

An Introduction to SAS® Visual Analytics for the Health Care Sector

Rhian Sian Pilling, Amadeus Software, Oxford, England

ABSTRACT

SAS Visual Analytics software is an in-memory analytic tool offering an intuitive, drag and drop interface enabling users of all abilities to explore their data effortlessly. Insights are shared easily leading to increased collaboration and improvements in the decision making process.

This paper presents an introduction into the SAS Visual Analytics software. Topics include, managing the Visual Analytics environment, the SAS Visual Analytics Data Builder, the SAS Visual Analytics Explorer and the SAS Visual Analytics Reporter. Readers will learn how to load and unload LASR tables, build queries to manipulate tables, create data items, apply filters to data, build interactive charts and tables and learn how to apply chosen properties and styles to their output.

This paper is aimed at all individuals who wish to gain a broad understanding of the SAS Visual Analytics software.

INTRODUCTION

What is SAS Visual Analytics?

SAS Visual Analytics software is a high performance analytics tool developed for organisations' in any industry working with data. Users are able to quickly explore and gain insight from their data by identifying patterns and trends using the simple-to-use descriptive and predictive analytics tasks available, with no coding required.

How can this help users working in the pharmaceutical Industry?

With increasing costs associated with drug discovery, development and regulatory compliance, pharmaceutical companies need a way to analyse their data quickly and efficiently to ensure that business decisions are reliable and safe products are released out onto the market as soon as possible.

When SAS Visual Analytics is deployed an analytic platform referred to as the SAS LASR Analytic Server is created. This Server provides simultaneous multi-user access to data that is loaded into memory, making exploring vast amounts of data instantaneous. Insights made on this data are up-to-date and immediately available ensuring that any business decisions made as a result of these findings are timely and accurate.

Pharmaceutical companies need to share and disclose findings with regulatory boards, colleagues and the public promptly. Visual Analytics provides the functionality to create reports that can be easily distributed via the web, mobile devices and Microsoft applications. This in turn, facilitates collaborative research and knowledge sharing.

SAS Visual Analytics Environment

SAS Visual Analytics Interface

SAS Visual Analytics is a single working environment with the following interfaces where work can be carried out:

- Data Builder
- Analytics Explorer
- Analytics Designer

Interfaces that are available to each user are dependent on the individuals' assigned role.

Roles

All users are assigned to roles and groups in SAS Visual Analytics. These roles and groups define how each user can interact with the Visual Analytics Environment by specifying access rights to specific tasks and data.

Pre-defined roles that can be assigned are listed below:

- Administration Role: Assigned to users that are responsible for managing the security and the visual analytics environment. Global administrators are able to manage users, set up role membership, start and stop LASR servers and load or unload data.

- Data Administration: Assigned to users responsible for ensuring that data is secure, readily available and in the correct structure to use for other visual analytic users.
- Analysis: Assigned to users who require access to all tasks that facilitate data exploration and reporting.
- Report Viewing: Assigned to users who only require access to view reports created in Visual Analytics.

Managements Console

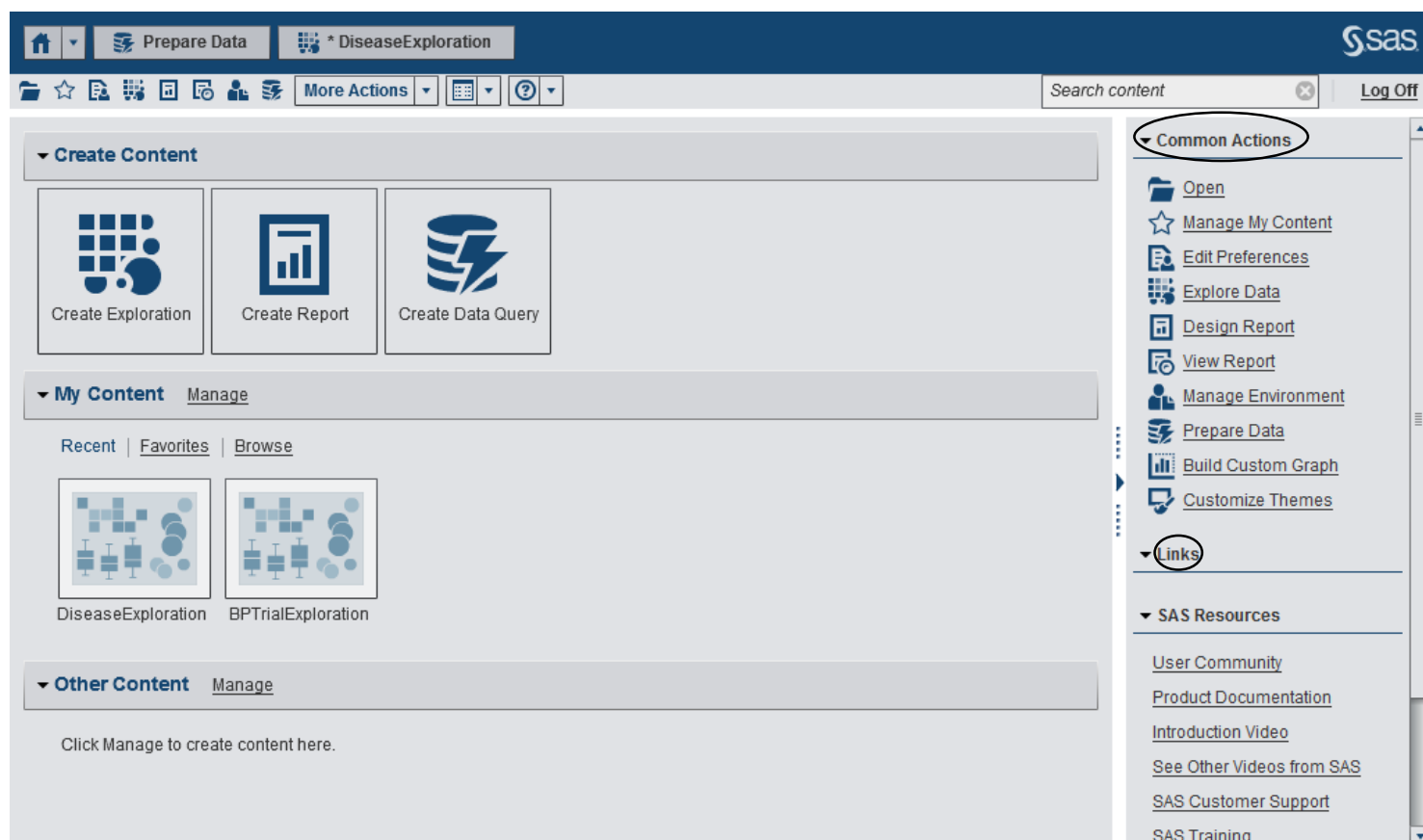
The management console is included in the SAS Visual Analytics package and is used to define users, roles and groups and store information on access rights assigned. However, this topic is beyond the scope of this paper.

Accessing SAS Visual Analytics

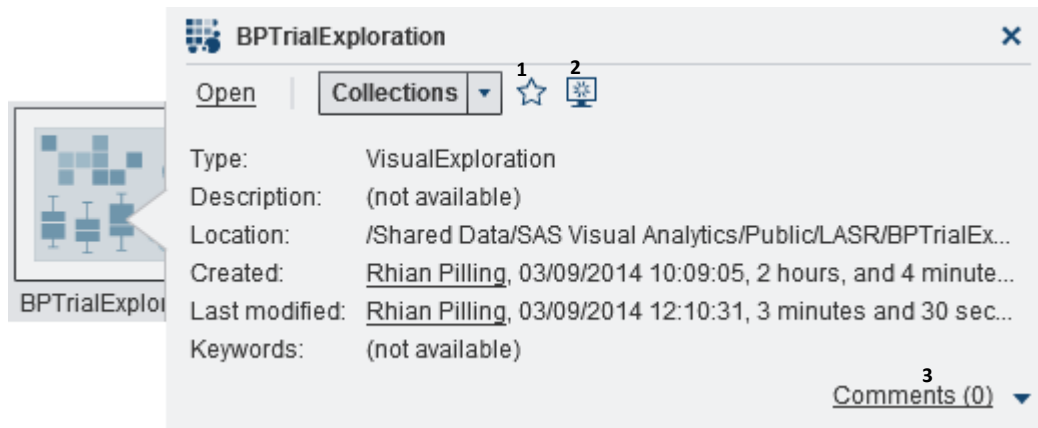
The SAS Visual Analytics Hub is the focal point where each environment is shown. It is the first interface a user will see when logging onto SAS Visual Analytics. As mentioned previously, the Hub is role based and tasks available to each user are dependent on the individuals' assigned role. For example, an analyst will have the access rights to the Analytics Explorer and Designer, whereas a data administrator will also have access to the Data Builder.

Interface of the Visual Analytics Hub

The SAS Visual Analytics hub is presented in two sections. To the right hand side, users can access the common actions pane, links panel and SAS resource panel. The main area of the screen presents the Visual Analytics applications that have been authorised for use, recent items and links to other SAS content.



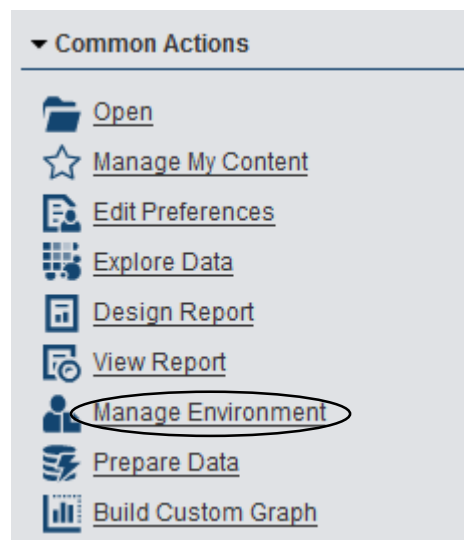
My Content contains a link to browse for explorations or reports and also displays items that have recently been accessed or that have been defined as favourites. Selecting on these items launches the object inspector window. The Object Inspector provides information on when the item was created, who created it, when it was last modified and further additional options to add the item to favourites¹, set as initial screen² and add/view comments³.



Manage the Environment

Administrator Role

Users that have been assigned the role Global administrator will have access to the Manage Environment application accessed under the Common Actions pane. Within this application, administrators can start and stop LASR servers, load and unload data, control access and set row level permissions on LASR tables, and manage devices and alerts sent to users.



All LASR tables and their status are shown within manage environment. The status of each table is viewed by selecting the icon . Tables that have been loaded onto a LASR server are represented by a green circle, those not loaded are represented by a red square. Tables are loaded and unloaded by selecting the appropriate table, right clicking and selecting the options load/unload as appropriate.

Table	Location	Description	Status	Rows	Columns	Library	Server	Loaded
BPTRIAL	/Shared Data/S...	Loaded on We...	●	73	16	Visual Analytic...	LASR Analytic S..	3 Sep 2014 09:...
DISEASE	/Shared Data/S...	Loaded on We...	●	793	8	Visual Analytic...	LASR Analytic S..	3 Sep 2014 09:...
SYSTOLICMON...	/Shared Data/S...	Loaded on We...	●	661	4	Visual Analytic...	LASR Analytic S..	3 Sep 2014 09:...
TRIALDEMOGR...	/Shared Data/S...	Loaded on We...	■			Visual Analytic...	LASR Analytic S..	
DRUGTRIAL	/Shared Data/S...	Loaded on We...	●	69	8	Visual Analytic...	LASR Analytic S..	3 Sep 2014 09:...

By selecting LASR in the menu bar followed by Manage Servers administrators can start and stop LASR servers. An additional tab will appear presenting information for each LASR server. Selecting a server status and right clicking, servers are stopped or started as necessary. Administrators are also able to load data directly onto the LASR server by selecting the load a table option.

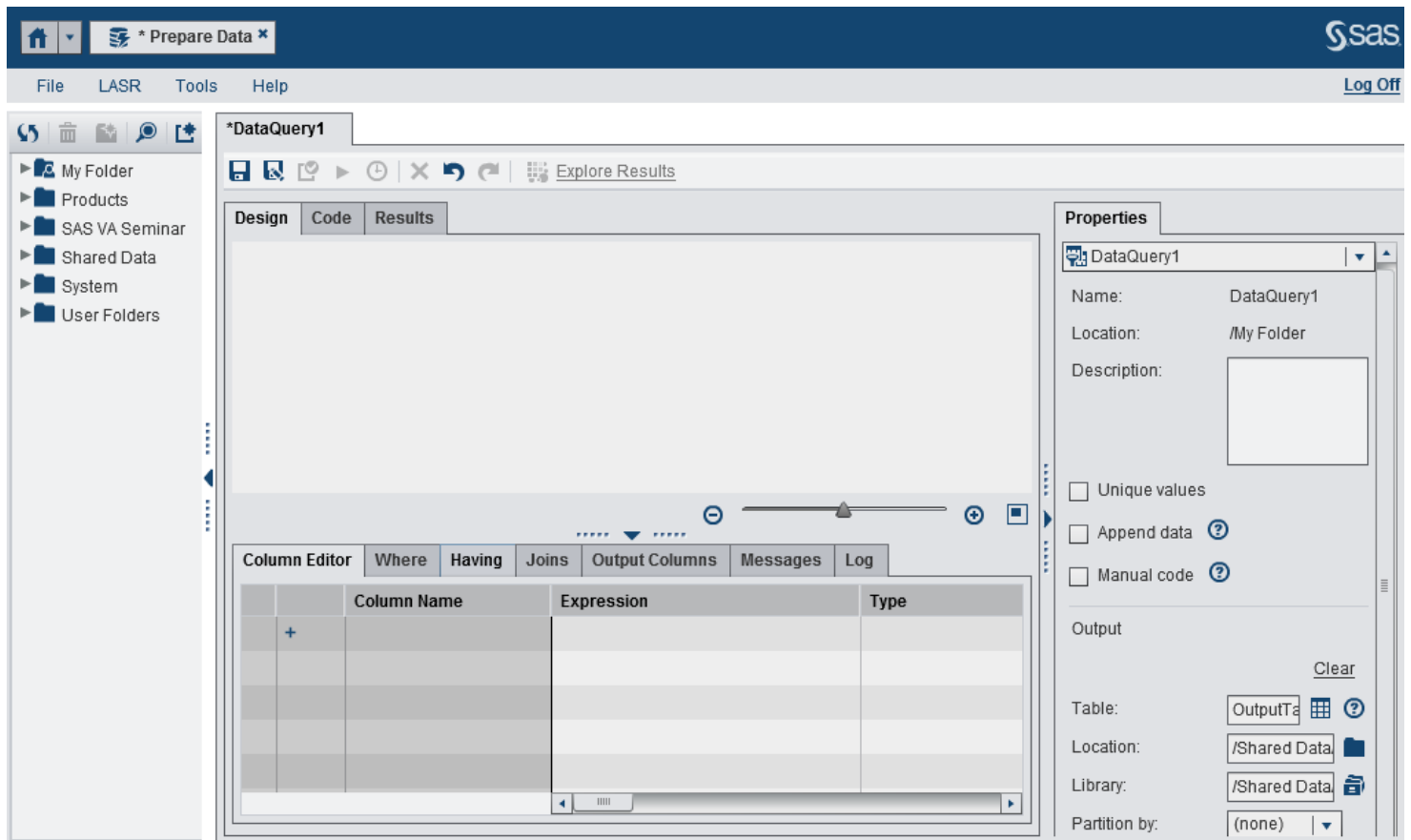
Server	Status	Host	Port	Started By	Description	Started
LASR Analytic Server - avsas94va	●	avsas94va.am...	10010	Theodora.Eleft...	SAS LASR Anal...	4 Mar 2014
Public LASR Analytic Server -	■	avsas94va.am...	10031		Visual Analytics...	

Within Tools in the menu bar further options are available to view devices that have been given rights to view/receive data. These mobile devices are managed through the Mobile device tab and include the options to blacklist or whitelist devices. Right clicking on the device, the appropriate options can be selected.

The SAS Visual Analytics Data Builder

Users that have been assigned to a Data Administrator role will have access to the SAS Visual Data Builder. This application is used to load data to the LASR server and prepare this data ready for exploration and analysis.

Access to the Visual Data Builder is through the Create Data Query icon located in the hub. The following interface is launched:



The interface contains a number of separate panes. The left pane (navigation pane) is used to select data that will either be used as input data sources when creating queries or loaded directly to the LASR server. This data must have already been loaded into metadata using the management console. The right pane (properties pane) specifies the general properties of the query along with information on the location, library and output table name. The lower pane, contains all the options available that can be used to prepare the data. The query is built in the workspace pane, which has a drag and drop facility.

How to Load Data?

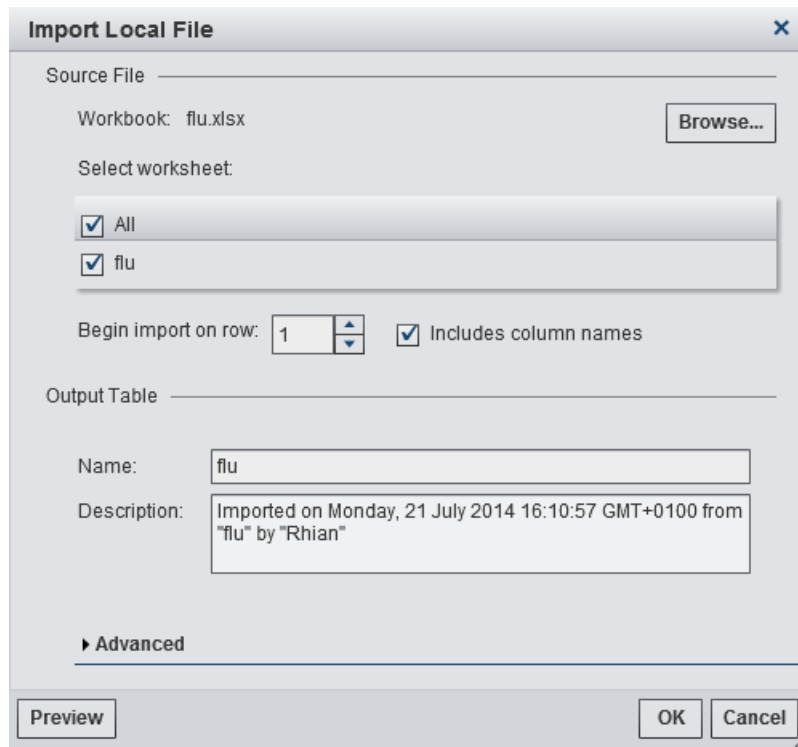
To load data to a server, the appropriate server should be started through the manage environment task. Data can then be loaded onto a LASR server in several ways.

- Importing files available locally to the end user;
- Tables can be registered in SAS Metadata
 - Loaded through a Query
 - Directly loaded onto the Server

A default LASR library is created when SAS Visual Analytics is deployed. The default library name given is Visual Analytics LASR.

Loading External Data

External files are imported through the option **File>Import>Local File**. Users control the method in which the data file is read in through options available on the Import window. Advanced options allow the user to specify the metadata location (default LASR Library created) and the location of the table created.



Import Local File

Source File _____

Workbook: flu.xlsx Browse...

Select worksheet:

- ☒ All
- ☒ flu

Begin import on row: 1 ☒ Includes column names

Output Table _____

Name: flu

Description: Imported on Monday, 21 July 2014 16:10:57 GMT+0100 from "flu" by "Rhian"

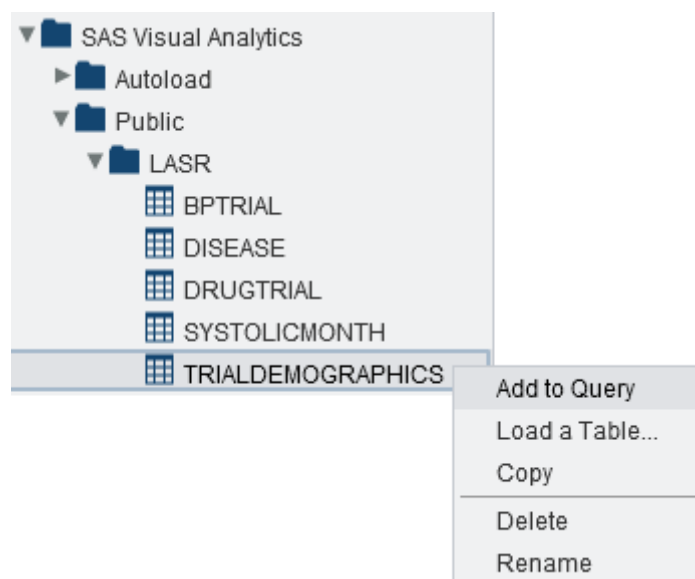
► Advanced

Preview OK Cancel

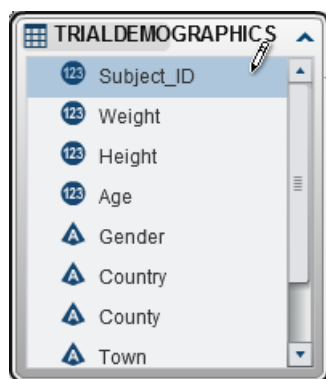
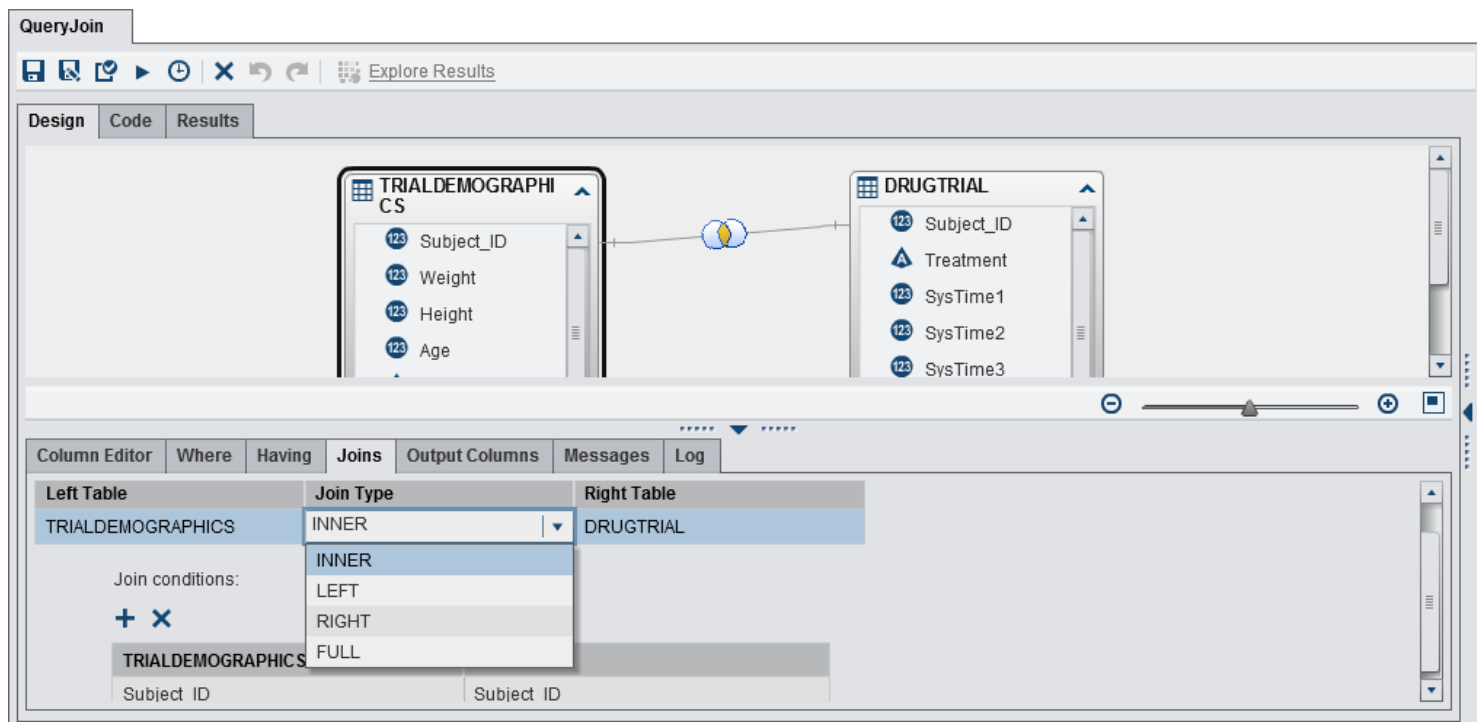
Loading through a query

A query is a metadata object that allows the user to load data into the LASR Server. As previously mentioned, this data must have already been loaded into metadata using the management console. A query is used when users wish to manipulate tables prior to their load. For example, to combine several data files, create derived variables, edit columns or perform data manipulations.

Tables are added to a query by selecting the table in the navigation pane, right clicking the table and selecting add to query. Any number of tables can be added to a single query.

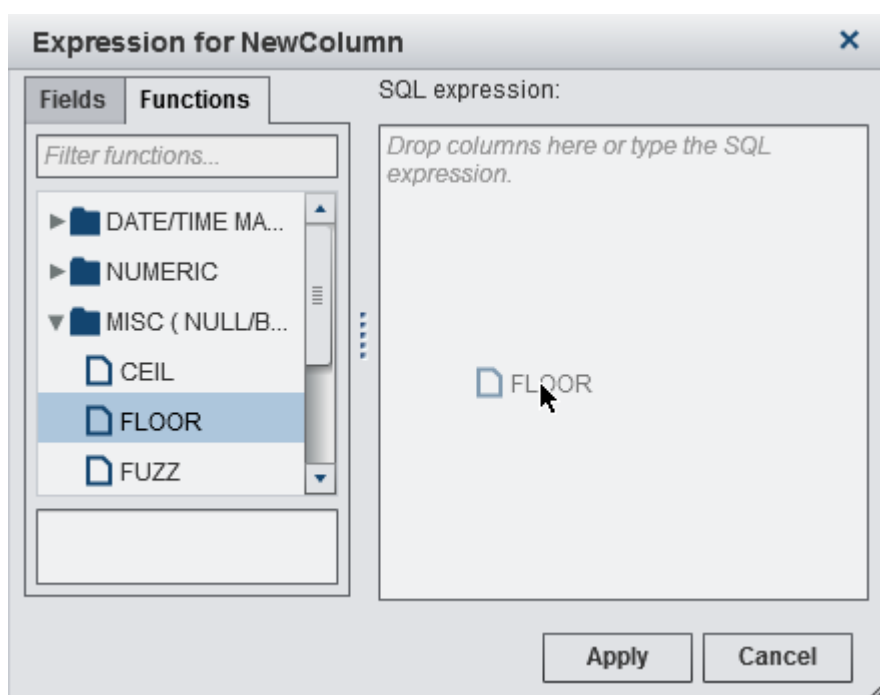


When more than one table is selected, an automatic join (if feasible) between the tables is created. By selecting the join tab in the lower pane and join type these can be altered. In addition, further joins can be created by selecting the + icon.



Variables to output to the new table must also be specified. This is achieved by selecting the appropriate variables in each table in the design pane. A pencil icon appears allowing you to do this.

Filtering options are used to subset data. There are two options available to users to filter the data. The having clause and the where clause. Users can create expressions for each under the corresponding tabs. Both clauses, create an SQL expression. The where clause applies this filter to the input data set, whereas the having clause also applies this filter to any columns that have been calculated.



Within the column editor users can create derived columns. To create a calculated column, the user must first create a new column in the column editor using the + option, then select the expression icon in the column row. The expression window opens, here the user can use the drag and drop feature to create the SQL expression. Expressions are made up of data fields and functions and can be as complex or as simple as the user wishes.

Validating and Running a Query

A query is validated though the icon . Any information relating to the query run will then be written to the log. When the logic is accurate, the query can be saved and run. Prior to saving a query, the user must assign query properties (right pane), including where the output table is to be saved and the metadata location.

Loading data directly

Occasionally data that has been registered in metadata, needs no further processing and can be loaded onto the LASR analytic server directly.

Within the navigation pane, users right click on the table, and select the option Load a table. A window launches where the user specifies the LASR table name, the location of the table library and the location of the metadata library. All source table information is automatically filled in.

LASR Server

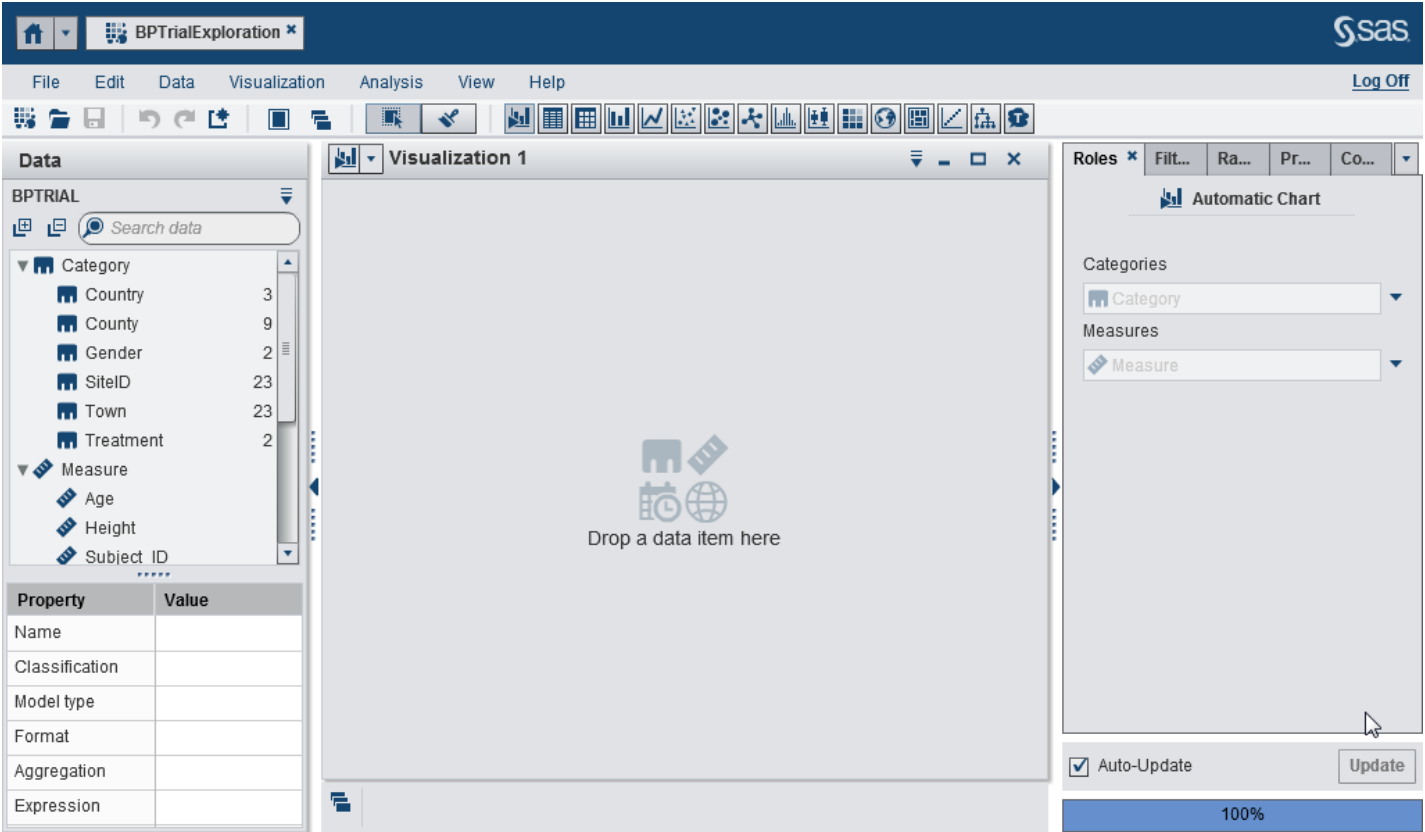
Data that has been loaded into the LASR Server remains in memory until it is unloaded or the server is terminated. The amount of tables that can be stored in the LASR server is restricted by the hardware and memory available.

Once data has been loaded into a LASR Server it is then available to use in other applications including SAS Visual Analytics Explorer and SAS Visual Analytics Report.

The SAS Visual Analytics Explorer

The SAS Visual Analytics Explorer is used as a data discovery and visualisation tool. Data is explored through visualisations which present an interactive view of the data. Types of visualisations created include charts, tables and analysis tasks including forecasting, correlation and decision trees. A collection of visualisations is referred to as an exploration.

The SAS Visual Analytics Explorer is accessed through the create exploration icon on the hub. The following interface is launched.

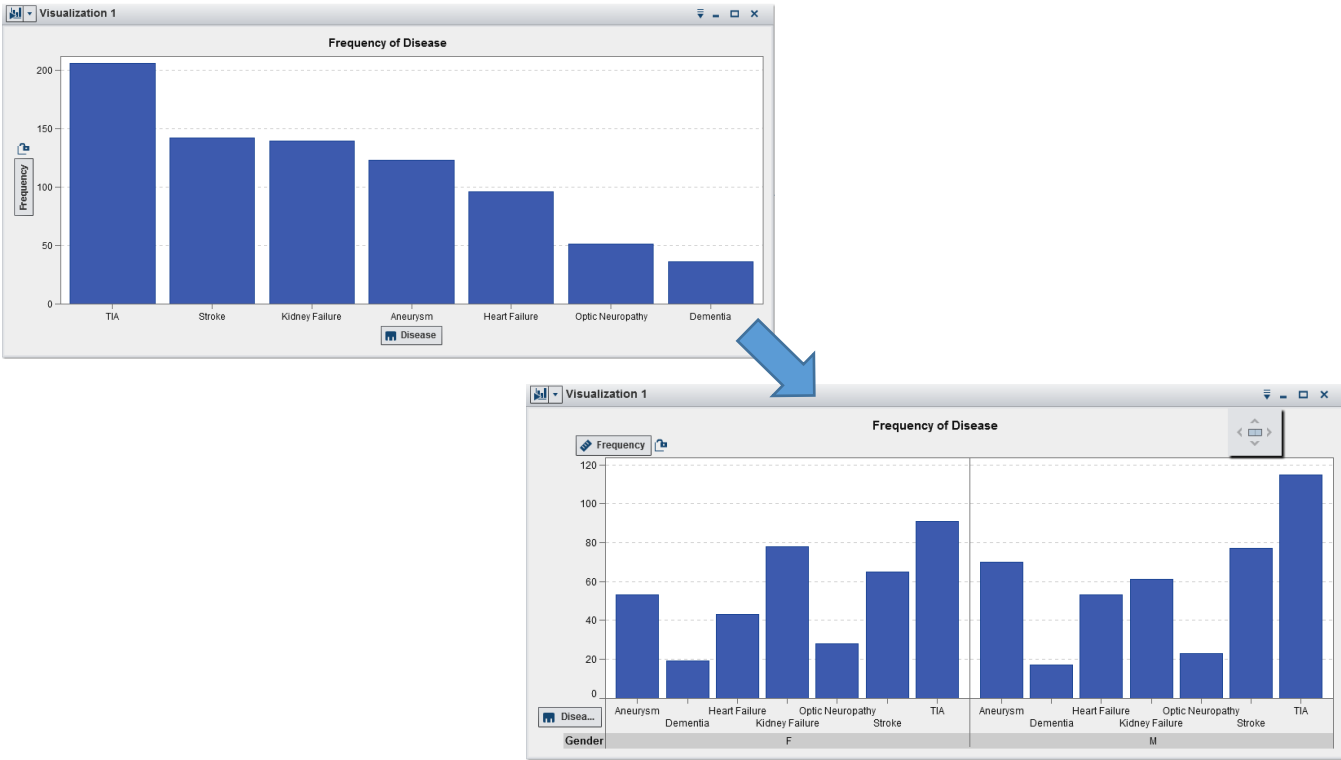


The left hand pane (data pane) contains information on variable properties contained in the data set and access to the data drop down menu . The right hand pane (properties pane) contains options to change the features and properties' of visualisations. Visualisations are created in the middle pane (workspace pane).

Creating Visualisations

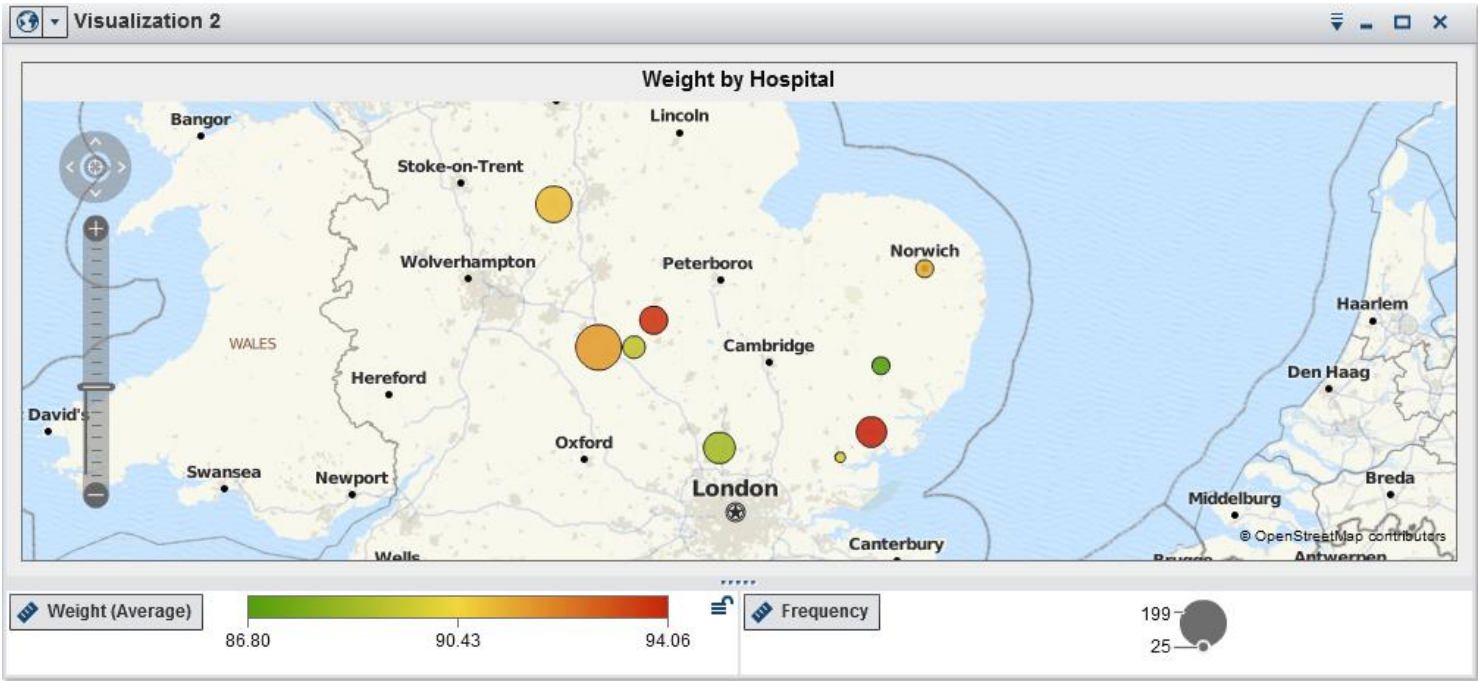
Visualisations are created quickly and easily through the drag and drop facility or by selecting specific charts and assigning the appropriate roles in the drop-down fields located in the properties pane.

By default, the automatic chart feature is selected. The automatic chart facility creates a chart that is applicable to the type of variables that have been selected. For example, when a categorical variable is selected a bar chart is automatically created. The chart will automatically update as more variables are added.



Alternatively, specific charts can be selected from the visualization drop down menu or menu bar. SAS Visual Analytics has very powerful geographical presentations.

If a user wishes view the spread of a particular continuous variable over specific locations, the appropriate plot to create is a geomap. Selecting geomap from the drop down option the appropriate visualisation template is created. Variables can then be dragged and dropped onto the appropriate positions or assigned through the roles tab in the properties pane.



It is important that when using geomap, that one of the data categories is defined as a geographic data item. This is achieved by right clicking on the data item and selecting geography.

A number of different options will be presented. Select the most appropriate for your data. For example, if in your data you have data items specifying the latitude and longitude of a location then the custom option can be selected. The following window appears. Select the data items representing latitude and longitude and select ok.

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, Province) Names
Subdivision (State, Province) SAS Map ID Values
Custom...
US State Names
US State Abbreviations
US ZIP Codes

Custom Geography

Latitude
Longitude
Coordinate Space

OK

Cancel

The new geographic data item will be shown in the data pane. This item can now be used in the geomap.

Data

DISEASE

+

+

Search data

Disease

7

Gender

2

Geography

Hospital

9

Measure

Age

Minimising any visualisations will send the visualization to the lower dock pane. New visualizations are created by selecting the icon  or alternatively through the visualisation tab on the menu bar.

Data Pane and Variable Properties

Only one data set can be explored at a time in the SAS Visual Analytics Explorer. Once that data has been loaded into the application, information on the variables is presented in the data pane.

Data

BPTRIAL

+

+

Search data

Country

3

County

9

Gender

2

SiteID

23

Town

23

Treatment

2

Measure

Age

Height

Subject_ID

SysTime1

Property	Value
Name	Age
Classification	Measure
Model type	Continuous
Format	Numeric
Aggregation	Sum

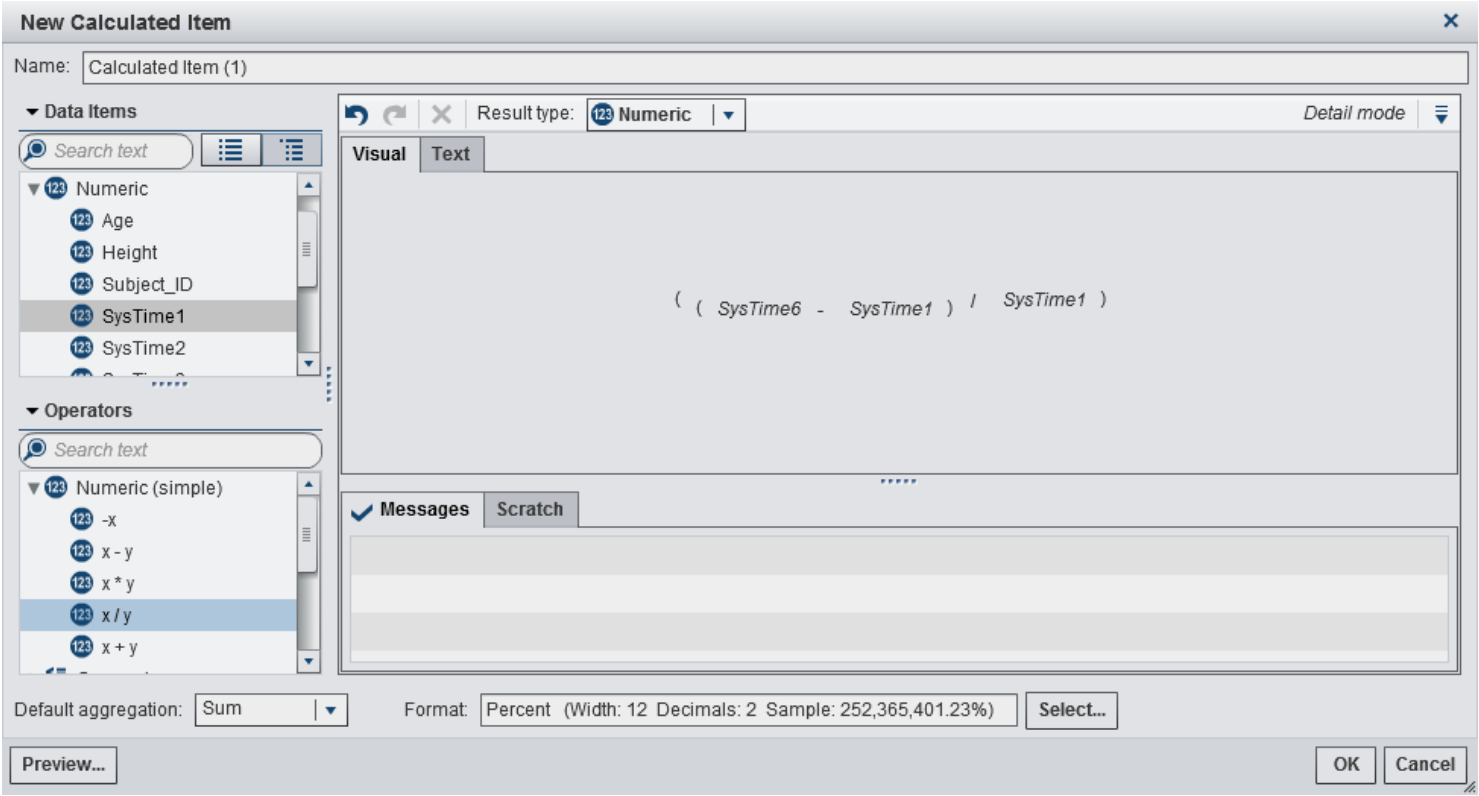
The data pane is split into categorical and numeric variables. Users can change the properties of any selected variable including the variable name and role. For numeric variables changes can also be made to their modelling type, format and aggregation. Changes are made by selecting the black triangle located in the bottom right hand corner.

In many situations, users also require the ability to calculate new data items. Within the explorer new data items can be calculated using either raw or aggregated data.

Both options are available to the user through the data drop down menu.

- **New calculated item** will create a new data item based on raw data.
- **New aggregated measure** will calculate a new data item on aggregated data.

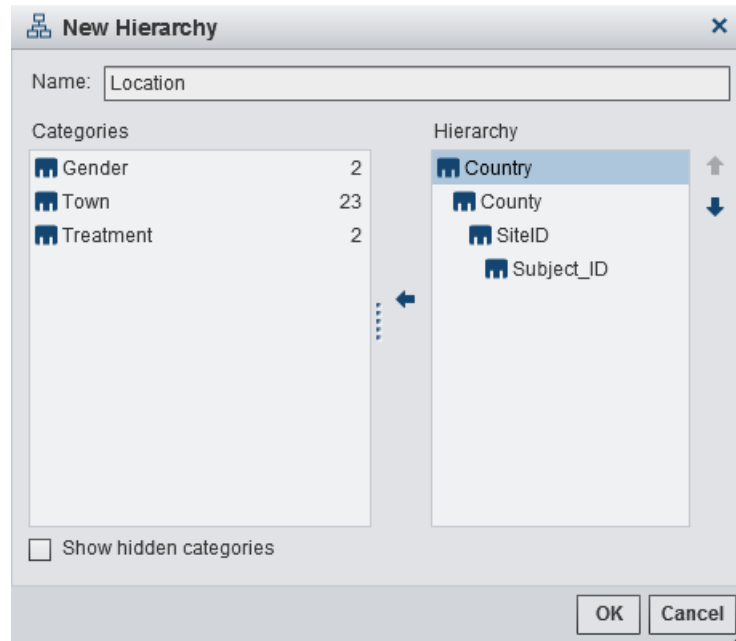
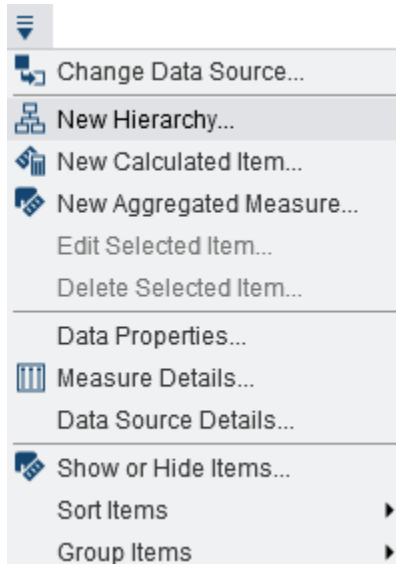
In both situations, the following window appears. Users can create an expression by dragging specific operators and data items from the data and operator panes. Any calculated items will then be presented in the data pane with the following icon .



Hierarchies

The interactive nature of SAS Visual Analytics Explorer allows users to visually explore any data that follows a hierarchical structure. By creating hierarchies, users are able to explore and investigate the different levels in the data. A hierarchy is made up of categorical data items and usually starts with a higher overall view of the data, followed by measures that are more detailed.

To create hierarchies, users' select the option create new hierarchy within the data drop down menu located in data pane. All categorical variables present in the data are listed. Users drag and drop variables into the hierarchy pane or use the arrows presented to create hierarchical structures.



Hierarchies created are presented under a separate hierarchies section within the data pane. Right clicking on the hierarchies, users are able to delete, rename or edit them.

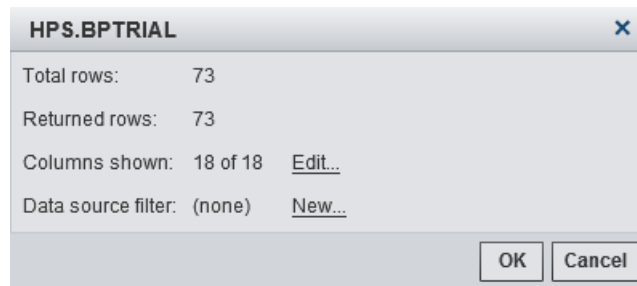
Filter Data

Within SAS Visual Analytics Explorer, there are a number of ways users can filter their data to select certain observations.

- Local filter: Applies to the current visualization
- Global Filter: Applies to all visualizations
- Data filter: Applies to variable(s) and affect all visualizations.

Global and Local filters are created by selecting the filter tab in the right pane and either dragging specific variables onto the tab or by right clicking on the variable in the data pane and selecting the filter type option required.

Data filters are applied through the pull down menu > **Data Source Details>New**. A filter window opens, with drag and drop features to create an expression as seen previously.



Comments

Users with the appropriate permissions are able to read and write comments on visualizations and explorations. Adding comments to visualizations provides the facility for collaboration between colleagues.

To add a comment to a specific visualization, users must first select the visualization and navigate to the comments tab in the right pane. To add a comment to an exploration, users must select **file > exploration comments** in the menu bar. In both instances the following window pane is viewed.

The image shows a web interface for searching and posting comments. At the top, there is a search bar with a magnifying glass icon and the text "Search within comments". Below the search bar, a message states "No comments were found." The bottom section of the interface contains a text input field with the placeholder text "Enter a topic name". To the left of this field is a paperclip icon for attaching files. To the right of the field are two buttons: "Post" and "Cancel". At the bottom right of the entire interface is a "Close" button.

Users can search for previous comments and add additional comments. To add a comment the user must first enter a topic name.

Exporting Visualisations and Data

Visualisations or the data used to create these visualisations can be exported from SAS Visual Analytics Explorer. Right clicking on a visualisation these options are given. Exported images are saved as Portable Network Graphic (PNG) files and data is saved as Comma Separated Value (CSV) files.

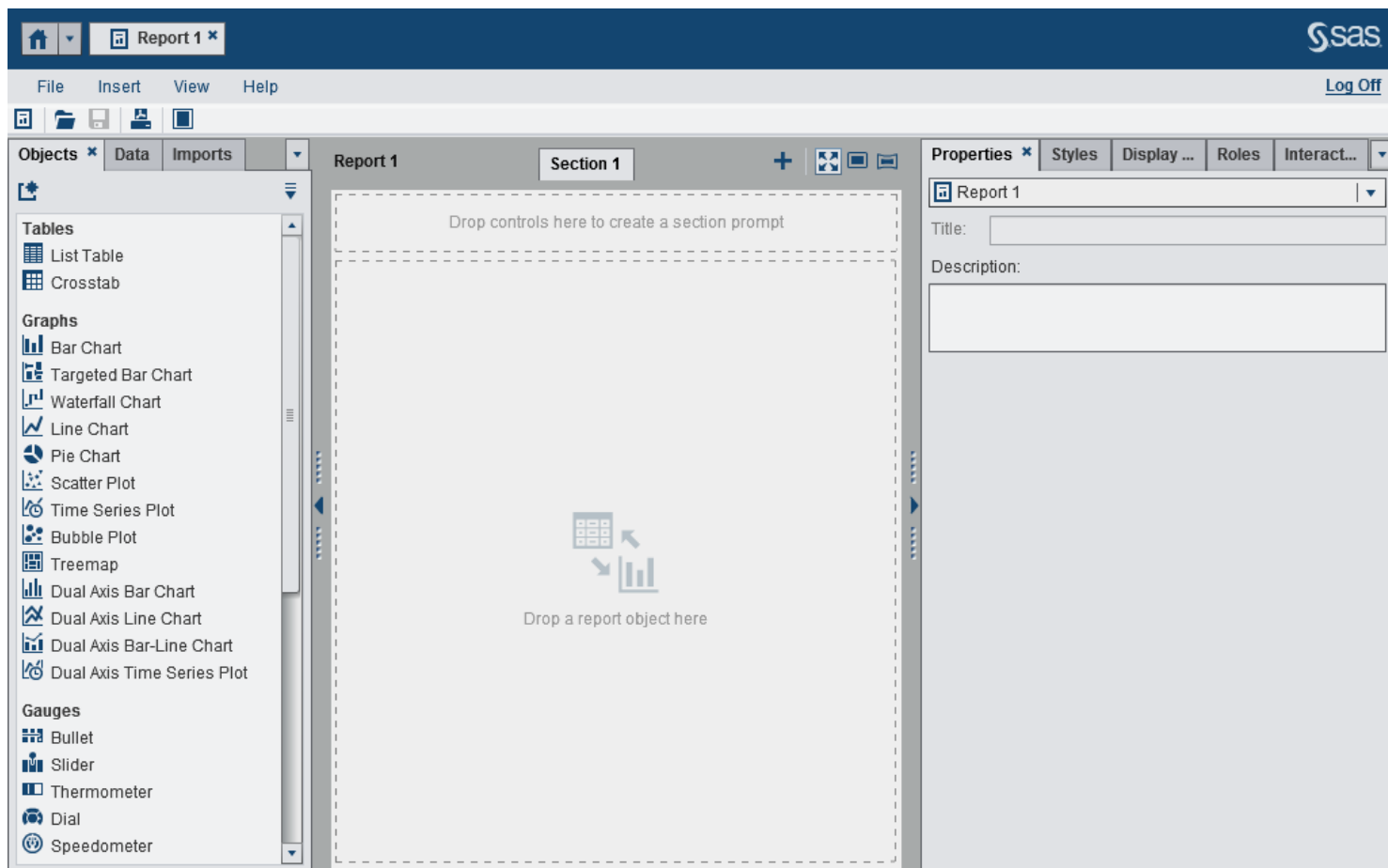
In addition, the export wizard provides the facility to export explorations as reports or pdf files. Explorations can be exported as a whole, or users are able to select specific visualisations. The export wizard is located under **File > Export**.

SAS Visual Analytics Designer

There are similar features in both the explorer and the designer. However, the explorer was primarily designed as an exploratory tool and therefore contains more statistical features in contrast, the designer was designed as a tool to create reports and dashboards and contains more presentation and style options.

Reports created in the Visual Analytics Designer can be made up from multiple sections and unlike the explorer, the designer can use multiple input sources within each section.

The Visual Analytics Designer can be accessed through the icon Create Report on the hub. The following interface launches.

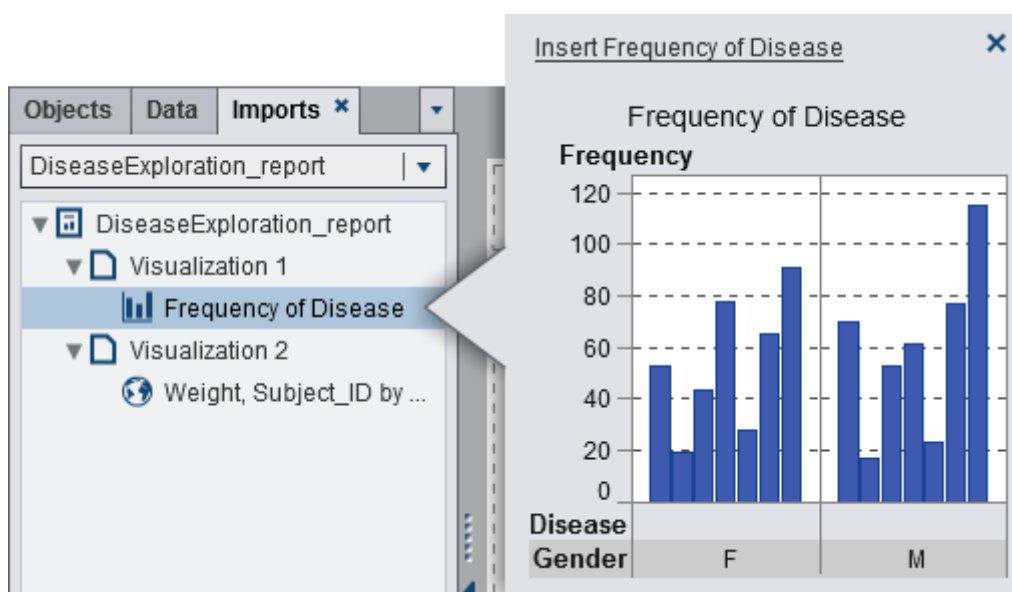


The left hand pane contains three tabs. The object tab, contains charts, plots and filters that can be added to a report. The data tab is used to select between data sources used in the report. Selecting data sources is achieved through either;

- The arrow in the Select a Data source field;
- Through the Add data source  icon.

As seen previously in the explorer, the data pane also contains a list of the variables present in the data set split into categorical and numerical sections. Properties for these variables can also be altered by changing the properties in the bottom portion of the window.

Through the import tab, users are able to open previously viewed or saved reports and any visualizations created. When a report is selected, users are able to select the specific objects they wish to import from the report.



The right pane also contains a number of tabs. A description about the report can be written in the properties tab. In addition the styles and display rules can all be modified to change the presentation and layout of selected reports. When selected on an object, these tabs will be linked back to the object, allowing the user to make changes to presentations and styles.

The centre pane is where objects are created. Unlike the explorer, the designer does not have an automatic chart facility, therefore all charts and tables are created by dragging the appropriate object onto the workspace panel. Multiple objects can be dragged onto the workspace.

Layout

By default these objects will be placed in a tiled layout format. Users have the option to change this to a precision format, allowing the user more flexibility in the positioning of objects. This option is accessed via selecting the section tab at the top of the workspace, and then selecting the properties tab and the layout drop down option.

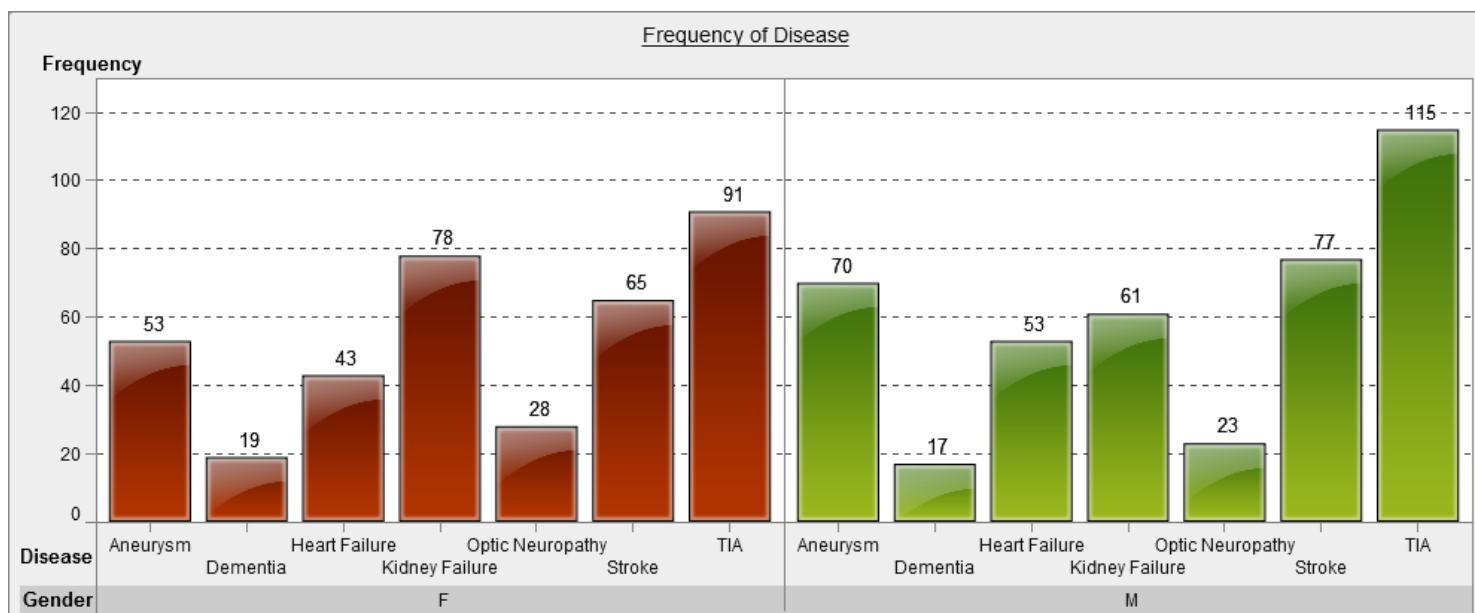


Data items are assigned to the object roles in a similar fashion to the explorer, whereby variables can be dragged and dropped onto the appropriate areas in the visualisation or roles can be assigned using the drop down options in the right hand pane under the roles tab. When objects are selected, the right hand pane and all its tabs relate to the selected object.

Report Presentation

There are a number of ways that users can enhance the reports created in the designer.

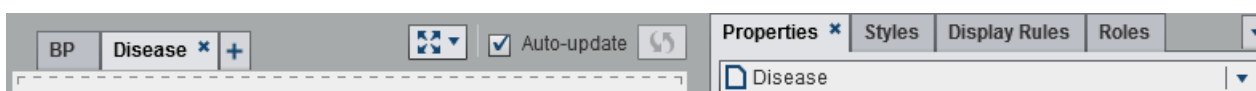
Users are able to modify frame and text styles as well as the colours used for any object under the tab Styles. Under the properties tab, users can also specify titles, and alter additional properties specific to the object including for example adding data labels and altering object layout. All properties and styles available will vary according to the object.



Creating Multiple Sections

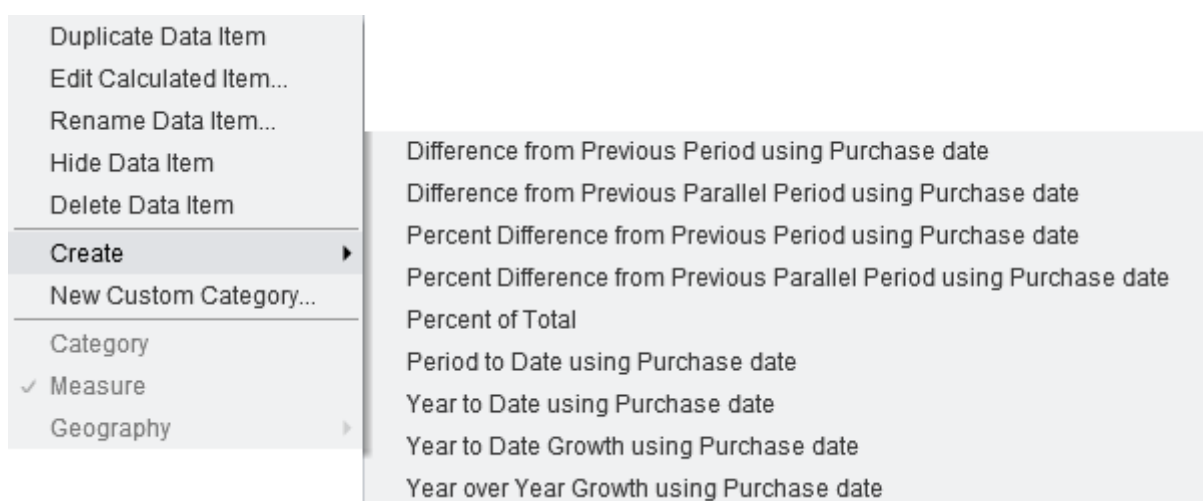
Users can create multiple sections within a report and there is no limit to the amount of sections that can be created. Sections are useful in order to group together similar views of the data.

To create multiple sections the + option at the top of the pane next to the first section report is selected. By default each section will be called section1 through to sectionn. This is altered through the section properties tab in the right hand pane.



Calculated items and hierarchies

As seen in the explorer, new items and hierarchies can be created. However, within the designer users are able to easily create a number of aggregated measures including a percentage of total value for a numeric variable. This item is derived by right clicking on the measure and selecting **Create > Percent of Total**. Other aggregated measures created are also shown below. Only those that are applicable can be selected. These options will be highlighted in bold.



Filtering

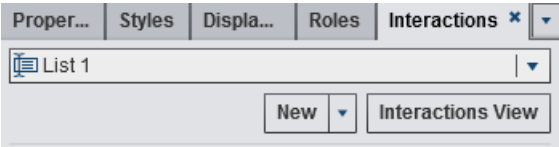
As in the explorer, users may wish to filter data to restrict the observations shown in a report. Data can be filtered through, data item filters, section prompts or using control objects and interactions.

Data item filters are global and affect all objects in the report. Data item filters are added by right clicking on the variable in the data pane and selecting New Data Item Filter. Users are then able to specify a filter range for continuous variables or for discrete or categorical variables select values presented in a list.

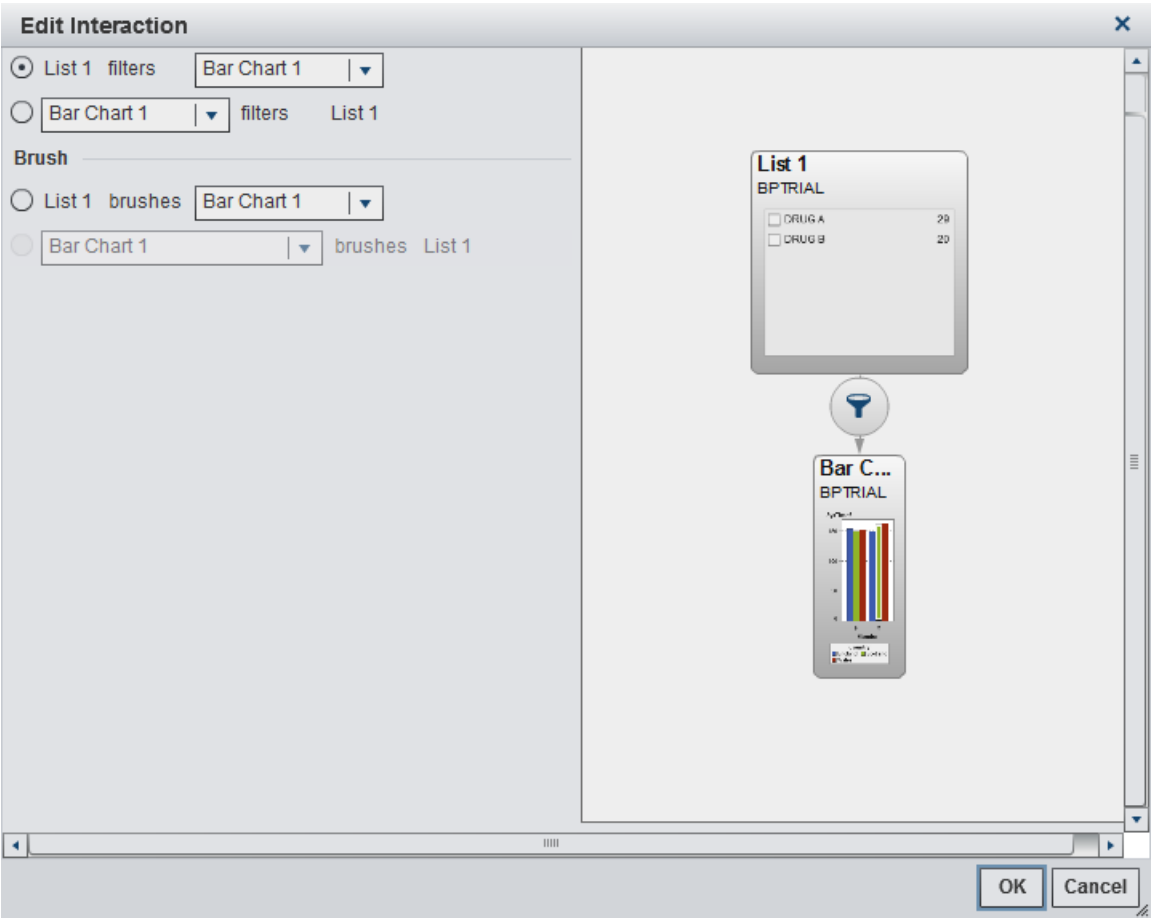
To create section prompts, control objects found in the objects tab are dragged to the upper section of a report. Variables to be used as filters are then dragged on top of the appropriate control filter. Section prompts are local to the section in which they have been specified and restrict the data shown.

Control objects can also be used to filter values when used as part of an interaction. Users can drag the control object onto the workspace and then define an interaction between the control object and chosen report objects.

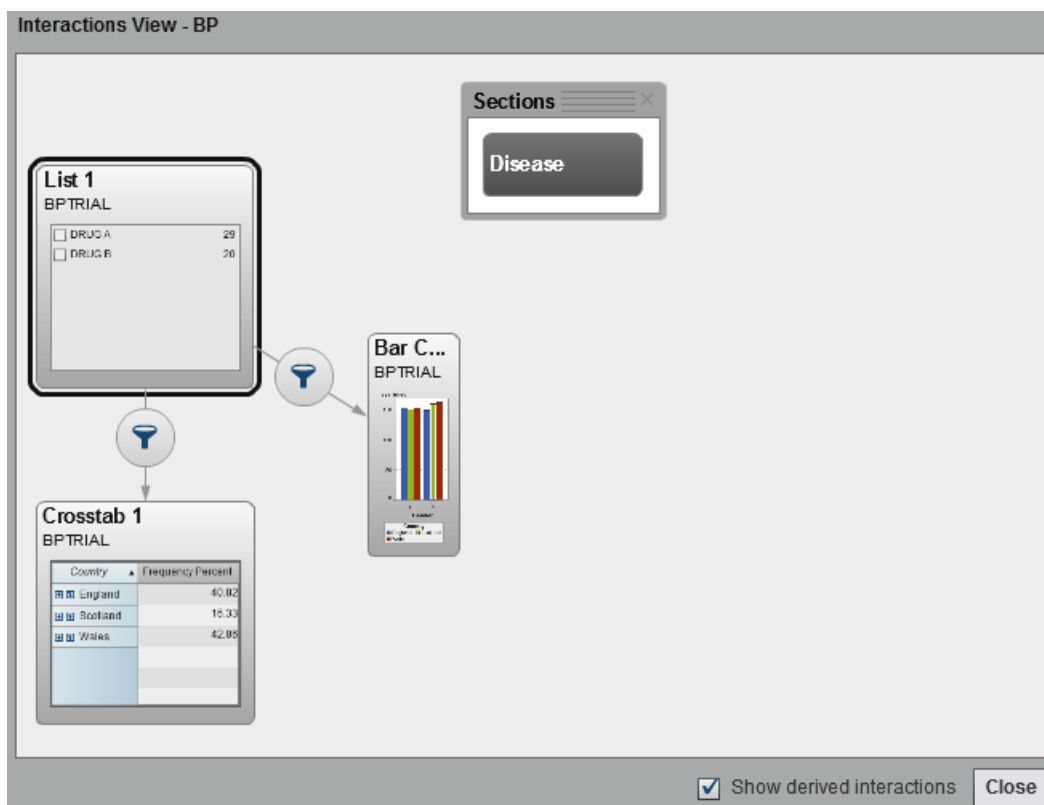
Interactions are included by first selecting the Interactions tab using the drop down arrow on the right pane. Users can then select the buttons new or Interactions view.



Selecting new interaction the following window appears:

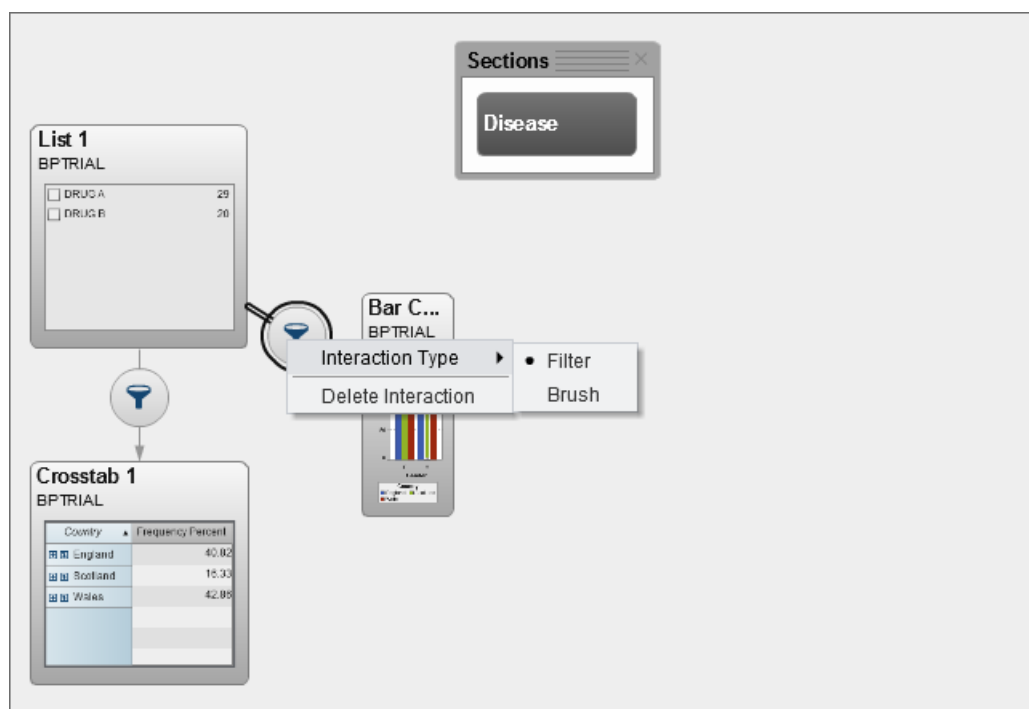


Within this window, interactions can be modified, created and altered. By selecting Interactions view the following window is shown.

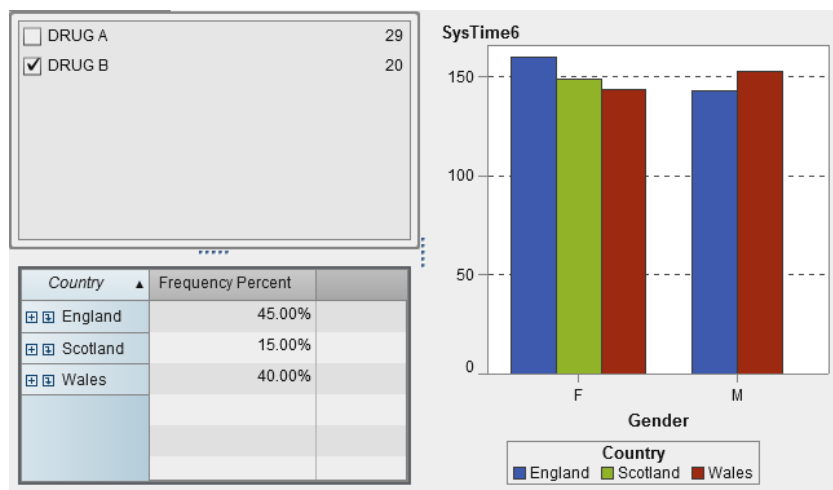
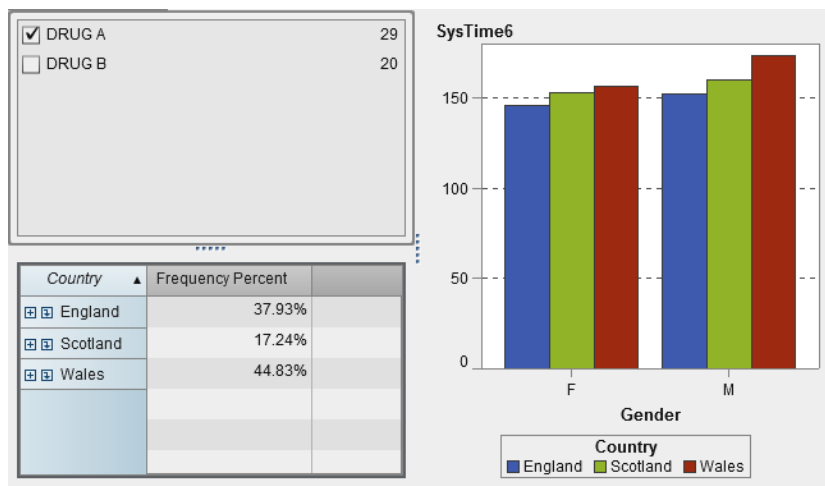


By opening up the interaction view on the interaction tab, interactions between objects can also be viewed and modified. Interactions are created by drawing a connection between the source to the target report object.

There are two types of interactions that can be created. A filter interaction and a brush interaction. By default the interaction is a filter, to change this to brush right click and select brush. The filter interaction is used to restrict the data that is displayed. The brush filter, enables users to highlight sections of data and to select these observations simultaneously in other objects.



In this example, the filter interaction is selected. List 1 filters each chart and restricts each output to represent the values selected. This is shown below:



CONCLUSION

In conclusion, SAS Visual Analytics in-memory technology ensures users of all abilities are able to explore large volumes of data quickly and efficiently.

Individuals working in the pharmaceutical company can quickly understand and monitor their data helping to identify key insights and areas of concern or improvement. Its fast in-memory analysis allows data to be explored instantaneously and with additional features including the ability to create hierarchies, calculate new measures and apply filters this data can be explored from a number of different perspectives ensuring any decisions made are informed, timely and accurate.

SAS Visual Analytics provides the functionality to ensure that these results and findings can be shared whenever and wherever needed thus promoting data transparency.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Rhian Sian Pilling
Amadeus Software Limited
Mulberry House
9 Church Green
Witney, Oxfordshire
OX28 4AZ
Work Phone: 01993 848010
Email: rhian.pilling@amadeus.co.uk
Web: www.amadeus.co.uk
Brand and product names are trademarks of their respective companies.