab, then the filter name on the **Filter** tab reflects that change. For example, if you create a post-aggregate filter for a data item called *Sales*, it is initially displayed on the **Filters** tab as *Sales* `(Sum)`. In the data item table, you change the aggregation to **Average**. On the **Filters** tab, the filter name is displayed as *Sales* `(Average)`. The post-aggregate filter attempts to keep the same range of data values that you originally selected for the aggregation.

> **TIP** Use the arrow to the left of the filter name on the **Filters** tab to expand or collapse the filter details when you are working with multiple filters.

5 (Optional) Click beside the filter name to change the operator. Select **Condition Type**, and then select an operator.

6 (Optional) Change the lower and upper values for the post-aggregate filter by clicking the value, and then entering a new value.

7 (Optional) If your data contains missing values, and you want to exclude those missing values from your report, then clear the **Include missing values** check box.

8 (Optional) If you cleared the **Auto-update** check box above the report canvas while you worked on your filter, then select it when you are ready to apply your filter changes.

Edit a Post-Aggregate Report Filter

To edit a post-aggregate filter:

1 If it is not already selected, select the report object in the canvas that you want to filter. The report object must have at least one data item assigned.

2 In the right pane, click the **Filters** tab.

3 Click beside the filter name. Then, select **Edit Filter**. The Edit Filter window is displayed.

4 Edit the expression for your post-aggregate filter.

- You can drag and drop conditions and operators onto the expression on the **Visual** tab in the right pane.
- You can enter the expression on the **Text** tab in the right pane.
- You can create an expression using both the **Visual** and **Text** tabs in the right pane.

For information, see Appendix 6, “Conditions for Filters,” on page 607.

Note: The AND and OR Boolean operators can accept more than two conditions. To add a condition to the operator, drag and drop a condition onto the operator name in the right pane. For example, to add a third condition to an AND operator, drag and drop the condition onto AND in the expression.

TIP Right-click the AND or OR operator in the expression, and then select **Add** ▸ **New Condition**.

5 Click **OK** to apply the post-aggregate filter.

Delete a Post-Aggregate Report Filter

To delete a post-aggregate filter, click ✕ beside the post-aggregate filter on the **Filters** tab.

46

Working with Report Interactions

Overview of Report 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 interactions view in SAS Visual Analytics Designer (the designer) enables report designers to specify which interactions they would like to add to tables, graphs, gauges, and controls in a report.

There are these types of report 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 controls. 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.

Tables, graphs, and gauges can be the source of an interaction, with the exception of time series plots. Report objects that use detail data cannot be the source of an interaction. Controls that are used on the canvas can also be the source of an interaction. Controls that are used as section prompts are treated as automatic filters and are not displayed in the interactions view.

The **Interactions** tab in the right pane enables you to create interactions and links.

Here are some key points about interactions:

Note: Interactions are available only for report objects in the same section.

Note: Report linking is a type of interaction. The report section that is the target of the link is subset by the values selected in the linked report object. For more information, see “Overview of Report Links” on page 449.

Creating a Report Interaction

Create a Filter Interaction

You can create interactions using the interactions view or the **Interactions** tab in the right pane.

The following objects that are imported from the explorer can be the source of a filter interaction:

- Box plots
- Heat maps with a relational category measure
- Forecast plots
- Geo maps

Imported correlation matrices, histograms, scatter plots with fit lines, heat maps without a selectable relational category measure, and box plots without a selectable relational category measure cannot be the source of an interaction. Hierarchy data items are not supported for interactions.

Here are some key points about creating filter interactions:

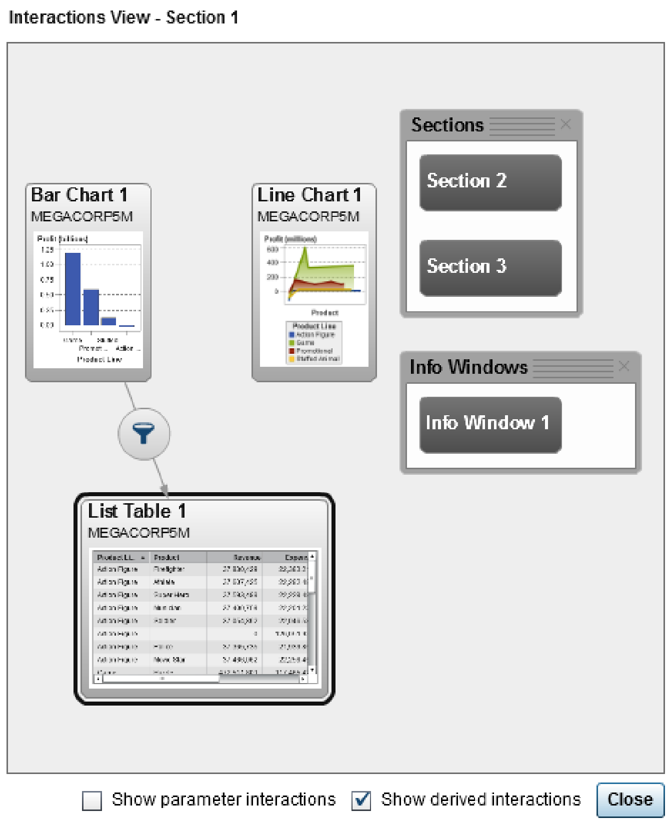
- You cannot create interactions from list tables, bubble plots, and scatter plots that use detail data. However, a list table that uses detail data can be the target of an interaction.
- Scatter plots cannot be the source of an interaction.
- Bubble plots can be the source of an interaction only if they have a **Grouping** role assigned.
- Sliders can be the target of an interaction. However, a slider that has the **Set fixed range** property set will not do anything when it is filtered because its data is fixed.
- A slider that is assigned to an aggregated measure data role cannot be the source of an interaction.
- Prompt containers enable you to delay the execution of interactions to report objects that are outside of the prompt container. However, interactions between report objects that are inside a prompt container are never delayed.

To create a filter interaction using the interactions view:

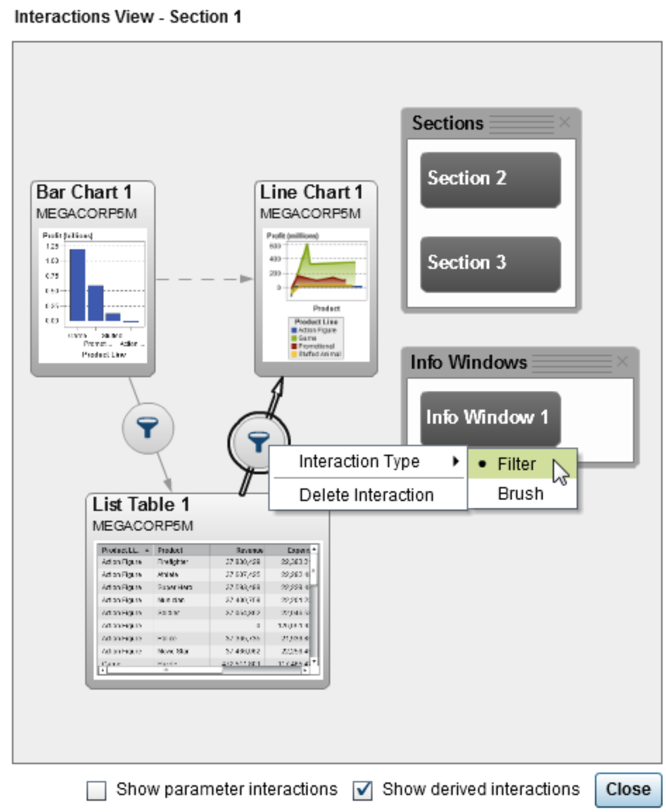
1. Add the report objects that you want to use to the canvas. For example, you might have a bar chart, a line chart, and a list table.
2. Click **Interactions View** on the **Interactions** tab to switch to the interactions view. You can also select **View** ▸ **Show Interactions**.

3 In the interactions 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):

Note: The source and target of an interaction should be based on the same data source. If you have multiple data sources, you are prompted to map the data sources to create the interaction. For more information, see "Map Data Sources" on page 444.



4 Right-click . Then, select **Interaction Type** ▶ **Filter**.



5 (Optional) Select the **Show parameter interactions** check box. (Parameter interactions are represented by dashed lines in the Interactions View. If the parameter is affecting a report object in another section, a link is drawn to that section. If a section contains a parameter that is targeting a report object in the current section, then a link is drawn from that section to the report object in the current section.)

6 (Optional) Clear the **Show derived interactions** check box. (Derived interactions are represented by dashed lines in the Interactions View.)

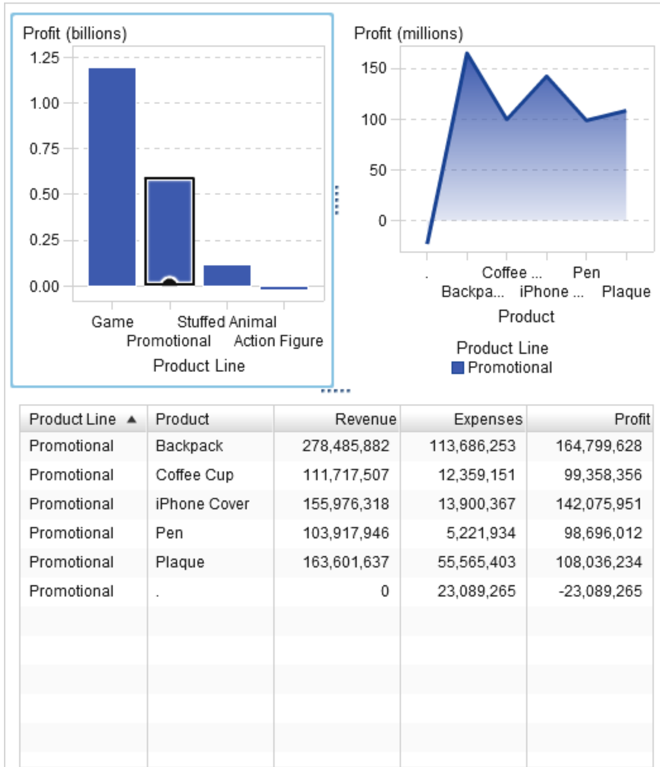
7 (Optional) Add other filter interactions.

8 Click **Close**. The new interaction (or interactions) appears on the **Interactions** tab in the right pane.

The following example shows the **Interactions** tab for the filtering example shown in Figure 46.1 on page 436.

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.

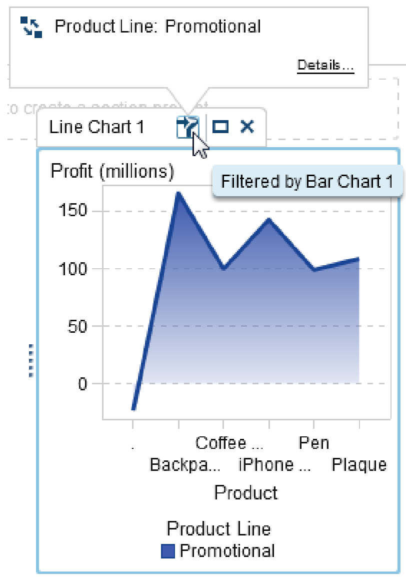
Figure 46.1 *A Filtering Example with a Bar Chart, a Line Chart, and a List Table*



Product Line ▲	Product	Revenue	Expenses	Profit
Promotional	Backpack	278,485,882	113,686,253	164,799,628
Promotional	Coffee Cup	111,717,507	12,359,151	99,358,356
Promotional	iPhone Cover	155,976,318	13,900,367	142,075,951
Promotional	Pen	103,917,946	5,221,934	98,696,012
Promotional	Plaque	163,601,637	55,565,403	108,036,234
Promotional	.	0	23,089,265	-23,089,265

Clicking on another selection applies the filter based on your new selection. For example, in the report above, you can click on the game product line to change the filter. An icon is added to the report object that is filtered.

Here is the line chart that was used in the previous example. The incoming filter icon appears above the report object when it is selected in the canvas. Click to get details about the filter that has been applied. Click **Details** to find out more about the filter. The Visual Element Filters window is displayed. It shows the exact query that was sent to the server.



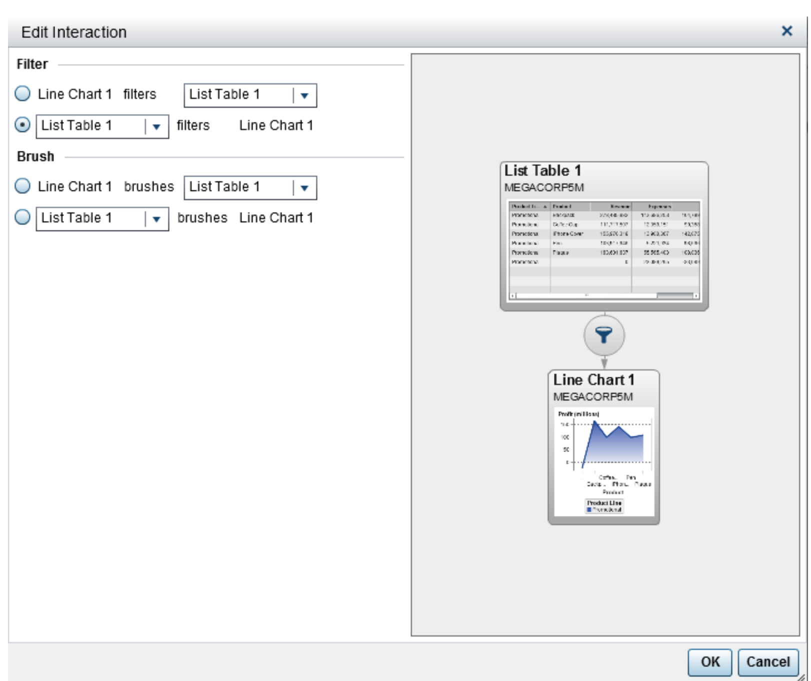
TIP To clear the selection and reset the filter (or filters), press Ctrl+click while viewing the original report object.

To create an interaction using the **New** button on 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 **New**, and then select **Interaction**. The **Edit Interaction** window is displayed.



4 In the Edit Interaction window, select the report object that you want to use as the source of a filter or brush interaction. For example, if your report has a list table, and you want it to filter a line chart, the first line should read *List Table 1* filters *Line Chart 1*.

The Edit Interaction window disables options for filter or brush interactions if they are not possible.

5 When you are finished working with interactions, click **OK**. The new interaction is displayed in the **Interactions** tab.

Create a Data Brushing Interaction

The following objects that are imported from the explorer cannot be the source of a brush interaction:

- Box plots
- Heat maps with a relational category measure
- Forecast plots
- Geo maps

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 Select the **Interactions** tab.

3 Click **Interactions View** on the **Interactions** tab to switch to the interactions view. You can also select **View** ▸ **Show Interactions**.

4 In the interactions view, draw a connection between the source and target report objects. When you have created an interaction, the is displayed between the report objects.

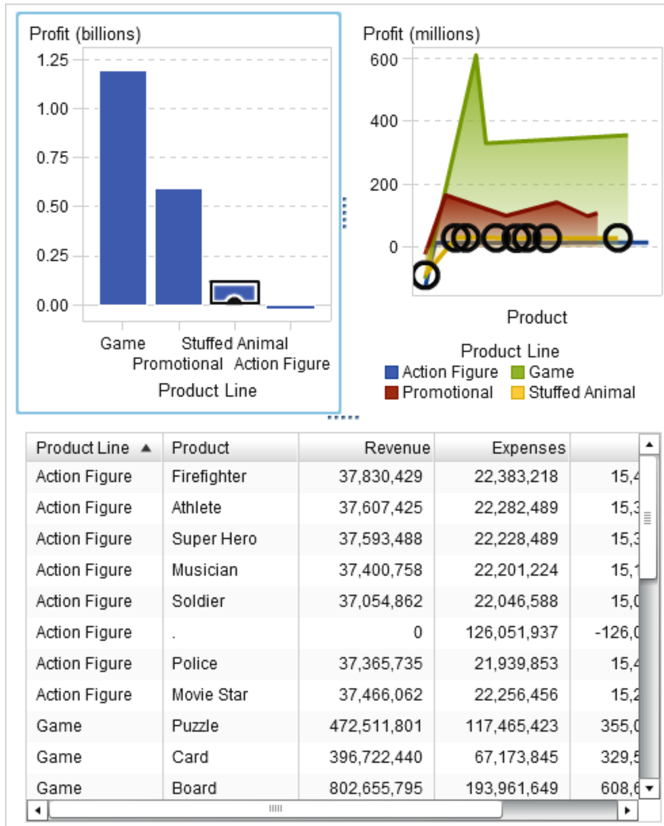
Note: The source and target of a data brushing interaction should be based on the same data source. If you have multiple data sources, you are prompted to map the data sources to create the interaction. For more information, see "Map Data Sources" on page 444.

5 Right-click . Then, select **Interaction Type** ▸ **Brush**. The filter icon changes to .

6 Click **Close**. The new interaction (or interactions) appears on the **Interactions** tab in the right pane.

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 line chart and the list table highlight the same data.

Figure 46.2 A Data Brushing Example with a Bar Chart, a Line Chart, and a List Table

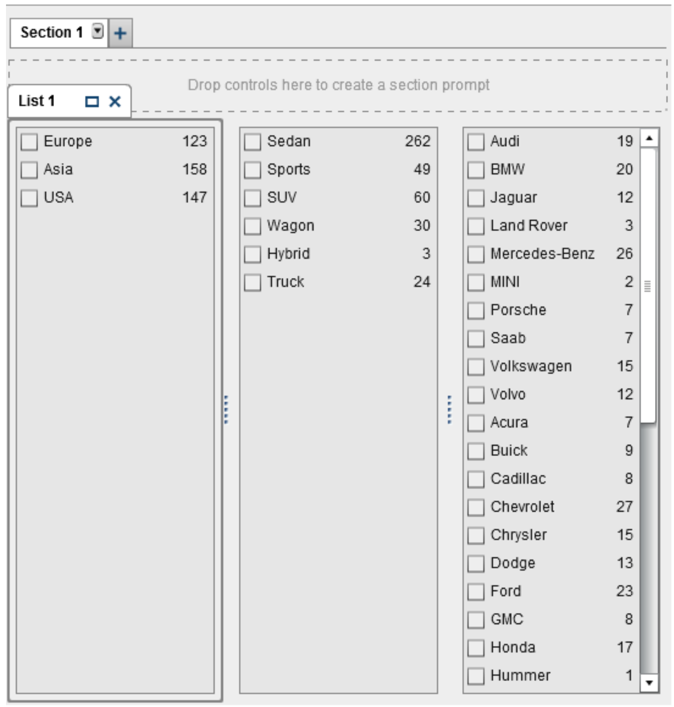


Note: If a graph contains a frequency measure, then a crosshatch pattern indicates when it is selected or brushed.

The brush interaction between controls behaves differently than a brush interaction between tables and graphs. When there is a brush interaction between controls, instead of highlighting or selecting the values in the other control, the values that match are moved to the top of the control, and the values that do not match are grayed out. If a frequency data item exists, then the frequency value is updated.

Here is an example of a report with three list controls that have brush interactions:

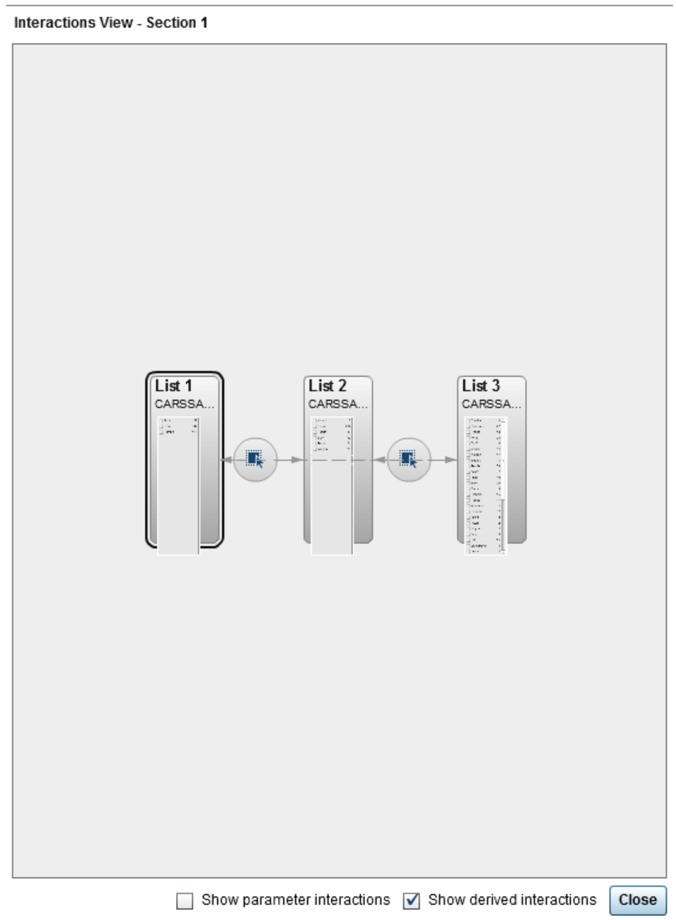
Figure 46.3 *Example: Three List Controls in a Report*



The first list contains three countries listed as the `Origin`. The second list contains six types of vehicles. The third list contains 38 car manufacturing companies. There is a brush interaction between the first list and the second list. There is a second brush interaction between the second list and the third list.

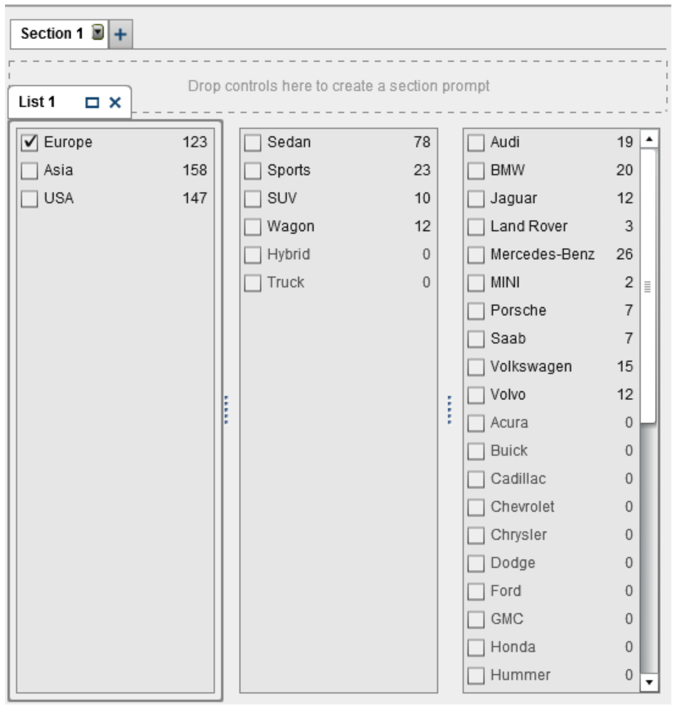
Here is an example of the interactions view for the report:

Figure 46.4 Three List Controls with Brush Interactions



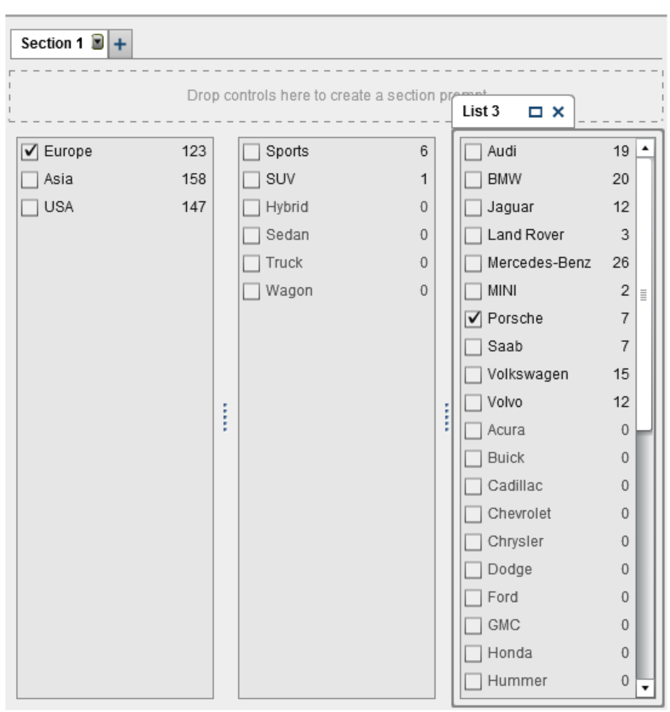
In step 1, the user has selected `Europe` as the country of `Origin` in the first list. Notice the changes in the second list. `Hybrid` and `Truck` are grayed out, and their frequency values have both changed to `0`. In the third list, all of the car manufacturers that are not located in Europe are grayed out, and their frequency values have all changed to `0`.

Figure 46.5 *Example: Three List Controls in a Report with One Brush Interaction*



In step 2, the user has selected **`Porsche`** as the car manufacturing company in the third list. Notice the changes in the second list. **`Hybrid`**, **`Sedan`**, **`Truck`**, and **`Wagon`** are grayed out, and their frequency values have all changed to `0`.

Figure 46.6 *Example: Three List Controls in a Report with Two Brush Interactions*



Map Data Sources

The source and target of an interaction or link should be based on the same data source. You might be prompted to map data sources so that an interaction or link works properly in these situations:

- When you try to create an interaction or link between a report object that uses one data source and another report object that uses a different data source.
- When you try to create a link between a report object that uses one data source and a report section that uses a different data source.
- When you try to create a link between a report object that uses one data source and a report that uses a different data source.

Here are some considerations for mapping data sources:

- For mappings created from a report link, the target data item cannot be a calculated item.
- Data source mapping is not supported from links in a text object.

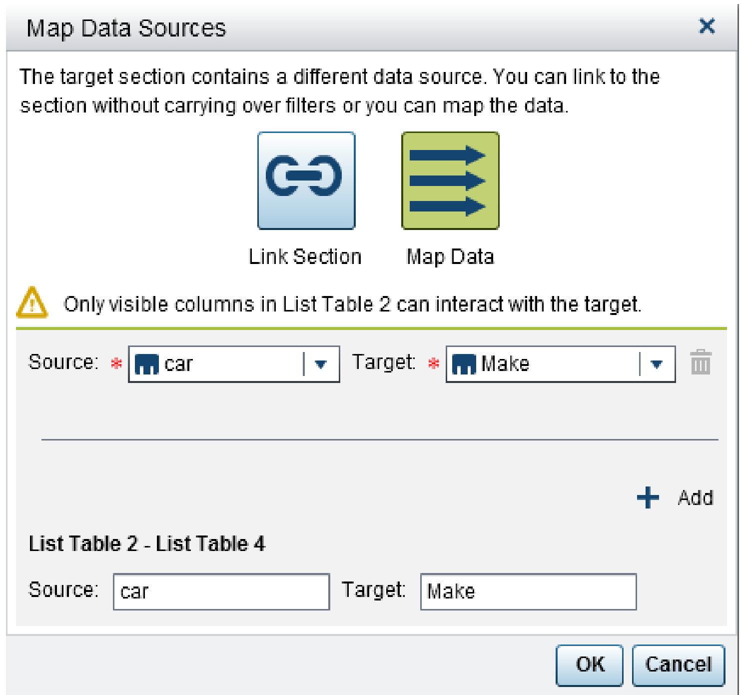
Here are some key points about mapping columns in data sources:

- A column in one data source can be mapped only once to another data source. If you need to map a column more than once, then the column needs to be duplicated in the data source.
- Mapped columns must share the same format for filters to work. For example, if the format of the source column is MMDDYYYY and the format of the target column is DDMMYYYY, then a filter will not work.

To map data sources using the Map Data Sources window:

1. Use the **Source** drop-down list to select a column from the first data source.
2. Use the **Target** drop-down list to select a corresponding column from the second data source.

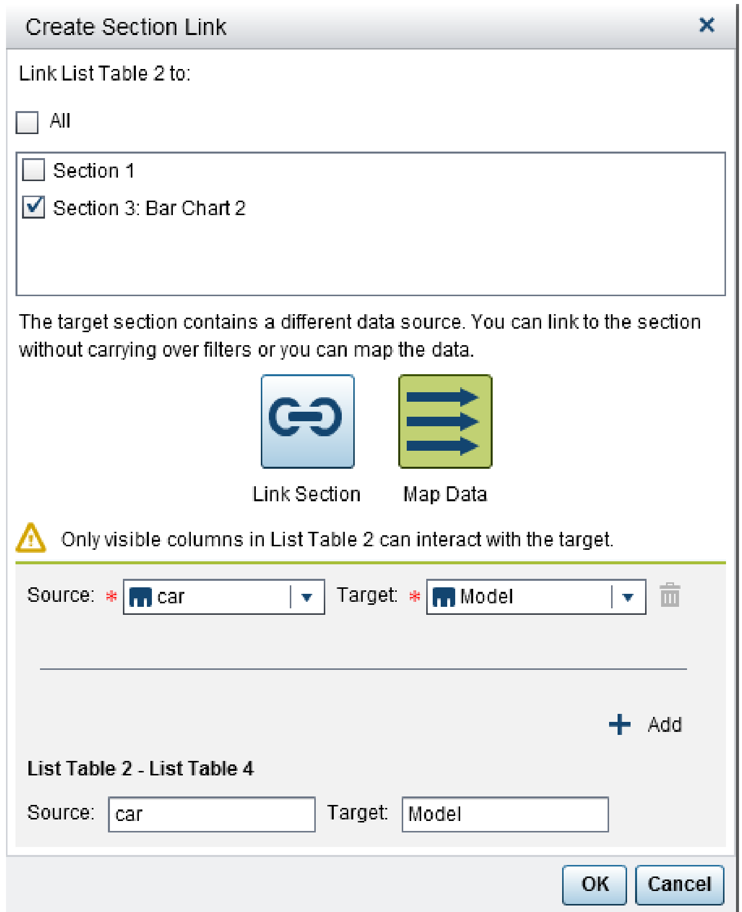
Here is an example of the Map Data Sources window for an interaction between two list tables that have different data sources:



TIP If there are multiple data sources, and you do not want to link to all of them, you can select the **Enable data source mapping** check box. For more information, see “Map Data Sources” on page 444.

For section links, you are not required to map the data sources. If you choose not to map the data sources, then the filters are not carried over. For more information, see “Create a Link to One or More Sections in a Report” on page 451.

Here is an example of the Create Section Link window with the map data sources information.



3 Map additional data items. Click +. A new row is displayed in the Map Data Sources window.

Note: If you create interactions between multiple objects on the same data source, but on different columns in that data source, then you must map each and every column in a Map Data Sources window. If you do not do this, then a subsequent mapping for the data source overrides a previous mapping.

4 The lower half of the window shows you how the mappings are applied to the objects shown in the **Source** and **Target** relationship. Depending on the interaction or link, you might need to map additional **Source** and **Target** objects.

5 Click **OK**.

Delete a Report Interaction

You can delete an interaction using the interactions view or the **Interactions** tab in the right pane.

To delete either a filter or data brushing interaction using the interactions view:

1. Select the **Interactions** tab.
2. Click **Interactions View** to switch to the interactions view. You can also select **View** ▶ **Show Interactions**.
3. Right-click or the , depending on which type of interaction you want to delete. Then, select **Delete Interaction**.
4. Click **Delete** in the confirmation message that is displayed.
5. (Optional) Remove other interactions.

You can delete an interaction using the **Interactions** tab. Click next to the interaction that you want to delete. Then, click **Delete** in the confirmation message that is displayed.

47

Working with Report Links

Overview of Report Links

Report links enable single-step access to a report or web page that is related to a report that you are currently viewing. For example, you might be looking at a bar chart that has sales information for each geographical region of your company. If you click the bar for the Northeast region, then a report link associated with the graph could take you to a different report that provides information about employees in each region. You can click in the top left corner of a destination report to return to the previous report.

Using SAS Visual Analytics Designer (the designer), you can add a link from a report object to another report, to a specific section or an info window in the current report, or to an external URL. If a destination report contains multiple sections, then you are able (when defining the link) to choose the initial section of the destination report that you want to open first.

When a report has an info window, you can provide additional information to a user who is viewing the report. For example, you might want a list table to provide additional information for a bar chart, or you might want to provide additional text about what is displayed in the data for a particular report object. In the SAS Visual Analytics Viewer (the viewer), a user double-clicks the data (for example, a bar, a bubble, a pie slice, a table row, and so on) in a report object that has an info window, and then the additional information is displayed in a new window in the viewer. For more information, see "Overview of Report Sections and Info Windows" on page 471.

Linking has elements of both a filter and an interaction. A report section that is the target of a link is filtered by the values that are selected in the linked report object. And, like interactions, objects that display detail data cannot be the source of a link.

For information about how links appear in the viewer, see "View Links in Reports" on page 544.

Imported box plots and heat maps with a relational category measure and imported forecast plots can be the source of section, report, or external links.

Starting in the 7.4 release, you can synchronize prompt values and parameters across linked reports. For example, suppose that you have two reports, *Report 1* and *Report 2*. When you follow a link from *Report 1* to *Report 2*, all of the prompts and their values are displayed in *Report 2*, and their states are synchronized. The same is true when you move from *Report 2* to *Report 1*.

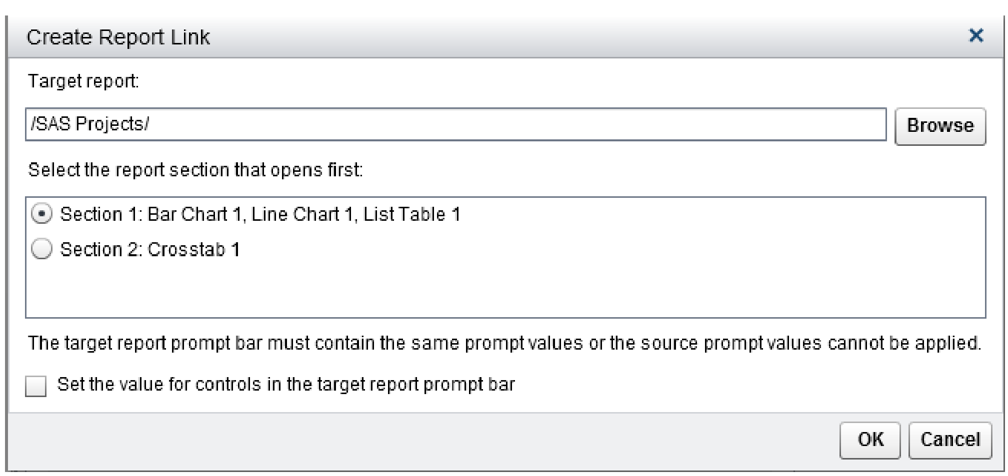
Creating Report Links

Create a Link to Another Report

Note: The following steps do not apply to text objects. For more information, see "Create a Link from a Text Object" on page 454.

To add a link from a report object or an image to another report:

1. If a report object is not already selected, select the object that you want to link from.
2. Click the **Interactions** tab.
3. Click **New**, and then select **Report Link**. The Choose an Item window is displayed.
4. Select a target report, and then click **OK**. The Create Report Link window is displayed.



5. Verify the name of the target report. If you want to select a different target report, click **Browse**.

Note: The source and target of report link should be based on the same data source. If you have multiple data sources, you are prompted to map the data sources to create the report link. For more information, see "Map Data Sources" on page 444.

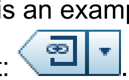
If the target report contains more than one section, then select the report section that you want to open first.

If you want to set values in the target report, select the **Set the value for controls in the target report prompt bar** option.

Note: This option sets values only on the controls that use the same data item as the source object or only on data items that filter the source object.

6 Click **OK**. The new report link is displayed on the **Interactions** tab in the right pane.

Double-click on an item in the report object (for example, a bar, a row in a table, a pie slice, and so on) to activate the report link. You are prompted to save changes to the report object before you can follow the link to the target report. To return to the source report object, use the button above the canvas.

Here is an example of the button that enables you to return to the source report object: .

Create a Link to One or More Sections in a Report

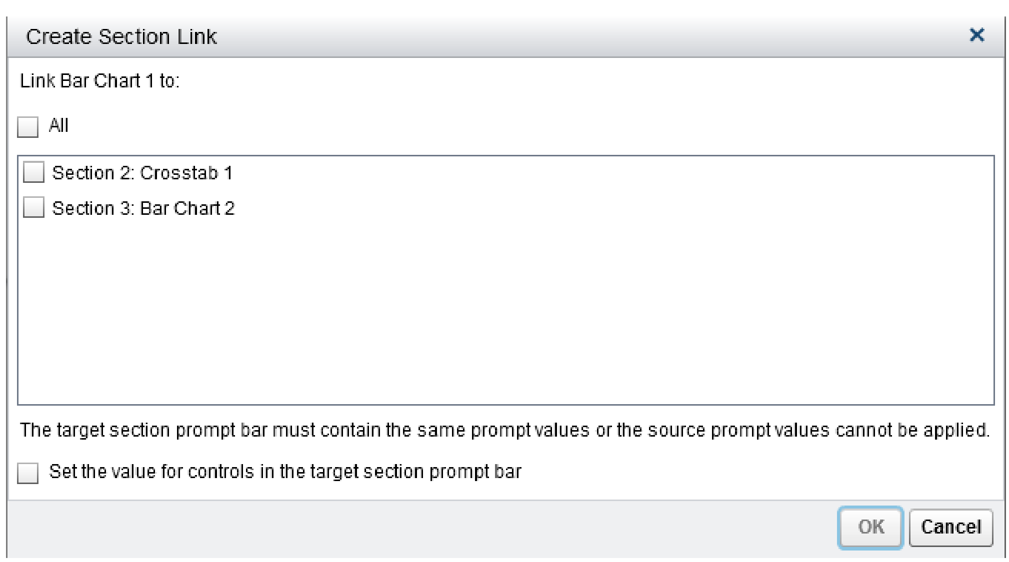
Note: The following steps do not apply to text objects. For more information, see "Create a Link from a Text Object" on page 454.

To add a link from a report object or an image to a specific section in the same report:

1 If a report object is not already selected, select the object that you want to link from.

2 Click the **Interactions** tab.

3 Click **New**, and then select **Section Link**. The Create Section Link window is displayed.



4 Select **All** to link to all of the sections in the report or select a specific section (or sections) in the report.

Note: The source and target of section link should be based on the same data source. If you have multiple data sources, you are prompted to map the data sources to create the section link. For more information, see "Map Data Sources" on page 444.

If you want to set values in the target section, select the **Set the value for controls in the target section prompt bar** option.

Note: This option sets values only on the controls that use the same data item as the source object or only on data items that filter the source object.

5 Click **OK**. The new section link is displayed on the **Interactions** tab in the right pane. If you selected multiple sections, those links are displayed on the **Interactions** tab.

Alternatively, you can create a section link using the interactions view.

Note: If you use a section link to navigate to another section, and then you change the value of a report prompt, the section link filter is removed.

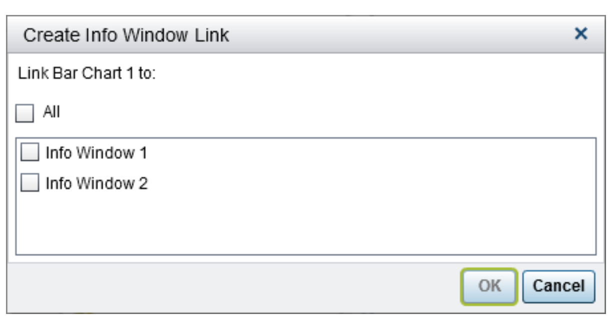
Create a Link to One or More Info Windows in a Report

Note: The following steps do not apply to text objects. For more information, see "Create a Link from a Text Object" on page 454.

To add a link from a report object or image to an info window in the same report:

1 Click the **Interactions** tab.

2 Click **New**, and then select **Info Window Link**. The Create Info Window Link window is displayed.



3 Select **All** to link to all of the info windows in the report, or select a specific info window in the report.

Note: The source and target of the info window link should be based on the same data source. If you have multiple data sources, you are prompted to map the data sources to create the info window link. For more information, see "Map Data Sources" on page 444.

4 Click **OK**. The new info window link is displayed on the **Interactions** tab in the right pane. If you selected multiple info windows to link to, those links are displayed on the **Interactions** tab.

Create a Link to an External URL

You can add a link from a report object, text, or image to an external URL. The `http://` part of the link is provided. The link can be relative to the current web application server. For example, you can link to a stored process, which means that you would not need to specify `http://`*server-name*:*port* because you are already logged on to that server.

TIP Click the **Link to Stored Process** button in the Create External Link window to link to a stored process without having to enter the required link information.

Note: UTF-8 is supported for external URL links.

Note: The following steps do not apply to text objects. For more information, see "Create a Link from a Text Object" on page 454.

To add a link to an external URL:

1 Click the **Interactions** tab.

2 Click **New**, and then select **External Link**. The Create External Link window is displayed.

3 Enter a **Label** and a **URL** for the link.

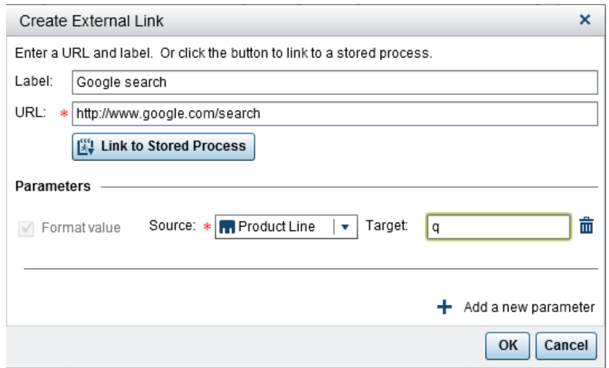
   TIP Do not add parameters to the **URL** field. Click + to specify additional parameters, such as a **Target**. These additional parameters are automatically added to the URL.

4 (Optional) Click **Link to Stored Process**. The Open window is displayed. Select a stored process, and then click **Open**. The **Label** field shows the name of the stored process, and the **URL** field shows the link information for the stored process.

   Note: If the stored process accepts prompts, they are automatically added to the **Parameters**.

5 (Optional) Click + to specify additional parameters for the link. You can specify the **Format value** option, a **Source**, and a **Target**. The **Format value** option applies only to dates and numbers. You can add multiple parameters.

   For example, you might have a report for your company's product lines. When a user views the report, you want them to be able to double-click a product line in the line chart, and then link to a Google search for that product line. In the Create External Link window for this example, you would enter *http://www.google.com/search* as the **URL**, select `ProductLine` as the **Source**, and then enter *q* as the **Target**.



6 Press Enter or Tab to activate the **OK** button. Click **OK**. The new external link is displayed on the **Interactions** tab in the right pane.

Double-click on a report object to link to the external URL.

Create a Link from a Text Object

You can add a link from text to an external URL, to another report, or to a specific section or an info window in the current report.

Note: Only static text can have links. Links are not available for dynamic text.

To add a link from a text object:

1 Double-click inside a text object on the canvas and enter the text.

2 Selecting some or all of the text, and then click in the floating toolbar. Select **Hyperlink**. The Link Setup window is displayed.

> **TIP** To change the default color of the hyperlinked text, you need to select a leading blank space and a trailing blank space around the hyperlinked text.

> **TIP** You can use the pop-up menu to cut, copy, and delete text. However, you have to use the keyboard shortcut (Ctrl+V) to paste text. Do not copy or paste text with hyperlinks. Instead, duplicate or import the text object.

3 Select the link **Type**. The options are:

External URL
The link target is an external URL. For example, you might want to link some text to your company's web page or Twitter feed. The link can be relative to the current web application server. Enter a **URL**. The `http://` part of the link is provided.

Report Link
The link target is another report. Enter the name of the **Target report** or browse to find it. If the target report has multiple sections, you can select which report section opens first.

Section Link
If your report has multiple sections, then the link target can be a section in the report. Select a **Target section** from the list.

Info Window Link

If your report has one or more sections and one or more info windows, then the link target can be an info window in the report. Select a **Target info window** from the list.

4 Click **OK**. The text that you selected is a link in the text object.

Edit a Report Link

To edit a report link or an external link:

1 Click next to the link on the **Interactions** tab. For a report link, the Edit Report Link window is displayed. For an external link, the Edit External Link window is displayed.

2 Make changes to the link.

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

Note: You cannot edit a section link.

Delete a Report Link

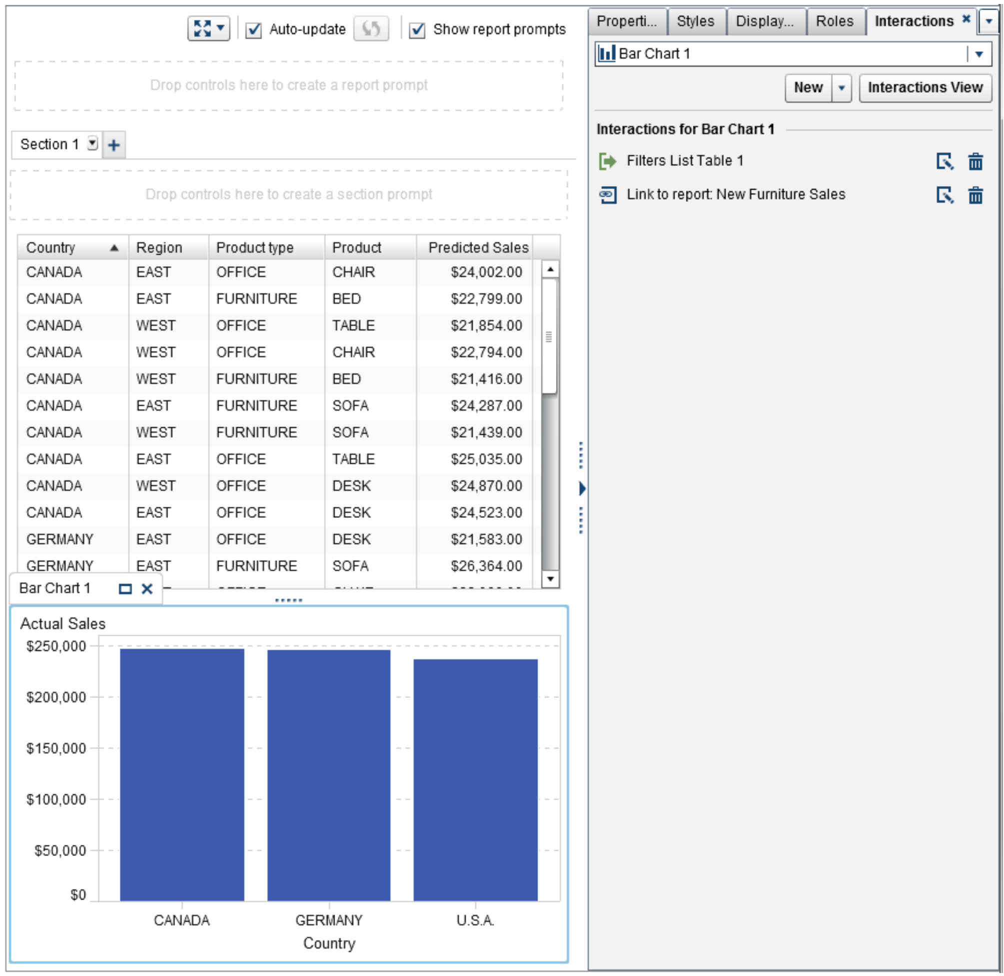
You can delete a section, report, or external link using the **Interactions** tab. Click next to the link that you want to delete. Click **Delete** in the confirmation message that is displayed.

To delete a link from a text object, select the linked text, and then click . Select **Delete Link**.

Example: How Report Links and Report Interactions Work Together

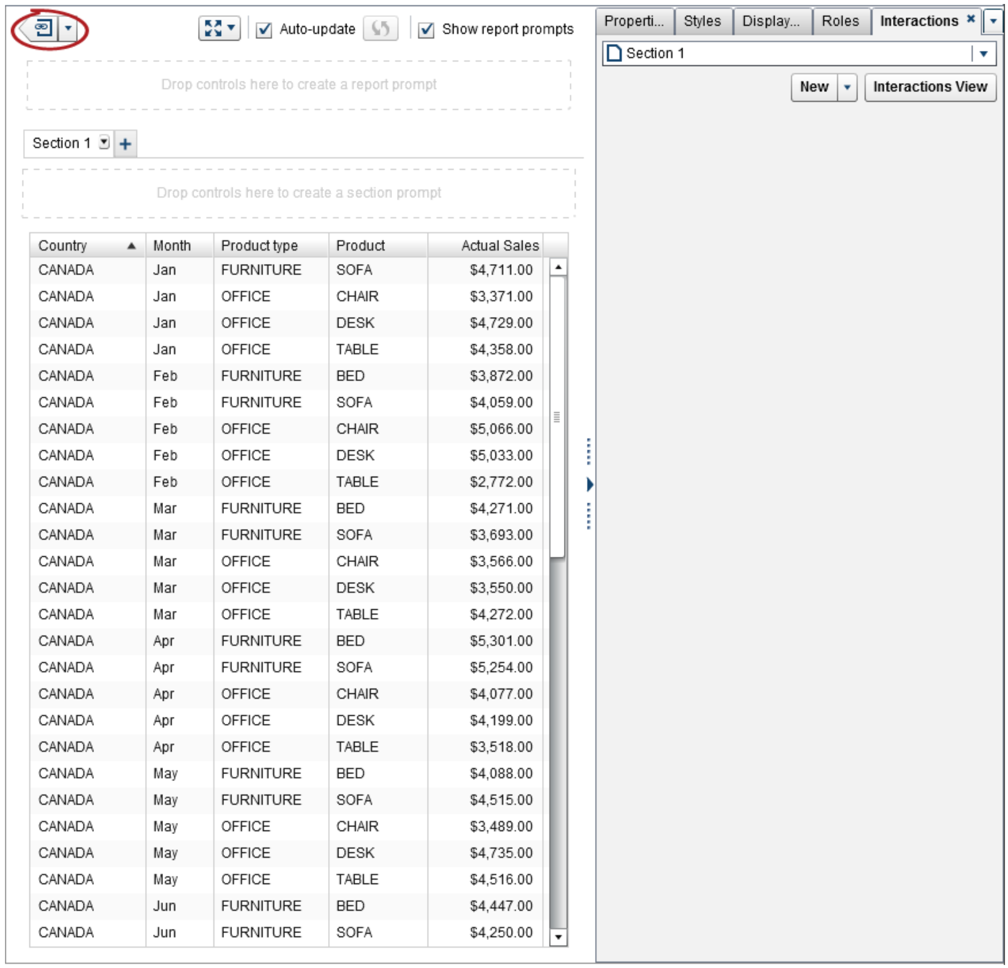
Suppose that you have two reports about furniture sales. The first report contains an interaction between a bar chart and a list table. The bar chart links to a second report. The filter interaction and the link are displayed on the **Interactions** tab. Currently, `CANADA` is selected in the bar chart, so the list table is filtered by that country. Here is an example of the first report:

Figure 47.1 Example: Furniture Sales Report



When you double-click the `CANADA` bar in the first report, the second report (which is the target of the report link) is displayed. The filter interaction for country in the first report is automatically passed to the second report when you link to it. Here is an example of the second report:

Figure 47.2 Example: Furniture Sales Report



You can remove the applied filter by clicking beside the in the upper left corner. Select the **Apply report link filters** option. After you remove the applied filter, you can choose to apply it again by selecting the **Apply report link filters** option.

You can return to the source report by clicking .

For information about how links appear in the viewer, see "View Links in Reports" on page 544.

48

Ranking Values in Reports

Overview of Ranking in Reports

Using SAS Visual Analytics Designer (the designer), you can rank the data in a report object to show the top (greatest) value or bottom (least) value for a category that is based on a measure. For a list table, you can also rank across a set of categories for the top value or bottom value in the set. A rank filters the values of a category based on the aggregated measure by the top or bottom of the values. A rank greatly reduces the visible categories to make it easier to focus on the top value or bottom value that interests a user.

For example, you might create a rank of the top 10 countries by frequency to select the 10 countries that are most represented in your report. As another example, you might create a rank of the top 10 countries by population to select the 10 countries with the greatest populations.

Ranks can accept parameters. For more information, see Chapter 49, “Working with Parameters in Reports,” on page 463.

Note: Rank descriptions are displayed for report objects that are imported from SAS Visual Analytics Explorer (the explorer) and that use on-demand data. You cannot edit or delete the ranks for imported objects.

Add a New Rank

You can use the **Ranks** tab to create ranks to subset the data in your reports. You can also right-click on your graphs and tables to add a rank to limit the number of bars, rows, and so on. Controls and gauges support ranks.

A single category rank ranks the top of a single category.

To add a rank:

1 Click the **Ranks** tab.

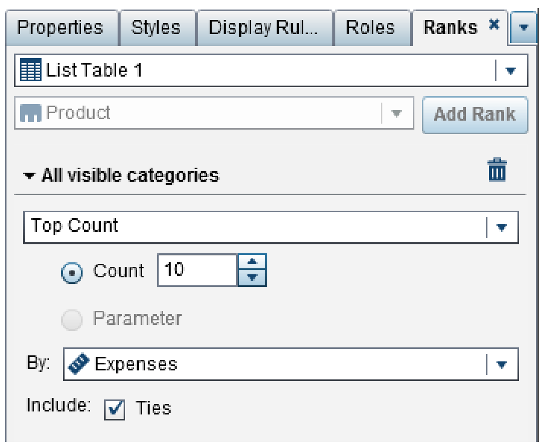
2 Select a data item. You can select any category or geography data item, regardless of whether it is assigned to the current report object.

For list tables only, you can select **All visible categories**. This rank ranks across the intersection. The top or bottom combination of the visible categories is displayed. For example, you select the region and product data items. You rank on the top 10 by profit. The top 10 region and product combinations are displayed. In this case, the column is no longer considered a single column, it is considered the crossing of the columns.

Note: If a list table has a rank for **All visible categories**, then that is the only rank that it can have.

The **Detail rank** option is available for list tables, bubble plots, and scatter plots that show detail data. If the report object has a detail rank, then it is the only rank that it can have. Otherwise, report objects can have multiple ranks.

3 Click **Add Rank**. The **Ranks** tab expands.



4 Select the type of rank from the drop-down list. These types are available:

- **Top Count** specifies that the rank selects the greatest values.
- **Bottom Count** specifies that the rank selects the least values.
- **Top Percent** specifies that the rank selects the greatest percentages.
- **Bottom Percent** specifies that the rank selects the least percentages.

5 If you selected **Top Count** or **Bottom Count** for the rank, then specify one of the following:

- Select **Count**, and then select a number for the rank. For example, if you select `5`, then the rank selects the five greatest values.
- Select a **Parameter**. If there is more than one parameter available, use the drop-down list to select the one that you want. For more information, see Chapter 49, “Working with Parameters in Reports,” on page 463.

6 If you selected **Top Percent** or **Bottom Percent** for the rank, then specify one of the following:

- Select **Percent**, and then enter a number between 0.1 and 99.9 in the **%** field.
- Select a **Parameter**. If there is more than one parameter available, use the drop-down list to select the one that you want. For more information, see Chapter 49, “Working with Parameters in Reports,” on page 463.

7 From the **By** drop-down list, select the measure that is used to create the rank. You can select any measure.

8 (Optional) Select **Ties** to include ties in the rank.

If you select **Ties**, then the rank selects as many values as necessary to include all of the ties. If you do not select **Ties**, then the rank selects only the number of values that are specified by the rank parameters.

For example, if your rank selects the top three values, but there are five values tied for the greatest value, then the number of values that are selected by the rank depends on the **Ties** option. If you select **Ties**, then the rank includes all five of the tied values. If you do not select **Ties**, then the rank includes only three of the tied values.

9 (Optional) For category-specific ranks, you can select **All Other** to show the measurements for the categories that did not qualify as a top or bottom value. This option is not available if the report object is showing only detail values.

Here are some key points about the **All Other** option:

- The option is available when you are ranking visible categories in certain types of report objects.
- The option applies only to the category that is being ranked. For example, suppose that you have *Region* and *Product* categories assigned to a list table. Then, you apply a *Region* rank with the **All Other** option set, so the "All Other" value might appear as a *Region* value, but not as a *Product* value.

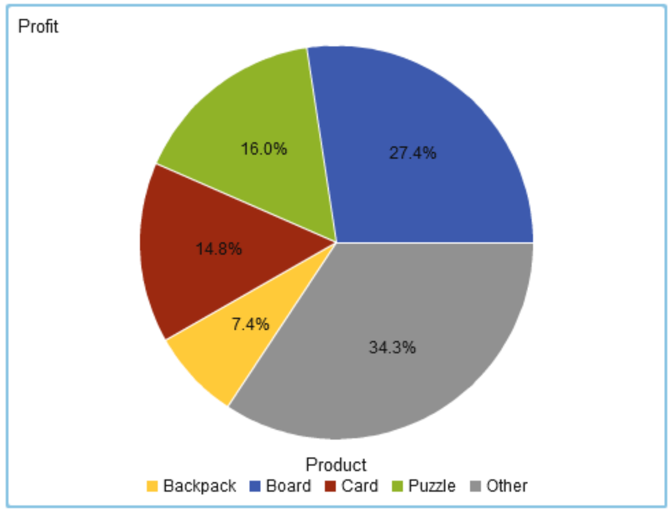
  If you want to combine category values that are excluded by rank into "All Other", then you need to use the **All Other** property. For more information, see "Use the Combine Excluded Rows (or Cells) into "All Other" Properties" on page 327.
- The option is not available if the rank is on a prompt control, geo bubble map, geo coordinate map, or geo region map.
- When the option is specified for a report object, the total, subtotal, and percent of total show values with respect to all of the data, rather than data just relative to data qualifying under the rank. The data that does not fit into the top or bottom of the rank is aggregated in the **All Other** category.
- You can use this option with a ranking to reduce the number of slices in a pie chart. However, this means the "Other" slice must be removed from the pie chart.
- If this option is selected, then any **All Other** value that appears in a table or graph cannot be selected. This means that the **All Other** value cannot be the source value for an interaction.
- If you do not select **All Other**, then the rank shows only the data as filtered by the category values that qualify as the top or bottom value.

By default, your new rank is applied automatically to the report object.

You can add more than one rank to a report object, as long as the first rank is not a **Detail rank**, or an **All visible categories** rank.

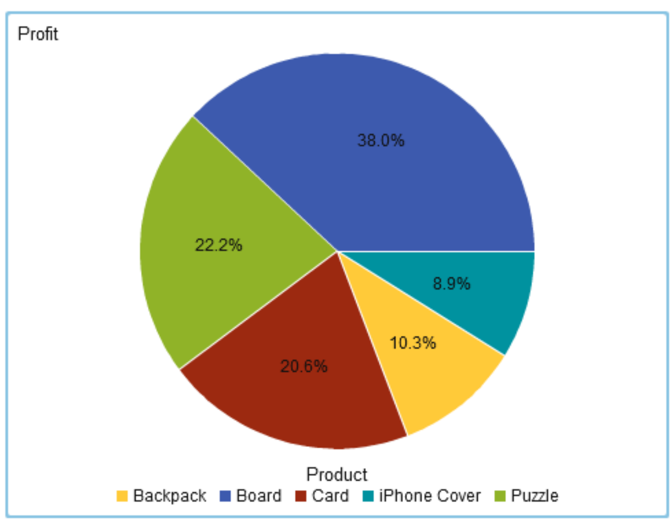
Here is an example of a pie chart that shows the profit for multiple product lines before a rank is applied:

Figure 48.1 *Pie Chart with Profits for Each Product Line Displayed*



Here is the same pie chart after a rank of the top five product lines is added. The **All Other** option was selected, so the “Other” slice is not displayed.

Figure 48.2 *Pie Chart with Profits for the Top Five Product Lines Displayed*



Delete a Rank

To delete a rank, click on the rank on the **Ranks** tab.

You can also right-click on the report object, and select **Delete Rank**, as long as there is only one rank for the report object.

49

Working with Parameters in Reports

Overview of Parameters

A *parameter* is a variable whose value can be changed and that can be referenced by other report objects. The SAS Visual Analytics Designer (the designer) supports parameters for controls in reports. If a control has an associated parameter, then when the value of the control changes, the parameter is assigned that changed value. When the value of the parameter changes, any report objects that reference the parameter detect the change accordingly.

Starting in the 7.4 release, you can create parameters for date and datetime data items. You can also create parameters that store multiple values.

Here are some key points about parameters:

- The type of parameter is required to match the type of data that is assigned to the control.
- If you duplicate a control with a parameter, the parameter is not copied from the original control.
- You cannot delete a parameter that is being used by a calculation, rank, filter, display rule, or text object.

Whenever a parameter value is updated, then all display rules, ranks, calculations, and filters that use that parameter are updated. Any report object in the report that uses the display rule, rank, calculation, or filter is updated accordingly. For an example of how parameters can be used in a complex report, see “Example: Using Parameters in a Report” on page 467.

Where Parameters Can Be Used in the Designer

You can use parameters in calculations, display rules, filters, and ranks. You can create, modify, and manage parameters using the **Data** tab in the left pane.

The following controls in the designer support parameters:

- The drop-down list control accepts any single-value parameter.
- The list control accepts multiple-value parameters only when multiple-selection is enabled.
- The button bar control accepts any single-value parameter.
- The text input field control supports character or numeric parameters. It does not support date or datetime parameters.
- The slider (single-point only) accepts numeric, date, and datetime parameters. It does not accept month or date and time format parameters.

You can assign a numeric parameter to a slider control or text input control. However, if the category data assigned to the control is also a numeric type, then you can assign a numeric parameter to a drop-down list control, list control, or button bar control. A numeric category data item is one that started as a measure data item, but then its classification property changed from Measure to Category. For more information, see “Example: Using a Numeric Parameter in a Report” on page 468.

Parameters can be used with the following features in the designer:

calculations
: calculated items and aggregated measures allow parameters. Parameters are supported wherever it makes sense to have a numeric, character, date, or datetime value. Multiple-value parameters can be used with only the In and NotIn operators.

display rules
: numeric parameters can be specified as the value of an expression rule.

  Note: Display rules cannot use character, date, datetime, or multiple-value parameters.

filters (detail filters, aggregated filters, and data source filters)
: a numeric, character, date, or datetime value parameter is supported wherever it makes sense to have a numeric, character, date, or datetime value. Multiple-value parameters can be used with only the In and NotIn operators.

  The data source filter is a special case. Ordinarily, a data source filter applies to all of the report objects on the canvas. However, if the data source filter contains a parameter, then the filter is not applied to the control that has that parameter assigned to it.

ranks
: for a rank, the parameter can be included for the *n* value of the rank. It is supported for **Top Count** or **Bottom Count** and **Top Percent** or **Bottom Percent**.

Note: Ranks cannot use character, date, datetime, or multiple-value parameters.

URLs

parameters in a report URL can be modified.

For example, suppose that you have the following URL for a report: http://*host*/SASVisualAnalyticsDesigner/?reportPath=%2FUser%20Folders%2Fsasdemo%2 FMy%20Folder&reportName=Parameterized%20Calculations%20DR&type=Report.BI&*Origin%20Parameter*=Europe&*Cost%20of%20gas*=3.35

The first parameter is called `Origin Parameter`, which enables you to specify a different country name. The second parameter is called `Cost of gas`, which enables you to specify different costs to see how different gas prices change the report.

Note: A parameter cannot be used to pass credentials. However, credentials can be used as the value of a parameter.

Note: You cannot provide a value for a multiple-value parameter using a URL parameter.

Create a New Parameter for a Report

To create a new parameter using the **Data** tab in the left pane:

1 Click , and then select **New Parameter**. The Create Parameter window is displayed.

2 (Optional) Modify the **Name** of the parameter.

3 Select the **Type** for the parameter. You can select one of the following:

- **Numeric**
- **Character**
- **Date**
- **Datetime**
- **Numeric list**
- **Character list**
- **Date list**
- **Datetime list**

For character parameters, you can specify a **Current value**.

For numeric parameters, specify a **Current value**. You must also specify the following options:

- **Minimum value** (which is required).
- **Maximum value** (which is required).
- **Format**.

Click **Select** to open the Select Format window, where you can select a **Format type**, **Width**, and **Decimals**. Click **OK** in the Select Format window.

For date, datetime, numeric list, date list, and datetime list parameters, select a **Format**. The date and datetime parameters also require the following:

- **Minimum value**
- **Maximum value**
- **Current value**

4 Click **OK** in the Create Parameter window. The icon identifies the new parameter in the list of data items on the **Data** tab.

You can create and edit parameters from the filter window and the calculated item window.

Edit a Parameter for a Report

To edit a parameter:

1 Right-click the parameter on the **Data** tab, and select **Edit Parameter**. The Edit Parameter window is displayed.

2 Make changes to the parameter.

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

Note: You cannot change a single-selection parameter to a multi-selection parameter, or vice versa.

Delete a Parameter for a Report

To delete a parameter that is not currently being used by any object:

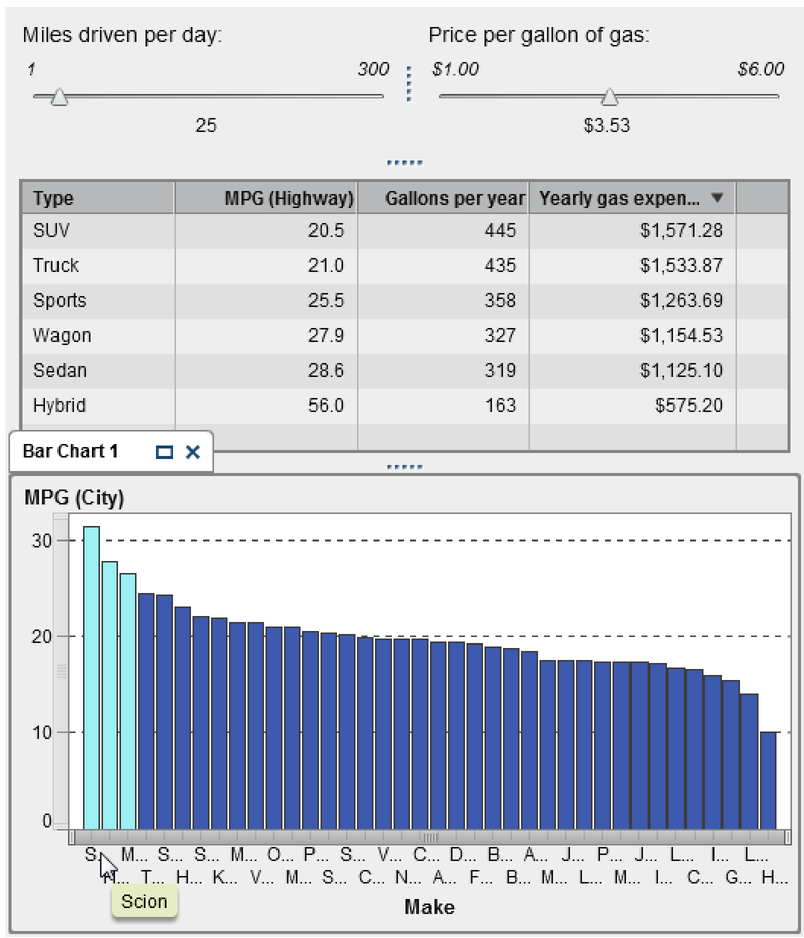
1 Right-click the parameter on the **Data** tab, and select **Delete Parameter**.

2 Click **Delete** in the confirmation message that is displayed. The parameter is removed from the list of data items on the **Data** tab and from any calculations, filters, display rules, or ranks that were using it.

Example: Using Parameters in a Report

You can use parameters to design complex reports. Here is an example of a complex report with parameters:

Figure 49.1 *Example: A Report with Parameters*



The report has two sliders, a list table, and a bar chart. One slider lets you adjust the number of miles that you drive each day. The other slider lets you adjust the price per gallon of gasoline. The results of the selections in the sliders are calculated, and the values are used in a calculated data item that is displayed in the list table. For example, you can see what happens to the list table when you change the parameter for *Miles driven per day* from *25* to *100*.

The values of the calculated data item are also used in the bar chart. The bar chart has a display rule, (*MPG (City)* > *Miles per day*), so that vehicles that have a fuel economy value that is better than the number of miles that you drive per day are highlighted.

Example: Using a Numeric Parameter in a Report

You can use a numeric parameter with a drop-down list control, list control, or button bar control. This example explains how to create a numeric parameter and use it with a button bar control in a report.

Suppose that you have a report for the sales of cars at a dealership. You have a data source that has a column named *Cylinders*. When your SAS Visual Analytics administrator loaded the data source, the *Cylinders* data item was defined as a measure. You want to add a button bar control to the report that shows the values for *Cylinders*, and has a parameter with the *Cylinders* values assigned to it.

Follow these steps to use a numeric parameter in your report:

1 Drag a button bar control from the **Objects** tab, and drop it onto the canvas.

2 Convert the *Cylinders* data item by right-clicking on it in the **Data** tab, and then selecting **Category**.

3 Create a numeric parameter by following these steps:

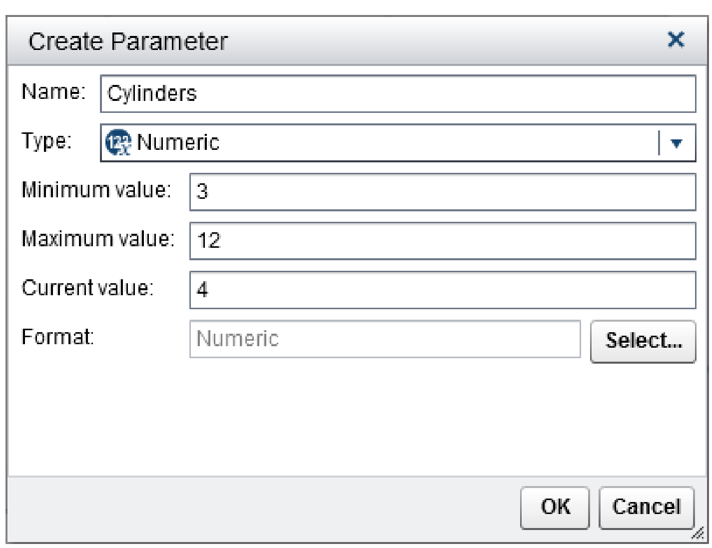
   a Right-click the *Cylinders* data item in the **Data** tab, and then select **Create Parameter from Data Item**. The Create Parameter window is displayed.

   b (Optional) Modify the **Name** of the parameter.

   c Select **Numeric** for the **Type** for the parameter.

   d Specify the **Minimum value**, **Maximum value**, and **Current value** for the numeric parameter.

   Here is an example of the Create Parameter window:



   e Click **OK**. The new numeric parameter is added to the **Data** tab, under the **Parameter** heading.

4. Drag the *Cylinders* data item, and drop it onto the button bar control.
5. Drag the *Cylinders* parameter, and drop it onto the button bar control. The button bar is populated with the values for the numeric parameter.

50

Maintaining Multi-Section Reports

Overview of Report Sections and Info Windows

Any report in SAS Visual Analytics Designer (the designer) can have multiple sections. (Sections are like pages.) 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. You can have one or more report objects in a section. There is no limit to the number of sections in a report.

Using the designer, you can add a link from a report object to a specific section in the current report. For more information, see “Create a Link to One or More Sections in a Report” on page 451.

Section prompts are controls that are placed in the special row area above the report objects on the canvas. A section prompt automatically filters all of the other report objects in the same section, as long as the report object uses the same data source as the section prompt control. For more information, see “Use a Control to Create a Section Prompt” on page 335.

When a report has an info window, you can provide additional information to a user who is viewing the report. For example, you might want a list table to provide additional information for a bar chart, or you might want to provide additional text about what is displayed in the data for a particular report object. In the SAS Visual Analytics Viewer (the viewer), a user double-clicks the data (for example, a bar, a bubble, a pie slice, a table row, and so on) in a report object that has an info window, and then the additional information is displayed in a new window in the viewer.

Info windows can be created only in the designer. A report can have one or more info windows, which are displayed as private tabs in the designer. Info windows are identified by the icon on the tab above the canvas. Info windows can have one or more report objects. You must provide a link to an info window from an existing report object to display a new window in the viewer.

Note: Section prompts are not allowed in info windows.

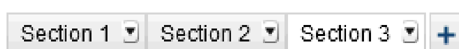
Add a Section to a Report

You can add sections to any report, including reports that were created in SAS Visual Analytics Explorer (the explorer).

To add a new section to a report, click to the right of the first section tab (or the last section tab that was added) in the report. The new tab appears to the right of the existing tab (or tabs).

Here is an example of multiple section tabs:

Figure 50.1 Multiple Section Tabs



You can use the **Properties** tab for the report to add a new section. In the **Sections** area, click . The new tab appears to the right of the existing tab (or tabs).

Add an Info Window to a Report

You can change any report section to an info window as long as the section does not have section prompts.

Note: If you change an existing report section to an info window, and there are links to that specific report section, then those links will no longer open that section.

Here is an example of multiple **Info Window** tabs:

To add an info window, click on the **Section** tab, and then select **Display as Info Window**. The tab name changes to **Info Window**. The **Properties** tab also updates.

For information about adding a link to an info window, see "Create a Link to One or More Info Windows in a Report" on page 452.

Reorder Report Sections or Info Windows

To change the order in which sections or info windows are displayed:

1 In the right pane, click the **Properties** tab.

2 Select the report name in the drop-down list.

3 In the **Sections** area, select the name of a section or an info window. Click  or  to reorder the sections or info windows.

Rename a Report Section or an Info Window

To rename a section or an info window:

1 Right-click the section tab or the info window tab that you want to rename. Select **Rename**. The existing name is highlighted.

2 Type a new name, and then press Enter. The new name also appears on the **Properties** tab in the right pane.

Alternatively, you can change the name on the **Properties** tab.

Duplicate a Report Section or an Info Window

To duplicate a section or an info window:

1 Right-click the section tab or the info window tab that you want to duplicate.

2 Select **Duplicate**. All of the report objects and the data in the original section or info window are copied to the duplicate section or info window. The duplicate appears on a new tab. For example, if the original section's name is `Marketing Reports`, then the duplicate section is named `Marketing Reports (1)`.

3 Rename the duplicated section or info window using the pop-up menu or the **Properties** tab.

Move a Report Object to Another Section or Info Window

To move a report object from one section in a report to another section or an info window, right-click the report object that you want to move. Select **Move <ReportObject> to ▶ <SectionName>**, where `<ReportObject>` is the name of the report object and `<SectionName>` is the name of the section tab or the info window.

> **TIP** You can drag an object from one section and drop it onto the tab for another section to move it to that section.

Note: You cannot move a section prompt from a section to an info window.

If you move an object to a section with precision layout, then you must manually move the object to its proper location. All objects are put in the top left corner by default.

> **TIP** If you move a section prompt from one section to another section in a report and there are multiple data sources, you must edit the data source mapping for an interaction to work. Right-click the control, and select **Edit Data Source Mapping**. For more information, see "Map Data Sources" on page 444.

Delete a Report Section or an Info Window

To delete a section or an info window from a report, right-click the tab that you want to remove. Select **Delete**.

Alternatively, click on the section tab or info window tab, and then select **Delete <Section Name or Info Window Name>**, where `<Section Name or Info Window Name>` is the name of the section or the info window. (For example, `Section 1`, `Info Window 1`, and so on.)

Note: The **Delete** menu option is not available if the report has only one section.

51

Sharing Reports with Other Users

Overview of Sharing Reports

Depending on your role and capabilities, the SAS Visual Analytics Designer (the designer) enables you to complete many tasks related to sharing reports. You can do the following:

- e-mail reports
- print reports
- export images or data from report objects
- add comments to reports
- distribute reports on a schedule
- localize (or translate) labels, tooltips, and other descriptive text in reports

E-mail a Report

You can e-mail a link to a report to someone using the designer. Here are some key points about what happens when a recipient clicks the link to a report:

- If the recipient has already signed in to SAS, the report is displayed when the report link is clicked.
- If the recipient is not signed in to SAS, they are prompted for a user ID and password. Starting in the 7.3 release, recipients can sign in as a guest by clicking **Guest** on the standard sign-in window for SAS applications.
- If the recipient opens the e-mail message on a PC or Mac with a Flash-enabled viewer, and then clicks the e-mail link, then the report is displayed in SAS Visual Analytics Viewer (the viewer).
- If the recipient opens the e-mail message on a mobile device, and then clicks the e-mail link, then the report is displayed using the SAS Visual Analytics App (previously called SAS Mobile BI) if the user has it installed.

TIP Recipients who use SAS Home (the home page) can use the **Application Shortcuts** setting to specify which application opens the report when they click on an e-mail link. For more information about settings, refer to the online Help that is available for the home page.

TIP Recipients who use the classic SAS Visual Analytics home page can use the **Order the actions associated with a content type** preference to specify which application opens the report when you click on an e-mail link. For more information, see “Specifying Your Preferences for the Classic SAS Visual Analytics Home Page” on page 637.

To e-mail a report from the designer:

1 Select **File** ▸ **E-mail**. The new e-mail message opens in your default e-mail application.

  The designer automatically adds a subject line with the report name. The URL for the report is automatically generated and appears in the e-mail message.

  Note: Only locally installed e-mail clients are supported. If your default e-mail application is web-based (for example, Gmail), you cannot e-mail a report.

2 Add recipients.

3 (Optional) Revise the subject line.

4 (Optional) Revise the e-mail message.

5 Send the e-mail message.

Printing Reports

You can create a PDF for a report, a report section, or a report object. The PDF can then be printed.

General Considerations for Printing

Here are some considerations for printing:

- Info windows, report prompts, section prompts, and prompt containers are not included in the PDF. However, any filters that are applied to report objects in the prompt container are displayed in the appendix.
- Vertical containers print only the currently selected report object. Horizontal and stack containers might print more than the currently selected report object if there is enough space.
- Select **Expand clipped content** to see the all of the content for tables, crosstabs, gauges, and containers with content that is only partially available in the layout of the report section. Each report object is displayed on a separate page at the end of the report.

> **TIP** In the generated PDF, click  to see the content that was clipped in the layout of the report section in the designer.

- If a list table or a crosstab has scroll bars, then some of the scrolled content might not be printed. However, more columns or rows might be printed than are visible in the designer. The print feature attempts to fill the space available on the page in the PDF.
- A list table or crosstab that is printed from the designer might show more rows than when the same list table or crosstab is printed after the report has been distributed. This happens because the state of the report in the designer is taken into consideration when it is printed. The print feature attempts to match what the user sees in the designer. When you print a report that has been distributed, there is no report state for the print feature to match.
- If a report object has a high-cardinality filter, and the **Show appendix information** option is selected, then the filter description might be truncated in the printed appendix.
- A crosstab with more than 5,000 selected cells cannot be printed.
- There can be differences in the fonts in a printed report if the fonts that are used to create a report do not match the fonts that are available on the server that generates the PDF.
- If you want to print from the designer using Mozilla Firefox, you must have Firefox configured to open a link in a new tab instead of in a new window. Instructions for making this change might vary between Firefox releases, so refer to the Firefox documentation. As an alternative, you can use a different supported browser, such as Microsoft Internet Explorer or Google Chrome.

Considerations for Printing a Stored Process or Stored Process Report

Here are some key points about printing output from a stored process or a stored process report:

- Output from a stored process starts on a new page.
- The stored process must use the %STPBEGIN and %STPEND macros. These macros provide standardized functionality for generating and delivering output from a stored process. For more information, see “Using the %STPBEGIN and %STPEND Macros” in *SAS Stored Processes: Developer’s Guide*.
- PDF security must be turned off for a stored process.
- The stored process report must be configured to produce PDF output. In the stored process that is referenced by the stored process report, set the _ODSDEST variable to `PDF`. Or, define a stored process parameter for the _ODSDEST variable and configure it with a value of `PDF` for the stored process report. It is recommended that you also set the _ODSOPTIONS variable to `notoc` to save processing time and to reduce the size of the PDF produced by the stored process and the final PDF. For more information about ODS options, see “Using the %STPBEGIN and %STPEND Macros” in *SAS Stored Processes: Developer’s Guide*.

Print a Report, a Report Section, or a Report Object

By default, an appendix is included in the PDF if the report, report section, or report object has descriptions, filters, warnings, errors, or display rule legends.

To create a PDF that can be printed for a report:

1. Select **File** ▶ **Print to PDF**, or click . The Print window is displayed.
2. Select the **Paper size**, **Orientation**, and **Margins**. Your selections are saved for that report.

   If you select **Custom** for the paper size, you can specify the **Height**, and **Width** of the page. For custom paper sizes, you cannot change the page orientation.
3. (Optional) Select the **Options** for your report.

   Here are some details about the printing options:

   - If you select **Include a cover page** check box, then the cover page provides the name of the report (or the label if the report has not been saved), the date of the print request, the user name, and the number of pages. Click **Add a description** to enter additional text for the cover page in the Cover Page Description Text window.

     Note: The description for the cover page is not saved if you click **Cancel** in the Print to PDF window.
   - If you clear the selection for the **Include a cover page** check box, any additional text that you added for the cover page is discarded.

- The **Show empty rows and columns in tables** option is for both list tables and crosstabs.
- The **Show appendix information** option is selected by default if the report, report section, or report object has parameters, descriptions, filters, warnings, errors, or display rule legends.

  When the **Show appendix information** option is specified, and the report, report section, or report object has parameters, descriptions, filters, warnings, errors, or display rule legends, then an appendix is created. Each report object that is selected to be included in the PDF is automatically assigned a value so that you can reference that report object in the appendix. For example, suppose that you select two report objects in the **Select the items to be printed** list. The first report object is assigned the value `A1.1` and the second report object is assigned the value `A1.2`. Suppose that the first report object has a description and the second report object has a filter that has been applied. The first report object's description is displayed in the appendix under the `A1.1` heading. The second report object's filter is displayed in the appendix under the `A1.2` heading.

4 Make your selections in the **Select the items to be printed** list. You can select the report, the report sections, or the report objects.

5 Click **Print**. The PDF is displayed in a browser.

6 (Optional) If you want to save any changes that you made in the Print window, then save the report.

> **TIP** The print options that you save with the report are used when you distribute the report to other users. Saved print options can impact the appearance of the report.

To print a report section, click on the **Section** tab, and then select **Print <*section-name*> to PDF**.

To print a report object, right-click on the report object in the canvas, and then select **Print <*report-object-name*> to PDF**. A stored process does not have the same pop-up menu selection.

The following table lists the icons that can appear in the appendix:

Icon	Description
	Indicates that a report object contains a filter.
	Indicates that a report object contains a warning.

Exporting Content from the Designer

Overview of Exporting from the Designer

You can export an image or data from a report object using the designer.

All users who have the Export Data capability can use the designer to export data from report objects to Microsoft Excel format for future viewing or printing. This exported output can be saved locally on disk, and then opened in Microsoft Excel. Or you can choose to create a delimited text file Tab-Separated Values (*.tsv) or Comma-Separated Values (*.csv) data file.

Here are some key points about exporting from the designer:

- When you export a graph from the designer, you are exporting the data, not the visual graph representation.
- Not all report objects in the designer support the exporting feature. For example, you cannot export data from 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 on the object.
- The designer does not preserve leading blanks in displayed data or exported data. However, you can filter for values that contain leading blanks.
- The designer uses the locale of the SAS LASR Analytic Server to export data.

Export an Image

You can export an image for any list table, crosstab, graph, or gauge in a report. This is useful if you want to include the image of a report object in a presentation, such as in Microsoft PowerPoint. Images are saved as PNG files.

Here are key points about exporting an image:

- An exported image does not show ranks or filters that are associated with the report object.
- The image defaults to the size of the report object in the report.
- The following characters in an image name are converted to an underscore when the image is exported: `/ \ : * ? " < > | %`

To export an image:

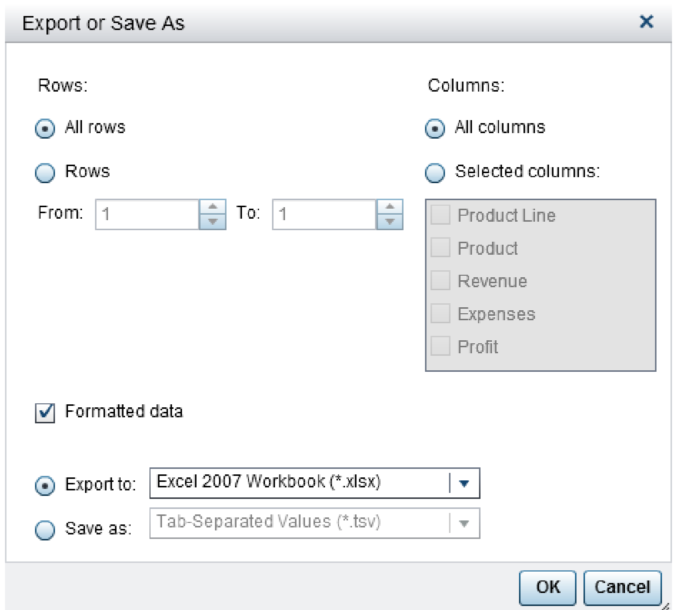
1. If the list table, crosstab, graph, or gauge that you want to export is not already selected, select it.
2. Right-click the report object, and select **Export Image**. The Save Image window is displayed.
3. Click **Save**. A download window is displayed.
4. Enter a name for the PNG file. Click **Save**.

Export Data from a List Table

To export data from a list table:

1 If the list table that you want to export is not already selected, select it.

2 Right-click the list table, and select **Export <listTableName>**, where `<listTableName>` is the name of the report object. The Export or Save As window is displayed.



3 Choose to export **All rows** or a range of **Rows**.

4 Choose to export **All columns** or **Selected columns**. If you choose **Selected columns**, select the check box (or check boxes) 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 **Export to**, then the only option is **Excel 2007 Workbook (*.xlsx)** to create a Microsoft Excel spreadsheet.

7 If you select **Save as**, choose either **Tab-Separated Values (*.tsv)** or **Comma-Separated Values (*.csv)** to create a data file.

8 Click **OK**.

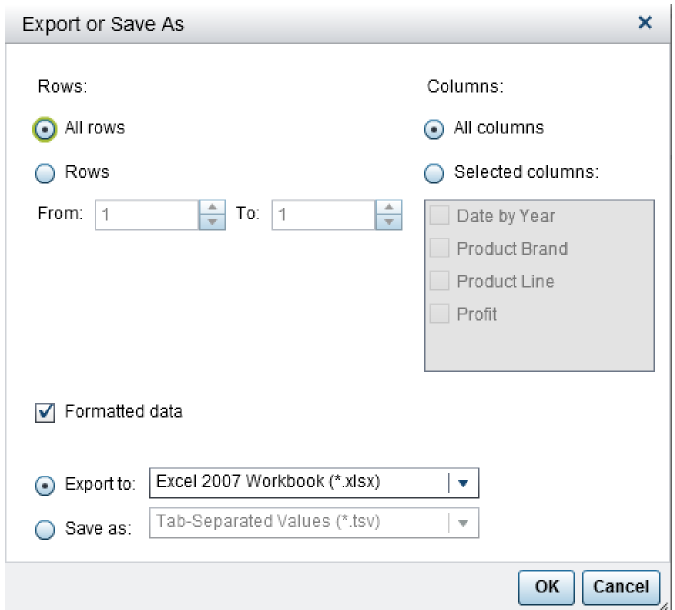
9 When you are prompted, choose either to open the file or to save it.

Export Data from a Crosstab

To export data from a crosstab:

1 If the crosstab that you want to export is not already selected, select it.

2 Right-click on the crosstab, and select **Export <crosstabName>**, where `<crosstabName>` is the name of the report object. The Export or Save As window is displayed.



3 Choose to export **All rows** or a range of **Rows**.

4 Choose to export **All columns** or **Selected columns**. If you choose **Selected columns**, select the check box (or check boxes) 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 **Export to**, then the only option is **Excel 2007 Workbook (*.xlsx)** to create a Microsoft Excel spreadsheet.

7 If you select **Save as**, choose either **Tab-Separated Values (*.tsv)** or **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.

Note: The designer exports data from a crosstab into a list table. This is different from SAS Visual Analytics Explorer (the explorer), which exports data from a crosstab into a crosstab.

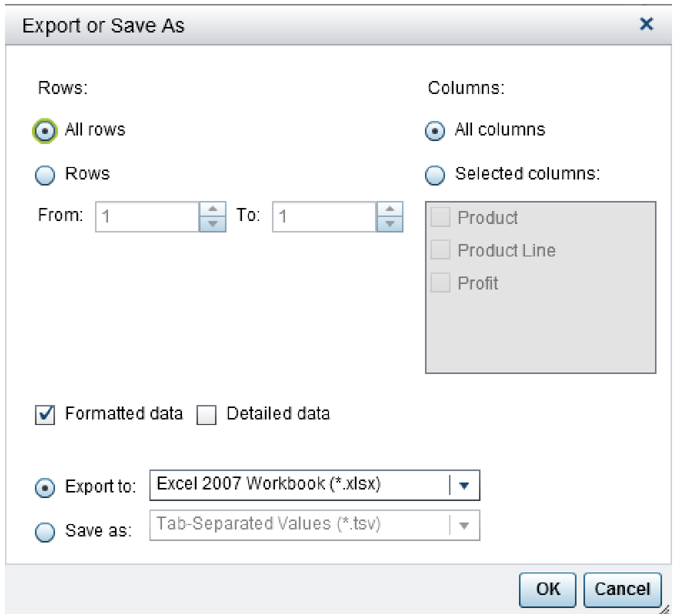
Export Data from a Graph

Note: You cannot export the actual visual graph. Only the data can be exported.

To export data from a graph:

1 If the graph that you want to export is not already selected, select it.

2 Right-click on the graph, and select **Export <graphName>**, where `<graphName>` is the name of the report object. The Export or Save As window is displayed.



3 Choose to export **All rows** or a range of **Rows**.

4 Choose to export **All columns** or **Selected columns**. If you choose **Selected columns**, select the check box (or check boxes) 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) Choose whether the exported data is formatted, either select or clear the **Formatted data** check box. This check box is selected by default.

6 (Optional) Choose whether detail data is exported for the graph. If you select **Detailed data**, then you can select the columns that you want to export.

Note: The **Detailed data** option is not available for imported report objects, custom graph objects that have multiple data definitions, or for any graphs in which detail data is not allowed.

7 If you select **Export to**, then the only option is **Excel 2007 Workbook (*.xlsx)** to create a Microsoft Excel spreadsheet.

8 If you select **Save as**, choose either **Tab-Separated Values (*.tsv)** or **Comma-Separated Values (*.csv)** to create a data file.

9 Click **OK**.

10 When you are prompted, choose either to open the file or to save it.

Add Comments to a Report

In the designer, you can add comments to the whole report, but you cannot add comments to the individual report objects.

TIP You must save a report before you can add comments.

Here are some key points about adding comments to a report:

- If comments have been added to an individual report object using the home page, the explorer, or the viewer, then those comments cannot be displayed or edited in the designer.
- If you have the Add or View Comments capability, then you can add or view comments. You can edit your own comments and respond to other comments.

Distributing Reports

Overview of Distributing Reports

If you have the Distribute Reports capability, then you can distribute reports using the designer. Distributing reports automates the process of delivering updated content to report users. You can distribute reports once or at recurring intervals, such as daily, multiple times daily, weekly, or monthly. In addition, distributing reports gives you the ability to create a time event to generate reports during non-peak hours.

CAUTION! Use only the designer to schedule and distribute SAS Visual Analytics reports. If you try to use the Schedule Manager plug-in to SAS Management Console and set an option that is not available in the designer, the report job might be incompatible with the designer.

When you distribute a report, you are considered the report job owner.

Distributing reports requires two steps. First, you define a *report job*, which contains a list of one or more reports, a list of one or more recipients, and an optional e-mail message. Both the report job owner and the recipient must be registered and have e-mail addresses stored in metadata because the report generates using the recipients' data access rights. Second, you create a *distribution*, which contains a report job and a time event. The time event specifies when and how often the distribution occurs. When a distribution runs, a history record is created. After you create a distribution, you can schedule, unschedule, run, delete, and show the history for your distribution.

Recipients receive an e-mail message with one or more reports attached as PDF files. The e-mail message contains a link to the report.

If you distribute a multi-section report that uses multiple data sources, and one of the data sources is not accessible to a recipient, then a warning message about the non-accessible data source is appended to the PDF.

TIP The print options that you save with a report are used when you distribute the report to other users.

TIP If you distribute reports to multiple locales or if you receive distributed reports, then you can specify your **User locale** in the global SAS preferences. For more information about setting global SAS preferences, see “Specifying Your Preferences” on page 9.

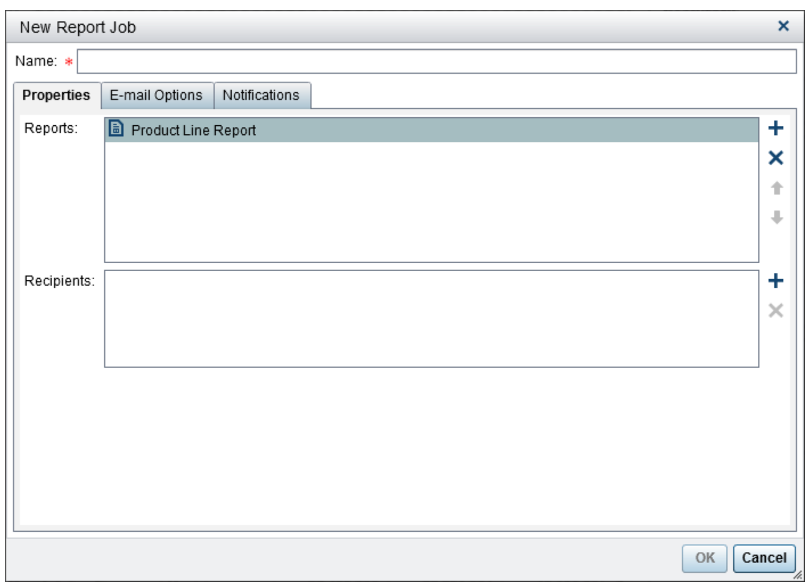
Add a New Report Job

To create a new report job:

1. Select **File** ▸ **Distribute Reports**. The Distribute Reports window is displayed.

   Note: Only your report jobs are displayed in the New Report Job window. If another user has already used the name that you enter for the report job, a warning message is displayed. You must enter a different name for your report job.

2. Click **New Report Job**. The New Report Job window is displayed.

3. Enter a **Name**. The report job name cannot contain blanks, white space, or these characters: `! " $ % & ' ( ) * . + , / : ; < = > ? @ [ \ ] ^ ` ` | ~ -`

   Note: Report jobs cannot be renamed.

4. On the **Properties** tab:

   - Click + beside the **Reports** field. Select one or more reports from the list that is displayed. When you select a report, it is added to the **Reports** list.

     TIP If the report that you want is not displayed in the pop-up list, click **Select another report**, and use the Open window to find the report that you want to distribute.

- Click + beside the **Recipients** field. The Add Recipients window is displayed. Select one or more users. Click **OK** to return to the New Report Job window.

  Note: A user's e-mail address must be stored in metadata to be displayed on the **View All** tab in the Add Recipients window.

5 On the **E-mail Options** tab:

- Enter a **Subject** for the e-mail message. If you leave this field blank, the name of the first report in the list is used as the subject.
- (Optional) Enter a **Message** for the body of the e-mail message.
- The **Show the report job owner in the message** option enables you to have your name appear in the e-mail message. Because the mail service cannot display the report job owner's name in the **From** field of the e-mail message, this is the only way to inform recipients who created the report job.

6 Use the **Notifications** tab to specify that you want to receive the report job owner notifications. You can select **Send e-mail notification when the report job starts**, **Send e-mail notification when the report job ends**, or both.

Note: If there is a warning (for example, when a recipient or a report cannot be found), then you automatically receive a notification with the details when the report job ends.

7 Click **OK**. The report job is displayed in the table in the Distribute Reports window.

Create a Report Distribution

To create a distribution:

1 Select **Distributions** in the left pane of the Distribute Reports window.

2 Select a report job, and click **Schedule**. The Schedule window is displayed.

3 Click **New** to create a time event. The Specify Schedule Details window is displayed.

4 Specify whether the distribution should happen **One time only** or **More than once**. By default, a new time event is defined to occur only once at a time that is approximately five minutes later than the time that you opened the Specify Schedule Details window.

If you select **One time only**, then select an **Hour** and a **Minute** for the distribution.

If you select **More than once**, then specify the details about when the distribution should occur. The specific fields that are available depend on the recurrence interval (**Hourly**, **Daily**, **Weekly**, **Monthly**, or **Yearly**) that you select.

- If you select **Hourly**, then the time is calculated from hour zero on a 24-hour clock. For example, if you leave **Interval hour** set to 1, then the data query runs at hour zero (midnight) and runs each hour.

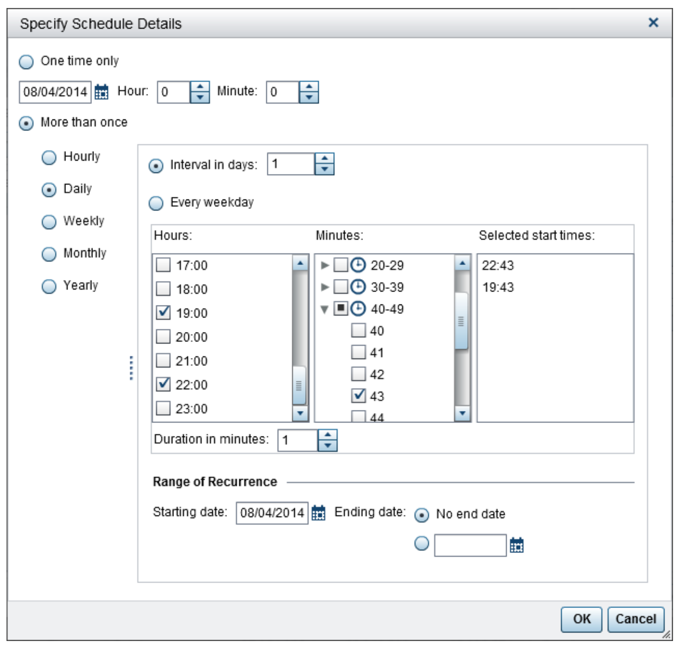
  Select an **Interval hour**, and then select a **Starting hour** and **Minute**.

The **Duration in minutes** field is used to specify the maximum number of minutes after the specified time has been reached that the dependency is kept in an open state. This value is used when a report job has multiple dependencies. It specifies a window of time in which distribution remains open so that other time events can be met. If the report job does not have multiple dependencies, then it is recommended to use the default duration of 1 minute.

If needed, specify the range of recurrence (the start date and end date for distribution). The default is to start at the current date and to not have an end date.

- If you select **Daily**, **Weekly**, **Monthly**, or **Yearly** for **More than once**, then you can select start times.

Here is an example of a **Daily** recurrence interval.



The following options are specific to the recurrence interval that you select:

Recurrence Interval	Options
Daily	Select a number for **Interval in days** or select **Every weekday** for the report job to run.
Weekly	Select a number for **Interval in weeks**, and then select the day (or days) that you want the report job to run.
Monthly	Select a number for **Interval in months**, and then select a **Day number** or a **Week** and **Day** that you want the report job to run.
Yearly	Select a number for **Interval in years** and the **Month**. Then, select a **Day number** or a **Week** and **Day** that you want the report job to run.

The following options are available for all recurrence intervals:

- Use the **Hours** and **Minutes** check boxes to select the times. The **Minutes** area contains groupings of 10-minute intervals. Selecting a check box for a minute grouping selects all of the minutes in that grouping. Your selections are displayed in the **Selected start times** area.

  TIP To select individual minutes, expand the grouping.

- The **Duration in minutes** field is used to specify the maximum number of minutes after the specified time has been reached that the dependency is kept in an open state. This value is used when a report job has multiple dependencies. It specifies a window of time in which the distribution remains open so that other events can be met. If the report job does not have multiple dependencies, then it is recommended to use the default duration of 1 minute.
- If needed, specify the range of recurrence (the start date and end date for distribution). The default is to start at the current date and to not have an end date.

For more information, see *Scheduling in SAS*.

5 Click **OK** to return to the Schedule window. The new schedule details are displayed under the **Schedule** heading.

6 Click **OK** to return to the Distribute Reports window. The new schedule details are displayed in the **Schedule** column.

7 (Optional) Click **Run Now** to have the distribution run immediately. Click **Close** in the information window.

8 Click **Close**.

Manage Report Distributions

The Distribute Reports window enables you to schedule, unschedule, run now, delete, and show history for distributions that you own. Only distributions that contain at least one report job that you own are displayed. (Some distributions might contain report jobs that other users own, but they are not displayed.)

TIP You can delete a distribution if it contains only report jobs that you own.

To see the history record for a selected distribution that has been run, click **Show History**. The Show History window is displayed. You can see the **Start Time**, **Finish Time**, **State**, and **Status Message**. Click **Close** to return to the Distribute Reports window.

Localize Reports

The designer enables you to localize (or translate) the labels, tooltips, and other descriptive text that is part of your report. You can export localizable text from your report so that you can apply one or more translations. You can translate your report into a new language without removing the text from the designer.

There is no limit to the number of translations that you can add to the report.

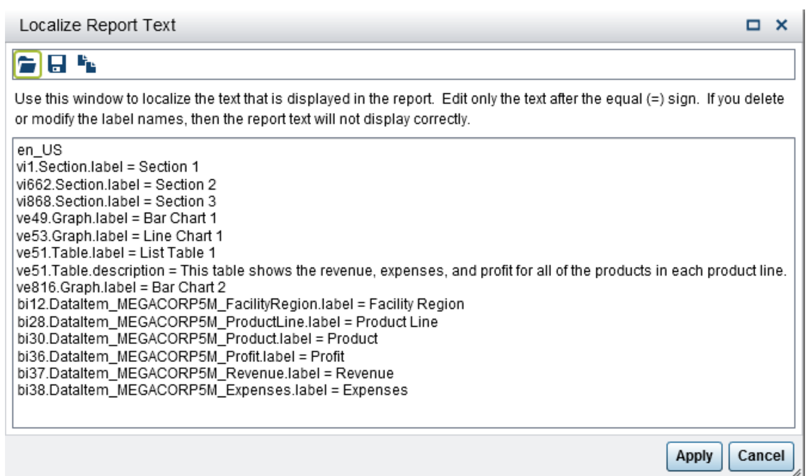
TIP If you localize a report and want to see the results immediately, it is recommended that you set the **User locale** setting or preference. For information about settings and the modern home page, refer to the online Help. For more information about preferences, see “Specify Global Preferences Using the Classic Home Page” on page 637.

To localize a report:

1 Select **File** ▸ **Save As** to save a copy of the report. This enables you to have a backup copy of your report.

2 Select **File** ▸ **Localize Report**. The Localize Report Text window is displayed. Identifiers, which provide context information, are displayed beside the original text. The text that is displayed after the equal sign (=) is the text that you can localize.

Note: You cannot change the original text in the report using the Localize Report Text window.



The first line in the Localize Report Text window contains the locale identifier, which is specified by the ISO 693 two-letter language code. If a language is created for a specific country or region, the language code can be followed by an underscore (_) and the ISO 693 two-letter country code. In the preceding screen capture, the `en_US` country code is displayed.

3 Localize the original text strings. You have the following options:

- Save the text in the Localize Report Text window to a file, modify the file, and then open the modified file in the Localize Report Text window.
- Modify the text in the Localize Report Text window.
- Copy and paste the text from the Localize Report Text window into an editor.

CAUTION! All report text is encoded as Unicode characters using the UTF-8 format. If you export the localizable text to an editor, make sure that the editor is capable of saving the data as UTF-8.

4 Click **Apply** to import the translations.

CAUTION! Your report is immediately updated and saved. You cannot revert to a previous version of the localized text or undo your changes.

The report text is displayed in the browser's locale or in the **User locale** that you specified for your global settings or preferences. For information about settings and the modern home page, refer to the online Help. For more information about preferences, see "Specify Global Preferences Using the Classic Home Page" on page 637.

For step-by-step instructions, examples, and a list of advanced features, see *One Report, Many Languages: Using SAS Visual Analytics to Localize Your Reports*, which is available on the SAS Visual Analytics page on the SAS support site.

Part 7

Creating Custom Graph Objects

52

Creating and Using Custom Graph Objects

About the Graph Builder

SAS Visual Analytics Graph Builder (the graph builder) enables you to create custom graph objects, which then become available in the SAS Visual Analytics Designer (the designer) for use in reports.

You can create custom graph objects for all graph types that are used in the designer. In your custom graph object, you can change the layout, add new graph elements, modify roles, change the visual properties, and so on. Your custom graph object is displayed under the **Custom** heading on the designer's **Objects** tab. For more information, see "Save a Custom Graph Object So It Appears in the Designer" on page 507.

You do not associate real data with the graph objects in the graph builder. Rather, you build graph objects or templates using sample data that is shipped with the graph builder. Report designers assign data when they include your graph objects in their reports.

The graph objects that you build have a consistent appearance that is compatible with the designer's graph objects. The graph builder enables you to produce a wide array of graph objects with more options for layouts and visual properties. Using the graph objects that you build, report designers can create simple or complex graphical views of their data.

You can do the following in the graph builder:

- drag and drop graph elements onto the canvas. Graph elements include the plots and charts that are available from the **Graph Elements** pane.

  For a definition and a picture of each available graph element, see "Graphs, Charts, and Plots" on page 555.

- use a gallery of predefined graph objects as templates from which you can build and customize custom graph objects. For more information, see "About the Graph Template Gallery" on page 496.

- combine and arrange the graph elements as needed. For example, you can layer multiple graph elements in a cell, create a lattice of cells in rows and columns, or combine these layouts.

  There are two types of lattice graphs:

  - A data-driven lattice is a multi-cell graph in which the cell data is determined by the values of one or more class variables. The number of cells is determined by the unique values of the class variables. Each cell of the lattice has the same graph elements. However, these graph elements have different values of the class variables.

  - A user-defined lattice is a multi-cell graph in which you create each cell independently. Each cell can contain different types of plots.

- determine and configure which data roles become available for assignment. For more information, see "Working with Roles" on page 498.

- specify the visual attributes of the graph elements. You can specify properties for the graph, the graph cells, for individual graph elements, for the axes, and for legends. For more information, see "Working with Properties" on page 517.

About the Graph Template Gallery

The graph builder includes a gallery that contains predefined graph objects. Using these predefined graph objects as templates, you can build and customize custom graph objects. The templates provide a useful starting point for creating interesting or complex graph objects.

> **TIP** All of the graph objects in the gallery are available as is for use in reports. However, these graph objects are not all displayed in the designer by default. Use the designer's Show or Hide Objects window to display the graph objects on the designer's **Objects** tab.

Here are the main features of the gallery:

- By default, the gallery is displayed when you start the graph builder. You can disable this feature by selecting the **Don't show this window at start-up** check box in the window. You can also change the default setting in your preferences. For more information, see "Specify Your Preferences for the Graph Builder" on page 508.
- If the gallery is not displayed, select **File** ▸ **New from Gallery** to display it.
- In the gallery, select a category in the left pane to see the available graph objects. You can then select the appropriate graph object from the right pane. For more information, see "Build a Custom Graph Object" on page 497.
- Some of the gallery graphs are created as data-driven lattices. When you select one of these graphs from the gallery, row and column lattice roles are automatically created. If you want to create a user-defined lattice from one of these gallery graphs, you must first delete the row and column lattice roles.

Build a Custom Graph Object

To build a custom graph object:

1. Do either of the following:
   - Drag and drop a graph element from the **Graph Elements** pane onto the blank canvas.
   - Select a graph object from the gallery.
     1. If the gallery is not displayed, select **File** ▸ **New from Gallery** to display it.
     2. Select a category in the left pane to see the available graph objects. The **All** category shows all of the graph objects that are in the gallery.
     3. When you find the graph object that you want, select the object, and click **OK**.

     Note: Some of the gallery graphs are created as data-driven lattices. For more information, see "About the Graph Template Gallery" on page 496.

   Sample data is used in the graph object. The report designer assigns real data when he or she includes the graph object in a report.
2. (Optional) On the **Role Definitions** tab, you can define roles for the graph object.

   You have the following options:
   - Provide more descriptive names for the default roles.
   - Click **Add Role** to add more roles.
   - Create a data-driven lattice by adding a lattice role. For more information, see "Creating a Data-Driven Lattice " on page 505.
3. (Optional) You can add additional graph elements to the graph object. See "Adding a Graph Element to an Existing Graph Object" on page 500.

4 (Optional) You can change the properties for the graph object, the cell, the graph element, the legend, or one or more axes. Select the item, and make your changes on the **Properties** tab.

5 Save the graph object. Depending on where you save the graph object, it might be immediately displayed on the **Objects** tab in SAS Visual Analytics Designer (the designer). For more information, see "Save a Custom Graph Object So It Appears in the Designer" on page 507.

Working with Roles

About Roles

When you create a custom graph object, you determine which data roles become available for assignment. However, you do not assign data to those roles in the graph builder. When the graph object becomes available to report designers for use in their reports, the report designers assign data items to the available roles.

Create Optional Roles

When you drag and drop a graph element onto the canvas, the required roles for that graph element are created on the **Role Definitions** tab. You can create additional, optional roles.

To create an optional role:

1 On the **Role Definitions** tab, click **Add Role**. The Add Role window is displayed. The contents of the Add Role window vary depending on the current graph element.

2 Select the type of role that you want to create. For more information, see "Types of Roles" on page 499.

3 (Optional) You can change the default role name to provide a more meaningful name.

4 Select the data type, or classification, for the role. For information about the data types, see "Working with Data Items in a Report" on page 363.

5 (Optional) Select **Required** to require report designers to assign data to the role. If this option is not selected, then report designers can choose whether to assign data to the role.

6 If the **Allow multiple data assignments** check box is available, then you can select or clear the check box. When this option is selected, the role can have more than one measure assigned. For example, in a line chart, if multiple measures are assigned to the Y axis, then the graph shows a line for each measure.

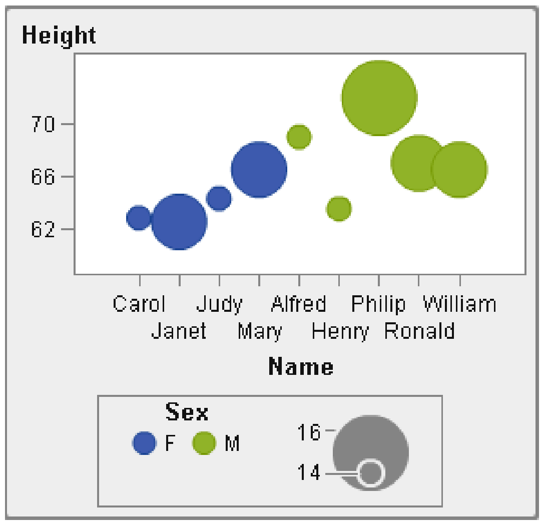
7 Click **OK**.

Types of Roles

In addition to the standard roles, such as Time, X, or Category, you can add the following roles by clicking **Add Role** on the **Role Definitions** tab. The roles that are available depend on the type of graph element.

- The Color role specifies a data column that is used to color the data.

  The following example shows the heights of a group of students. A student's gender is specified for the Color role.

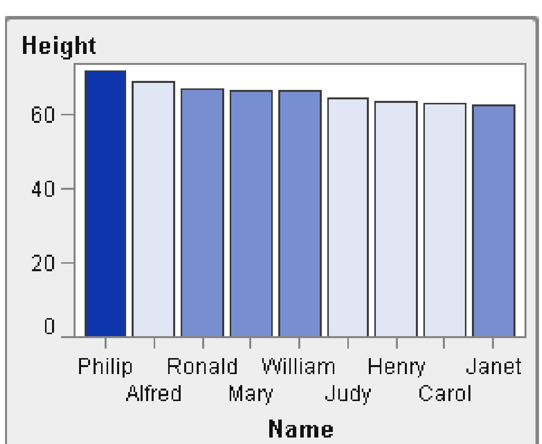


The Color role can be any data type. In this example, the Color role is assigned a Category data type.

Note: When you add the Color role to a graph element in a data-driven lattice, the Measure data type is not always available. The Measure data type is not available for a bar chart, a waterfall chart, or a scatter plot.

In some graph elements, such as a bar chart, the Group role also colors the graph based on the values of the group variable. In those graph elements, the Color role is used to color the data across a color gradient as shown in the following example. In this case, the Color role must be a Measure data type.

In this example, the color gradient indicates the age range of the students.



- The Group role specifies a data column that is used to group the data. The graph elements for each unique group value are automatically distinguished by different visual attributes, such as color, line style, and so on. The Group role can be a Category data type or a Datetime data type.
- The Data Label role specifies a data column that is used for data labels.
- The Data Tip role specifies a data column that is used for data tips. The Data Tip role must be numeric.
- The Target role specifies a data column that is used to generate a targeted bar chart. A targeted bar chart compares data that is summarized by the value of a category in reference to target values.

Adding a Graph Element to an Existing Graph Object

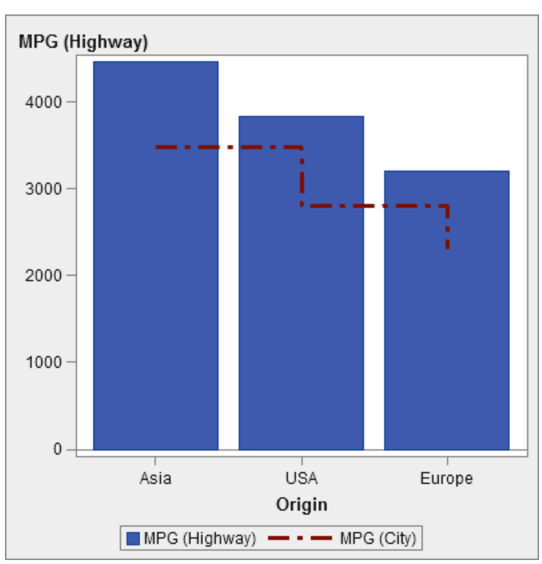
You have two options for adding a graph element to an existing graph object. You can overlay the new graph element on top of an existing graph element. Or, you can add the graph element so that it creates a cell in a new row or column.

Add an Overlay

In an overlay, one graph element is superimposed on another graph element. You can add an overlaid graph element to a single-cell graph as well as to a lattice.

In this example, a step plot is overlaid on a bar chart. The example shows the graph element with data assigned to it.

Figure 52.1 Step Plot Overlaid on a Bar Chart



To create an overlay:

1 Drag and drop a graph element from the **Graph Elements** pane onto the canvas.

2 Drag and drop a second graph element directly onto the existing graph element.

   If the graph elements are compatible, they are automatically assigned a shared role. This enables both graph elements to have a common X axis role. You can unshare the shared role if needed.

   Some graph elements, such as the pie chart and the treemap, cannot be layered in an overlay. A message is displayed when you try to drag and drop these types of graph elements.

   For more information, see the following topics:

   - "Incompatible Graph Elements" on page 503
   - "Sharing Data Roles" on page 520

3 (Optional) On the **Role Definitions** tab, you can change the default role names to provide more meaningful names.

> **TIP** The graph elements are layered in the order in which they were added to the canvas. You can change the order by moving a graph element to the front or the back. For more information, see "Change the Order of Graph Elements in a Cell" on page 513.

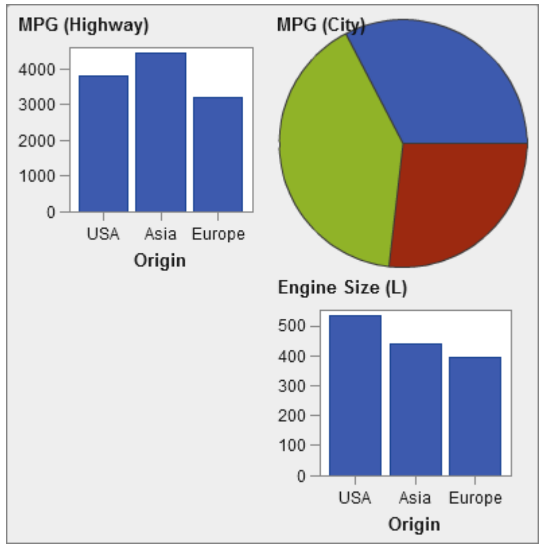
Add a Graph Element to a New Row or Column

A graph element can be added so that it creates a cell in a new row or column. This arrangement results in a graph object with more than one cell. It is called a user-defined lattice. In a user-defined lattice, each cell is created independently and can contain different types of graph elements. Cells are aligned in rows and columns that you create by dragging and dropping graph elements.

Note: You cannot create new rows or columns in a data-driven lattice. For a description of a data-driven lattice, see "About Data-Driven Lattices " on page 505.

This example shows a user-defined lattice with four cells. One of the cells is empty. The example shows the graph element with data assigned to it.

Figure 52.2 User-Defined Lattice

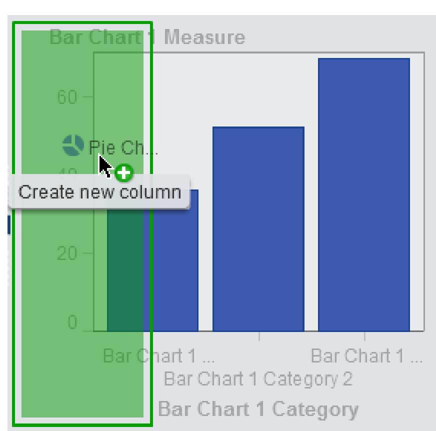


To create a user-defined lattice:

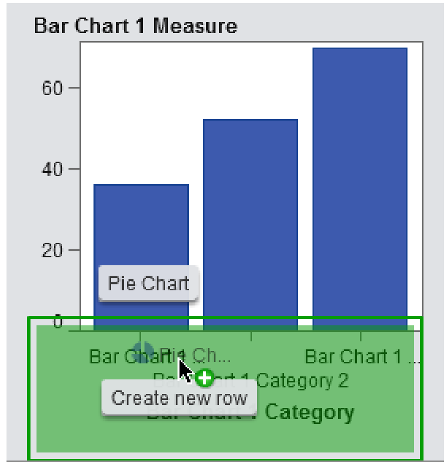
1 Drag and drop a graph element from the **Graph Elements** pane onto the canvas. The graph element is created in the center of the graph area of the canvas.

2 Drag and drop another graph element onto the canvas as follows:

- Drag and drop the graph element onto the right or left border of the existing graph to create a new column. The area becomes shaded to indicate the location of the new column.



- Drag and drop the graph element onto the top or bottom border of the existing graph to create a new row.



3 To add more cells, continue dragging and dropping graph elements onto the borders of the existing graph to create new columns or rows. You can drag and drop a graph element onto the area between two columns to insert a new column. Or, you can drag and drop a graph element onto the area between two rows to insert a new row. If the existing graph contains a lattice with an empty cell, you can drag and drop the graph element into that empty cell.

   You can create up to 10 rows and 10 columns.

4 (Optional) On the **Role Definitions** tab, you can change the default role names to provide more meaningful names.

Incompatible Graph Elements

Some graph elements can be layered in an overlay with one graph element superimposed on another. Other graph elements, however, are incompatible. They cannot be layered in the same cell. The following table shows which graph elements cannot be layered in the same cell.

Note: Pie charts and treemaps are not compatible with any other graph elements. They are not listed in the table.

Table 52.1 *Graph Element Type Incompatibility*

Graph Element	Not Compatible
Band plot	Bar chart, horizontal Schedule chart

Graph Element	Not Compatible
Bar chart, horizontal	Band plot Bar chart, vertical Line chart Needle plot Schedule chart Vector plot Waterfall chart
Bar chart, vertical	Bar chart, horizontal Schedule chart Vector plot
Bubble plot	(Compatible with all graph elements except pie charts and treemaps)
Line chart	Bar chart, horizontal Schedule chart Vector plot
Needle plot	Bar chart, horizontal Schedule chart
Scatter plot	(Compatible with all graph elements except pie charts and treemaps)
Schedule chart	Band plot Bar chart, horizontal Bar chart, vertical Line chart Needle plot Step plot Time series plot Vector plot Waterfall chart
Series plot	(Compatible with all graph elements except pie charts and treemaps)
Step plot	Schedule chart
Time series plot	Schedule chart
Vector plot	Bar chart, horizontal Bar chart, vertical Line chart Schedule chart Waterfall chart

Graph Element	Not Compatible
Waterfall chart*	Bar chart, horizontal Schedule chart Vector plot Waterfall chart

* The waterfall chart requires the X axis to have a shared role to avoid unpredictable results.

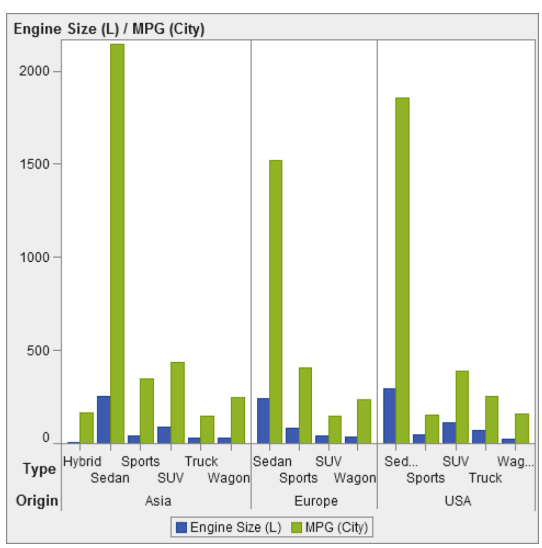
Creating a Data-Driven Lattice

About Data-Driven Lattices

A data-driven lattice is a multi-cell graph in which the cell data is determined by the values of one or more class variables. Class variables are specified in the designer when the report designer creates a report that includes the lattice. The number of cells is determined by the unique values of the class variables. Each cell of the lattice has the same graph elements. However, these graph elements have different values of the class variables.

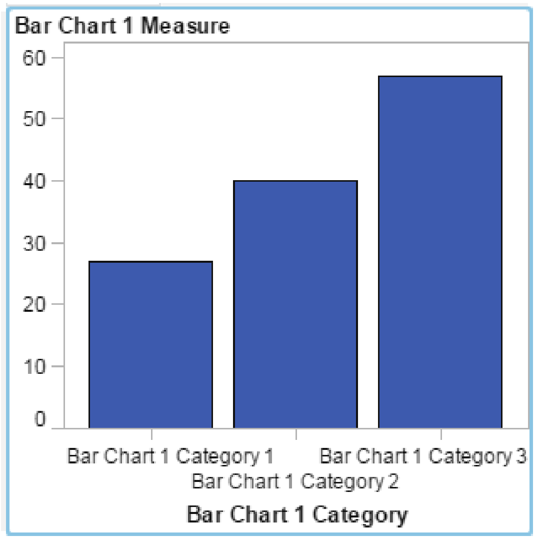
The following example shows a graph with data assigned to it. The custom graph object includes a data-driven lattice with a columnar role. The graph object allows multiple data assignments for the Measure role.

Figure 52.3 Data-Driven Lattice Multi-Cell Graph



In the report, the graph object includes a three-cell lattice of bar charts showing engine size and miles per gallon for different types of automobiles. The country of origin is specified as the class variable. As a result, a cell is created for each country of origin.

The following example shows the graph object that was created in the graph builder. The lattice role was specified for columns.



Here are the characteristics of data-driven lattices:

- You can define up to two lattice roles: one role for the rows and another role for the columns. When the report designer uses the lattice in a report, the class variables are assigned to the lattice roles. All class variables must come from a single data source. The number of rows and columns in the graph is determined by the unique values of the class variables.
- If you specify that the role is required, then a report designer must assign data to the role.

  If this option is not selected, then report designers can choose whether to assign data to the role. In that case, it is possible for a graph object to have no lattice.
- Graph elements can be overlaid the same way they are overlaid in single-cell or user-defined multi-cell graphs. When the report designer adds a lattice to a report, the overlaid graph element is displayed in every cell.
- Properties that you specify for graph elements are applied to all of the cells of the graph object.

Create a Data-Driven Lattice

To create a data-driven lattice:

1. Create the graph object that you want as the basis for your lattice. For more information, see “Build a Custom Graph Object” on page 497.
2. (Optional) On the **Role Definitions** tab, you can change the default role names to provide more meaningful names.
3. Specify the role for a class variable. In the **Data-Driven Lattice** section of the **Role Definitions** tab, click **Add Lattice Role**.

   The Add Role window is displayed.

The **Data-Driven Lattice** section is not available in either of the following cases:

- the graph object already contains more than one cell
- a Color role with a Measure data type has been added to the bar chart, waterfall chart, or scatter plot

In addition, data-driven lattices do not allow graph elements to be overlaid unless their aggregations are compatible. Graph elements with compatible aggregations either perform no aggregation at all or they use the same set of shared Category roles.

4 For the **Role Type**, select **Row** or **Column** to specify the layout for the cells.

5 Specify a role name or use the default name.

6 Select the data type, or classification, for the role. For information about the data types, see "Working with Data Items in a Report" on page 363.

7 (Optional) Select **Required** to require report designers to assign data to the role. If this option is not selected, then report designers can choose whether to assign data to the role. In that case, it is possible for a graph to have no lattice.

8 If the **Allow multiple data assignments** check box is available, then you can select or clear the check box. When this option is selected, the role can have more than one measure assigned. For example, in a line chart, if multiple columns are assigned to the Y axis, then the graph shows a line for each measure.

9 Click **OK**.

Repeat the previous steps to specify a second role. For example, if you specified a role for a multi-row layout, you can specify an additional role for columns.

Save a Custom Graph Object So It Appears in the Designer

Custom graph objects are saved when you select either **File** ▸ **Save** or **File** ▸ **Save As**. Graph objects that are saved in the `My Folder` location are automatically displayed in the list of **Custom** objects on the designer's **Objects** tab.

Graph objects that you save in public folders are not automatically displayed on the **Objects** tab. However, you can add the graph objects manually. For more information, see "Add a Graph Object to the Designer's Objects Tab" on page 508.

Add a Graph Object to the Designer's Objects Tab

Custom graph objects that are saved in the `My Folder` location are automatically added to the designer's **Objects** tab. Custom graph objects that you save in public folders are not automatically displayed on the **Objects** tab. However, you can add the graph objects manually.

TIP All of the graph objects in the gallery are available in the designer. However, by default, not all of these graph objects are displayed on the designer's **Objects** tab. Use the designer's Show or Hide Objects window to display any of these graph objects on the designer's **Objects** tab.

To add a graph object to the designer's **Objects** tab:

1. On the **Objects** tab in the designer, click , and select **Show or Hide Objects**. The Show or Hide Objects window is displayed.
2. Click **Select Custom**. The Choose an Item window is displayed.
3. Navigate to the graph object that you want to add.
4. (Optional) To search for a graph object, do the following:
   a. Click in the toolbar. The Search window is displayed.
   b. Enter the name of the graph object that you want to find.
   c. Click **Search**.

   TIP As an alternative, you can search for all graph objects. Specify **Graph template** for **Type**.
5. Select the graph object, and click **OK**.

   The selected graph object is now displayed with a check mark in the Show or Hide Objects window.
6. Click **OK**. The **Objects** tab is updated.

Specify Your Preferences for the Graph Builder

To specify preferences that are specific to the graph builder:

1. In the graph builder, select **File** ▸ **Preferences** to open the Preferences window.
2. Select **SAS Visual Analytics Graph Builder**.

3 Select or deselect **Show template gallery at start-up**. The gallery contains predefined graph objects as templates from which you can build and customize custom graph objects.

4 Click **OK**.

To specify global SAS preferences, see “Specifying Your Preferences” on page 9. To specify general preferences for SAS Visual Analytics, see “Specify Settings Using SAS Home” on page 10.

53

Modifying Custom Graph Objects

Change a Custom Graph That Has Been Saved

When you save a custom graph object, depending on where you save it, it might be immediately displayed on the **Objects** tab in SAS Visual Analytics Designer (the designer). For more information, see “Save a Custom Graph Object So It Appears in the Designer” on page 507.

To change a graph that has been saved, do either of the following:

- Right-click the graph object in the designer, and select **Edit Custom Graph**.
- Select **File** ▶ **Open**, and then locate and open the file for your graph object.

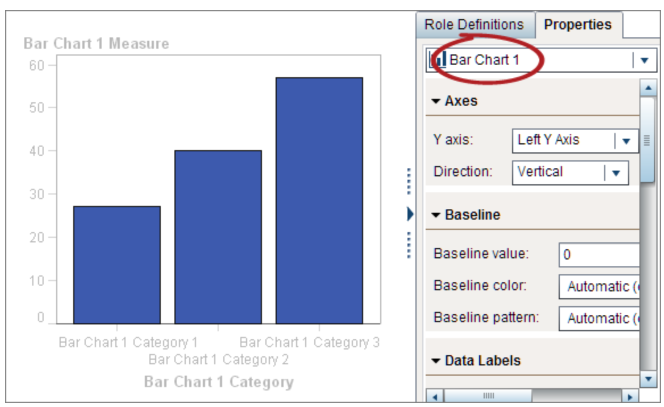
Note: Any changes that you make to the graph object do not appear in existing reports that were created from the original graph object. However, new reports that are created after a modified graph object has been saved reflect the changes to the graph object.

Select Components and Elements of a Graph Object

To change the properties of a graph element such as a bar chart, you must first select the graph element. Similarly, to change the properties of a legend or an axis, you must first select the component.

When you select a component in a graph object, the selected component is highlighted. The component is also displayed in the drop-down list at the top of the **Properties** tab. The available properties vary depending on the selected component.

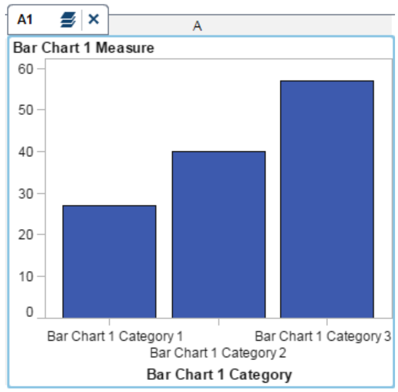
The following display shows a highlighted bar chart. The **Properties** tab displays the properties for the bar chart, which is identified as **Bar Chart 1**.



TIP You can click ▼ beside the drop-down list on the **Properties** tab, and then select the component that you want to modify.

Here are some suggestions for selecting graph elements:

- To select a line chart or a time series plot, click a data point on a line.
- To select the entire graph object, click above the graph element or cell area.
- To select a cell in a multi-cell graph object, click the background area just above the graph element. The cell is highlighted, and a tab appears in the upper left corner of the cell. From the cell's tab, you can delete and reorder graph elements. You can move graph elements to other cells.



In a multi-cell graph object, cells are identified by a combination of a letter and a number, such as A1, B1, A2, and so on. The letters identify columns and the numbers identify rows. The following table represents the labeling for a multi-cell graph object that contains three columns and three rows:

A1	B1	C1
A2	B2	C2
A3	B3	C3

Change the Order of Graph Elements in a Cell

When you overlay a graph element on top of another, the graph elements are layered with the last one added on top. For example, suppose that you create a graph object with a line chart, and then you add a bar chart. Because the bar chart was added last, it appears in front of the line chart. Depending on the data that is assigned to the graph elements in the designer, the bar chart might obscure part of the line chart. You can reorder the graph elements so that the line chart appears in front of the bar chart.

To change the order of the graph elements in a cell:

1 From the drop-down list at the top of the **Properties** tab, select the cell that you want to change. Cells are identified by a combination of a letter and a number, such as A1, B1, A2, and so on. For more information, see "Select Components and Elements of a Graph Object" on page 512.

   In the **Graph Elements** area on the **Properties** tab, the graph elements in the cell are listed in the order in which they were added to the cell.

2 Drag a graph element to a new position.

Remove a Graph Element

To remove a graph element from a graph object, right-click the graph element, and select **Remove `graph element name`**.

Note: In a user-defined lattice, you can remove graph elements using the cell's tab. For more information, see "Remove a Cell or a Graph Element in the Cell" on page 516.

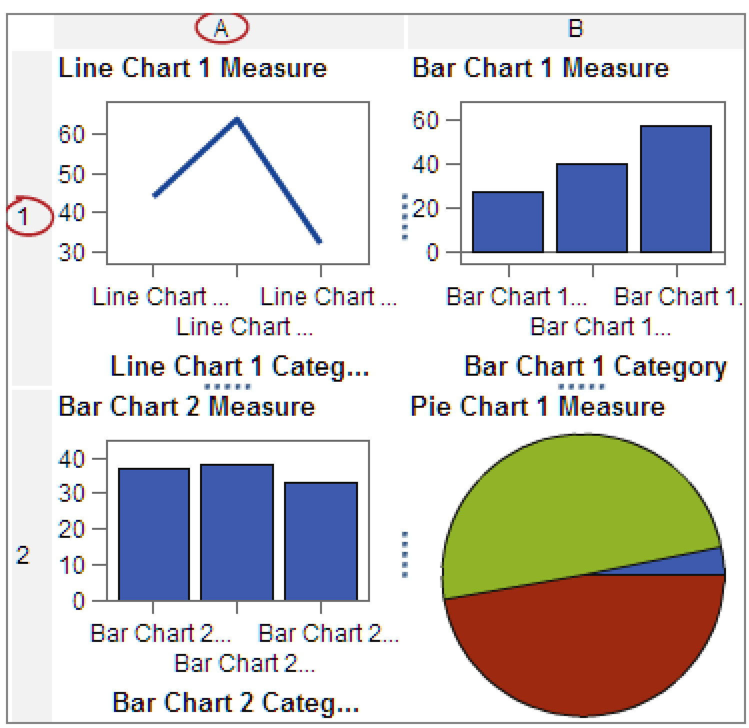
Working with User-Defined Lattices

About User-Defined Lattices

A user-defined lattice is a multi-cell graph in which each cell is created independently. And, each cell can contain different types of graph elements. Cells are aligned in rows and columns that you create by dragging and dropping graph elements.

The following example shows four cells arranged in two rows and two columns. Letters identify the columns and numbers identify the rows.

Figure 53.1 User-Defined Lattice



When you select a cell, a tab appears in the upper left corner of the cell. From the cell's tab, you can delete the cell and its contents by clicking ✕. If you click

, the graph elements in the cell are listed. From this list, you can delete and reorder the graph elements. You can move graph elements to other cells.

For more information about selecting cells and how they are labeled in the graph, see "Select Components and Elements of a Graph Object" on page 512.

Move a Graph Element from One Cell to Another Cell

To move a graph element to a different cell:

1. Select the cell whose graph element you want to move. For more information, see "Select Components and Elements of a Graph Object" on page 512.
2. In the cell's tab, click to list the graph elements in the cell.
3. Drag and drop a graph element from one cell into the target cell.

   You can also drag and drop the graph element so that it creates a new row or new column. For more information, see "Add a Graph Element to a New Row or Column" on page 501.

Resize a Row or a Column

You can change the width of a column and the height of a row in a user-defined lattice. For example, you might want a particular column to be wider than the others.

To resize a row:

1. Select between the row that you want to change and an adjacent row.
2. Drag the handle upward or downward to change the row height.

To resize a column:

1. Select between the column that you want to change and an adjacent column.
2. Drag the handle left or right to change the column width.

Add a Row or a Column to a Lattice

To add a new row or column to a lattice, drag and drop a graph element to the location where you want the new row or column. For more information, see "Add a Graph Element to a New Row or Column" on page 501.

You can also move a graph element from a row or column to another row or column. For more information, see "Move a Graph Element from One Cell to Another Cell" on page 515.

Remove a Cell or a Graph Element in the Cell

You can remove a graph element from the cell. You can also remove the cell itself.

1 Select the cell.

2 To remove a graph element from the cell:

   a Click to list the graph elements in the cell.

   b Click next to the graph element.

   You can also right-click the graph element, and select **Delete**.

3 To remove the cell, click on the cell's tab.

   The cell and its contents are removed. If the cell is the only cell in a row or column, then the respective row or column is removed as well.

Move or Delete a Row or a Column

In a graph, a heading contains a letter that identifies the column and a number that identifies the row.

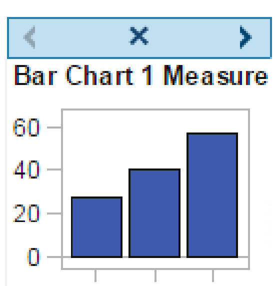
From the headings, you can move and delete columns and rows. When you move a column or row, the column or row switches places with the adjacent column or row. For example, if you move column B to the left, column B switches places with column A.

To more or delete a column or row:

1 Select the graph.

2 Position the cursor over a heading. A button and arrows are displayed on the heading.

   The following example shows a column heading that contains a button to delete the column and arrows to move the column. The button to move the column to the left is unavailable.



3 To delete the column or row, click .

4 To move the column or row, click an arrow.

   - Click or to move a column left or right.
   - Click or to move a row up or down.

Some arrows might not be available. For example, the leftmost column can be moved only to the right, as is the case in the previous example. Only › is available for the column.

Working with Properties

About Properties

The **Properties** tab lists the properties for the selected graph component. Components that have properties include graph elements, cells, axes, legends, and the entire graph. Properties enable you to control the overall appearance of your graph, from the graph's background color to the tick value interval on the axis. Report designers can override some properties in their report.

To change the properties for a graph object, cell, graph element, legend, or one or more axes, select the item, and make your changes on the **Properties** tab.

See Also

"Select Components and Elements of a Graph Object" on page 512

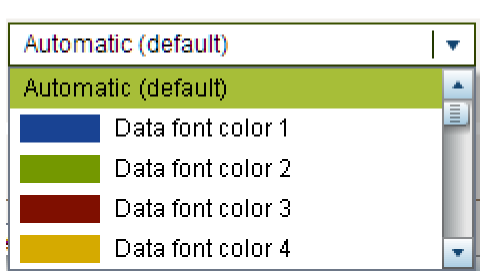
About Style Properties

The color schemes and visual attributes for a graph come from the active report theme, which is a named collection of style elements. Each style element contains attributes such as fill color, marker symbol, line style, font face, and many others. Each part of a graph element, such as a marker, a bar, a line, or a label, derives its visual attributes from a specific style element from the active report theme. Similarly, the attributes for axes and legends come from a style element.

The style elements are developed to produce effective graphics without changing the default settings. However, you can change the default settings by changing style properties on the **Properties** tab. When you specify a style property, such as a color or font style, you are changing the style element that is applied to the graph component.

The following display shows the available style elements for bar chart labels.

The list of style elements is displayed when you click ▾ beside the **Data label font color** drop-down list on the **Properties** tab.



Each color is shown in the list beside its style element. **Automatic** indicates the default value for the graph element. **Automatic** enables the graph element to cycle through a list of colors provided for group colors and multi-response colors.

Visual attributes that are referenced by a style element provide consistency within the report theme. If you later change the report theme, the graph element is compatible with the new theme. In addition, some of the properties that you specify can be overridden by report designers when they use your graph objects in their reports.

Create an Overview Axis

The overview axes that you create in the graph builder are similar to the overview axes that you create in the designer. In addition, you can do the following:

- select specific graph elements from the main graph to show in the overview axis
- select completely different graph elements from the main graph to show in the overview axis

To create an overview axis:

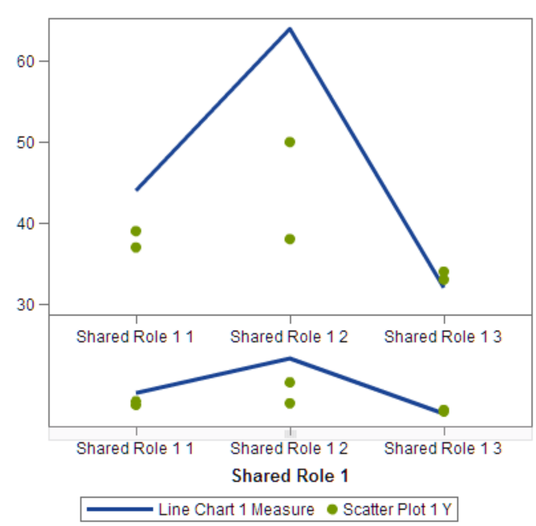
1. Drag and drop a graph element from the **Graph Elements** pane onto the canvas.

   You can drag and drop another graph element to create an overlay. However, do not create a new cell for this additional graph element. You cannot create an overview axis when the graph object contains more than one cell.

2. Select the graph object.

3. On the **Properties** tab, select the **Show overview axis** check box.

   The graph object displays the overview axis beneath the existing graph elements. By default, the existing graph elements are displayed in the overview axis, as shown in the following example:



4. To display different graph elements in the overview axis, do the following:

   a. On the **Properties** tab, select **Specify graph elements** from **Contents**.

   b. Drag and drop one or more graph elements from the **Graph Elements** pane onto the overview axis area of the canvas.

The graph element names that you added are displayed in the **Graph elements** section on the **Properties** tab.

c If you added more than one graph element to the overview axis area, you can change the order of the graph elements by dragging and dropping a graph element to a different position.

Change Attribute Rotation

When you apply a group role to your graph element, by default, the graph builder rotates through predefined style elements to determine how to present each unique group value. This type of rotation also occurs when multiple response roles are assigned to the graph element or when graph elements are overlaid.

When the graph element is rendered, the attributes for colors, line patterns, and marker symbols are rotated for the values in your group. Whether you change the attributes or keep the default attribute values, the manner in which the attributes are combined is determined by the attribute rotation priority.

To change the attribute rotation priority:

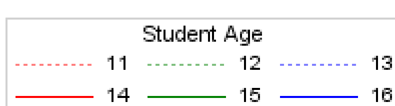
1 Select the graph object. For more information, see "Select Components and Elements of a Graph Object" on page 512.

2 On the **Properties** tab, select an option from **Attribute rotation priority**.

You have two options for the attribute rotation priority:

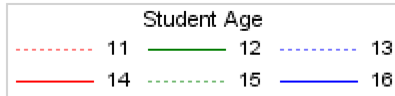
Rotate color only until all colors used	attributes such as marker symbols and line patterns are held constant while each color in the list is applied exhaustively to the graph element.
Rotate all attributes	each attribute rotates through its own list to generate a unique combination for each group value. Attributes include colors, marker symbols, line patterns, and so on.

For example, the following figure shows an example rotation for an age grouping. The dotted-line pattern is held constant while red, green, and blue colors are applied to the dotted lines for the consecutive age group values. If there are more group values, red, green, and blue colors are applied to the solid lines.



If you change the rotation pattern, and specify **Rotate all attributes** for the attribute rotation priority, the contrasting colors and line patterns are rotated at the same time.

A red dotted-line pattern is applied to the first group crossing. Then, a green solid-line pattern is applied to the second crossing, a blue dotted-line pattern is applied to the third crossing, and so on.



Sharing Data Roles

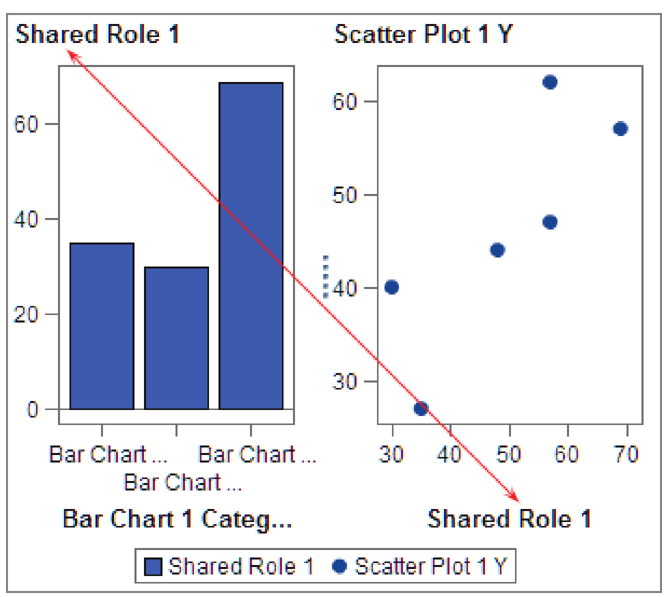
About Shared Data Roles

If your graph object contains multiple graph elements, you can specify that some data roles be shared across two or more graph elements. You share data roles when you want the graph elements to use the same data role for either of their axes. When report designers who use the graph object in their reports assign roles, a single data column is assigned to all roles that use the shared role.

Note: If the graph element types are compatible, overlaid graph elements are automatically assigned a shared role. (See “Incompatible Graph Elements” on page 503.) You can unshare a role if needed.

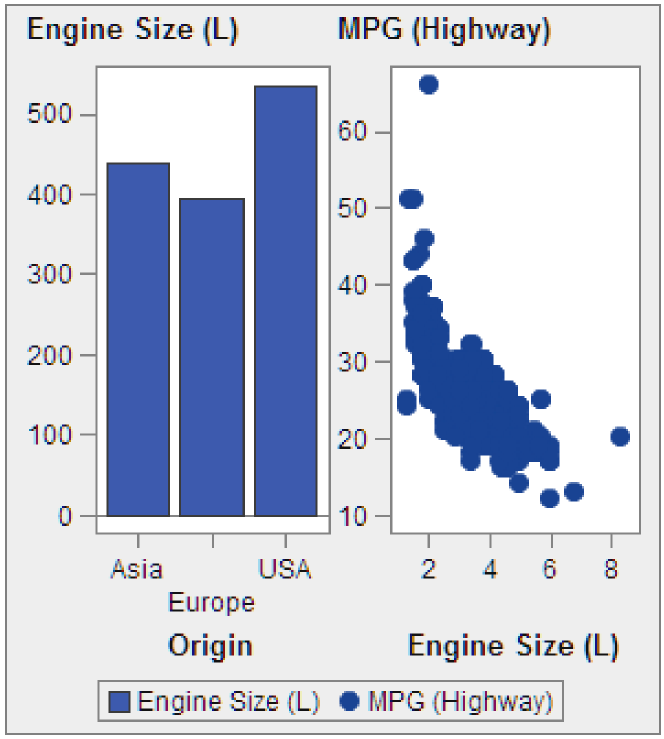
Shared roles are very useful in user-defined lattices. They enable you to manage data roles across columns and rows. For example, you might share data roles if you want to analyze relationships in your data.

In this example, the X axis of the scatter plot shares a role with the Measure axis of the bar chart.



When the graph object is used in the designer, the data might be applied as shown in the following example.

The bar chart shows engine sizes for different regions of the world. (The engine size has been changed from a Sum to an Average measure.) The scatter plot shows the miles-per-gallon trend for the different engine sizes.



Share or Unshare a Data Role

To share a data role, on the **Role Definitions** tab, click ▾ next to the role that you want to share. Select **Create Shared Role With Another Role** and then select the name of a role.

You can also click **Add Shared Role**, and then create and share the data role.

To unshare a data role, click ▾ next to the shared role, and select **Unshare**.

54

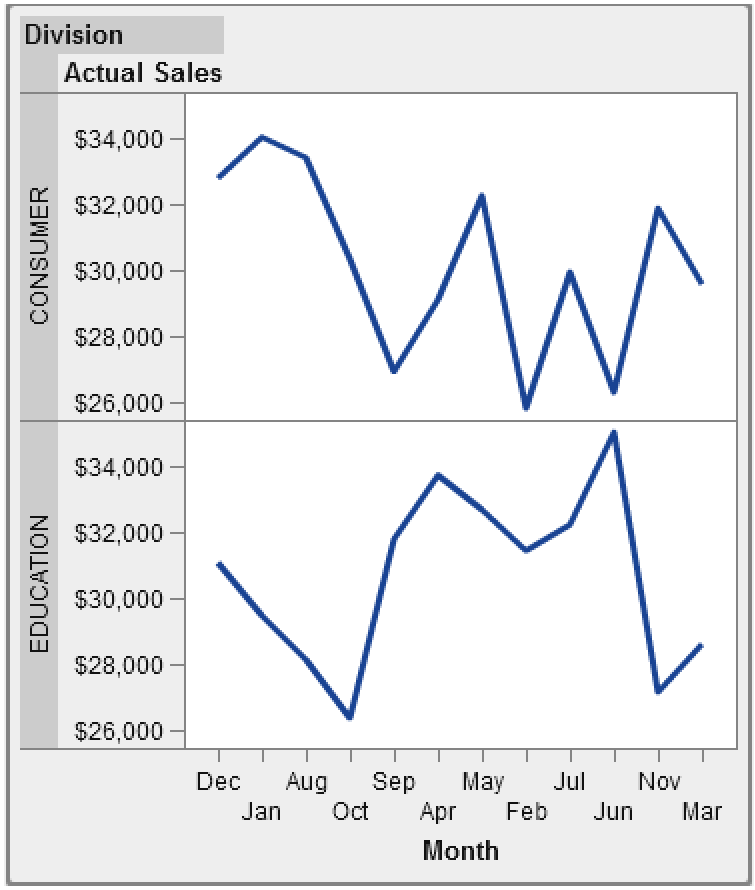
Graph Builder Examples

Example: Data-Driven Lattice

About the Data-Driven Lattice Example

This example shows the sales for a line of retail products over a period of time. The custom graph object specifies a row-based lattice in which the rows correspond to the corporate divisions that are responsible for the sales. The corporation has two divisions, resulting in a graph with two rows.

Figure 54.1 Example Data-Driven Lattice



Build the Graph Object for the Data-Driven Lattice Example

1. In the graph builder, drag and drop a line chart from the **Graph Elements** pane onto the canvas.
2. On the **Role Definitions** tab, click ▾ next to **Category**, and select **Edit Role**. The Edit Role window is displayed.
3. Select **Datetime** for **Classification**.

   Click **OK**.

   Specifying **Datetime** forces the report designer to assign date and time data to the Category role.
4. Specify the role to which a class variable is assigned.
   a. In the **Data-Driven Lattice** section on the **Role Definitions** tab, click **Add Lattice Role**. The Add Role window is displayed.
   b. Select **Row** for **Role Type**.
   c. Select **Category** for **Classification**.

d Select **Required**. With this option selected, the report designer must assign data to this role to render the report with data assigned.

e Click **OK**.

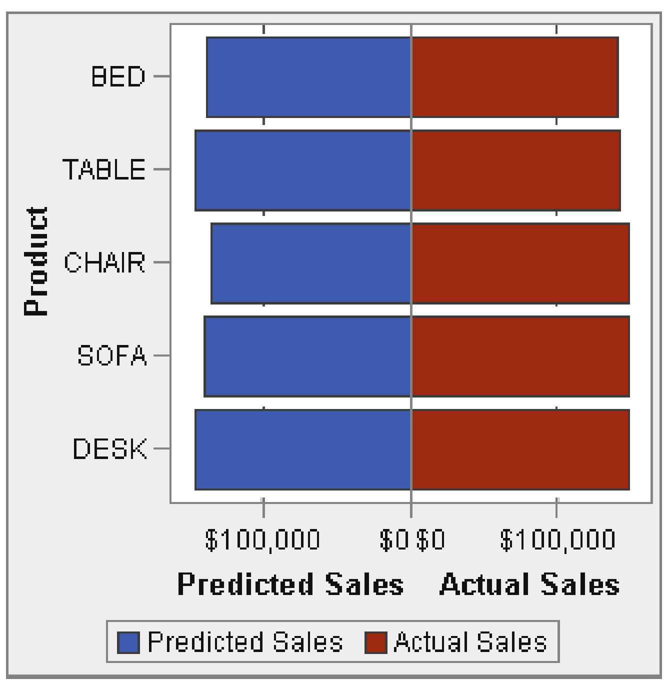
5 Save the graph object. See “Save a Custom Graph Object So It Appears in the Designer” on page 507.

Example: User-Defined Lattice (Butterfly Chart)

About the User-Defined Lattice Example

This example uses a butterfly chart to show the actual sales compared to the predicted sales for a line of retail products. The butterfly chart is useful for comparing two unique values. In this chart, the two values are arranged on each side of the Y axis.

Figure 54.2 *Example User-Defined Lattice*



Build the Graph Object for the User-Defined Lattice Example

TIP As a shortcut, you can select a butterfly chart from the graph gallery.

1 In the graph builder, drag and drop a bar chart from the **Graph Elements** pane onto the canvas.

2 Drag and drop a second bar chart from the **Graph Elements** pane onto the left edge of the canvas. This action creates a new column for the second bar chart.

3 Share the category roles. On the **Role Definitions** tab, click ▼ next to the **Category** for either bar chart. Select **Create Shared Role With Another Role** and then select the other category role.

The Add Shared Role window is displayed.

Click **OK**.

4 Change both bar charts to a horizontal layout.

a Select a bar chart.

b On the **Properties** tab, click ▼ next to **Direction**, and select **Horizontal**.

c Repeat the previous two steps for the other bar chart.

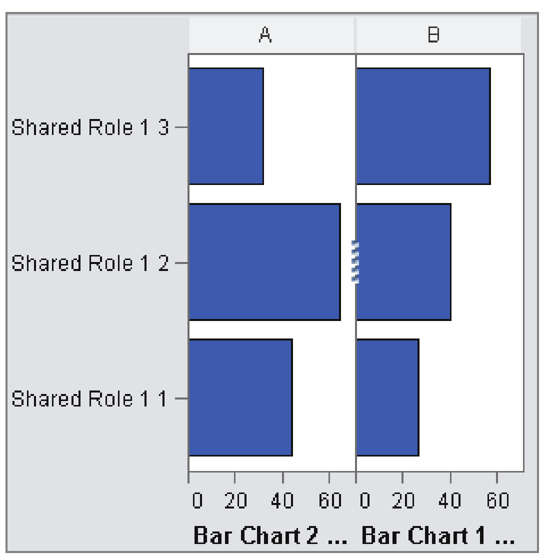
5 Specify a uniform column and row axis.

a Select the full graph. (**Custom Graph** should be displayed on the **Properties** tab).

b On the **Properties** tab, click ▼ next to **Y axis range (left axis only)**, and select **Same within each row**.

c On the **Properties** tab, click ▼ next to **X axis range**, and select **Same for all cells**.

The bar charts should resemble the following:

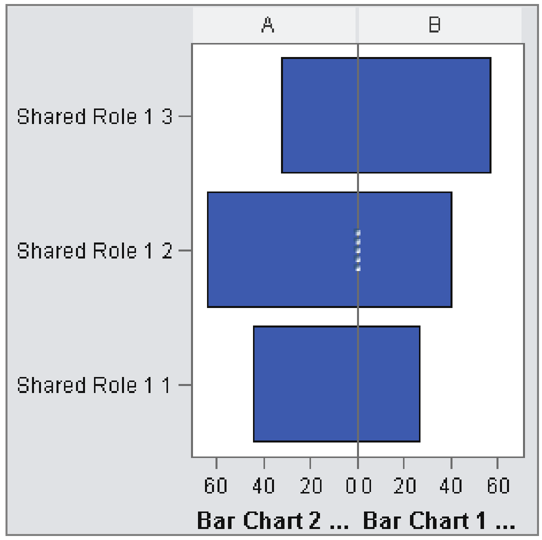


6 Reverse the order of the X (Measure) axis for the left bar chart.

a Click the horizontal axis for the left bar chart.

b On the **Properties** tab, select **Reverse order**.

The bar charts resemble the following:



7 Change the color of the right bar chart to distinguish it from the left bar chart.

a Select the right bar chart.

b On the **Properties** tab, click ▾ next to **Fill color**, and select **Data color 3**. The right bar chart changes to a different color. In the default theme, the color is deep red.

8 Display the grid lines for the X axes.

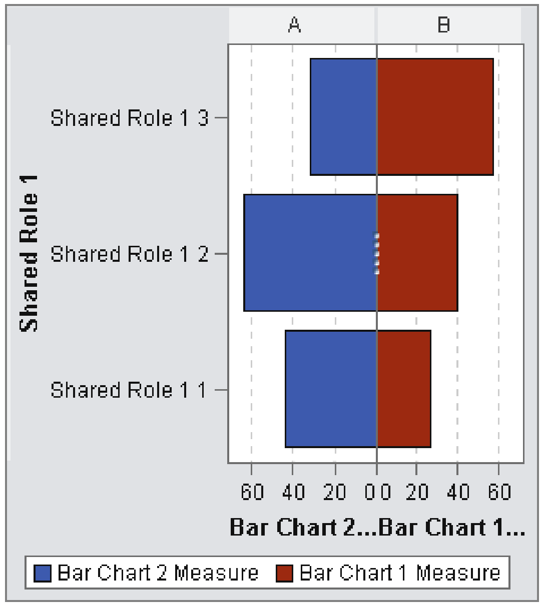
a Click the horizontal axis for the left bar chart.

b On the **Properties** tab, select **Show grid lines**.

c Repeat the previous two steps for the right bar chart.

9 Save the graph object. See “Save a Custom Graph Object So It Appears in the Designer” on page 507.

The final graph object, complete with the legend, resembles the following:

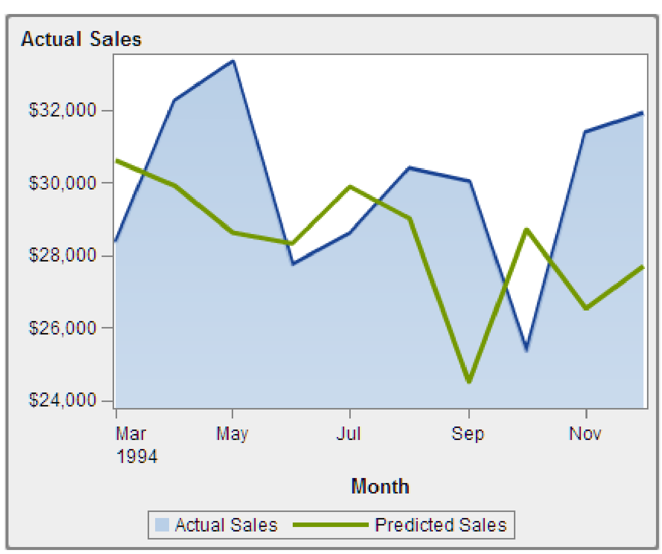


Example: Filled Overlay

About the Filled Overlay Example

This example uses two time series plots to show actual sales compared to predicted sales for a line of retail products. The graph is enhanced by applying a filled area to one of the overlaid plots.

Figure 54.3 Example Filled Overlay



Build the Graph Object for the Filled Overlay Example

1. In the graph builder, drag and drop a time series plot from the **Graph Elements** pane onto the canvas.
2. Drag and drop a second time series plot from the **Graph Elements** pane onto the first time series plot. This action creates an overlaid plot.

   Note: When you overlay the plots, the Time role is automatically shared between the plots.
3. Specify a fill area for the first time series plot.
   a. Select the first time series plot. (On the **Properties** tab, the plot might be identified as Time Series Plot 1.)
   b. On the **Properties** tab, click ▾ next to **Grouping style**, and select **Overlay Filled**.
   c. Click ▾ next to **Fill color**, and select **Prediction limits color**.
4. Save the graph object. See "Save a Custom Graph Object So It Appears in the Designer" on page 507.

Example: Vector Plot

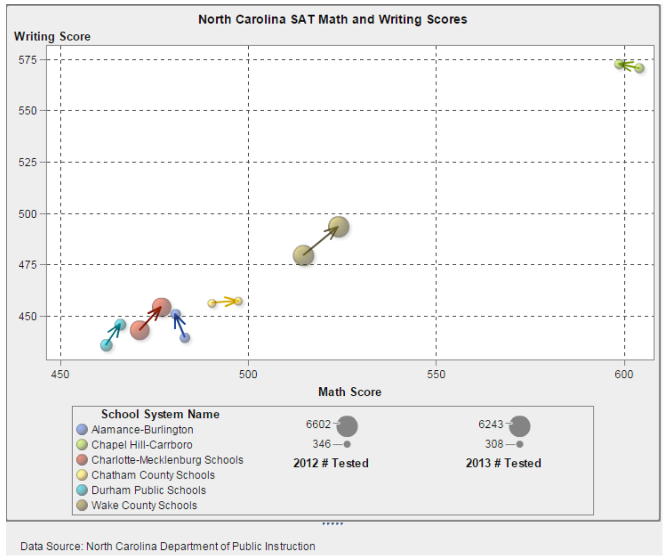
About the Vector Plot Example

This example uses a vector plot to show the changes in SAT exam scores for different North Carolina school systems from 2012 to 2013. The graph is enhanced by adding bubble plot overlays.

Here are the main features of this graph:

- Vector lines show the changes in the average math and writing scores between 2012 and 2013.
- Bubbles represent the average score for each year. The size of the bubble represents the number of students who took the test.
- To modify the appearance, the arrow line thickness was reduced and the transparency for the bubbles was increased. In addition, grid lines and a sheen data skin were applied to the graph. (These changes were made in the designer.)
- To reduce the number of school systems for comparison, a filter was applied in the designer on the School System Name category.

Figure 54.4 Example Vector Plot with Bubble Plots



Build the Graph Object for the Vector Plot Example

TIP As a shortcut, you can select a bubble change plot from the graph gallery.

1 In the graph builder, drag and drop a vector plot from the **Graph Elements** pane onto the canvas.

2 Drag and drop a bubble plot from the **Graph Elements** pane onto the vector plot.

3 Drag and drop a second bubble plot from the **Graph Elements** pane onto the vector plot.

4 On the **Role Definitions** tab, share the roles. This action joins the vector origin point to the first bubble, and the vector termination point to the second bubble.

a Click ▾ next to the **Bubble Plot 1 X** role. Select **Create Shared Role With Another Role ▸ Vector Plot 1 X Origin**.

In the Add Shared Role window, specify **Xstart** as the name of the shared role. Click **OK**.

b Click ▾ next to the **Bubble Plot 1 Y** role. Select **Create Shared Role With Another Role ▸ Vector Plot 1 Y Origin**.

In the Add Shared Role window, specify **Ystart** as the name of the shared role. Click **OK**.

c Click ▾ next to the **Bubble Plot 2 X** role. Select **Create Shared Role With Another Role ▸ Vector Plot 1 X**.

In the Add Shared Role window, specify **Xend** as the name of the shared role. Click **OK**.

d Click ▼ next to the **Bubble Plot 2 Y** role. Select **Create Shared Role With Another Role** ▶ **Vector Plot 1 Y**.

In the Add Shared Role window, specify **Yend** as the name of the shared role. Click **OK**.

5 On the **Role Definitions** tab, add a group role to all three plots.

a In the **Bubble Plot 1** section, click **Add Role**. The Add Role window is displayed. The **Group** type is selected by default.

Click **OK**.

b Repeat the previous step for the **Bubble Plot 2** section.

c In the **Vector Plot 1** section, click **Add Role**. The Add Role window is displayed.

Select **Group** for the type of role. Click **OK**.

6 Share the **Group** role among all three plots.

a In the **Bubble Plot 1** section, click ▼ next to the **Group** role. Select **Create Shared Role With Another Role** ▶ **Bubble Plot 2 Group**.

In the Add Shared Role window, specify **Color** as the name of the shared role. Click **OK**.

b In the **Vector Plot 1** section, click ▼ next to the **Group** role. Select **Use Shared Role** ▶ **Color**.

7 Specify that the plots start with the same color.

a On the **Properties** tab, select **A1** (the cell).

b Click ▼ next to **Overlaid plot colors**, and select **All graphs start with the same color**.

8 Remove redundancy from the legend.

a On the **Properties** tab, select the **Discrete Legend**.

b In the **Display in legend** box, deselect the **Vector Plot 1** and the **Bubble Plot 2** check boxes.

9 Specify the bubble size. This is an optional change made only to improve the appearance of the bubble sizes with respect to the vector lines.

a On the **Properties** tab, select **Bubble Plot 1**.

b Next to **Radius of smallest bubble (pixels)**, enter the number **5**.

c Next to **Radius of largest bubble (pixels)**, enter the number **12**.

d Repeat the previous steps for Bubble Plot 2.

10 Save the graph object. See "Save a Custom Graph Object So It Appears in the Designer" on page 507.

Part 8

Viewing Reports

55

Viewing Reports on a Mobile Device

What Are the SAS Visual Analytics Apps?

The SAS Visual Analytics Apps (previously called SAS Mobile BI) are free mobile apps. Using these apps, you can view and interact with SAS Visual Analytics reports, as well as share comments and observations with others. The apps support all charts and graphs that are available in SAS Visual Analytics.

For more information, see the SAS Visual Analytics Apps Documentation.

Where Can I Find SAS Visual Analytics Apps?

You can download the apps from the following locations:

- Apple App Store
- Google Play
- Microsoft Store

56

Viewing Reports in Modern SAS Visual Analytics Viewer

Overview of Viewing Reports in SAS Visual Analytics Viewer

As an alternative to viewing reports in a native mobile app, you can use the web viewer. For users with a Report Viewing role, the SAS Visual Analytics Viewer (the viewer) enables them to view report content.

Open a Report in the Modern Viewer

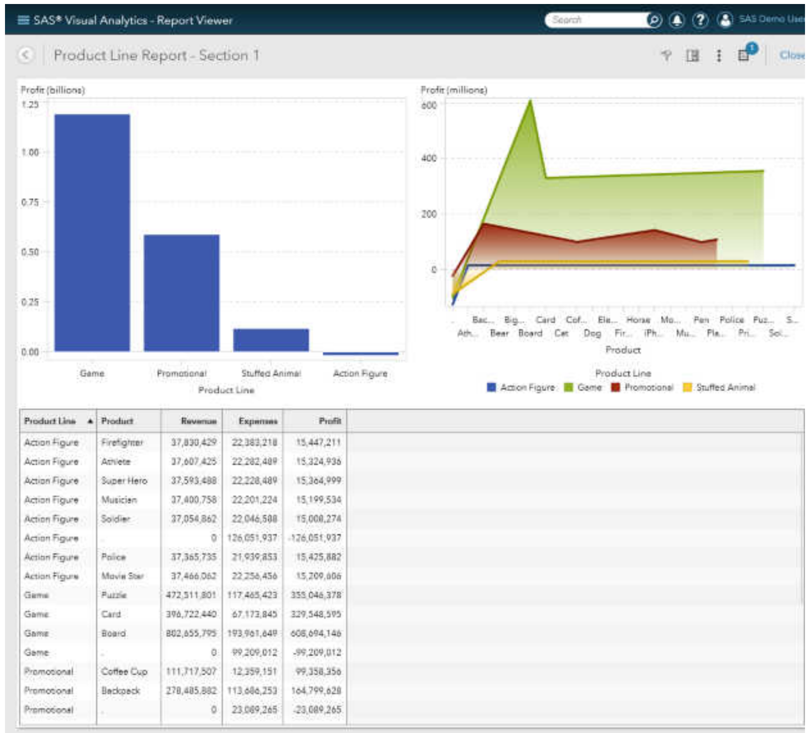
To open a report in the modern viewer:

- On SAS Home (modern), click next to a report, and then click **View**.
- On the classic home page, use the object inspector to view details about the report, and then click **View**. For more information about the object inspector, see "Discover Details Using the Object Inspector on the Classic Home Page" on page 633.
- In the modern viewer, click **Browse** or click on a report in the **Recent** list.

Note: The layout of some charts is dependent on the size of the display area. This means that the same treemap might appear slightly different in the viewer than it does in SAS Visual Analytics Designer (the designer) or in a native mobile app.

Here is an example of a report in the modern viewer:

Figure 56.1 *A Report in the Modern Viewer*



Product Line	Product	Revenue	Expenses	Profit
Action Figure	Firefighter	37,830,429	22,383,218	15,447,211
Action Figure	Athlete	37,607,425	22,282,489	15,324,936
Action Figure	Super Hero	37,593,488	22,228,489	15,364,999
Action Figure	Musician	37,400,758	22,201,224	15,199,534
Action Figure	Soldier	37,054,862	22,046,588	15,008,274
Action Figure	.	0	126,051,937	-126,051,937
Action Figure	Police	37,365,735	21,939,853	15,425,882
Action Figure	Movie Star	37,466,062	22,256,456	15,209,606
Game	Puzzle	472,511,801	117,465,423	355,046,378
Game	Card	396,722,440	67,173,845	329,548,595
Game	Board	802,655,795	193,961,649	608,694,146
Game	.	0	99,209,012	-99,209,012
Promotional	Coffee Cup	111,717,507	12,359,151	99,358,356
Promotional	Backpack	278,485,882	113,686,253	164,799,628
Promotional	.	0	23,089,265	-23,089,265

Here are some key points about the modern viewer:

- You cannot edit stored process prompts in the modern viewer. In the modern viewer, stored processes are executed using default prompt values.
- The web viewer is not supported on mobile devices. Mobile users are redirected to SAS Visual Analytics Apps (previously called SAS Mobile BI). For more information, see "What Are the SAS Visual Analytics Apps?" on page 535.
- The modern viewer does not support right-to-left (RTL) languages. If you are using RTL languages, be sure to specify **Classic** for the **Default Appearance** setting.

For more information about using the modern viewer, refer to the online Help that is available for the viewer. For more information about using the classic viewer, see Chapter 57, "Viewing Reports in Classic SAS Visual Analytics Viewer," on page 541.

View a Report Object with SAS Graphics Accelerator

What is SAS Graphics Accelerator?

Starting in the 7.4 release, you can view some types of report objects with SAS Graphics Accelerator.

SAS Graphics Accelerator is a Google Chrome extension that enables users with visual impairments or blindness to explore data visualizations. It supports alternative presentations of data visualizations that include enhanced visual rendering, text descriptions, tabular data, and interactive sonification. Sonification uses non-speech audio to convey important information about the graph.

Installation

For information about installing SAS Graphics Accelerator, see http://support.sas.com/software/products/graphics-accelerator/index.html.

Supported Report Objects

The following report objects support SAS Graphics Accelerator:

- bar chart
- bubble plot
- line chart
- time series plot
- pie chart
- scatter plot

Open SAS Graphics Accelerator

To view a report object with SAS Graphics Accelerator, position your cursor over the button for the report you want to duplicate, and then click .

SAS Graphics Accelerator displays the report object in a new Google Chrome tab.

The SAS Graphics Accelerator tab contains a link to the *SAS Graphics Accelerator: User's Guide*.

About Guest Access in the Viewer

SAS Visual Analytics system administrators can configure support for guest access. Users with guest access can access only the home page and the viewer. Guest access uses a shared account, so it does not provide individualized features, such as history, favorites, preferences, or alerts. If guest access is configured, it is available at a special URL such as http://*host*/SASVisualAnalyticsViewer/guest.jsp

Accessing SAS Visual Analytics as a guest is useful if you do not have a metadata identity. This enables you to view reports that are widely available under a generic, shared account. You can also view reports that are available to the public on the Internet.

Note: Some features such as comments and alerts are not available with guest access.

57

Viewing Reports in Classic SAS Visual Analytics Viewer

Open a Report in the Classic Viewer

Starting in the 7.3 release, the modern viewer is the default. To switch to the classic viewer, you must modify your settings. Click your name, and select **Settings**. Under **SAS Report Viewer**, click **Default Appearance** to specify the appearance of the viewer. Select **Classic**. You must sign out and sign in for the change to take effect.

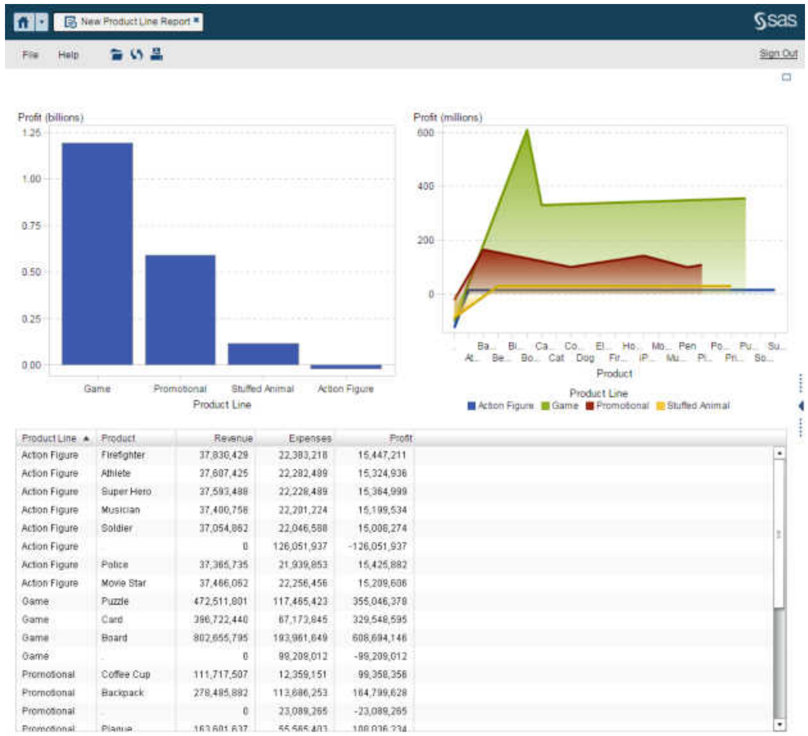
To open a report in the classic viewer:

- On SAS Home (modern), click i next to a report, and then click **View**.
- On the classic home page, use the object inspector to view details about the report, and then click **View**. For more information about the object inspector, see "Discover Details Using the Object Inspector on the Classic Home Page" on page 633.
- In the classic viewer, select **File** ▶ **Open**.

Note: The layout of some charts is dependent on the size of the display area. This means that the same treemap might appear slightly different in the viewer than it does in SAS Visual Analytics Designer (the designer) or in the SAS Visual Analytics Apps (previously called SAS Mobile BI).

Here is an example of a report in the classic viewer:

Figure 57.1 *A Report in the Classic Viewer*



If you have the Create Report capability, then you can select **File** ▶ **Edit Report** in the current report. The designer is displayed, and then you can edit the report.

You can e-mail and print reports to share with other users, and export data and images of report objects. For more information about sending reports to other users, prerequisites and conditions for printing, and exporting, see Chapter 51, “Sharing Reports with Other Users,” on page 475.

Note: Info windows are not included in the PDF that is generated when you print. An info window can be printed separately by clicking **Print report to PDF** in the info window.

View Report Object Information in the Classic Viewer

You can select a report object to display icons that provide more information about that report object. Depending on which type of report object you select, you might see the following icons:

Icon	Description
ⓘ	Click to display the title and description of the report object.

Icon	Description
	Click to maximize the report object. This icon is displayed only if there is more than one report object being displayed.
	Click to restore the report object to its original size. This icon is displayed only if you previously maximized the report object.
	Click to display incoming filter information for a report object. This icon is displayed only if the data for the report object has been filtered as a result of selecting data in another report object.
	Click to display the prompt dialog box for a stored process. This icon is displayed only for prompted stored processes. You can use the prompt dialog box to change prompt values for a stored process and to re-execute the stored process.

Note: Report authors can disable object selection, so you might not be able to select some report objects.

Add Comments to a Report in the Classic Viewer

If you have the Add Comments capability, then you can add or view comments. Comments can be added to a report or to an object within a report. The comments that you add are automatically saved with the report.

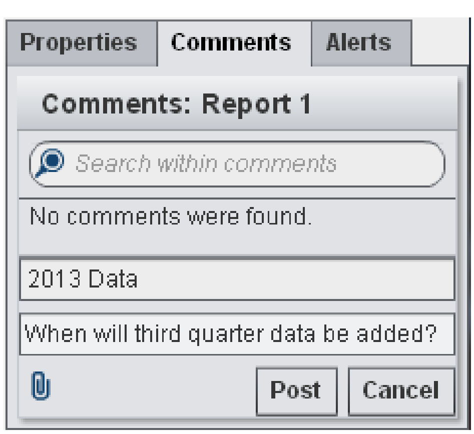
To add a comment to a report or to an object within a report:

1 Expand the right pane. Click the **Comments** tab. (If you are adding a comment to a report object, then select the report object.)

2 Enter a topic name and a comment.

Here is an example:

Figure 57.2 Adding Comments in SAS Visual Analytics Viewer



3 (Optional) Click  to attach a file to your comment.

4 Click **Post** to add your comment. Your comment is added to the **Comments** tab in the right pane.

To respond to an existing comment:

1 Expand the right pane. Click the **Comments** tab. Select an existing comment, and then enter a reply.

2 (Optional) Click to attach a file to your reply.

3 Click **Post**. Your reply is added to the **Comments** tab in the right pane.

Note: To edit another user's comments or to delete comments, you must have the **Comments:Administrator** predefined role.

To search for a comment:

1 Expand the right pane. Click the **Comments** tab. Enter the word or phrase that you want to search for in the **Search within comments** field on the **Comments** tab.

2 (Optional) To clear the **Search within comments** field, click .

Interacting with Reports in the Classic Viewer

Filter, Brush, and Drill in Reports

Depending on the interactions that were defined by the report designer, you might be able to interact with your report in the following ways:

filter
: restricts the data that is returned from a query to a data source. Click on data in the source report object to filter data in the target report object (or objects). Clicking on different data applies the filter based on the new data. To clear the selection, right-click, and select **Clear Selection**.

brush
: enables you to show the same data selected simultaneously in two or more tables, graphs, or both. Click on data in the source report object to brush data in the target report object (or objects). To clear the selection, right-click, and select **Clear Selection**.

drill down
: enables you to move from summary information to more detailed data. If the data contains hierarchies, you can double-click on the data to drill down the hierarchy into detailed information. When you drill down the hierarchy, breadcrumbs at the top of the report object enable you to drill back up the hierarchy.

View Links in Reports

Report objects can link to other report sections or whole reports. And, they can link to external links. To view a link from a report object, double-click on the report object. If there are multiple links or interactions from a report object, then a list appears that enables you to select a link or interaction. When you are viewing a link, the following buttons are displayed in the top left corner: .

If a report object that you are linking to shares a data source with the current report object, then the target report object is filtered based on the data value that you double-clicked. If you do not want the target report object to be filtered, you can click , and deselect **Apply report link filters**. If the data source is not shared between the two report objects, then no additional filtering takes place in the target report object.

To go back to the original report object, click .

A report object, text, or image can link to an info window within the same report. The info window provides additional information. For example, a list table might provide additional information for a bar chart. Or, there might be additional text about what is displayed in a report object. When you double-click the data (for example, a bar, a bubble, a pie slice, a table row, and so on) in a report object that has an info window link, the info window is displayed as a new window.

Subscribe and Unsubscribe to Alerts in the Classic Viewer

You can subscribe to existing alerts for report objects and receive notifications when the alert criteria is met.

To subscribe or unsubscribe to alerts for a report object:

1 Expand the right pane. Click the **Alerts** tab. This tab contains a list of all alert conditions for all report objects within the report.

2 Select or deselect the **Subscribe** check box next to an alert to subscribe or unsubscribe to that alert.

TIP You can specify a preference for receiving notifications via e-mail or text message. For more information, see “Specify General Preferences for the Designer” on page 305.

Specify Your Preferences for the Classic Viewer

To specify preferences that are specific to the classic viewer:

1 Select **File** ▸ **Preferences** to open the Preferences window.

2 Select **SAS Visual Analytics Viewer**.

3 If the **SAS High Contrast** theme is selected as a global preference, then you can select the **Override report theme when High Contrast Theme is**

selected check box to ensure that reports are displayed using the SAS High Contrast theme. This overrides any theme settings that are made in the designer.

4 Select **Modern**, **Classic**, or **Administrator default** to specify the default appearance of the viewer. This change takes effect after you sign out and sign back in.

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

To specify your preferences for receiving notifications, see "Specify General Preferences for the Designer" on page 305. To specify general preferences for SAS Visual Analytics, see "Specify Settings Using SAS Home" on page 10. To specify global SAS preferences, see "Specifying Your Preferences" on page 9.

Part 9

Appendixes

Appendix 1

Keyboard Shortcuts for SAS Visual Analytics

The following table contains many of the keyboard shortcuts for SAS Visual Analytics. In SAS Visual Analytics, some shortcuts are displayed within parentheses in tooltips and menu item labels. Some are also displayed in the Keyboard Shortcuts window (press F9 to open that window).

Note: When you use a keyboard shortcut to activate a button, move 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 Help button, you must first move the focus to the table before you press Ctrl+?.

Table A1.1 Keyboard Shortcuts

Action	Keyboard Shortcut
Open the Keyboard Shortcuts window.	F9 **Note:** The Keyboard Shortcuts window might not contain all of the shortcuts for your application.
Open a Help pop-up window from the **Help** button.	Ctrl+? **Note:** This shortcut does not work on some keyboards (for example, the Italian keyboard).
Zoom in.	Ctrl+plus sign
Zoom out.	Ctrl+minus sign
Reset the zoom state.	Ctrl+0
Maximize view (collapses the category pane and the tile pane and hides the status bar and the application bar, which includes the menu bar). or Exit maximized view (expands the category pane and the tile pane and shows the status bar and the application bar).	Ctrl+Alt+Shift+M

Action	Keyboard Shortcut
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.
Open the Landmarks window.	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 the 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 the 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 between table headings and table content.	For a two-dimensional table, make sure that the focus is on the table and that you are not in Edit mode. Press Ctrl+F8 to switch the focus between column headings and table cells. Use the arrow keys to navigate from heading to heading. For a multidimensional table, make sure that the focus is on a table cell and that you are not in Edit mode. Press Ctrl+F8 to switch the focus between column headings, row headings, and table cells. Use the arrow keys to navigate from heading to heading.
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.

Action	Keyboard Shortcut
Sort columns in a table.	To sort a single column, navigate to its column heading (press Ctrl+F8). Press the 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.
Change the width of the current column.	Navigate to the column heading (press Ctrl+F8). Then, press Ctrl+left arrow or Ctrl+right arrow to change the width of the column.
Move the current column.	Navigate to the column heading (press Ctrl+F8). Then, press Shift+left arrow to move one column to the left, and press Shift+right arrow to move one column to the right.
Automatically re-size the current column to fit its contents.	Navigate to the column heading (press Ctrl+F8). Then, press Enter.
Open a drop-down list or drop-down menu.	Make sure that the focus is on the control, and press Ctrl+down arrow.
Exit a single application in the SAS Visual Analytics home page.	Tab to the application’s button at the top of the browser window, and press Delete.

Appendix 2

Gallery of Report Objects

Tables

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. A list table cannot use a hierarchy or a percentage of subtotals.

You can add sparklines to a column (if the data source contains a date data item) when aggregated data is displayed in the list table.

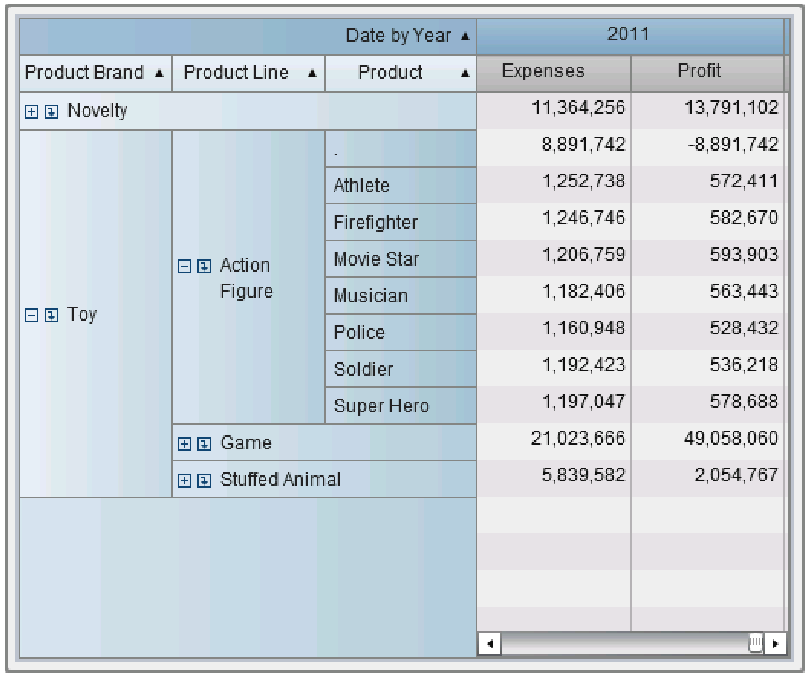
Figure A2.1 *A List Table*

Product Line	Revenue	Expenses	Profit ▼	Profit Trend
Game	1,671,890,035	477,809,929	1,194,080,107	
Promotional	813,699,290	223,822,374	589,876,916	
Stuffed Animal	276,990,966	159,548,680	117,442,285	
Action Figure	262,318,761	281,390,254	-19,071,493	

Crosstabs

A *crosstab* (also known as a crosstabulation table) shows an aggregate metric for the intersections of two or more categories. Crosstabs often have two or more categories assigned to both the rows and columns, forming a matrix. Crosstabs can be easier to read than list tables because they often use less space, and they always collapse repeating values for outer category data items into one unique value, which is known as grouping. A crosstab can use a hierarchy.

Figure A2.2 A Crosstab



		Date by Year ▲	2011	
Product Brand ▲	Product Line ▲	Product ▲	Expenses	Profit
Novelty			11,364,256	13,791,102
Toy	Action Figure	.	8,891,742	-8,891,742
		Athlete	1,252,738	572,411
		Firefighter	1,246,746	582,670
		Movie Star	1,206,759	593,903
		Musician	1,182,406	563,443
		Police	1,160,948	528,432
		Soldier	1,192,423	536,218
		Super Hero	1,197,047	578,688
	Game		21,023,666	49,058,060
	Stuffed Animal		5,839,582	2,054,767

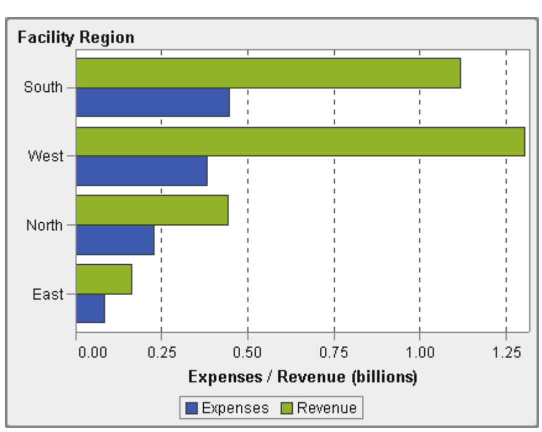
Graphs, Charts, and Plots

Bar Charts

A *bar chart* consists of vertical or horizontal bars that represent quantitative data. Use bar charts to compare data that is aggregated by the distinct values of a category.

You can apply grouping and create data-driven lattices. You can filter or rank your data based on a specified number of top or bottom values.

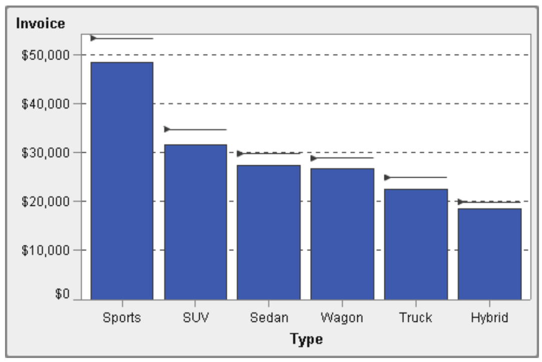
Figure A2.3 A Bar Chart



Targeted Bar Charts

A *targeted bar chart* is a variation of the bar chart that has target values. A target value is represented as a triangle with a line at the target value that is determined by the target role.

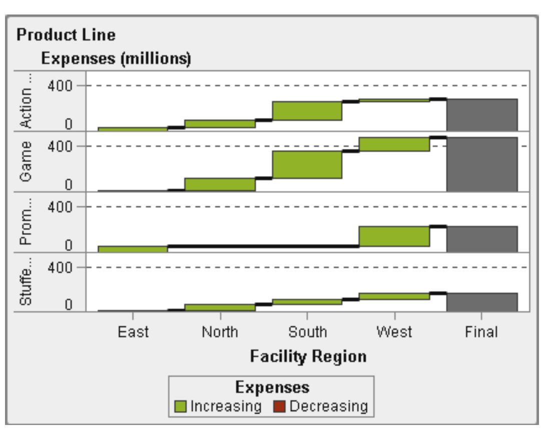
Figure A2.4 A Targeted Bar Chart



Waterfall Charts

A *waterfall chart* (also known as a progressive bar chart) shows how the initial value of a measure increases or decreases during a series of operations or transactions. 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 indicate the magnitude and type (positive or negative, for example) of the operation or transaction. The resulting chart is a stepped bar showing how incremental changes lead to the final value of the measure.

Figure A2.5 A Waterfall Chart

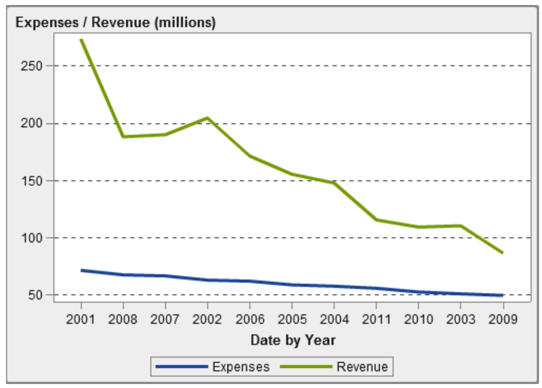


Line Charts

A *line chart* shows the relationship of one or more measures over some interval, such as time or a series of ranges. You can measure a single measure (univariate analysis), or you can show the relationships among multiple measures (multivariate analysis), such as the leading or lagging relationship between advertising and sales over time. The category on the X axis of a line chart is discrete; the category on the X axis of a time series plot is continuous.

You can apply grouping and create lattices.

Figure A2.6 *A Line Chart*



Pie Charts

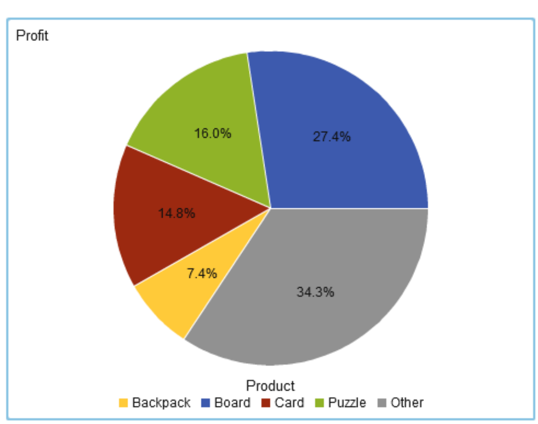
A *pie chart* displays a part-to-whole relationship in a circle divided into multiple slices for each value of a category data item based on a single measure data item. Each slice represents the relative contribution of each part to the whole. In a pie chart, the legend is sorted by contribution.

In SAS Visual Analytics Designer (the designer), a pie chart does not show a slice with a missing or zero response.

Effective pie charts limit the number of slices to 5 or 6. In the designer, you can use a rank to reduce the number of slices in a pie chart. For more information, see "Add a New Rank" on page 459.

Note: The Other slice does not display data tip values. In addition, the Other slice always sums the included values, regardless of the aggregation method selected for the measure. For example, if the aggregation method selected is Count, then the Other slice displays the sum of the individual counts.

Figure A2.7 *A Pie Chart*

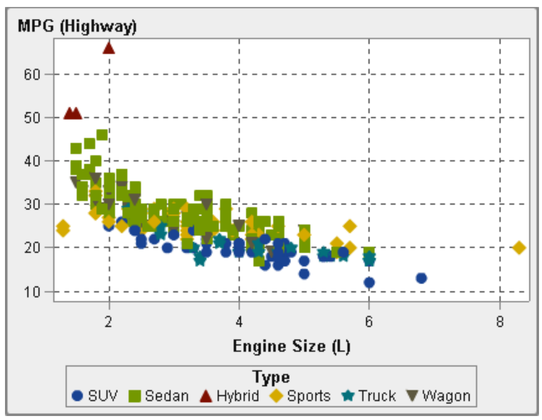


Scatter Plots

A *scatter plot* is a two-dimensional plot that shows the relationship of two measure data items. Each marker (represented by a symbol such as a dot, a square, or a plus sign) represents an observation. The marker's position indicates the value for each observation. Use a scatter plot to examine the relationship between numeric data items. You can apply grouping by assigning a category to the **Color** role.

Scatter plots do not use aggregated data.

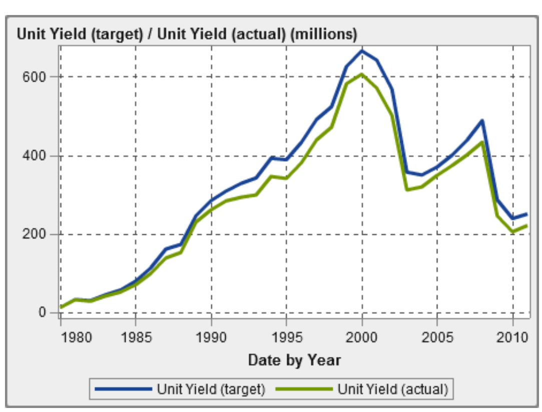
***Figure A2.8** A Scatter Plot*



Time Series Plots

A *time series plot* shows an ordered sequence of values that are observed at equally spaced time intervals. A time series plot requires a date, datetime, time, or hierarchy data item that is continuous on the X axis.

***Figure A2.9** A Time Series Plot*

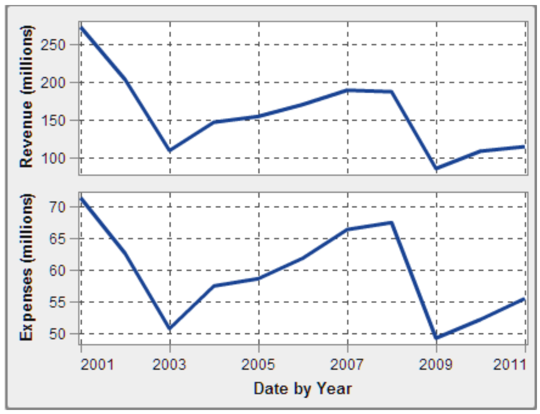


Comparative Time Series Plots

A *comparative time series plot* uses line segments to chart two measures on different scales over time. A comparative time series plot requires a date, datetime, time, or hierarchy data item that is continuous on the X axis. The X axis is shared across both plots.

TIP This report object is not displayed in the designer by default. You can select what you want to display using the designer's Show or Hide Objects window.

Figure A2.10 A Comparative Time Series Plot



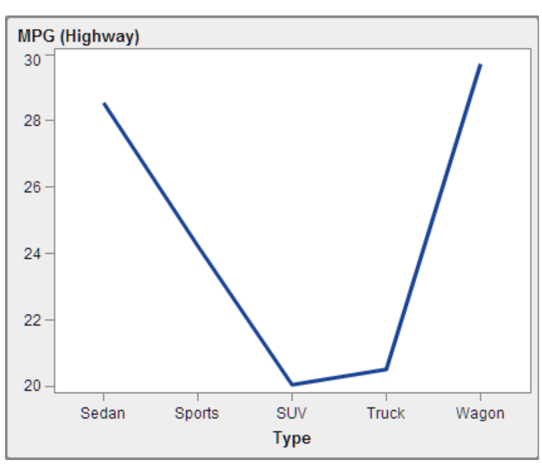
Series Plots

A *series plot* displays a series of line segments that connect observations of input data. A series plot can use numeric or character data on the X axis.

Note: You must create and save this custom graph object in the SAS Visual Analytics Graph Builder (the graph builder) before the object is available for use in reports.

The following example shows MPG averages for different types of vehicles:

Figure A2.11 A Series Plot

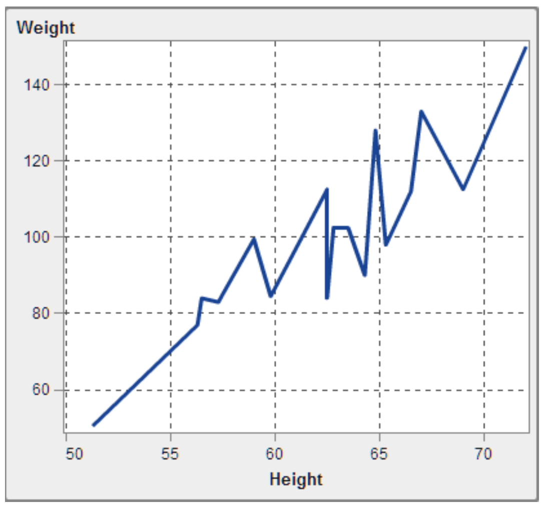


Numeric Series Plots

A *numeric series plot* shows the relationship of one or more measures across a series of numeric values. A numeric series plot requires numeric data on the X axis.

TIP This report object is not displayed in the designer by default. You can select what you want to display using the designer's Show or Hide Objects window.

Figure A2.12 A Numeric Series Plot

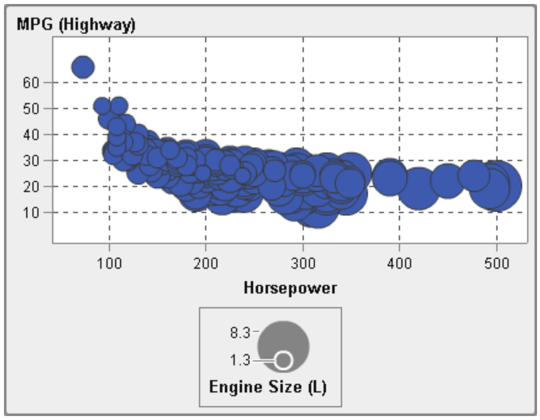


Bubble Plots

A *bubble plot* is a variation of a scatter plot in which the markers are replaced with bubbles. A bubble plot displays the relationships among at least three measures. Two measures are represented by the plot axes, and the third measure is represented by the size of the bubbles. Each bubble represents an observation. A bubble plot is useful for data sets with dozens to hundreds of values. You can add categories to the **Grouping** and **Lattice** roles.

Note: A bubble's size is scaled relative to the minimum and maximum values of the size variable. The minimum and maximum sizes are illustrated in the plot legend. The actual value for each bubble is displayed as a data tip. For example, the legend that is displayed in Figure A2.13 on page 561, the minimum size is 1.3 and the maximum size is 8.3.

Figure A2.13 A Bubble Plot

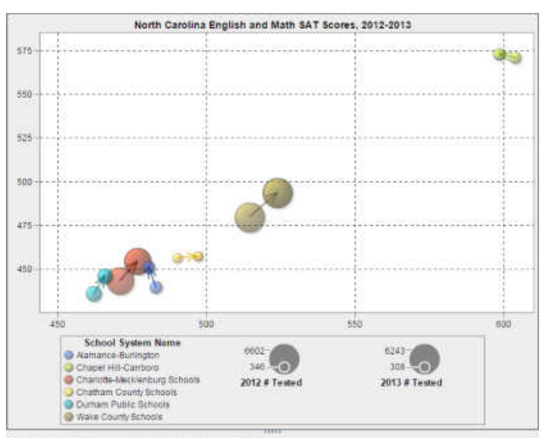


Bubble Change Plots

A *bubble change plot* shows changes in two sets of measures using bubbles and directed line segments.

TIP This report object is not displayed in the designer by default. You can select what you want to display using the designer's Show or Hide Objects window.

Figure A2.14 A Bubble Change Plot



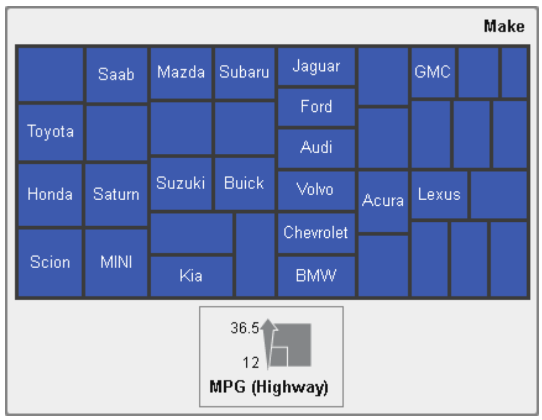
Treemaps

A *treemap* displays your data as a set of rectangles (called tiles). Each tile represents a category node or a hierarchy node. The color of each tile represents the value of the first measure. The size of each tile represents the value of the second measure. (There are two data roles for measures in a treemap—**Size** and **Color**.) For example, a sales data treemap might have tile sizes that represent the number of orders, and it might have tile colors that are derived from color gradients that represent sales.

The layout of the tiles in the treemap is dependent on the size of the display area because it uses a space-filling algorithm to lay the tiles down. This means that the same treemap might appear slightly different in the designer than it does in the viewer or on a mobile device because the aspect ratio and size available in those viewers might be different from what the original report designer sees in the designer.

Note: Treemaps allow only one category data item or hierarchy data item.

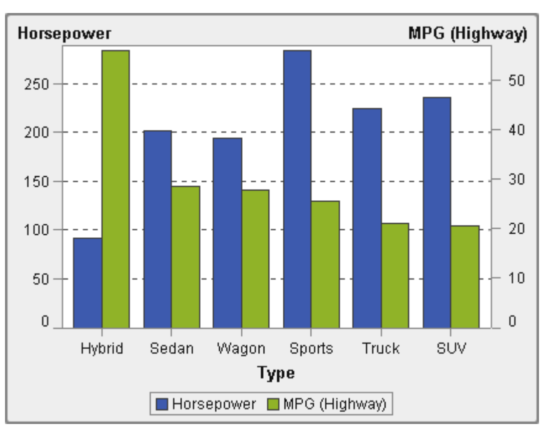
Figure A2.15 A Treemap



Dual Axis Bar Charts

A *dual axis bar chart* is a variation of the bar chart that has two measures. A measure is on each axis.

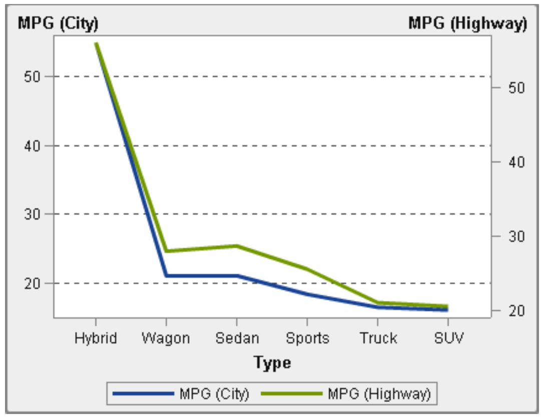
Figure A2.16 A Dual Axis Bar Chart



Dual Axis Line Charts

A *dual axis line chart* is a variation of the line chart that has two measures. A measure is displayed on both the left and right side of the Y axis. The relationship between two measures can be examined on two different scales in a dual axis line chart.

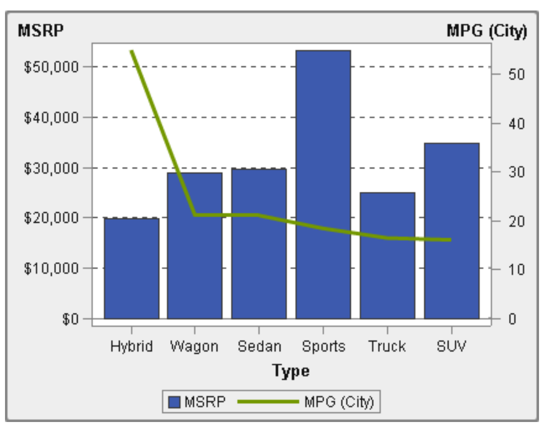
Figure A2.17 *A Dual Axis Line Chart*



Dual Axis Bar-Line Charts

A *dual axis bar-line chart* is a variation of the bar chart that has two measures. A measure is on each axis, and the bar chart is overlaid by a line chart.

Figure A2.18 *A Dual Axis Bar-Line Chart*

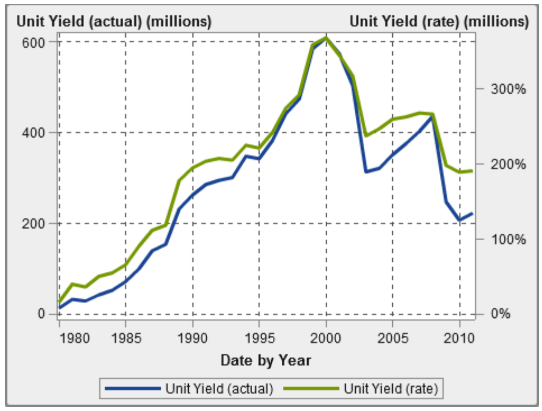


Dual Axis Time Series Plots

A *dual axis time series plot* is a variation of the time series plot that has two measures. A measure is displayed on both the left and right side of the Y axis.

For example, a dual axis time series plot can be useful when you need to display two measures that have the same unit of measurement and different scales, such as quantity ordered and returns, or when you need to display two measures that have different units of measurement, such as sales and quantity ordered.

Figure A2.19 A Dual Axis Time Series Plot



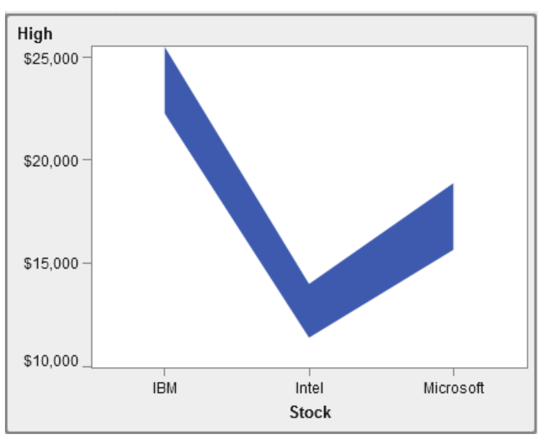
Band Plots

A *band plot* draws a horizontal band with two Y values for each X value. Or, it draws a vertical band with two X values for each Y value. A band plot is typically used to show confidence, error, prediction, or control limits. The points on the upper and lower band boundaries can be joined to create two outlines. The area between the boundaries is filled.

Here are some key points about band plots:

- You must create and save this custom graph object in the graph builder before the object is available for use in reports.
- A band plot does not support display rules in the designer.

Figure A2.20 A Band Plot



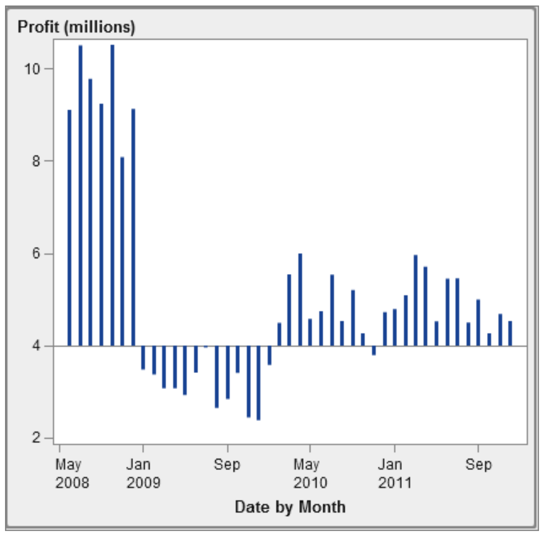
Needle Plots

A *needle plot* is a plot in which data points are connected by a vertical line that connects to a horizontal baseline. The baseline intersects the 0 value or the minimum value on the vertical axis.

TIP This report object is not displayed in the designer by default. You can select what you want to display using the designer's Show or Hide Objects window.

The following example shows profits during a particular time period. The example specifies an optional baseline value on the Y axis.

Figure A2.21 *A Needle Plot*



Step Plots

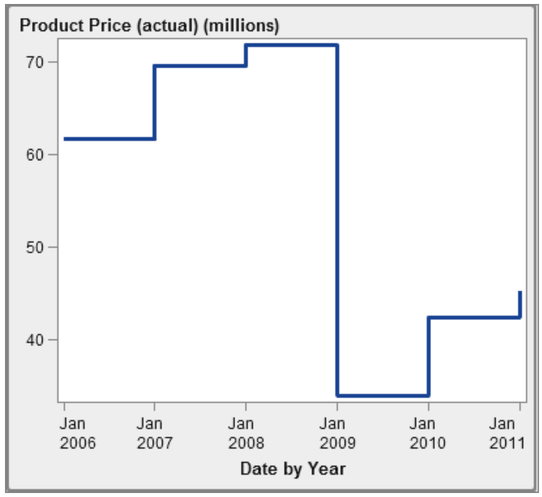
A *step plot* consists of a series of horizontal and vertical line segments (giving the appearance of steps) that connect observations of input data.

Note: A step plot does not support display rules in the designer.

TIP This report object is not displayed in the designer by default. You can select what you want to display using the designer's Show or Hide Objects window.

The following example shows the price trend during a particular time interval:

Figure A2.22 *A Step Plot*

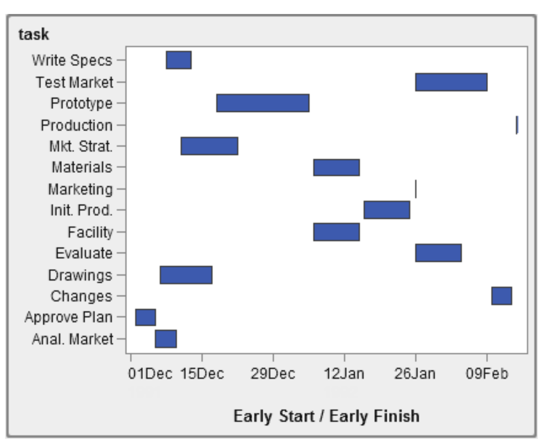


Schedule Charts

A *schedule chart* makes it easy to visualize time lines by representing tasks, start dates, durations, and end dates in cascading horizontal bar charts.

TIP This report object is not displayed in the designer by default. You can select what you want to display using the designer's Show or Hide Objects window.

Figure A2.23 *A Schedule Chart*



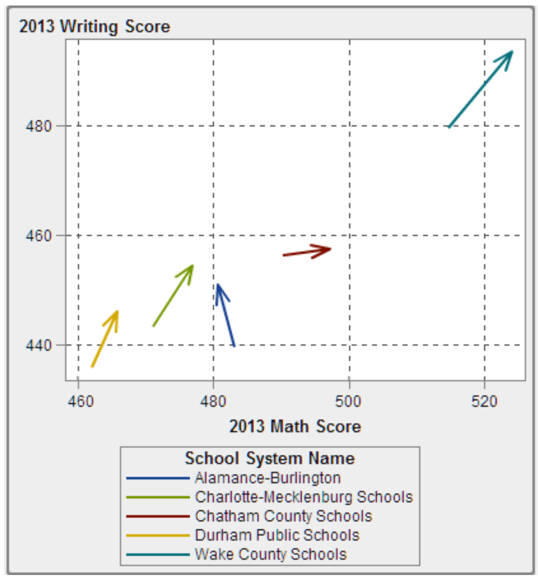
Vector Plots

A *vector plot* shows the change in value of a measure using directed line segments, or vectors, to represent both direction and magnitude at each point.

TIP This report object is not displayed in the designer by default. You can select what you want to display using the designer's Show or Hide Objects window.

The following example shows the changes in exam scores for different North Carolina school systems from 2012 to 2013:

Figure A2.24 A Vector Plot



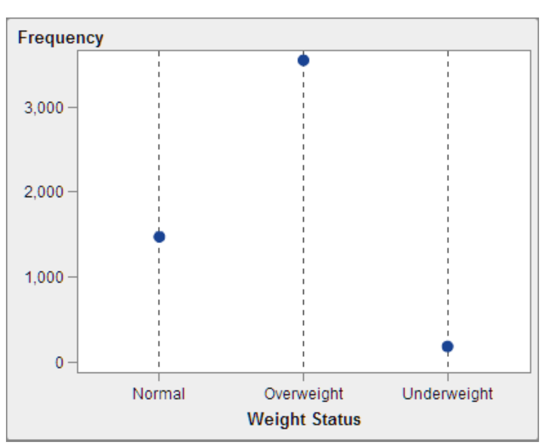
For an example that shows a vector plot enhanced with bubble plots, see "Example: Vector Plot" on page 529.

Dot Plots

A *dot plot* compares data that is aggregated by the value of a category.

TIP This report object is not displayed in the designer by default. You can select what you want to display using the designer's Show or Hide Objects window.

Figure A2.25 A Dot Plot

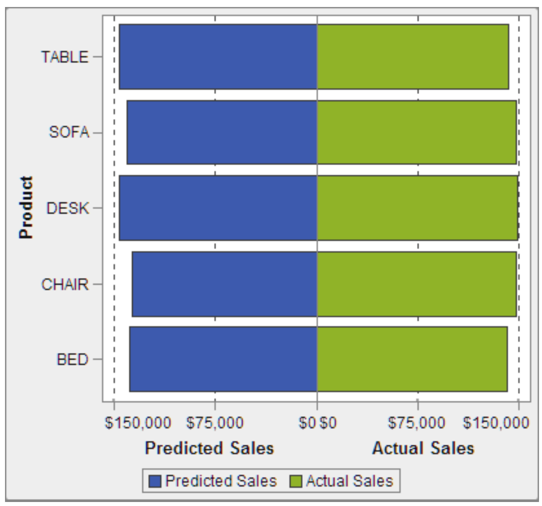


Butterfly Charts

A *butterfly chart* compares two measures for a category of values.

> **TIP** This report object is not displayed in the designer by default. You can select what you want to display using the designer's Show or Hide Objects window.

Figure A2.26 A Butterfly Chart

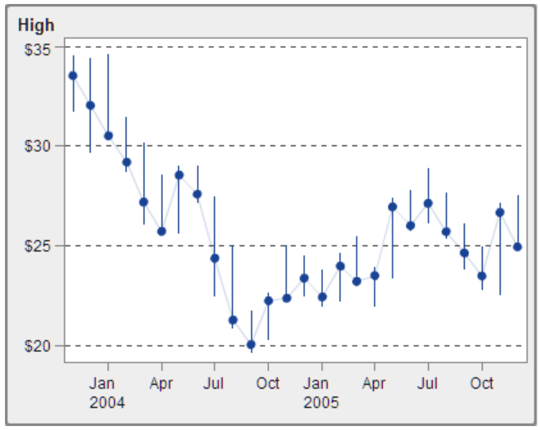


Stock High-Low Plots

A *stock high-low plot* tracks changes in the price of a tradable asset over time. This plot creates a display of floating vertical lines that represent high and low stock values. The plot also displays stock closing values as markers.

> **TIP** This report object is not displayed in the designer by default. You can select what you want to display using the designer's Show or Hide Objects window.

Figure A2.27 *A Stock High-Low Plot*



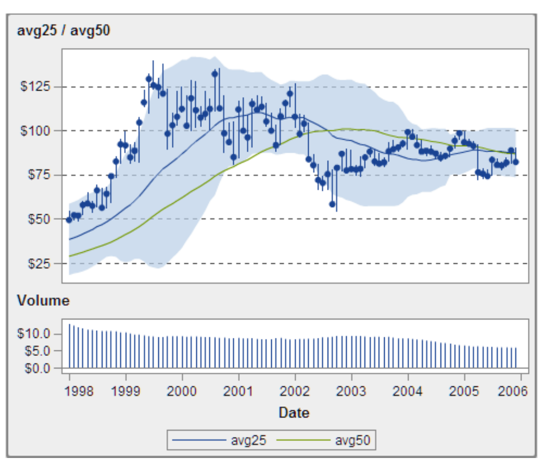
Stock Volume and Volatility Plots

A *stock volume and volatility plot* tracks changes in the price of a tradable asset over time with additional context.

This plot creates a display of floating vertical lines that represent high and low stock values. The plot also displays stock closing values as markers. The plot shows the stock's moving average and Bollinger upper and lower bands.

TIP This report object is not displayed in the designer by default. You can select what you want to display using the designer's Show or Hide Objects window.

Figure A2.28 *A Stock Volume and Volatility Plot*



Controls

A control is a report object that filters or narrows the scope of the data that user is viewing. A control enables a report designer to select a category that he or she wants the report viewer to see.

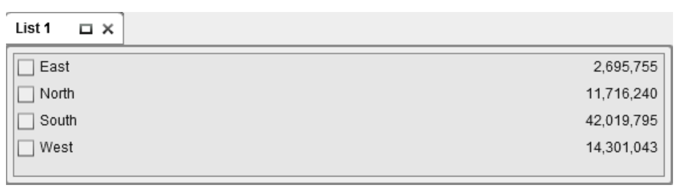
The following controls are available in the designer:

- drop-down lists

  Figure A2.29 *A Drop-down List Control*

- lists

  Figure A2.30 *A List Control*

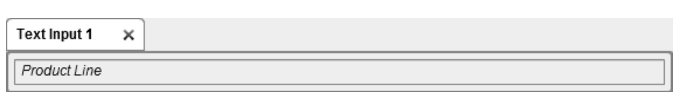


- button bars

  Figure A2.31 *A Button Bar Control*

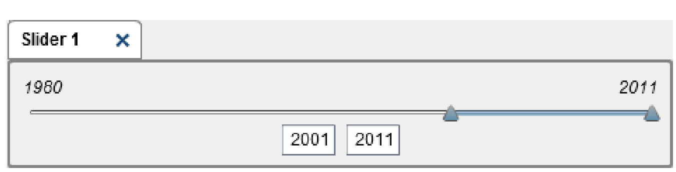
- text input fields

  Figure A2.32 *A Text Input Control*



- sliders

  Figure A2.33 *A Slider Control*

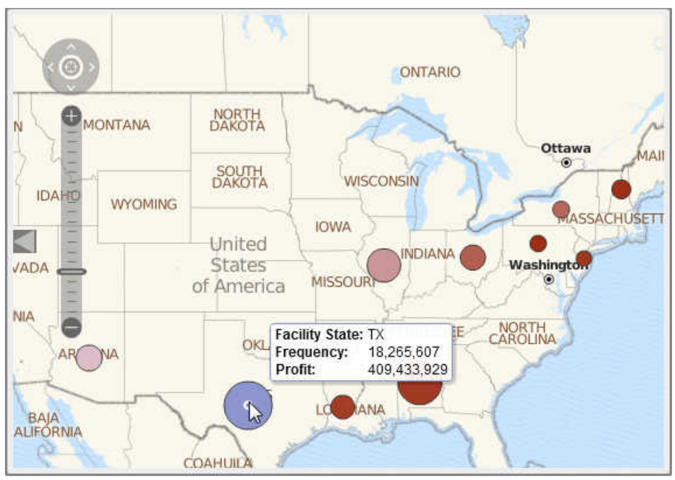


Other Report Objects

Geo Bubble Maps

A geo bubble map is 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. The bubbles are automatically colored based on the location. You provide a measure data item that determines the size of the bubble. A geo bubble map requires a data item that contains geographical information and is assigned to a geography role.

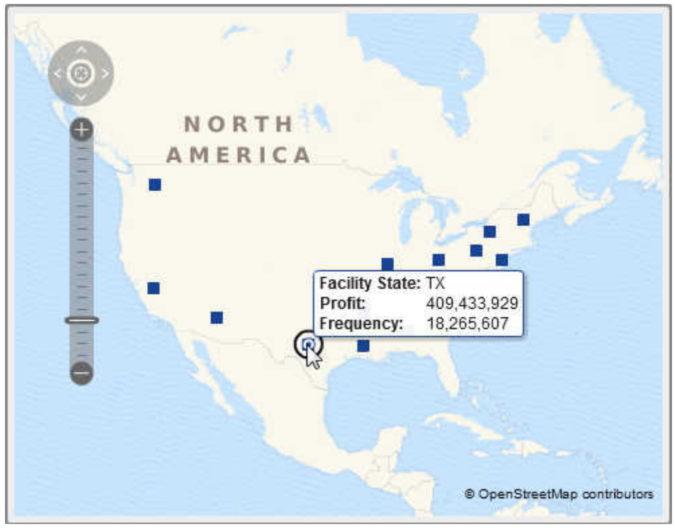
Figure A2.34 Geo Bubble Map



Geo Coordinate Maps

A geo coordinate map is a simple scatter plot that is overlaid on a geographic map. A geo coordinate map is used in place of a geo bubble map when the cardinality is too high and the geo bubble map cannot display the number of points. Each point in a geo coordinate map is located at the center of a geographic region or at the coordinates of a location. A geo coordinate map requires a data item that contains geographical information and is assigned to a geography role.

Figure A2.35 Geo Coordinate Map

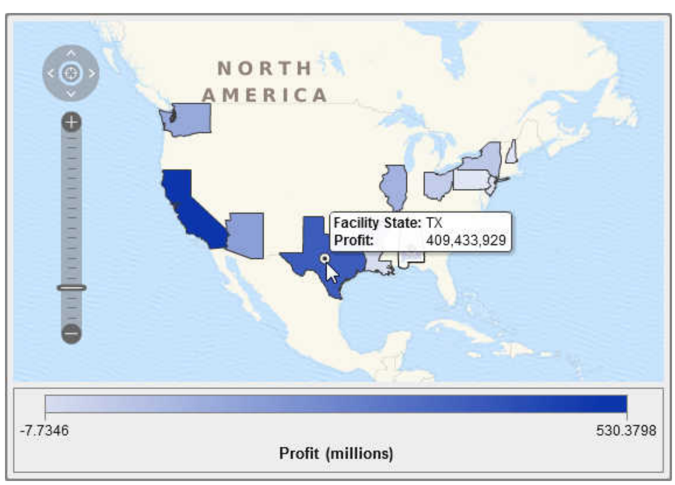


Geo Region Maps

A geo region map (also known as a choropleth map) is a two-dimensional map that uses color combinations to represent different categories or levels of magnitude. You can fill geographical boundaries (for example, a country or a state) on a map with color, based on measure values that are aggregated to the level defined by a geographical boundary.

Geo region maps do not support ZIP code data.

Figure A2.36 Geo Region Map



Gauges

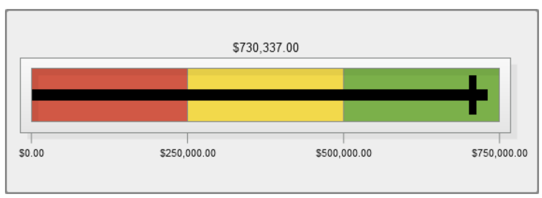
Bullet Gauges

A bullet gauge is a dashboard indicator that compares an actual value to a target value and compares them in intervals. The actual value of the primary measure is indicated by an inset horizontal bar.

Note that the scale of a bullet gauge often begins at zero, but it can contain both positive and negative values if both types of values apply to the primary measure, such as profit. The inset horizontal bar should always begin at zero so that comparing multiple bullet graphs is not confusing.

The bullet gauge requires a primary measure and a range display rule, and the target measure is optional. The default orientation is horizontal with an option to display the gauge vertically.

Figure A2.37 *A Bullet Gauge*



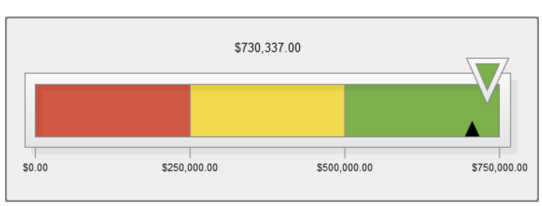
Slider Gauges

A slider gauge is a dashboard indicator that compares an actual value to a target value and compares them in intervals. The actual value of the primary measure is indicated by a downward-facing arrow. The target value is indicated by a small upward-facing arrow.

Like a bullet gauge, a slider gauge is oriented horizontally by default, but you can change the gauge's orientation to vertical. You should use a slider gauge when the numeric scale does not start at zero.

The slider gauge supports dates (which are continuous) and numeric categories.

Figure A2.38 *A Slider Gauge*



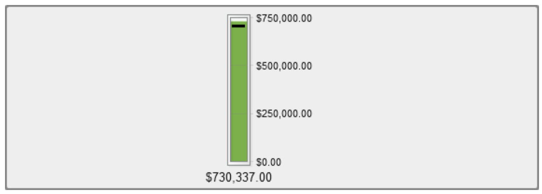
Thermometer Gauges

A thermometer gauge is a dashboard indicator that compares an actual value to a target value and compares them in intervals. The actual value of the primary measure is indicated by a vertical bar. The target value is indicated by a small black line.

The thermometer gauge requires a primary measure value and a range-based display rule. A target measure value is optional. The entire vertical bar is colored conditionally based on one color from the display rule.

The base of a thermometer bar should always start at zero. You can set this by defining your first range display rule to begin at zero. The designer always shows the base of the bar at the bottom of the thermometer.

Figure A2.39 A Thermometer Gauge



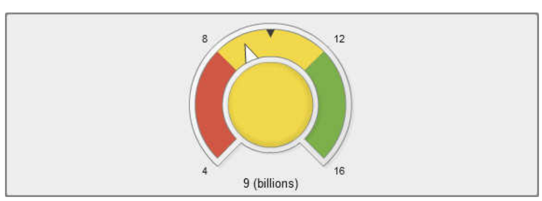
Dial Gauges

A dial gauge is an arc-shaped dashboard indicator that compares an actual value to a target value and compares them in intervals. The actual value of the primary measure is indicated by an arrow that points outward from the inner circle. The target value is indicated by an arrow that points inward from the outer arc. The color of the center circle is the color associated with the primary measure value's range interval.

The dial gauge requires a primary measure value and a range-based display rule. For more information, see "Adding Gauge-Level Display Rules" on page 415.

A target measure value is optional.

Figure A2.40 A Dial Gauge

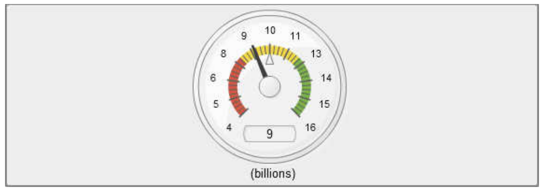


Speedometer Gauges

A speedometer gauge is a circular dashboard indicator that compares an actual value to a target value and compares them in intervals. The actual value of the primary measure is indicated by the larger pointer. The target value is indicated by a small triangle along the quantitative scale, either pointing inward or outward, depending on the **KPI skin** option for the gauge.

A speedometer gauge requires a primary measure value and a range-based display rule. A target measure value is optional.

Figure A2.41 A Speedometer Gauge



Word Clouds

A word cloud displays a set of category values as text, grouped in a cloud-like shape. Depending on the word cloud data roles, the size of the text indicates the frequency of a category value or the value of a measure that corresponds to a category value.

Figure A2.42 Word Cloud

Appendix 3

Editing a Data Expression in Text Mode

Overview of Text Mode

In SAS Visual Analytics, you create and edit filters, calculated items, and aggregated items by using an expression editor. The **Text** tab of the expression editor enables you to edit the expression as text.

You can add operators and data items to your expression by dragging and dropping them onto the expression or by entering the names of the operators or data items.

As you enter text, a list of suggestions appears. For example, if you enter `date`, then a drop-down list appears and enables you to select any of the operators and data items whose names begin with “date.”

When you make changes to your expression, it is automatically evaluated to determine whether it is valid. If the expression is not valid, then an error appears on the **Messages** tab, and the **OK** button is disabled. There might be a brief delay as your expression is evaluated.

Specifying Operator Parameters

When you add an operator to the expression, any parameters that are required by the operator are represented between braces { }. For example, if you add the **x – y** operator, then your expression appears as {Number} — {Number}.

Each parameter value that you enter should replace the entire string between the braces, including the brace characters. For example, you might replace {Number} with 12 or with a data item such as Expenses.

You can automatically select the next operator in the expression by pressing Ctrl + Shift + spacebar.

Specifying Data Item Names and Global Parameter Names

Data item names and global parameter names can be entered as plain text and are not case sensitive. You can enter names formally by using the format '*data-item-name*'n for a data item, or '*parameter-name*'p for a global parameter. If you switch to the **Visual** tab, then all of your names are converted to the formal format.

Note: If a name contains quotation marks, then you must use the \ character to escape the quotes.

Note: If a name contains spaces, then you must use the formal format.

Note: If a data item or global parameter has the same name as an operator, then you must use the formal format. For example, if you have a category named Year, then enter the name as **`'Year'n`** to avoid conflict with the **Year** operator.

Using Formatted and Unformatted Values

By default, category values and discrete numeric and date values are evaluated as formatted values. Continuous numeric values are evaluated as unformatted values.

To override this default behavior, you can add `[raw]` (to use unformatted values) or `[formatted]` (to use formatted values) to the right of the data item.

For example,

```
'Expenses'n[formatted]
```

specifies that the Expenses measure is evaluated as a formatted value.

Specifying String Values

To enter a string value, you can enter the string between single quotes or double quotes. If your value contains a quotation mark, use the \ character to escape the quote. If you use double quotes to enclose the string, then you do not need to escape the single quotes. If you use single quotes to enclose the string, then you do not need to escape the double quotes. For example, `"O'Reilly"`, `'O\'Reilly'`, and `'"Hello"'` are all valid.

To enter a string that contains a newline character, use **`\r`**, **`\n`**, or both to specify the newline character.

Specifying Date, Time, and Datetime Values

For date, datetime, and time values, specify a formatted value in quotes, followed by the letter “d” for a date value, the letters “dt” for a datetime value, or the letter “t” for a time value.

For time values, a value with a leading underscore is a reference to a specific time of the day.

A time value without a leading underscore specifies a period of elapsed time.

Here are some examples of date, datetime, and time values:

Table A3.1 Example Date, Datetime, and Time Values

Date	'23JUN2013'd 'JUN2013'd '2013'd 'q32013'd

Datetime	'23JUN2013_5:23:55’dt '23JUN2013_5:23'dt '23JUN2013'dt
Time	'_12:23:55't '_12:23't
Elapsed Time	'44:23:55't '00:01:34't

Specifying Aggregated Values

For aggregated values, specify the format, *aggregation-type* [*context*] (*value*), where *context* specifies one of the following aggregation contexts:

ByGroup
: calculates the aggregation for each subset of the data item that is used in a visualization. For example, in a bar chart, an aggregated measure with the **ByGroup** context calculates a separate aggregated value for each bar in the chart.

ForAll
: calculates the aggregation for the entire data item (after filtering). For example, in a bar chart, an aggregated measure with the **ForAll** context uses the same aggregated value (calculated for the entire data item) for each bar in the chart.

See “Aggregated (Simple) Operators” on page 589 for a list of the aggregation types that are available.

For example,

```
sum [bygroup] ('cost'n)
```

aggregates the sum of the measure COST for each BY-group value.

Specifying a Missing Value

Use a period character (.) to specify a missing numeric or date value. Use empty quotes ('') to specify a missing string value.

Appendix 4

Aggregations for Measures

The aggregation that is assigned to a measure determines how its values are summarized in a visualization or report object.

For example, in a bar chart of Sales by Quarter, each bar represents the aggregated values of the Sales measure for a specific quarter. If the aggregation for Sales is **Sum**, then the bars represent the sum (total) of sales for each quarter. If the aggregation for Sales is **Average**, then the bars represent the average sales for each quarter.

Note: Some aggregation types can override the data format that is used to display values in a visualization or report object. For example, if a measure has the Currency format with zero decimal places of precision, and you apply the **Variance** aggregation, then the values are displayed using the Comma format with two decimal places of precision instead.

You can specify the following aggregations for your measures:

Sum
: calculates the sum (total) of the values of a measure.

Average
: calculates the average (mean) value of a measure.

Standard Deviation
: calculates the standard deviation of a measure.

Standard Error
: calculates the standard error of the mean of a measure.

Variance
: calculates the variance of a measure.

Count
: calculates the total number of nonmissing values of a measure.

Number Missing
: calculates the number of missing values in a measure.

Minimum
: calculates the smallest value of a measure.

First Quartile
: calculates the first quartile of a measure.

Median
: calculates the median value of a measure.

Third Quartile
: calculates the third quartile of a measure.

Maximum
calculates the largest value of a measure.

Skewness
calculates the skewness of a measure. Skewness indicates the distribution of values. A positive value indicates that the distribution is heavier for values greater than the mean. A negative value indicates that the distribution is heavier for values less than the mean.

Kurtosis
calculates the kurtosis of a measure. The kurtosis value indicates how peaked the distribution is. A larger value indicates a more sharply peaked distribution. A smaller value indicates a flatter distribution.

Coefficient of Variation
calculates the coefficient of variation of a measure. The coefficient of variation is the ratio of the standard deviation to the mean.

Uncorrected Sum of Squares
calculates the uncorrected sum of squares of a measure. The uncorrected sum of squares is the sum of the squared values.

Corrected Sum of Squares
calculates the corrected sum of squares of a measure. The corrected sum of squares is the sum of the squared deviations from the mean.

T-statistic (for Average = 0)
calculates the Student's *t* statistic for a measure, assuming a mean value of zero.

P-value (for T-statistic)
calculates the probability of observing the *t* statistic value or a more extreme value. A small value indicates that the mean is likely not equal to zero.

Appendix 5

Operators for Data Expressions

Overview of Operators for Data Expressions

In the explorer and the designer, you can calculate data items and create filters by using expressions that contain operators.

Numeric (Simple) Operators

-x

returns a value with the opposite sign of the input value.

For example, `--1` returns 1 and `-1` returns -1.

x - y
: subtracts the second value from the first value.

  For example, `2 - 1` returns 1.

x * y
: multiplies the first and second values together.

  For example, `2 * 3` returns 6.

x / y
: divides the first value by the second value.

  For example, `6 / 2` returns 3.

x + y
: adds the first and second values together.

  For example, `1 + 2` returns 3.

Comparison Operators

BetweenExclusive
: returns true if the first value is within the range defined by the second and third values (excluding the bounding values).

  For example, `X BetweenExclusive(50, 100)` returns true if X is greater than 50 and less than 100.

BetweenInclusive
: returns true if the first value is within the range defined by the second and third values (including the bounding values).

  For example, `X BetweenInclusive(50, 100)` returns true if X is greater than or equal to 50 and less than or equal to 100.

In
: returns true if the first value is in the list specified by the second parameter. Select your list by choosing the values from the drop-down list or the selector window.

  For example, `X In ('A', 'B', 'C')` returns true when the value of X is either A, B, or C.

  Note: This operator cannot be used to compare measures.

Missing
: returns true if the value is a missing value.

  For example, `X Missing` returns true if the value of X is missing.

NotBetweenExclusive
: returns true if the first value is outside the range defined by the second and third values (excluding the bounding values).

  For example, `X NotBetweenExclusive(50, 100)` returns true if X is less than 50 or greater than 100.

NotBetweenInclusive
: returns true if the first value is outside the range defined by the second and third values (including the bounding values).

For example, `X NotBetweenInclusive(50, 100)` returns true if X is less than or equal to 50 or greater than or equal to 100.

NotIn

returns true if the first value is not in the list specified by the second parameter. Select your list by choosing the values from the drop-down list or the selector window.

For example, `X NotIn ('A', 'B', 'C')` returns true when the value of X is not A, B, or C.

Note: This operator cannot be used to compare measures.

NotMissing

returns true if the value is not a missing value.

For example, `X NotMissing` returns true if the value of X is not missing.

x < y

returns true if the first value is less than the second value.

x <= y

returns true if the first value is less than or equal to the second value.

x <> y

returns true if the first value is not equal to the second value.

x = y

returns true if the first value is equal to the second value.

x > y

returns true if the first value is greater than the second value.

x >= y

returns true if the first value is greater than or equal to the second value.

Boolean Operators

AND

joins two conditions and returns true if both conditions are true.

For example,

```
(1 = 1) AND (2 = 2)
```

returns true, and

```
(1 = 1) AND (2 = 1)
```

returns false.

IF... ELSE

returns different values, depending on whether the condition is true. The first parameter specifies the condition. The second parameter specifies the value to return if the condition is true. The third parameter specifies the value to return if the condition is false.

For example,

```
if (X > Y) return X else Y
```

returns the value of X if X is greater than Y, but returns the value of Y otherwise.

Note: Starting in the 7.1 release, the IF... ELSE operator can also be used in report filters in the designer.

NOT
: returns true if the condition is false.

For example, `not (1 = 2)` returns true.

OR
: joins two conditions and returns true if either condition is true.

For example,

```
(1 = 1) OR (2 = 2)
```

returns true, and

```
(1 = 1) OR (2 = 1)
```

returns true.

Numeric (Advanced) Operators

Abs
: returns the absolute value of the input value.

For example, `Abs(-3)` returns 3.

Ceil
: rounds the input value up to the nearest integer.

For example, `Ceil(4.2)` returns 5 and `Ceil(-4.8)` returns -4.

Exp
: raises the constant *e* to the power specified by the input value.

For example, `Exp(5)` returns *e* to the 5th power (148.41).

Floor
: rounds the input value down to the nearest integer.

For example, `Floor(4.8)` returns 4 and `Floor(-4.2)` returns -5.

Ln
: returns the natural logarithm (base *e*) of the input value.

For example, `Ln(10)` returns the *e*th root of 10 (2.30...).

Log
: returns the logarithm of the first value, where the second value specifies the base.

For example, `64 Log 8` returns the base 8 logarithm of 64 (2).

Mod
: returns the remainder after dividing the first value by the second value.

For example, `5 Mod 2` returns 1.

Power
: raises the first value to the power of the second value.

For example, `5 Power 2` returns 5 to the 2nd power (25).

Root
returns the *n*th root of the first value, where the second value specifies *n* (the base of the root).

For example, `27 Root 3` returns the 3rd (cube) root of 27 (3).

Round
rounds the first value to the number of decimal places that is specified by the second value. Select the second value from the drop-down list.

For example, `7.354 Round 2` returns 7.35.

Note: If you select 0 decimal places, then the values are rounded to the nearest integer.

TreatAs
allows a numeric, date, or datetime value to be used as a different data type within other operators. Select one of the following:

Date
allows the value to be used as a date.

Datetime
allows the value to be used as a datetime value.

Number
allows the value to be used as a number.

Time
allows the value to be used as a time value.

The value is treated as a raw value instead of being converted. Date values are the number of days since 01JAN1960. Datetime values are the number of seconds since 01JAN1960. Time values are the number of seconds since midnight.

For example, `TreatAs(_Date_, 19600)` returns 30AUG2013 as a date value.

Note: The TreatAs operator is useful for calculating elapsed time between two datetime values. For example, `(TreatAs(_Number_, '23OCT2013'd) – TreatAs(_Number_, '15JAN2013'd))` calculates the number of days between 15JAN and 23OCT, which is 281.

Trunc
truncates the input value to an integer.

For example, `Trunc(8.9)` returns 8 and `Trunc(-8.9)` returns -8.

Date and Time Operators

Note: Date and time operators are not supported for aggregated items.

DateFromMDY
creates a date value from separate month, day, and year values. The first value specifies the month as a number from 1–12. The second value specifies the day as a number from 1–31. The third value specifies the year as a four-digit number.

For example, `DateFromMDY(1, 15, 2013)` returns 15JAN2013.

DateFromYQ
: creates a date value from separate year and quarter values. The first value specifies the year as a four-digit number. The second value specifies the quarter as a number from 1–4.

  For example, `DateFromYQ(2013, 1)` returns 01JAN2013.

  Note: The date is generated using the first day of each quarter.

DatePart
: converts a datetime value to a date value.

  For example, `DatePart('15JAN2013_17:15'dt)` returns 15JAN2013.

DateTimeFromDateHMS
: creates a datetime value from a date value and separate hour, minute, and second values. The first value specifies the date. The second value specifies the hour as a number from 0–23. The third value specifies the minute as a number from 0–59. The fourth value specifies the second as a number from 0–59.

  For example, `DateTimeFromDateHMS('15JAN2013'd, 17, 15, 23)` returns January 15, 2013 05:15:23 PM

DateTimeFromTimeMDY
: creates a datetime value from a time value and separate month, day, and year values. The first value specifies the time. The second value specifies the month as a number from 1–12. The third value specifies the day as a number from 1–31. The fourth value specifies the year as a four-digit number.

  For example, `DateTimeFromTimeMDY('_17:15:23'dt, 1, 15, 2013)` returns January 15, 2013 05:15:23 PM.

DayOfMonth
: returns the day of the month from a date value as a number from 1–31.

  For example, `DayOfMonth('15JAN2013'd)` returns 15.

DayOfWeek
: returns the day of the week from a date value as a number from 1–7 (1 is Sunday).

  For example, `DayOfWeek('15JAN2013'd)` returns 3 (Tuesday).

DayOfYear
: returns the day of the year from a date value as a number from 1–366.

  For example, `DayOfYear('15FEB2013'd)` returns 46.

Hour
: returns the hour from a time or datetime value as a number from 0–23.

  For example, `Hour('17:15:23't)` returns 17.

Minute
: returns the minute from a time or datetime value as a number from 0–59.

  For example, `Minute('17:15:23't)` returns 15.

Month
: returns the month from a date value as a number from 1–12.

  For example, `Month('15JAN2013'd)` returns 1.

Now
: creates a datetime value from the current date and time.

For example, `Now()` returns the current date and time.

Quarter
: returns the quarter from a date value as a number from 1–4.

For example, `Quarter('15AUG2013'd)` returns 3.

Second
: returns the second from a time or datetime value as a number from 0–59.

For example, `Second('17:15:23't)` returns 23.

TimeFromHMS
: creates a time value from separate hour, minute, and second values. The first value specifies the hour as a number from 0–23. The second value specifies the minute as a number from 0–59. The third value specifies the second as a number from 0–59.

For example, `TimeFromHMS(17, 15, 23)` returns 05:15:23 PM.

TimePart
: converts a datetime value to a time value.

For example, `TimePart('15JAN2013_17:15:23'dt)` returns 05:15:23 PM.

WeekNumber
: returns the week of the year as a number from 0–53, where week 1 begins on the first Sunday of the year. Dates before the first Sunday of the year return 0.

For example, `WeekNumber('04AUG2013'd)` returns 31.

Year
: returns the year from a date value as a four-digit number.

For example, `Year('15JAN2013'd)` returns 2013.

Aggregated (Simple) Operators

Avg
: calculates the average (mean) value of a measure.

Count
: calculates the total number of nonmissing values of a measure.

Distinct
: calculates the number of distinct values in a category. If the category contains missing values, then the distinct count is increased by one.

Max
: calculates the largest value of a measure.

Median
: calculates the median value of a measure.

Min
: calculates the smallest value of a measure.

NumMiss
: calculates the number of missing values in a data item.

Q1
: calculates the first quartile of a measure.

Q3
: calculates the third quartile of a measure.

StdDev
: calculates the standard deviation of a measure.

StdErr
: calculates the standard error of the mean of a measure.

Sum
: calculates the sum (total) of the values of a measure.

Var
: calculates the variance of a measure.

Aggregated (Advanced) Operators

CoefVar
: calculates the coefficient of variation of a measure. The coefficient of variation is the ratio of the standard deviation to the mean.

CSS
: calculates the corrected sum of squares of a measure. The corrected sum of squares is the sum of the squared deviations from the mean.

First
: calculates the first value of a measure based on chronological order. The first parameter specifies the measure. The second parameter specifies the sequence data item that is used to determine the chronological order. The sequence data item can be either a date or time data item or a numeric data item. The third parameter specifies whether missing values are included. Select _IncludeMissing_ to include missing values or select _ExcludeMissing_ to exclude missing values.

  Note: If there are multiple measure values for the first value of the sequence data item, then the minimum measure value is selected.

  Note: The First aggregation always calculates measure values by using the sequence data item that you specify. If your visualization or report object uses a different date or time data item, then the results might be misleading to viewers who do not know the expression for the aggregated data item.

Kurtosis
: calculates the kurtosis of a measure. The kurtosis value indicates how peaked the distribution is. A larger value indicates a more sharply peaked distribution. A smaller value indicates a flatter distribution.

Last
: calculates the last value of a measure based on chronological order. The first parameter specifies the measure. The second parameter specifies the sequence data item that is used to determine the chronological order. The sequence data item can be either a date or time data item or a numeric data item. The third parameter specifies whether missing values are included.

Select _IncludeMissing_ to include missing values or select _ExcludeMissing_ to exclude missing values.

Note: If there are multiple measure values for the last value of the sequence data item, then the minimum measure value is selected.

Note: The Last aggregation always calculates measure values by using the sequence data item that you specify. If your visualization or report object uses a different date or time data item, then the results might be misleading to viewers who do not know the expression for the aggregated data item.

Percentile
: calculates the specified percentile of a measure. Specify a number between 0 and 100. For example, 85 specifies the 85th percentile, the value for which 85% of the values are lower.

PvalT
: calculates the probability of observing the *t* statistic value or a more extreme value. A small value indicates that the mean is likely not equal to zero.

Skewness
: calculates the skewness of a measure. Skewness indicates the distribution of values. A positive value indicates that the distribution is heavier for values greater than the mean. A negative value indicates that the distribution is heavier for values less than the mean.

TStat
: calculates the Student's *t* statistic for a measure, assuming a mean value of zero.

USS
: calculates the uncorrected sum of squares of a measure. The uncorrected sum of squares is the sum of the squared values.

Periodic Operators

About Periodic Operators

Periodic operators aggregate values over a period of time.

If you assign a periodic aggregated item to a visualization or report object that contains dates, the aggregated item displays the aggregated values for each time period in the visualization or report object.

In a visualization or report object that does not contain dates, the aggregated item displays values that use today's date as a reference. If the date data item for the operator does not contain data for the interval that contains today's date, then the operator returns missing values.

Periodic operators are evaluated using time intervals. Intervals specify whether the aggregation is applied on a monthly basis, a quarterly basis, and so on. You can specify a specific interval, or you can specify that the interval is inferred. For an inferred interval, the aggregation is evaluated based on its context in the visualization or report object. For example, if your visualization contains a bar chart of sales by month, then the inferred interval is monthly.

Note: Periodic operators return a missing value in the following scenarios:

- Data does not exist for the specified time period.
- The date data item for the period calculation does not match the date data item in the visualization or report object. You must use the same date data item or a duplicate data item that is based on the same data item.
- The interval for the operator is smaller than the interval of the date format in the visualization or report object (for example, if your interval is by month, but the date format is Year).
- For operators that use inner and outer intervals, the inner interval is larger than the outer interval.
- The inferred interval is by week of the year or by an interval smaller than a day.
- The inferred interval is by day for any operator that has an offset other than 0.

  For the explorer, the ParallelPeriod and RelativePeriod operators always return a missing value when the inferred interval is by day.

The following periodic operators are available:

Table A5.1 *Periodic Operators*

CumulativePeriod	returns the aggregated value for a period of time and all of the previous periods of time within a larger period of time (for example, the year-to-date total of monthly values).
ParallelPeriod	returns the aggregated value for a period of time that is parallel to the current period of time (for example, the value for the same month of the previous year).
Period	returns the aggregated value for a period of time (for example, the value for the current month).
PeriodWithDate	returns the aggregated value for a specific, constant period of time (for example, the value for the month that includes 15OCT2013).
RelativePeriod	returns the aggregated value for a period of time that is relative to the current period (for example, the value for the previous month of the same year).

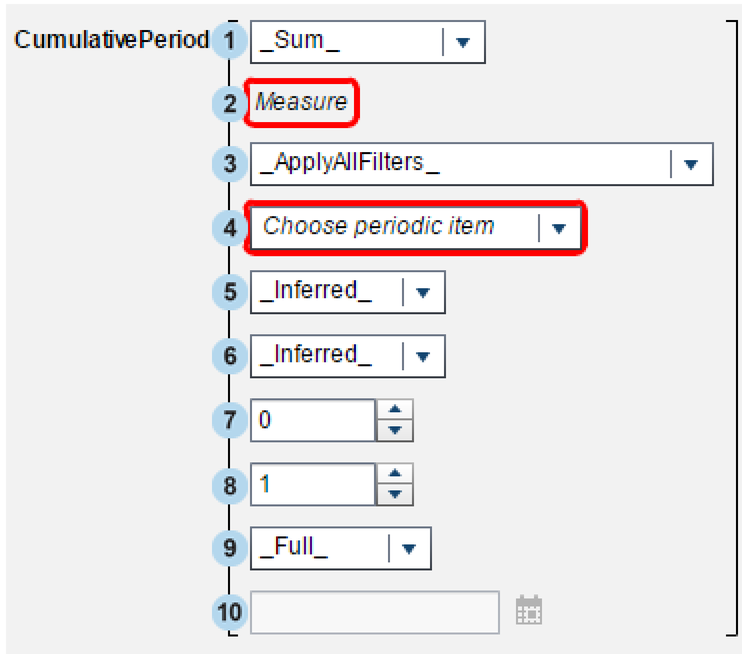
CumulativePeriod

The CumulativePeriod operator returns aggregated values for a period of time and all of the previous periods, within a larger period of time (for example, the year-to-date total of monthly values).

Note: The CumulativePeriod operator resets at the beginning of each calendar year. You cannot have a CumulativePeriod operator with a date range that overlaps two calendar years.

Specify the following parameters:

Figure A5.1 Parameters for the CumulativePeriod Operator



1 The aggregation that is applied to the measure.

2 The measure to aggregate over time.

3 Which time filters should be applied before processing the aggregated measure. Select one of the following:

ApplyAllFilters
: applies all filters before processing the aggregated measure.

IgnoreAllTimeFrameFilters
: ignores all filters that are based on the same date data item as the periodic operator.

IgnoreInteractiveTimeFrameFilters
: ignores all interactive filters (from prompts and interactions) that are based on the same date data item as the periodic operator.

Note: This parameter is not available in the explorer.

4 The date data item for the period calculation. Only data items whose formats specify year are available.

5 The inner interval (smaller time period) for which the values are aggregated. For example, specify _ByMonth_ as the inner interval and _ByYear_ as the outer interval to aggregate the year-to-date values for each month.

Select one of the following:

Inferred
: specifies that the interval is determined automatically from the visualization or report object that displays the aggregated item.

ByMonth
: specifies a monthly interval.

ByQuarter
specifies a quarterly interval.

ByYear
specifies a yearly interval.

6 The outer interval (larger time period) that provides the context for the cumulative aggregation. For example, specify _ByMonth_ as the inner interval and _ByYear_ as the outer interval to aggregate the year-to-date values for each month.

Select one of the following:

Inferred
specifies that the interval is determined automatically from the visualization or report object that displays the aggregated item.

ByMonth
specifies a monthly interval.

ByQuarter
specifies a quarterly interval.

ByYear
specifies a yearly interval.

7 The number of outer intervals to offset from the current period. 0 specifies that the period from the current outer interval is used. A negative value indicates a previous interval.

For example, if your inner interval is by month and your outer interval is by year, then -1 specifies the year-to-date monthly values for the previous year.

8 The starting point for each new outer period. For example, if your inner interval is by month and your outer interval is by year, then 3 specifies that each year begins in the third month.

Note: This parameter is not available in the explorer.

Note: If the outer interval is inferred or by year, then the value must be an interval between 1 and 12. If the outer interval is by quarter, then the value must be an interval between 1 and 4.

9 The scope for the period. The scope specifies how much of each period is aggregated.

Select one of the following:

Full
aggregates the values for the entire period.

ToDate
aggregates only the values up to a specific day of the outer interval.

ToToday
aggregates only the values up to the equivalent of today's position in the current interval. For example, if today is the 40th day of the quarter, and the outer interval is by quarter, then only the values up to the 40th day of each quarter are used.

The value for today is evaluated dynamically whenever the aggregated item is viewed in a visualization or report object.

10 If you select _ToDate_ as the scope, then select the date that is used to subset each period.

For example, if you select 09NOV2013, and the outer interval is by year, then only the values up to November 9 of each year are used in the aggregation.

For example,

```
CumulativePeriod(_Sum_, 'Expenses'n, _ApplyAllFilters_, 'Date'n, _ByMonth_, _ByYear_, 0, 1,
```

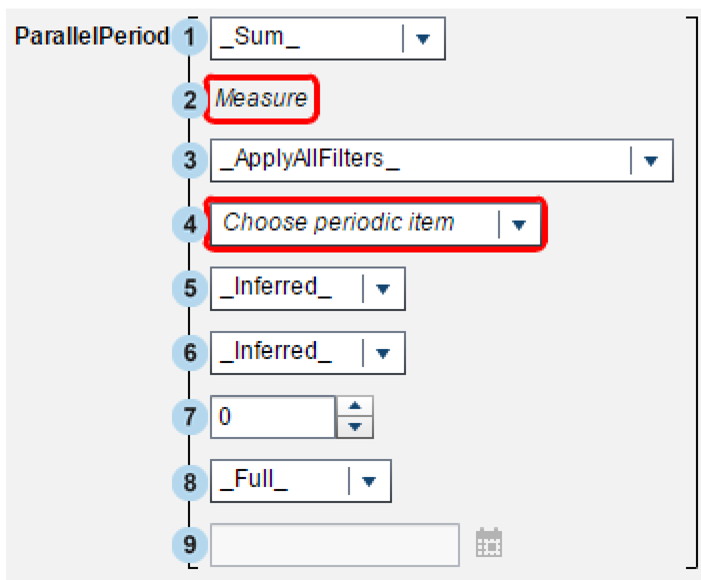
aggregates the sum of year-to-date monthly values for the Expenses measure using date values from the Date data item.

ParallelPeriod

The ParallelPeriod operator returns aggregated values for a period of time that is parallel to the current period (for example, the value for the same month of the previous year).

Specify the following parameters:

Figure A5.2 *Parameters for the ParallelPeriod Operator*



1. The aggregation that is applied to the measure.
2. The measure to aggregate over time.
3. Which time filters should be applied before processing the aggregated measure. Select one of the following:

   ApplyAllFilters
   : applies all filters before processing the aggregated measure.

   IgnoreAllTimeFrameFilters
   : ignores all filters that are based on the same date data item as the periodic operator.

   IgnoreInteractiveTimeFrameFilters
   : ignores all interactive filters (from prompts and interactions) that are based on the same date data item as the periodic operator.

   Note: This parameter is not available in the explorer.

4 The date data item for the period calculation. Only data items whose formats specify year are available.

5 The inner interval (smaller time period) for which the values are aggregated. For example, specify _ByMonth_ as the inner interval to aggregate the values for each month.

Select one of the following:

Inferred
: specifies that the interval is determined automatically from the visualization or report object that displays the aggregated item.

ByMonth
: specifies a monthly interval.

ByQuarter
: specifies a quarterly interval.

ByYear
: specifies a yearly interval.

6 The outer interval (larger time period) that provides the context for the parallel period aggregation. For example, specify _ByMonth_ as the inner interval and _ByYear_ as the outer interval to aggregate the monthly values for a different year.

Select one of the following:

Inferred
: specifies that the interval is determined automatically from the visualization or report object that displays the aggregated item.

ByMonth
: specifies a monthly interval.

ByQuarter
: specifies a quarterly interval.

ByYear
: specifies a yearly interval.

7 The number of outer intervals to offset from the current period. 0 specifies that the period from the current outer interval is used. A negative value indicates a previous interval.

For example, if your inner interval is by month and your outer interval is by year, then -1 specifies the monthly values for the previous year.

8 The scope for the period. The scope specifies how much of each period is aggregated.

Select one of the following:

Full
: aggregates the values for the entire period.

ToDate
: aggregates only the values up to a specific day of the outer interval.

ToToday
: aggregates only the values up to the equivalent of today's position in the current interval. For example, if today is the 40th day of the quarter, and the outer interval is by quarter, then only the values up to the 40th day of each quarter are used.

The value for today is evaluated dynamically whenever the aggregated item is viewed in a visualization or report object.

9 If you select _ToDate_ as the scope, then select the date that is used to subset each period.

   For example, if you select 09NOV2013, and the outer interval is by year, then only the values up to November 9 of each year are used in the aggregation.

For example,

```
ParallelPeriod(_Sum_, 'Expenses'n, _ApplyAllFilters_, 'Date'n, _ByMonth_, _ByYear_, -1, _Ful
```

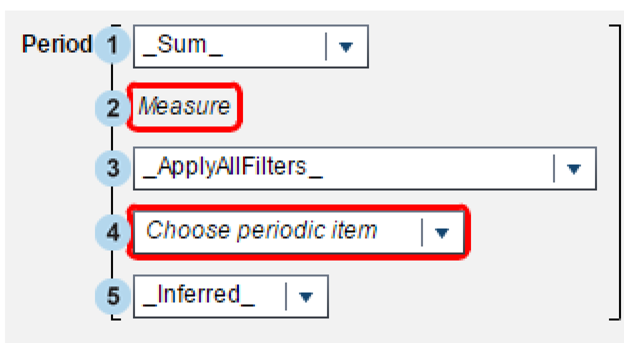
aggregates the sum of monthly values for the Expenses measure for the previous year using date values from the Date data item.

Period

The Period operator returns aggregated values for a period of time (for example, the value for the current month).

Specify the following parameters:

Figure A5.3 *Parameters for the Period Operator*



1 The aggregation that is applied to the measure.

2 The measure to aggregate over time.

3 Which time filters should be applied before processing the aggregated measure. Select one of the following:

   ApplyAllFilters
      applies all filters before processing the aggregated measure.

   IgnoreAllTimeFrameFilters
      ignores all filters that are based on the same date data item as the periodic operator.

   IgnoreInteractiveTimeFrameFilters
      ignores all interactive filters (from prompts and interactions) that are based on the same date data item as the periodic operator.

   Note: This parameter is not available in the explorer.

4 The date data item for the period calculation. Only data items whose formats specify year are available.

5 The interval for which the values are aggregated. For example, specify _ByMonth_ as the interval to aggregate the values for each month.

Select one of the following:

Inferred
: specifies that the interval is determined automatically from the visualization or report object that displays the aggregated item.

ByMonth
: specifies a monthly interval.

ByQuarter
: specifies a quarterly interval.

ByYear
: specifies a yearly interval.

For example,

```
Period(_Sum_, 'Expenses'n, _ApplyAllFilters_, 'Date'n, _ByMonth_)
```

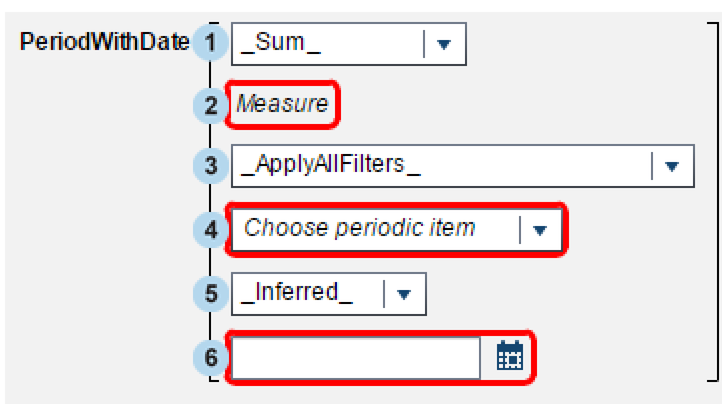
aggregates the sum of monthly values for the Expenses measure using date values from the Date data item.

PeriodWithDate

The PeriodWithDate operator returns aggregated values for a specific, constant period of time (for example, the value for the month that includes 15OCT2013).

Specify the following parameters:

Figure A5.4 Parameters for the PeriodWithDate Operator



1. The aggregation that is applied to the measure.
2. The measure to aggregate over time.
3. Which time filters should be applied before processing the aggregated measure. Select one of the following:

   ApplyAllFilters
   : applies all filters before processing the aggregated measure.

   IgnoreAllTimeFrameFilters
   : ignores all filters that are based on the same date data item as the periodic operator.

   IgnoreInteractiveTimeFrameFilters
   : ignores all interactive filters (from prompts and interactions) that are based on the same date data item as the periodic operator.

Note: This parameter is not available in the explorer.

4 The date data item for the period calculation. Only data items whose formats specify year are available.

5 The interval for which the values are aggregated. For example, specify _ByMonth_ as the interval to aggregate the values for each month.

Select one of the following:

Inferred
: specifies that the interval is determined automatically from the visualization or report object that displays the aggregated item.

ByMonth
: specifies a monthly interval.

ByQuarter
: specifies a quarterly interval.

ByYear
: specifies a yearly interval.

6 The reference date for the period aggregation.

For example,

```
PeriodWithDate(_Sum_, 'Expenses'n, _ApplyAllFilters_, 'Date'n, _ByMonth_, '15OCT2013'd)
```

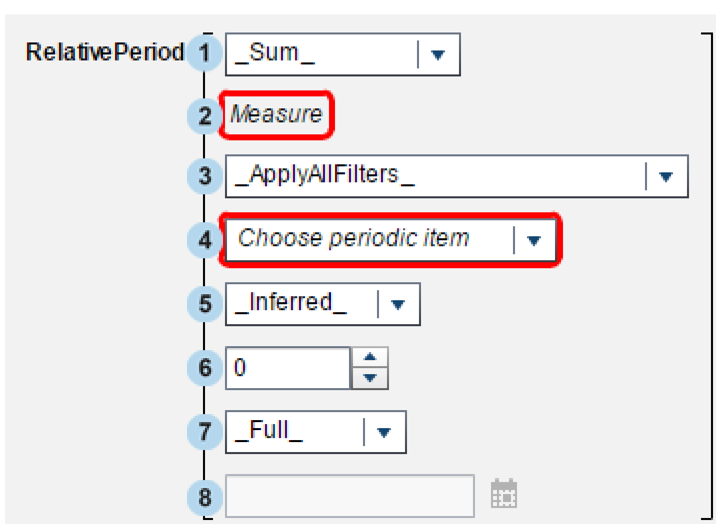
aggregates the sum of monthly values for the Expenses measure using date values from the Date data item.

RelativePeriod

The RelativePeriod operator returns aggregated values for a period of time that is relative to the current period (for example, the previous month of the same year).

Specify the following parameters:

Figure A5.5 Parameters for the RelativePeriod Operator



1 The aggregation that is applied to the measure.

2 The measure to aggregate over time.

3 Which time filters should be applied before processing the aggregated measure. Select one of the following:

ApplyAllFilters
: applies all filters before processing the aggregated measure.

IgnoreAllTimeFrameFilters
: ignores all filters that are based on the same date data item as the periodic operator.

IgnoreInteractiveTimeFrameFilters
: ignores all interactive filters (from prompts and interactions) that are based on the same date data item as the periodic operator.

Note: This parameter is not available in the explorer.

4 The date data item for the period calculation. Only data items whose formats specify year are available.

5 The interval for which the values are aggregated. For example, specify _ByMonth_ as the interval to aggregate the year-to-date values for each month.

Select one of the following:

Inferred
: specifies that the interval is determined automatically from the visualization or report object that displays the aggregated item.

ByMonth
: specifies a monthly interval.

ByQuarter
: specifies a quarterly interval.

ByYear
: specifies a yearly interval.

6 The number of intervals to offset from the current period. 0 specifies that the period from the current interval is used. A negative value indicates a previous interval.

For example, if your interval is by month, then -1 specifies the monthly values for the previous month.

7 The scope for the period. The scope specifies how much of each period is aggregated.

Select one of the following:

Full
: aggregates the values for the entire period.

ToDate
: aggregates only the values up to a specific day of the interval.

ToToday
: aggregates only the values up to the equivalent of today's position in the current interval. For example, if today is the 40th day of the quarter, and the outer interval is by quarter, then only the values up to the 40th day of each quarter are used.

: The value for today is evaluated dynamically whenever the aggregated item is viewed in a visualization or report object.

8. If you select _ToDate_ as the scope, then select the date that is used to subset each period.

   For example, if you select 09NOV2013 and the outer interval is by quarter, then only the values up to the 40th day of each quarter are used in the aggregation.

For example,

```
RelativePeriod(_Sum_, 'Expenses'n, _ApplyAllFilters_, 'Date'n, _ByMonth_, -1, _Full_)
```

aggregates the sum of monthly values for the Expenses measure for the previous month using date values from the Date data item.

Text (Simple) Operators

Note: All text operators are case sensitive.

Note: Text operators are not supported for aggregated items.

Concatenate
: appends the second input string to the first input string.

  For example, `Concatenate('A', 'B')` returns **`AB`**.

Contains
: specifies that a matching value must contain the specified string.

  For example, `'Catcher' Contains 'Cat'` returns true.

EndsWith
: specifies that a matching value must contain the specified string at the end of the value.

  For example, `'Catcher' EndsWith 'her'` returns true.

Format
: applies a SAS format to the input value. Click the format field to select the format that you want to apply. The output from the Format operator is a string.

  For example, `Format(1015.35, 'DOLLAR6.2')` returns **`$1,015.35`** as a string value.

  Note: Standard date formats in SAS Visual Analytics display date and datetime values in the locale of your browser. You can display date and datetime values in the locale of the data source by using national language formats. The names of national language formats begin with “NL.” For example, the NLDATE format displays date values by using the locale of the data source.

LowerCase
: changes all of the characters in a text string to lowercase.

  For example, `LowerCase('SAS INSTITUTE')` returns **`sas institute`**.

NotContains
: specifies that a matching value must not contain the specified string.

  For example, `'Catcher' NotContains 'Dog'` returns true.

Parse
: interprets a numeric or datetime value from the input string. Click the format field to select the format that is used to interpret the string. The output from the Parse operator is either a number or a datetime value, depending on the format that you select.

  For example, `Parse('15JAN2013', 'DATE9.')` returns 15JAN2013 as a date value.

StartsWith
: specifies that a matching value must contain the specified string at the start of the value.

  For example, `'Catcher' StartsWith 'Cat'` returns true.

UpCase
: changes all of the characters in the text string to uppercase.

  For example, `UpCase('sas institute')` returns **`SAS INSTITUTE`**.

Text (Advanced) Operators

Note: All text operators are case sensitive.

Note: Text operators are not supported for aggregated items.

FindChar
: finds the position of a character or any of a set of characters within a text string. The position of the first match is returned as a numeric value. If no matches are found, then 0 is returned. The first input string specifies the value to search within. The second input string specifies the list of characters to search for.

  For example, `FindChar('mystring', 'sz')` returns 3.

FindString
: finds the position of a string within another string. The position of the first match is returned as a numeric value. If no matches are found, then 0 is returned. The first input string specifies the value to search within. The second input string specifies the string to search for.

  For example, `FindString('mystring', 'st')` returns 3.

GetLength
: returns the length of an input string as a numeric value.

  For example, `GetLength('mystring')` returns 8.

GetWord
: returns a word from an input string where the words are separated by spaces, periods, or other special characters. The first parameter specifies the input string. The second parameter specifies the number of the word to return where 1 is the first word.

  For example, `GetWord('my test string', 2)` returns **`test`**.

RemoveBlanks
: removes space characters from the input string. The first parameter specifies the input string. The second parameter specifies which space characters to remove. Select one of the following:

All
: removes all spaces from the string.

Leading
: removes spaces at the beginning of the string.

LeadingAndTrailing
: removes spaces at the beginning and end of the string.

Trailing
: removes spaces at the end of the string.

For example, `RemoveBlanks('my test string', '_ALL_')` returns **`myteststring`**.

RemoveChars
: removes all instances of a set of characters from the input string. The first parameter specifies the input string. The second parameter specifies the list of characters to remove.

For example, `RemoveChars('my_test_string', '_')` returns **`myteststring`**.

RemoveWord
: removes a word from an input string where the words are separated by spaces or special characters. The first parameter specifies the input string. The second parameter specifies the number of the word to remove where 1 is the first word.

For example, `RemoveWord('my test string', 2)` returns **`my string`**.

Note: In addition to spaces, the following characters are used as delimiters in the input string: . < ( ) + & ! $ * ; ^ - / , % | ’

Replace
: replaces a substring within the input string with a replacement string. The first parameter specifies the input string. The second parameter specifies the substring to replace. The third parameter specifies the replacement string. The fourth parameter specifies which instances of the substring to replace. Select one of the following:

ALL
: replaces every instance.

FIRST
: replaces the first instance only.

LAST
: replaces the last instance only.

For example, `Replace('my test string test', 'test', 'new', '_ALL_')` returns **`my new string new`**.

ReplaceWord
: replaces a word from an input string where the words are separated by spaces, periods, or other special characters. The first parameter specifies the input string. The second parameter specifies the number of the word to replace where 1 is the first word. The third parameter specifies the replacement string.

For example, `ReplaceWord('my test string', 2, 'new')` returns **`my new string`**.

Reverse
: reverses the order of the characters in the input string.

For example, `Reverse('A B C')` returns **`C B A`**.

Substring
: returns a substring from the input string based on the position of the characters. The first parameter specifies the input string. The second parameter specifies the position of the first character to return. The third parameter specifies the number of characters to return.

For example, `Substring('my test string', 4, 3)` returns **`tes`**.

Update
: replaces a substring from the input string based on the position of the characters. The first parameter specifies the input string. The second parameter specifies the position of the first character to replace. The third parameter specifies the number of characters to replace. The fourth parameter specifies the replacement string.

For example, `Update('my test string', 4, 3, 'nex')` returns **`my next string`**.

URLDecode
: removes URL encoding from the input string. URL encoding replaces some characters with a % character followed by a two-digit hexadecimal code.

For example, `URLDecode('support.sas.com%2Fmy%20string')` returns **`support.sas.com/my string`**.

URLEncode
: applies URL encoding to the input string. URL encoding replaces some characters with a % character followed by a two-digit hexadecimal code.

For example, `URLEncode('support.sas.com/my string')` returns **`support.sas.com%2Fmy%20string`**.

Calculating Compound Annual Growth Rate

Compound annual growth rate (CAGR) is an investing and business term for the effective constant year-over-year rate of return that produces a target result value at the end of multiple years, assuming that the CAGR is compounded at the end of each year. For example, you might use CAGR to compare trends over multiple years in revenue or in the number of units sold. For SAS Visual Analytics, you can calculate the yearly CAGR using the expression builder.

This example compares trends in the growth rate for yearly sales amounts between different product types or regions.

The basic data items are:

`sales`
: This is a numeric measure with a currency format and a default aggregation of Sum.

`ProductType`
: This is a string category data item.

`RegionName`
This is a string category data item.

`TransactionDate`
This is a date data item with a format of Month, Day, Year (MMDDYYYY).

The duplicate data item is:

`TransactionDateYear`
This data item is a duplicate of the `TransactionDate` data item, but with the Year format.

The calculated data items are:

`BeginningYearNum`
This data item should be a numeric type with a `Float4.0` format and an aggregation of Minimum.

```
BeginningYearNum = Year('31DEC2010'd)
```

`EndingYearNum`
This data item should be a numeric type with a `Float4.0` format and an aggregation of Minimum.

```
EndingYearNum = Year(‘transactionDate’n)
```

The aggregated measure data items are:

`NumYears`
This data item has a `Float4.0` format.

```
NumYears = Min [_ByGroup_] ('EndingYearNum'n) - Min
[_ByGroup_] ('BeginningYearNum'n)
```

`BeginningValue`
This data item needs to be set to the same currency format as the `sales` data item.

```
BeginningValue = PeriodWithDate(_Sum_, 'sales'n,
'transactionDate'n, _ByYear_, '31DEC2010'd)
```

`EndingValue`
This data item needs to be set to the same currency format as the `sales` data item.

```
EndingValue = Period(_Sum_, 'sales'n, 'transactionDate'n,
_ByYear_)
```

`NormalizedRatio`
This data item has a `Float12.2` format.

```
NormalizedRatio = 'EndingValue'n / 'BeginningValue'n
```

`CAGR`
This data item has a `Percent` format.

```
CAGR = ('NormalizedRatio'n Power ( 1 / 'NumYears'n ) ) – 1
```

To use the CAGR, you should add `TransactionDateYear`, `CAGR`, and any other categories of interest (for example, `RegionName`, `ProductType`, and so on) to a list table, a crosstab, or a graph.

Appendix 6

Conditions for Filters

In SAS Visual Analytics, filters are based on expressions that contain operators. Conditions enable you to easily add the most common operators to your expression. Depending on the type of data that is used by the filter, you can select from the following categories of filter conditions:

Table A6.1 Conditions for Character Data

=	Specifies that a matching value must match one of the filter values exactly.
Contains	Specifies that a matching value must contain the filter value.
EndsWith	Specifies that a matching value must contain the filter value at the end of the value.
In	Specifies that a matching value is in the list that you select. To select your list, choose the values from the drop-down list.
Missing	Specifies that a missing value matches the filter.
NotContains	Specifies that a matching value must not contain the filter value.
NotIn	Specifies that a matching value is not in the list that you select. To select your list, choose the values from the drop-down list.
NotMissing	Specifies that a nonmissing value matches the filter.
StartsWith	Specifies that a matching value must contain the filter value at the start of the value.

Table A6.2 Conditions for Numeric Data and Date and Time Data

<>	Specifies that a matching value must not be equal to the filter value.
=	Specifies that a matching value must be equal to the filter value.

<	Specifies that a matching value must be less than the filter value.
<=	Specifies that a matching value must be less than or equal to the filter value.
>	Specifies that a matching value must be greater than the filter value.
>=	Specifies that a matching value must be greater than or equal to the filter value.
Between [exclusive]	Specifies that a matching value must be greater than the first filter value and less than the second filter value.
Between [inclusive]	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.
In	Specifies that a matching value is in the list that you select. To select your list, choose the values from the drop-down list. **Note:** This condition is not available for continuous numeric data.
Missing	Specifies that a missing value matches the filter.
NotBetween [exclusive]	Specifies that a matching value must be less than the first filter value or greater than the second filter value.
NotBetween [inclusive]	Specifies that a matching value must be less than or equal to the first filter value or less than or equal to the second filter value.
NotIn	Specifies that a matching value is not in the list that you select. To select your list, choose the values from the drop-down list. **Note:** This condition is not available for continuous numeric data.
NotMissing	Specifies that a nonmissing value matches the filter.

Appendix 7

Data Limits

Data Limits for SAS Visual Analytics Explorer

Some of the visualizations in SAS Visual Analytics Explorer (the explorer) have limits to the number of data values that they can display. The limit values are affected by the **Visualization data threshold** setting in the Preferences window. For more information about the Preferences window, see "Specify Your Preferences for the Explorer" on page 111.

The following table displays the data limits for visualizations in the explorer:

Table A7.1 *Data Limits for the Explorer*

Visualization Type	Variation	Behavior when Limit Is Exceeded	Default Threshold	Minimum Threshold	Maximum Threshold
Table	—	Paging is applied to the table.	10,000	1,000	100,000
Crosstab	—	An error message appears.	21,000,000 cells	1,000,000 cells	26,000,000 cells
Bar Chart	Not grouped or latticed.	The bar chart shows a subset that contains the first or last ranked values.	3,000	810	3,625
	Grouped or latticed.	An error message appears.	3,000	810	3,625

Visualization Type	Variation	Behavior when Limit Is Exceeded	Default Threshold	Minimum Threshold	Maximum Threshold
Line Chart	Single numeric or datetime category in the **Category** role.	The line chart shows a subset that contains the first or last ranked value.	10,000	5,630	11,250
	Single string category in the **Category** role.		4,000	1,380	4,750
	No string categories in the **Category** role, grouped or latticed.	An error message appears.	10,000	5,630	11,250
	One or more string categories in the **Category** role, grouped or latticed.		4,000	1,380	4,750
Scatter Plot	Two measures, not grouped.	The scatter plot is converted to a heat map.	40,000	9,375	48,750
	Two measures, grouped.	An error message appears.	40,000	9,375	48,750
	Three or more measures, not grouped.	The scatter plot is converted to a correlation matrix.	80,000 / number of measures	18,750 / number of measures	97,500 / number of measures
	Three or more measures, grouped.	An error message appears.	80,000 / number of measures	18,750 / number of measures	97,500 / number of measures
Bubble Plot	No categories.	The bubble plot is converted to a heat map.	25,000	7,500	30,000
	Grouped.	The bubble plot shows the top or bottom value by size.	500	150	600
	Latticed.	An error message appears.	1,050	4,900	50,000
	Grouped and latticed.		500	150	600
	Animated.		50,000	15,000	60,000
Network Plot	—	An error message appears.	8,000	100	10,000
			1,000 nodes, 1,000 links		
Sankey Diagram	—	An error message appears.	3,000 links	500 links	3,625 links
			8,020 rows	1,000 rows	10,000 rows

Visualization Type	Variation	Behavior when Limit Is Exceeded	Default Threshold	Minimum Threshold	Maximum Threshold
Histogram	—	—	None	None	None
Box Plot	—	An error message appears.	2,900 boxes	800 boxes	3,500 boxes
Heat Map	No categories.	—	None	None	None
	One or more categories.	An error message appears.	3,000	1,250	3,500
Geo Map	Bubble overlay.	The bubble overlay is converted to a coordinate overlay.	5,000	1,500	6,000
	Region overlay.	The region overlay is converted to a coordinate overlay.	5,000	1,500	6,000
	Coordinate overlay.	An error message appears.	82,000	10,000	100,000
Treemap	No additional levels.	The treemap shows a subset that contains the first or last ranked value.	4,900	1,050	6,000
	One or more additional levels.	The number of additional levels is reduced to 0. If the threshold is still exceeded, then the treemap shows a subset that contains the first or last ranked value.	4,900	1,050	6,000
Correlation Matrix	—	—	60 measures		
Decision Tree	Number of nodes.	An error message appears.	180	75	200
	Number of target values.	The decision tree displays truncated data.	50	30	50
	Number of predictor values.	An error message appears.	4,300	1,500	5,000

Visualization Type	Variation	Behavior when Limit Is Exceeded	Default Threshold	Minimum Threshold	Maximum Threshold
Word Cloud	—	The word cloud displays the top words by frequency, by measure value, or by topic weight.	100	100	100

In addition to the data thresholds that are set by the Preferences window, there are server data limits that can be set by the system administrator. For information about the server data limits, see the topic “Manage High-Cardinality Data” in the *SAS Visual Analytics: Administration Guide*.

High-Cardinality Thresholds for Report Objects

Client-side thresholds for report objects are documented in the following table. These thresholds affect the designer and the classic viewer. For information about adjusting high-cardinality thresholds for the modern viewer, printing, and SAS Visual Analytics Apps (previously called SAS Mobile BI), see the “Configuration Properties: SAS Mobile BI” topic in *SAS Intelligence Platform: Middle-Tier Administration Guide*.

Note: The second column indicates the maximum number of unique values.

Table A7.2 Client-Side Thresholds for Report Objects

Report Object	Rows
Gauges (bullets, sliders, thermometers, dials, and speedometers)	10
Word cloud	100
Bubble plots	1,500
Bar charts (regular, targeted, dual axis, and dual axis bar-line)	3,000
Waterfall charts	3,000
Line charts (regular and dual axis line)	4,000
Geo maps (bubble, coordinate, and region)	5,000
Step plots	10,000
Time series plots (regular and dual axis)	10,000
Treemaps	25,000

Report Object	Rows
Pie charts	40,000
Scatter plots	40,000
Tables (list tables and crosstabs)	40,000

Here are some key points about high-cardinality thresholds in report objects:

- In general, requests that exceed a client-side report object threshold cause an error message to be displayed. An exception is that for detail data, excess rows are truncated.
- Scatter plots always show detail data. List tables show details if the user selects the **Show detail data** check box. Bubble plots show details unless a category is assigned to the Group role.
- If a data source contains more than a million records, queries are blocked for data items whose cardinality is greater than 10,000 unique items in any particular category.

Appendix 8

Troubleshooting in SAS Visual Analytics Designer

Repairing Reports

Issue: The Repair Report window is displayed.

Resolution:

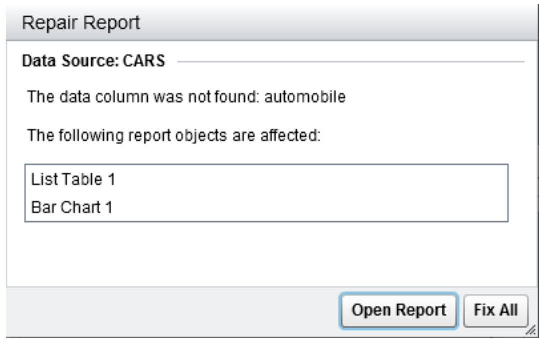
When you open a report in the designer, you might be prompted to repair the report. For example, this can happen when one or more columns have been removed from the data source. You might be prompted to repair the report when you refresh a report or change a data source. You might be prompted to repair a report if your system administrator is still deciding which data items should be in a table.

To repair a report:

1 Decide whether you want to fix individual data items or all of the data items in the report.

   Fixes can be made at the individual data item level. For example, if two data items are removed from a report, then the Repair Report window displays separate entries for each data item. The report objects that are impacted by each data item appear below each of the data items.

Here is an example of the **Repair Report** window that is displayed when a missing column affects two report objects:



2 Click **Fix** or **Fix All**. For example, if there is a column missing in the data source and you click **Fix**, then the missing column is removed from the report.

Note: If you click **Open Report**, then all of the data is removed from the report objects.

A confirmation message is displayed when the report has been fixed.

Displaying Alert Notifications

Issue: An expected alert notification has not been received.

Resolution:

There are many factors that can affect alert notifications. For example, system administrators for SAS Visual Analytics can set a property for how frequently the system checks to see whether the alert condition has been met. And, alert notifications are not sent if a server is down.

The timing of when alert notifications are sent can vary depending on when the alert condition is met. For example, suppose that you specify that alert notifications are sent every five days. The alert condition is met for the first time at 10:00am on September 2. After that, the system continues to check for the alert condition based on the hour or minute value that you specified for the **Query the system every** field. Because the frequency is set to five days, no alert notifications will be sent between 10:01am on September 2 and 9:59am on September 7. This is the case, no matter how many times the alert condition is met during that timeframe. At 10:00am on September 7, if the alert condition is met, then another alert notification is sent, and the five-day countdown starts again. On the other hand, if the alert condition has not been met at 10:00am on September 7, then no alert notification is sent. However, now that five days have passed, an alert notification is sent as soon as the alert condition is met again. Suppose that an alert notification is sent at 2:00pm on September 10. After the alert notification is sent, the five-day countdown starts again, and the earliest the next alert notification can be sent is 2:00pm on September 15. If you modify the notification frequency, then the notification history is reset.

You might not receive an alert notification as soon as the alert condition is met because of the hour or minute value that is set in the **Query the system every** field in the designer. For example, suppose that you set the value in the **Query**

the system every field to `1 day`. If the alert condition is met one hour after the system check is done, then it will be another 23 hours before the system checks for the alert condition again and sends an alert notification.

Contact your system administrator for additional assistance.

Exporting Data from Report Objects to Microsoft Excel 2007

Issue: An error occurs when data from a report object is exported using the Export to: Excel 2007 Workbook (*.xlsx) option or when there are character codes displayed in the exported Excel file.

Resolutions:

- If an error occurs during exporting, then in the Export or Save As window, select either the **Save as: Tab-Separated Values (*.tsv)** or the **Save as: Comma-Separated Values (*.csv)** option to create a data file.
- If you see character codes displayed in a file that has been exported to Excel, then re-open the exported file in Excel, and save it without any changes. The character codes are removed from the file.

Specifying Colors for Data Labels

Issue: The data label color for my graph is not black.

Resolution: If a graph object that you add to a report does not have the desired attributes, try creating a custom graph object in the SAS Visual Analytics Graph Builder (the graph builder). For example, suppose the data labels for a dual axis bar-line graph are not the colors that you want. You can create a dual axis bar-line graph in the graph builder and specify the font color for the data labels.

Appendix 9

Using URL Parameters to View a Report

If guest access is configured for the SAS Visual Analytics Viewer, you can use this URL: http://*host*/SASVisualAnalyticsViewer/VisualAnalyticsViewer/guest.jsp. However, to log on and display a single report using guest access, you can use a URL such as this one (which has parameters to specify the name and location of the report): http://*host*/SASVisualAnalyticsViewer/VisualAnalyticsViewer_guest.jsp?reportName=My+Report&reportPath=/Shared+Data/VA+Reports/&appSwitcherDisabled=true. This second URL is useful for tasks such as using an <iframe> tag to incorporate a single report within another web application because the SAS Visual Analytics application bar, menu bar, and toolbar are not included.

The following table lists the parameters that can be included in the second URL:

Parameter	Description
reportName	Specifies the name of the report that you want to view. This parameter is required if you are accessing a specific report through a URL.
reportPath	Specifies the path to the report that you want to view. This parameter is required if you are accessing a specific report through a URL.
appSwitcherDisabled	Disables the application bar, which enables you to return to the home page or to access other parts of SAS Visual Analytics and other SAS applications. The default value is false. **Note:** If you are displaying in an <iframe> tag, you must disable the application bar by specifying `appSwitcherDisabled=true`.
commentsEnabled	Enables the **Comments** tab in the right pane. The default value is true. If both propertiesEnabled and commentsEnabled are false, then the right pane is not added to the viewer.
propertiesEnabled	Enables the **Properties** tab in the right pane. The default value is true. If both propertiesEnabled and commentsEnabled are false, then the right pane is not added to the viewer.
reportViewOnly	Hides the SAS Visual Analytics banner, including all of the menus.

Parameter	Description
reportContextBar	Enables all of the menu options for the report. The default value is true. If reportContextBar is false, then it is implied that both propertiesEnabled and commentsEnabled are false, and the right pane is not added to the viewer. This parameter can be used only in the modern viewer. **Note:** If reportContextBar is false, you will also hide the back link, which does not allow section linking or report linking to work properly.

Parameters are specified in the URL as a sequence of name and value pairs using query string syntax. The URL specifies your server, an absolute path to SAS Visual Analytics, and the query string (following the question mark character). Each parameter name in the query string is separated from the next value by an equal sign (=). Multiple name and value pairs are separated by ampersand characters (&). In this example, `reportName=My+Report` is the parameter that specifies the name of the report. The second name and value pair (`reportPath=/Shared+Data/VA+Reports/`) is the parameter that specifies the path to that report. The third name and value pair (`appSwitcherDisabled=true`) disables the application bar. It must be present if you are displaying the report in an <iframe> tag.

There are special rules for formatting name and value pairs in a URL. Special characters (such as most punctuation characters, including spaces) in a value must be URL-encoded. A space can be encoded as a plus sign (+) or %20. Other characters are encoded using the %*nn* convention, where *nn* is the hexadecimal representation of the character in the ASCII character set. In the previous example, the value **`/Shared+Data/VA+Reports/`** actually identifies the report path **`/Shared Data/VA Reports/`**. The spaces in the names are encoded as plus signs (+).

The reportName and reportPath parameters must be specified in the URL. The reportName parameter specifies the name of the report that you want to view (for example, **`My Report`**). The reportPath parameter specifies the path to that report (for example, **`/Shared Data/VA Reports/`**).

Note: You can use these parameters with the standard viewer URL that requires login credentials (http://*host*/SASVisualAnalyticsViewer/VisualAnalyticsViewer.jsp). For example, you can use http://*host*/SASVisualAnalyticsViewer/VisualAnalyticsViewer.jsp?reportViewOnly=true to hide the banner.

URL parameters can be used to specify values for report parameters, such as display rules, filters, ranks, and aggregated measures.

Note: You must know the exact name of the parameter and the valid values for that parameter if you are going to include the parameter in your URL or change its values.

Appendix 10

Schema for Imported Tweets

The schema for the imported tweets is shown in the following table:

Column Name	Column Type	Length	Description
author	Character	128	The author's screen name.
authordescription	Character	1024	The author's description.
authorfavouritecount	Numeric	8	The number of tweets the author has indicated as a "favorite".
authorfollowercount	Numeric	8	The number of followers the author had when the tweet was imported from Twitter.
authorfriendcount	Numeric	8	The number of users the author was following when the tweet was imported from Twitter.
authorid	Numeric	8	Twitter's unique ID for the author.
authorimageurl	Character	1024	A link to the author's profile image.
authorlang	Character	2	The BCP 47 code for the author's self-declared user interface language.
authorlocation	Character	128	The author's self-declared location.
authorname	Character	128	The author's self-declared name in the author's profile.
authortimezone	Character	256	The author's self-declared time zone.
authorurl	Character	1024	A URL provided by the author in the author's profile.
body	Character	1024	The body of the tweet.
deviceinfo	Character	1024	The utility that was used to post the tweet. It is represented as an HTML-formatted string.

Column Name	Column Type	Length	Description
`docid`	Numeric	8	Twitter's unique ID for the tweet. You can specify this variable as a unique row identifier when working with text analysis and word cloud visualizations.
`doclatitude`	Numeric	8	The latitude value of the tweet's coordinates (if available). If it is not available, then a missing value is populated.
`doclongitude`	Numeric	8	The longitude value of the tweet's coordinates (if available). If it is not available, then a missing value is populated.
`isretweet`	Numeric	8	A value of 1 indicates that the tweet is a retweet. Otherwise, the value is 0.
`referenceauthor`	Character	128	Screen name of the user to whom this tweet was a reply. If the tweet is not a reply, then this field is empty.
`referenceauthorid`	Numeric	8	User ID of the user to whom this tweet was a reply. If the tweet is not a reply, then the value for this field is set to -1.
`publisheddatetime`	Numeric	8	The tweet's published date and time as a SAS datetime value (based on the number of seconds since January 1, 1960 at midnight).
`publisheddatetimestr`	Character	34	The tweet's published date and time in string format. Example: `October 24, 2013 6:56:25 PM GMT`.
`tags`	Character	150	A semi-colon separated list of hash tags that are mentioned within the tweet.
`listoflinks`	Character	1024	A semi-colon separated list of URLs that are included in the body of the tweet. Expanded links are used where available.
`mentionedusernames`	Character	256	A semi-colon separated list of names of twitter users mentioned within the tweet.
`mentionedusers`	Character	256	A semi-colon separated list of screen names of twitter users mentioned within the tweet.
`retweetcount`	Numeric	8	The number of times this tweet has been retweeted when the tweet was imported from Twitter.

The schema for tweets is different from the schema that Twitter uses for tweets, users, and entities. The following list provides URLs to the information from Twitter for comparison purposes:

- Tweets: https://dev.twitter.com/overview/api/tweets
- Users: https://dev.twitter.com/overview/api/users
- Entities: https://dev.twitter.com/overview/api/entities

Appendix 11

About the Classic SAS Visual Analytics Home Page

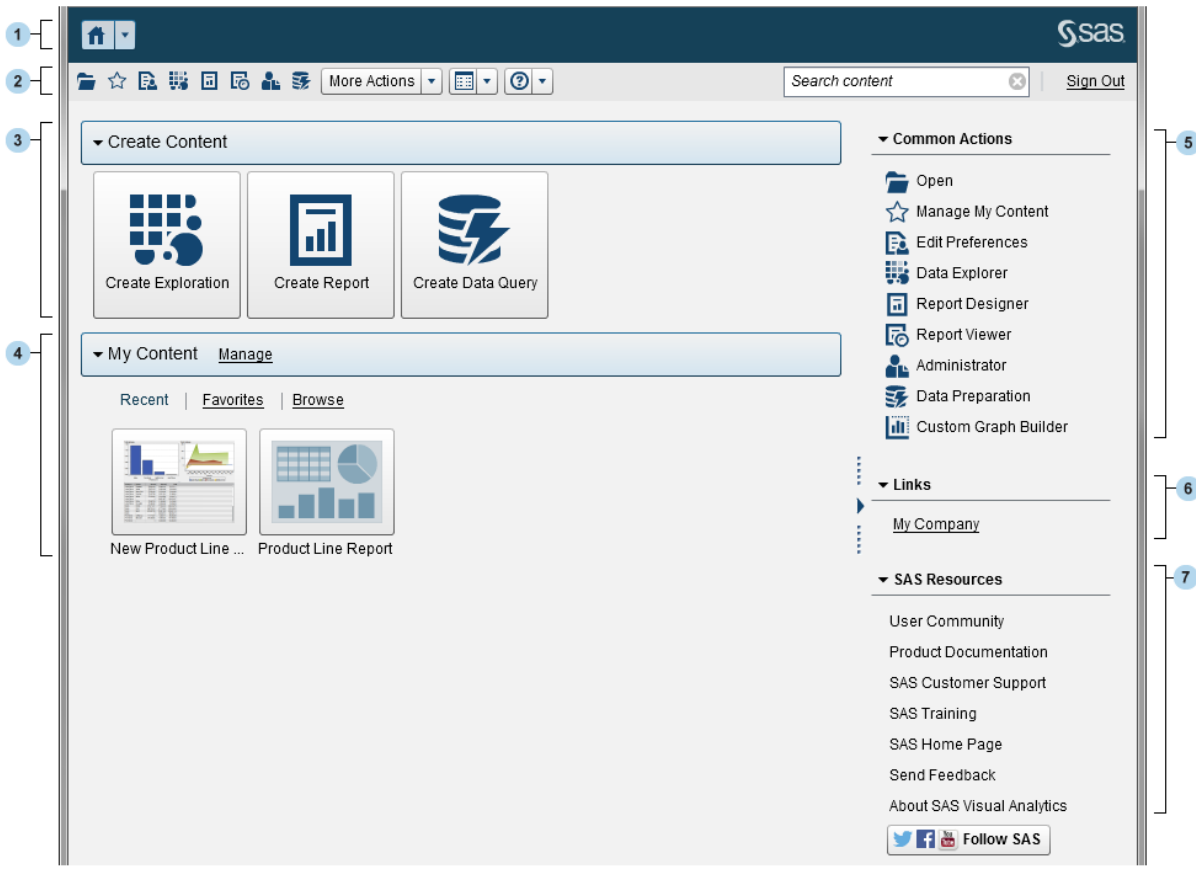
Your First Look at the Classic SAS Visual Analytics Home Page

The home page enables you to create new content in SAS Visual Analytics. In addition, it enables you to access content that you and others have created.

Starting in the 7.2 release, the home page has two appearances: modern and classic. The modern appearance is the default. You can specify the appearance in the modern home page settings or in the classic home page preferences.

Here are the features of the classic home page:

Figure A11.1 The Classic Home Page



1 The application bar enables you to return to the classic home page and to access other parts of SAS Visual Analytics and other SAS applications that integrate with the classic home page. You can access your recently created or viewed reports, explorations, stored processes, stored process reports, data queries, or other objects in your recent history. Buttons are displayed for each open application.

2 The menu bar enables you to access task options, such as creating reports, exploring data, managing your environment or favorites, setting your view, and getting help on using SAS Visual Analytics. You can search all SAS content from the menu bar, and you can sign out of SAS Visual Analytics.

3 The **Create Content** area provides icons to let you quickly explore data, create a new report, or prepare data, depending on your role, the associated capabilities, and your SAS software licenses. Other installed SAS applications might add actions to the **Create Content** area.

4 The **My Content** area lists any metadata objects that are created by a supported SAS application. For example, explorations, reports, queries, tables, stored processes, or stored process reports that you have opened or created recently are listed. It also lists any content that you have marked as a favorite or as part of a collection. Click **Browse** to explore folders to find a report, exploration, stored process, stored process report, table, or query.

Note: All tables are displayed because the classic home page does not distinguish between LASR tables and other tables.

5 The **Common Actions** section provides an alternate way for you to access features and other installed SAS applications. For more information, see “Working with the Right Pane on the Classic Home Page” on page 631.

6 The **Links** section provides links to pages that you have bookmarked. For more information, see “Working with the Right Pane on the Classic Home Page” on page 631.

7 The **SAS Resources** section provides links to the SAS website, the SAS Visual Analytics User Community, and to social media. For more information, see “Working with the Right Pane on the Classic Home Page” on page 631.

Note: Guest access does not provide individualized features on the classic home page, such as history or alerts. By default, a user with guest access has a Basic role and can access only the classic home page and the viewer. For more information, see “Guest Access” on page 7.

If you have the **Theme Designer for Flex: Administration** role in SAS Management Console, you can access the SAS Theme Designer for Flex from the **More Actions** item on the menu bar or in the **Common Actions** section.

Here is an example of the menu bar with the **More Actions** item.

Figure A11.2 Classic Home Page Menu Bar with the More Actions Item

Once you start using the classic home page, thumbnails enable you to open explorations, reports, stored processes, stored process reports, tables, queries, and folders that you have created or opened. The default view is secure, generic thumbnails that represent the content. An administrator can set a property that specifies the use of thumbnails that are shared and unique to each object. Shared thumbnails are unique for each individual report object, so each report looks different from other reports, and each exploration looks different from other explorations.

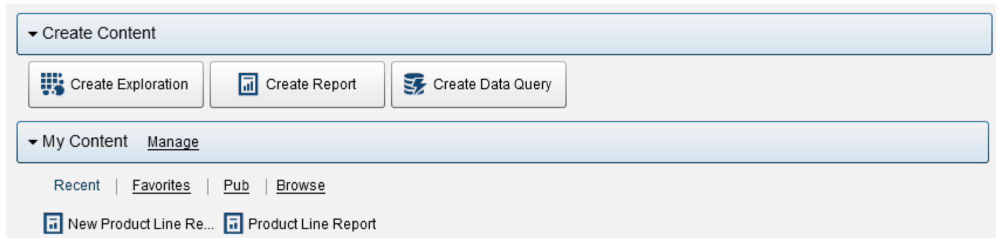
Generic thumbnails are distinguished by content type only. All reports look the same, but they appear different from explorations. Here are examples of the generic thumbnails that you might see on the classic home page:

Figure A11.3 Generic Thumbnails for an Exploration, a Report, and a Stored Process

You can choose to view the content on the classic home page as a list. To change your view, click  on the menu bar, and then select **Thumbnail** or **List**. The default is **Thumbnail**.

Here is an example of the list view for the classic home page:

Figure A11.4 List View



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

- "Your First Look at SAS Visual Data Builder" on page 38
- "Your First Look at the Explorer" on page 109
- "Your First Look at the Designer" on page 300

Manage Content on the Classic Home Page

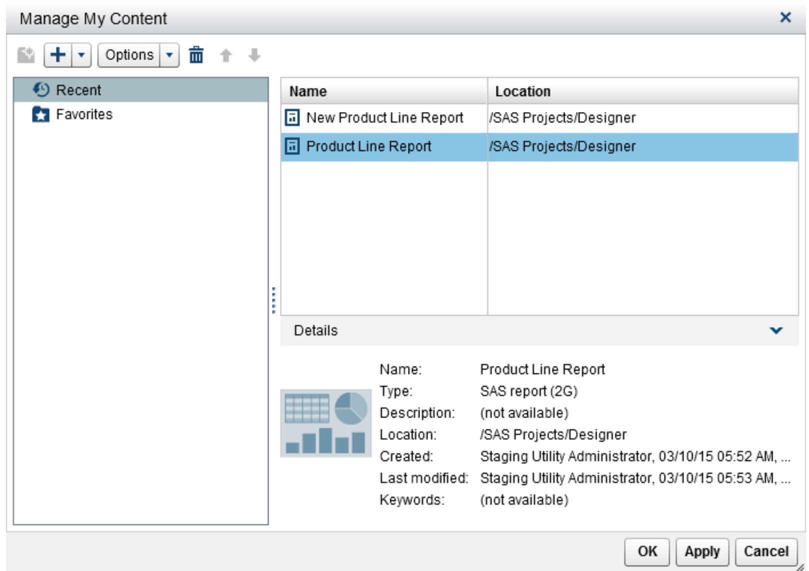
Overview of Classic Home Page Content

The classic home page displays recent reports, explorations, stored processes, stored process reports, tables, and queries, as well as favorites and collections under the **My Content** heading on the classic home page. A *favorite* is a report, an exploration, a stored process, a stored process report, a table, or a query that you can quickly access. A *collection* is a group of favorites. A collection can be shared with multiple users. A *favorites group* is a grouping of objects that you can add to a collection or to your **Favorites** list. A favorites group that is added to a collection has the same access rights that the collection has.

Note: A user with guest access cannot manage content.

To manage your content, click **Manage** to the right of the **My Content** heading. The Manage My Content window is displayed.

Figure A11.5 *Manage My Content Window*



The following icons are available in the Manage My Content window:

Icon	Description
	Enables you to create a new favorites group for your reports, explorations, stored processes, stored process reports, tables, and queries. Folder names cannot be longer than 60 characters. This icon is available only when **Favorites** is selected.
	Enables you to add a report, exploration, stored process, stored process report, table, or query to your list of favorites. You can also create or add a collection of favorites.
Options	Enables you to select either **Clear Recent History** or **Clear Favorites**.
	Enables you to delete favorite or recent content, one at a time. You can delete multiple items by pressing the Ctrl key. For a collection, you have a choice of permanently deleting the collection or simply removing it from under the **My Content** heading on the classic home page.
	Enables you to move favorites content up in the list to change what is displayed under the **My Content** heading on the classic home page.
	Enables you to move favorite content down in the list to change what is displayed under the **My Content** heading on the classic home page.

When you select a report, exploration, stored process, stored process report, table, or query in the Manage My Content window, the details such as the name and location of that item are displayed.

Create a Collection Using the Classic Home Page

A collection is a group of favorites that you can share with other users. If you have the Create Collections capability, then you can create collections. You can create a collection that points to your favorite reports, explorations, stored processes, stored process reports, tables, queries, folders, and favorite groups.

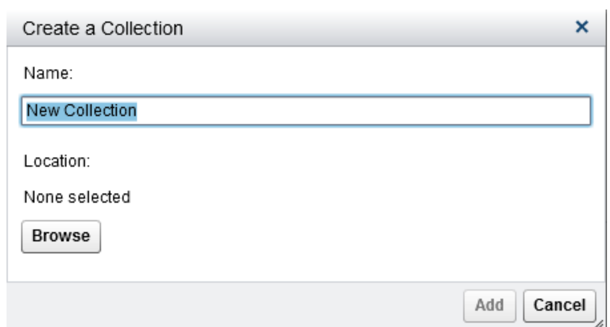
> **TIP** Collections that you create using the home page are available in the SAS Visual Analytics Apps (previously called SAS Mobile BI).

To create a new collection:

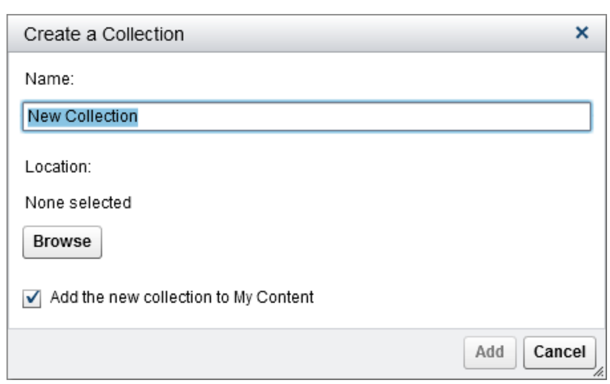
1 Do one of the following:

- On the classic home page, click **Manage** to the right of the **My Content** heading. The Manage My Content window is displayed.

  Click to open the menu, and then select **Create a Collection**. The Create a Collection window is displayed.

- In the object inspector, click **Collections**. Then, select **Create a Collection**. The Create a Collection window is displayed.

2 Enter a **Name** for the collection.

3 Click **Browse** to select a **Location** for the collection. The Choose a Location window is displayed.

  Select an existing folder or create a new one. Click **OK** to return to the Create a Collection window.

  Note: When you create a new collection, you can add items to it from within the Manage My Content window only if the collection is a root collection in **My Collections**.

4 (Optional) If you opened the Create a Collection window from the object inspector, then you can clear the **Add the new collection to My Content** check box. This means that the new collection is not displayed on the classic home page next to the **Favorites** link.

5 Click **Add**. If you are in the Manage My Content window, then the new collection is displayed on the left.

In the Manage My Content window, you can rename, remove, or permanently delete a collection by right-clicking it and selecting the applicable option.

Add a Favorites Group to a Collection or to Your Favorites List

You can use a favorites group to bring together your favorite reports, explorations, stored processes, stored process reports, tables, or queries. You can add a favorites group to a collection or to your **Favorites** list.

Here are some key points about favorites groups:

- Favorites groups are displayed only in the tree on the left side of the Manage My Content window.
- Favorites groups that you add using the classic home page are available in the SAS Visual Analytics Apps (previously called SAS Mobile BI).
- Favorites groups can be added only using the classic home page.

To add a favorites group:

1 On the classic home page, click **Manage** to the right of **My Content**. The Manage My Content window is displayed.

2 Select **Favorites** or a collection.

> **TIP** When you add a favorites group to a collection that is publicly shared, any favorites groups inside that collection are also publicly shared.

3 On the menu bar, click . A favorites group is identified by the icon.

In the Manage My Content window, you can rename or remove a favorites group by right-clicking it and selecting the applicable option.

Working with the Right Pane on the Classic Home Page

About the Right Pane of the Classic Home Page

The **Common Actions**, **Links**, and **SAS Resources** sections are displayed in the right pane on the classic home page. Someone who has the **Visual Analytics: Administration** role can control which sections are displayed in the right pane. For example, someone with the **Visual Analytics: Administration** role might hide all of the sections, two of the sections, or only one section. For any of the sections that are displayed, you can control which sections are visible in your SAS Visual Analytics session. You can also control the order in which the items in the sections are displayed.

The following sections can be displayed in the right pane:

Common Actions
: Provides an alternate way for you to access features, such as designing reports, exploring data, managing your environment or content. For example, you can click **Build Custom Graph** to access the graph builder. If you have a Data Building role, then a **Prepare Data** link is displayed. You might also see links to other SAS products, such as SAS Theme Designer for Flex, depending on your role and the SAS licenses that your site has.

Links
: Provides links to pages that you have bookmarked. Someone with the **Visual Analytics: Administration** role can also provide shared links for all users. For information about links, see "Manage Links in the Right Pane of the Classic Home Page" on page 633.

SAS Resources
: Provides links to the SAS website, the SAS Visual Analytics User Community, and to social media.

Hide Content in the Right Pane of the Classic Home Page

To hide content in the right pane, position your mouse pointer over the section title (**Common Actions**, **Links**, or **SAS Resources**), and click ✕.

To hide or reorder content in a specific section, position your mouse pointer over the section name, and click .

Note: You cannot hide or reorder the first three **Common Actions** (which are **Open**, **Manage My Content**, and **Edit Preferences**).

To hide content in the right pane using the menu bar:

1 Click ▾ beside on the menu bar.

2 Clear the selection for one or more of the following sections:

- **Common Actions**
- **Links**
- **SAS Resources**

The selected section is hidden in the right pane.

Note: If someone with the **Visual Analytics: Administration** role has hidden any of the sections in the right pane, then the sections are not displayed in the menu.

Show Content in the Right Pane of the Classic Home Page

To show content in the right pane:

1 Click ▾ beside on the menu bar.

2 Select one or more of the following sections:

- **Common Actions**
- **Links**
- **SAS Resources**

The selected section is displayed in the right pane.

To show or reorder content in a specific section, position your mouse pointer over the section name, and click .

Note: You cannot hide or reorder the first three **Common Actions** (which are **Open**, **Manage My Content**, and **Edit Preferences**).

If you have the Administer Hub capability, then you will also see a **Manage Shared View of Right Pane** menu item when you click beside on the menu bar. When you select this menu item, the Manage Shared View of the Right Pane window is displayed. Administrators can use this window to hide sections, restore sections, and add links in the right pane. Links that are added using the Manage Shared View of the Right Pane window are shared links that anyone can see.

Manage Links in the Right Pane of the Classic Home Page

Note: All URLs must start with `http://` or `https://`.

You can add, edit, and delete links in the **Links** section that is displayed in the right pane.

To add new links, position your mouse pointer over the **Links** section heading in the right pane to activate the +, which opens the Add Link window. Links that are added using the Add Link window are private and visible only to the user who created them. Click to open the Manage Links window, where you can add, delete, re-order, and hide links. You cannot delete shared links that were added using the Manage Shared View of the Right Pane window.

If you have an Administration role, then you can click beside on the menu bar, and select **Manage Shared View of Right Pane**. The Manage Shared View of the Right Pane window is displayed. You can use this window to add links to the right pane that are shared with other users.

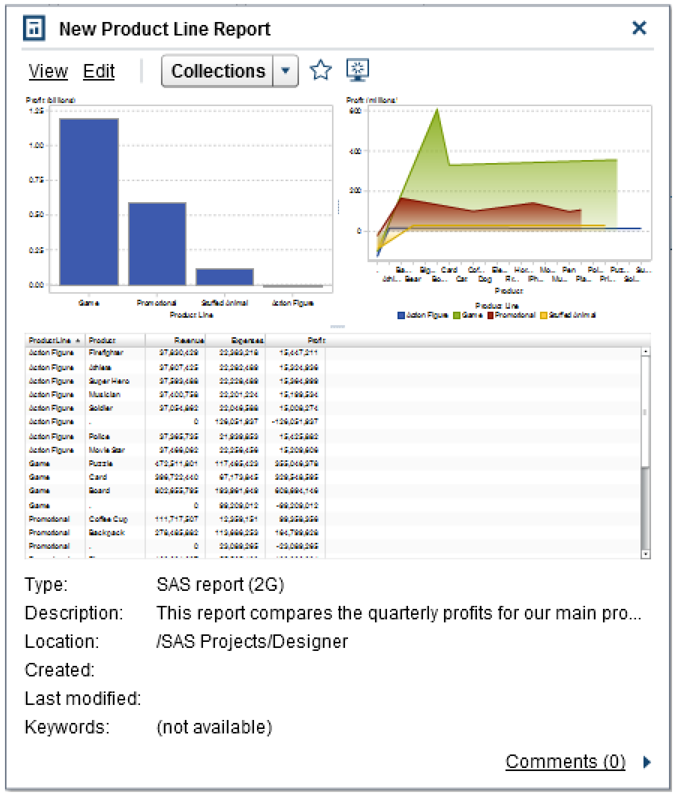
Discover Details Using the Object Inspector on the Classic Home Page

The classic home page provides an object inspector that enables you to quickly find more information about an object when you click it. You can see a description of the report, exploration, stored process, stored process report, table, query, folder, or collection; the location of the selected object; the author; when the object was created and modified; and a list of keywords.

Your role and capabilities determine the availability of actions on the toolbar (such as viewing, editing, or opening) in the object inspector. For more information about roles and capabilities, see the *SAS Visual Analytics: Administration Guide*.

Here is an example of what you might see in the object inspector for a report that has been viewed:

Figure A11.6 The Object Inspector on the Classic Home Page



You can use the toolbar in the object inspector to do one or more of the following tasks (which are available depending on your role and capabilities):

- **View** a report, which opens the report in SAS Visual Analytics Viewer (the viewer).
- **Edit** a report, which opens the report in SAS Visual Analytics Designer (the designer) so that you can edit or change the objects in the report.
- **Open** an exploration, which displays SAS Visual Analytics Explorer (the explorer). The **Open** link is also available for stored processes, stored process reports, queries, and folders.
- **Create Report** for a table, which displays the designer.
- **Explore Data** for a table, which displays the explorer.
- Click **Collections** to select an existing collection or to create a new collection. For more information, see "Create a Collection Using the Classic Home Page" on page 630.
- Click ☆ to add the object to your list of favorites. When the star icon is yellow, it indicates that the object is in your list of favorites. Click ★ to remove the object from your list of favorites.

- Click to set the object as the initial screen when you start SAS Visual Analytics. When the icon is blue, it indicates that the object is set as your initial screen. Click to remove this initial screen setting.

In the object inspector, you can also do the following:

- Click the user name beside the **Created** or **Last modified** labels to search for all of the SAS reports, queries, tables, stored processes, and stored process reports that have been created or modified by that user.
- Click **Comments** to add or view comments. Any user who has the Add or View Comments capability can add comments to any report, exploration, stored process, stored process report, table, or query. They can view existing comments about any report, exploration, stored process, stored process report, table, or query using the **Comments** link in the lower right of the object inspector. You cannot comment on folders, favorite groups, or collections. For more information, see “Add Comments to Objects on the Classic Home Page” on page 635.

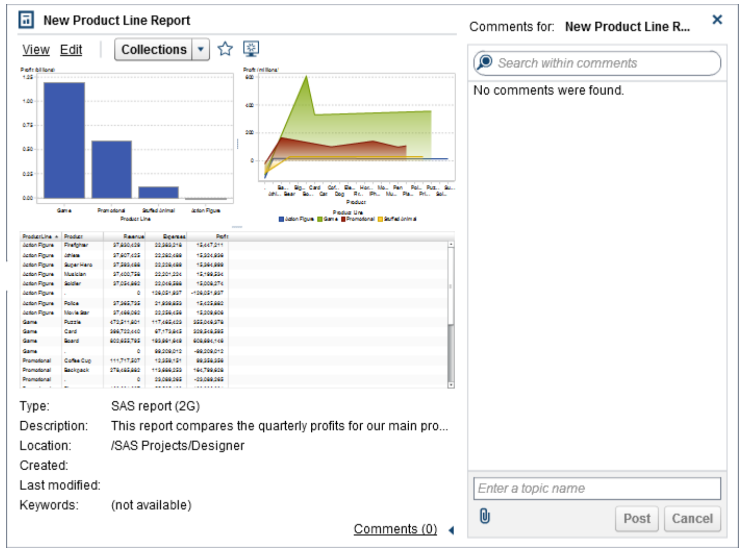
Add Comments to Objects on the Classic Home Page

If you have the Add or View Comments capability, then you can use the object inspector on the classic home page to add (or view) comments to objects on the home page. You can add comments to reports, explorations, stored processes, stored process reports, tables, and queries. You can also respond to existing comments or edit your own comments. You cannot comment on folders, favorite groups, or collections. You cannot see comments that were added to visualizations using the explorer.

To add a comment:

1. Click on a report, exploration, stored process, stored process report, table, or query to open the object inspector.
2. Click **Comments** in the lower right corner to expand the object inspector.

Figure A11.7 *Comments in the Object Inspector*



3 Enter a topic name and a comment.

4 Click **Post** to add your comment. The **Comments** link in the object inspector updates to show that there is a comment.

5 (Optional) Click to attach a file or an image to your comment.

To respond to an existing comment:

1 Click on a report, exploration, stored process, stored process report, table, or query to open the object inspector.

2 Click **Comments** in the lower right corner to expand the object inspector.

3 Select an existing comment. Then, enter a reply.

4 (Optional) Click to attach a file or image to your reply.

5 Click **Post** to add your comment. The **Comments** link in the object inspector updates to show that there is a comment.

Note: To edit another user's comments or to delete comments, you must belong to the predefined role **Comments:Administrator**. This role includes the capabilities of editing or deleting comments.

To search for a comment:

1 Enter the word or phrase that you want to search for in the search field. Press Enter.

2 (Optional) To clear your search, click . Then, you can enter another word or phrase in the search field.

Specifying Your Preferences for the Classic SAS Visual Analytics Home Page

All of your preferences persist between sessions. Preferences are not available for a user with guest access.

Specify Global Preferences Using the Classic Home Page

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** on the classic home page. The Preferences window is displayed.

   Note: If you are in the data builder, the explorer, the designer, or the viewer, then select **File** ▸ **Preferences** to open the Preferences window.

2. Click **Global Preferences** in the left pane.

3. Specify your preferences.

   - Select a **User locale** to specify your language and geographic region.

     Note: If you change the **User locale** preference, then you must sign out and sign in to SAS Visual Analytics for the change to take effect.

     TIP If you receive distributed reports, it is recommended that you set the **User locale** preference. The report distribution feature does not have access to the browser locale, so it depends on the locale specified for the **User locale**. For more information, see "Distributing Reports" on page 484.

   - 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 settings for focus indicator** to change the color, thickness, and opacity of the focus in your SAS web applications.
   - Click **Reset** to show all warning and confirmation messages.

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

Specify the SAS Visual Analytics General Preferences Using the Classic Home Page

Using the SAS Visual Analytics classic home page, you can specify your general preferences for SAS Visual Analytics. Preferences are saved on a per-user basis.

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 a value for **Show this number of recent items**. There is a minimum of zero and a maximum of 25. The default setting is 10 items.

   Click **Clear History** to reset your history.
4. Click **OK** to apply your changes.

Specify Your Preferences for the Classic Home Page

To specify preferences that are specific to the classic home page:

1. Click either or **Edit Preferences** to open the Preferences window.
2. Click **Home** in the left pane.
3. Select an option from the **Initial screen** drop-down list. The available options depend on your role and capabilities.

   > **TIP** If you select **A specific object**, then you can use the Choose an Item window to choose one content type (for example, SAS reports (2G), visual explorations, or a SAS Visual Statistics project) for your initial screen. (SAS Visual Statistics is licensed separately.) If you select a report in the Choose an Item window, then you might be able to click beside the **OK** button to select either **View (Visual Analytics Viewer)** or **Edit (Visual Analytics Designer)**. The available content types and the options for opening reports depend on the SAS products that your site has licensed and how they are configured.
4. Specify **Order the actions associated with a content type** preferences. Select a **Content type**, and then use the arrow keys to specify the **Order of actions**. This preference also affects the order of actions in the object inspector’s toolbar.

   This preference also determines which application opens when you click on a link to a report in an e-mail message. For example, suppose that you specify **Edit - SAS Visual Analytics Designer** as the first item in the list of actions for the **SAS report (2G)** content type. Then, you receive an e-mail message with a link to a report. When you click the link to the report, it is displayed in the designer.
5. Click **OK** to apply your changes.

About Searching from the Classic Home Page

You can search all reports, explorations, stored processes, stored process reports, tables, and queries that are on the metadata server from the menu bar on the classic home page. The list of items that you can search for depends on the SAS products that your site has licensed.

Note: There can be a delay between the time that an object (for example, a report or an exploration) is created or changed and when the search is updated.

The following fields in metadata are searched:

- Comment Title
- Comment Description
- Created
- Description
- Keywords
- Last Modified
- Title

The search field on the menu bar supports the following types of searches:

- single word
- multiple word
- in-word wildcards

  Note: A wildcard search has a limit of 200 words.

  Use the asterisk (*) to represent possible characters in a particular position in a word to generalize the word or to make it easier to find in a search. For example, if you specify *“cat*”*, then the search matches any word with the prefix *cat*, such as cats, category, catfish, and so on.
- spaces in strings that use quotation marks indicate that the words are considered together as a phrase for the search. For example, suppose that you search for *"Sample Report”*. The search matches *Sample Report 1* and *A Sample Report*, but not *Sample Values Report* or *Sample-Report*.
- the `AND` operator

  For example, if you specify *“A AND B”*, then only documents that contain both A and B are returned.
- plus (+) and minus (-) syntax

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

  - Specifying `2012 2013 2014` matches documents that contain any of the three terms.

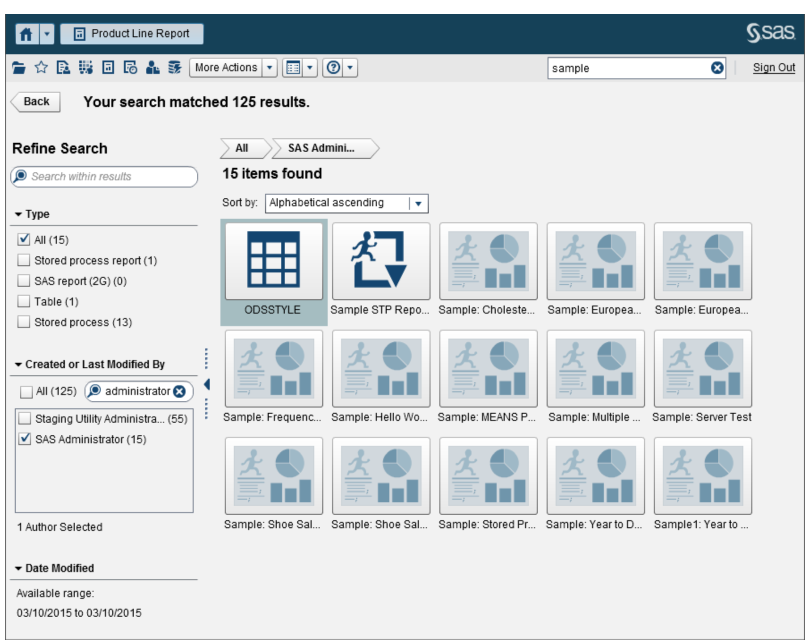
- Specifying `+2012 +2013 +2014` matches only documents that contain all three terms.
- Specifying `2012 2013 -2014` matches documents that contain `2012` or `2013`, but only if they do not contain `2014`.

Here are some considerations for searching from the classic home page:

- Single- and multiple-word searches that are plain text have a wildcard appended to them. For example, if you search for the word *sample*, then it is converted to **`sample*`**. This means that you get different results than if you entered *“sample”*, +*sample*, or **sample**.
- The search is not case sensitive.
- Up to 5000 items can be returned in a single search. If your search results exceed 5000 items, then the first 5000 items appear, and a message is displayed at the top of the Search Results window. The message tells you how many results matched your search and how many have been displayed. Narrowing your search using the search field enables the relevant search results to be displayed. When you submit a new search that returns fewer than 5000 items, you can see all the search results.
- Search results on the classic home page are not ranked. In the workspace, you can use the **Sort by** drop-down list to specify that you want the list of items to sort by **Alphabetical ascending**, **Alphabetical descending**, **Date ascending**, or **Date descending**. The default is **Alphabetical ascending**.
- A search string must be shorter than 4,000 characters.

Suppose that you search for the word *sample*. All words that contain *sample* at the beginning are displayed in the workspace. (For example, a word that starts with *samples* is displayed.) Here is an example of the search results:

Figure A11.8 Search Results on the Classic Home Page



Refining Your Search Results for the Classic Home Page

You can refine your search results using your original search on the metadata server. After you enter your original search using the menu bar on the classic home page, the search results are displayed in the workspace. Note that when you refine your search results, the initial 5000 item subset is not affected.

The refine search feature is dependent on your locale. The search is not case sensitive.

You can refine the search results displayed in the workspace using the **Search within results** field in the left pane. When you enter a term or terms, you can use quotation marks and spaces when you are searching within your results. Only the items that match your refined search appear in your workspace. Each search term that you enter must appear at least once in the object name, the description, the author's name, or keywords in the metadata.

You can also refine your search results by selecting (or clearing) check boxes in the left pane. Note that the totals in the parentheses in any pane are dependent on what has been selected in the other two panes.

The following panes are available:

- The **Type** pane shows the available object types. In SAS Visual Analytics, you can search for **Stored process report**, **SAS report (2G)**, **Table**, **Stored process**, **Visual data query**, or **Visual exploration**. Depending on the SAS products that you have licensed at your site, other SAS object types might appear in the **Type** pane. The default is **All** object types.

  The available types are determined by the capabilities that your system administrator has assigned. For example, you might see only the **All**, **SAS report (2G)**, **Table**, and **Stored process** check boxes in the **Type** pane.

  If you select all of the check boxes for the individual types, then the check boxes clear, and the **All** check box is selected automatically. If you clear the check boxes for all of the types, then the **All** check box is selected automatically.

  Note: If your role permits you to see only one type, then you will not see the **Type** pane.

- The **Created or Last Modified By** pane enables you to search for a specific user or to select one or more users by name.

  The number in parentheses beside the user's name identifies the number of objects that the user has created or modified, taking into account the filters that have been selected in the other panes.

  If you select all of the check boxes for the users in the list, then the check boxes clear, and the **All** check box is selected automatically. If you clear the check boxes for all of the users, then the **All** check box is selected automatically.

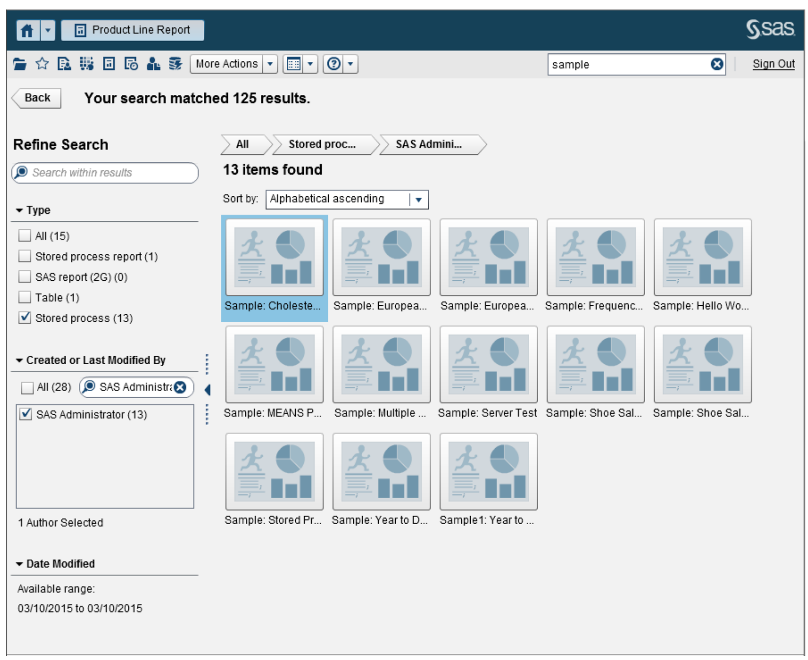
- The **Date Modified** pane lists the dates that the objects were last modified. You can refine the search to a specific date range by specifying a **From** date, a **To** date, or both. If the initial results set represents only a single day, then the **From** and **To** fields are hidden to conserve visual space.

The dates are time zone dependent. If your company has offices in multiple time zones, then the date modified time stamp on the object reflects the time at which it was saved using the current user's time zone. For example, if a report was modified by a user on January 3 at 2:00 a.m. in New York, another user in California sees that it was modified on January 2 at 11:00 p.m.

At the top of the workspace, you can see the breadcrumbs for the selections that you have made to refine your search. Breadcrumbs change as you modify the selections in the **Type**, **Created or Last Modified By**, and **Date Modified** panes. The breadcrumbs also show any search strings that you entered in the **Search within results** field.

Suppose that you search for the word *sample*. Then, you refine the search results by selecting **Stored process** for **Type**. All stored processes that contain the word *sample* are displayed in the workspace. Here is an example of the refined search results:

Figure A11.9 *Refined Search Results on the Classic Home Page*



Recommended Reading

Here is the recommended reading list for this title:

- *SAS Visual Analytics: Administration Guide*
- *SAS Visual Analytics: Getting Started with Data on Windows*
- *SAS Visual Analytics: Getting Started with Data Preparation*
- *SAS Visual Analytics: Getting Started with Exploration and Reporting*
- *SAS Visual Analytics: Getting Started with Analytical Models*
- SAS Technical Paper: One Report, Many Languages: Using SAS Visual Analytics to Localize Your Reports
- Help and tutorials integrated into SAS Visual Analytics Apps (previously called SAS Mobile BI)
- *An Introduction to SAS Visual Analytics: How to Explore Numbers, Design Reports, and Gain Insight into Your Data*
- Paper SAS4080-2016: *Designing SAS® Visual Analytics Reports: Write Once, View Anywhere*
- Paper SAS6321-2016: *If You Build It, Will They Understand? Designing Reports for the General Public in SAS® Visual Analytics*
- Paper SAS6361-2016: *Store Processes and SAS® Visual Analytics: Giving Users the Power to Load*
- Paper SAS3802-2016: *Carry-on Suitcases and Mobile Devices: Using SAS® Visual Analytics Designer for Creating Optimally Designed Reports for SAS® Mobile BI*
- SAS offers instructor-led training and self-paced e-learning courses to help you get started with SAS Visual Analytics. For more information about the courses available, see support.sas.com/training.

For a complete list of SAS publications, go to sas.com/store/books. If you have questions about which titles you need, please contact a SAS Representative:

SAS Books
SAS Campus Drive
Cary, NC 27513-2414
Phone: 1-800-727-0025
Fax: 1-919-677-4444
Email: sasbook@sas.com
Web address: sas.com/store/books

Glossary

Apache Hadoop (Hadoop)
an open-source framework that enables the distributed processing of large data sets, across clusters of computers, using a simple programming model.

calculated column
a column that does not exist in any of the tables that are accessed, but which is created as a result of a column expression.

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.

co-located data provider
a distributed data source, such as SAS Visual Analytics Hadoop or a third-party vendor database, that has SAS High-Performance Analytics software installed on the same machines. The SAS software on each machine processes the data that is local to the machine or that the data source makes available as the result of a query.

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 (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.

dependency
a trigger condition that must be met before a job can run in a scheduled flow.

deployed job
a job that has been saved in a deployment directory and can be scheduled.

deployment directory
the location for generated SAS DATA step programs that will be executed by the batch server as part of a scheduled flow.

file event
a file-related occurrence that is used as a trigger in a scheduled flow. For example, a file event occurs when a scheduling server determines that a specified file exists.

filter
a set of specified criteria that are applied to data in order to identify the subset of data for a subsequent operation, such as continued processing.

flow
a set of jobs and associated dependencies that is scheduled in the Schedule Manager plug-in in SAS Management Console.

Hadoop
See Apache Hadoop.

job
a collection of SAS tasks that can create output.

job event
a job-related occurrence that is used as a trigger in a scheduled flow. For example, a job event occurs when the scheduling server issues a command to determine whether a job ran successfully.

job flow
a group of jobs and their dependencies, including dependencies on other jobs, on files, or on specified dates and times. *See also* job.

join condition
a combination of join keys and a comparison operator.

L10N
See localization.

local data
data that is accessible through the file systems on a computer. This includes data on hard drives or available through network file systems.

locale
a setting that reflects the language, local conventions, and culture for a geographic region. Local conventions can include specific formatting rules for paper sizes, dates, times, and numbers, and a currency symbol for the country or region. Some examples of locale values are French_Canada, Portuguese_Brazil, and Chinese_Singapore.

localization (L10N)
the process of adapting software for a particular geocultural region (locale). Translation of the user interface, system messages, and documentation is a large part of the localization process.

query
a set of instructions that requests particular information from one or more data sources.

remote data
data that is not accessible through the file systems available to a computer. To use remote data, you must direct a SAS server to access the data that is available through file systems on the remote machine.

report
output that is generated by running custom SAS code against the data in your project.

role (user role)
a set of capabilities within an application that are targeted to a particular group of users.

SAS Management Console
a Java application that provides a single user interface for performing SAS administrative tasks.

SAS Stored Process (stored process)
a SAS program that is stored on a server and defined in metadata, and which can be executed by client applications.

SAS Workspace Server
a SAS server that provides access to SAS Foundation features such as the SAS programming language and SAS libraries.

scheduling server
a server that runs deployed jobs in a scheduled flow. The scheduling server determines whether the schedule criteria and dependencies have been met before a job is run.

source
See data source.

stored process
See SAS Stored Process.

subquery
a query-expression that is nested as part of another query- expression. Depending on the clause that contains it, a subquery can return a single value or multiple values.

time series
an ordered sequence of values of a variable that are observed at equally spaced time intervals.

Unicode
a 16-bit encoding that is the industry standard for supporting the interchange, processing, and display of characters and symbols from most of the world's writing systems.

Unicode Transformation Format 8
See UTF-8.

user role
See role.

UTF-8 (Unicode Transformation Format 8)
a method for converting 16-bit Unicode characters to 8-bit characters. This format supports all of the world's languages, including those that use non-Latin 1 characters.

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.

Index

D

E

F

J

L

M

N

O

S

U

V

W