

MashZone NextGen Visual Analytics User Guide

Innovation Release

Version 10.0

April 2017

This document applies to MashZone NextGen Visual Analytics Version 10.0 and to all subsequent releases.

Specifications contained herein are subject to change and these changes will be reported in subsequent release notes or new editions.

Copyright © 2006-2017 Software AG, Darmstadt, Germany and/or Software AG USA Inc., Reston, VA, USA, and/or its subsidiaries and/or its affiliates and/or their licensors.

The name Software AG and all Software AG product names are either trademarks or registered trademarks of Software AG and/or Software AG USA Inc. and/or its subsidiaries and/or its affiliates and/or their licensors. Other company and product names mentioned herein may be trademarks of their respective owners.

Detailed information on trademarks and patents owned by Software AG and/or its subsidiaries is located at <http://softwareag.com/licenses>.

Use of this software is subject to adherence to Software AG's licensing conditions and terms. These terms are part of the product documentation, located at <http://softwareag.com/licenses> and/or in the root installation directory of the licensed product(s).

This software may include portions of third-party products. For third-party copyright notices, license terms, additional rights or restrictions, please refer to "License Texts, Copyright Notices and Disclaimers of Third Party Products". For certain specific third-party license restrictions, please refer to section E of the Legal Notices available under "License Terms and Conditions for Use of Software AG Products / Copyright and Trademark Notices of Software AG Products". These documents are part of the product documentation, located at <http://softwareag.com/licenses> and/or in the root installation directory of the licensed product(s).

Use, reproduction, transfer, publication or disclosure is prohibited except as specifically provided for in your License Agreement with Software AG.

Table of Contents

Introducing MashZone NextGen Visual Analytics	5
Starting Visual Analytics.....	6
Visual Analytics user interface.....	7
Visual Analytics Explorer.....	7
Visual Analytics Console.....	10
Using Visual Analytics	13
Import data source.....	14
Create analysis.....	14
Select data source.....	14
Select analysis criteria.....	15
Partitioning data.....	15
Change criteria type.....	16
Select criteria values.....	16
Specify query filters.....	16
Delete criteria.....	17
Refresh criteria data.....	18
Calculate full result.....	18
Save analysis.....	19
Edit analysis properties.....	19
Set visualization.....	20
Open an analysis.....	20
Select chart type.....	20
Change axes assignment.....	23
Show trend line.....	24
Split the analysis view.....	24
Set data limit.....	25
Display data table.....	25
Save analysis.....	25
Freeze real-time data visualization.....	26
Share analyses.....	26
Create projects.....	26
Using the configuration panels.....	27
Which aggregated values are available?.....	28
Manage data sources	29
Manage historical data sources.....	30
Manage file-based data sources.....	30
Create file-based data source.....	31
Delete existing file-based data source.....	31
Update data of existing file-based data source.....	31

Update meta data of existing file-based data source.....	32
Manage cache-based data sources.....	32
Create cache-based data source.....	32
Delete existing cache-based data source.....	33
Update cache-base data source.....	33
Manage real-time data sources.....	34
Create real-time data source.....	34
Delete existing real-time data source.....	34
Update real-time data source.....	35
Real-time data source registration.....	35
Define date and time formatting pattern.....	36
How does Visual Analytics handle data sources?.....	37
Historical file-based data sources.....	37
Historical cache-based data sources.....	37
Real-time data sources.....	38
Visual Analytics Administration.....	41
Manage user privileges.....	42
Create user.....	42
Delete user.....	43
Change user password.....	43
Change user role.....	44
Create user password.....	44
Set SSL configuration.....	45
Enabling TLS.....	45
Setup of certificates and key material.....	45
Enabling TLS with JKS-based key material.....	46
Enabling TLS with PEM-based key material.....	46
Enabling TLS with PKCS#2-based key material.....	46
Apache Spark Integration.....	47
Advanced Configuration.....	47
Deployment of Spark on a separate machine.....	48
Importing larger data sources.....	48

1 Introducing MashZone NextGen Visual Analytics

- Starting Visual Analytics 6
- Visual Analytics user interface 7

MashZone NextGen Visual Analytics (short: Visual Analytics) provides the means to interactively visualize, explore, analyze, and blend streaming and historical data from multiple sources. Visual Analytics has a particular focus on Internet of Things (IoT) use cases, but the capabilities provided also support other use cases in retail, financial services - anywhere you have large amounts of data in motion. Visual Analytics allows operations analysts, maintenance specialists, researchers, developers, and data scientists to do their job more effectively.

Visual Analytics enables you to

- Explore and analyze real-time data sources and historic data using interactive visualizations
- Highlight and color measure visualizations using dimensions and other measures
- Capture, save, and share data explorations
- Quickly discover new data insights via automatic statistical profiles
- Filter data as you explore
- Auto-discover and deploy new data sets

Visual Analytics detects and analyzes automatically your deployed data sources. Based on the data source structure Visual Analytics provides you with all relevant analysis criteria. These include dimensions, measures and several aggregated criteria values, e.g., the average value or the standard deviation. Using the analysis criteria available you are able to configure your data queries and to explore the source data. And by means of various graphical elements you visualize interactively your analyses.

Starting Visual Analytics

You can start and use Visual Analytics as web application in your web browser.

Prerequisite

Visual Analytics has been installed.

Note: Software AG Installer enables you to install MashZone NextGen Visual Analytics. Detailed information on how to use Software AG Installer is available in the documentation **Using the Software AG Installer**.

Procedure

1. Start the Visual Analytics server. For this open the Windows Start menu and click **Start MashZone Next Gen Visual Analytics** in the **Software AG > Start servers** program group.
2. Start the Visual Analytics user interface.
 - a. Open a web browser.

- b. Enter the following address in the address bar of the web browser. The default port number is **9002**.

<computer name>.<dnsdomain>:<port>

e.g., localhost:9002

The Visual Analytics start page will be opened in your Web browser tab.

3. Click **Explore**.

The Visual Analytics login page will be displayed.

4. Enter your user name and password and click **Log In**. For example, **Administrator/manager** if you want to log in as administrator user with the corresponding administration privileges.

Visual Analytics will be started in your web browser tab.

[Manage user privileges](#)

Visual Analytics user interface

Visual Analytics provides you with a clear and easy to use graphical user interface (GUI).

Visual Analytics Explorer

In the Visual Analytics Explorer you are able to explore your data sources, create and save your analyses, and organize your analyses in projects.

The Visual Analytics Explorer is displayed by default when you start Visual Analytics. If required, you can open the Visual Analytics Explorer by clicking the  **Explorer** icon in the program bar.

Visual Analytics Explorer provides several views for organizing and accessing your data sources, analyses and projects.

- The **ALL** view contains all your registered data sources, created analyses, and projects.
- The **DATA SOURCES** view lists all your registered data sources.
- The **ANALYSES** view lists all your saved data analyses.
- The **PROJECTS** view lists all your analyses projects. A project groups a number of analyses under a specific name.

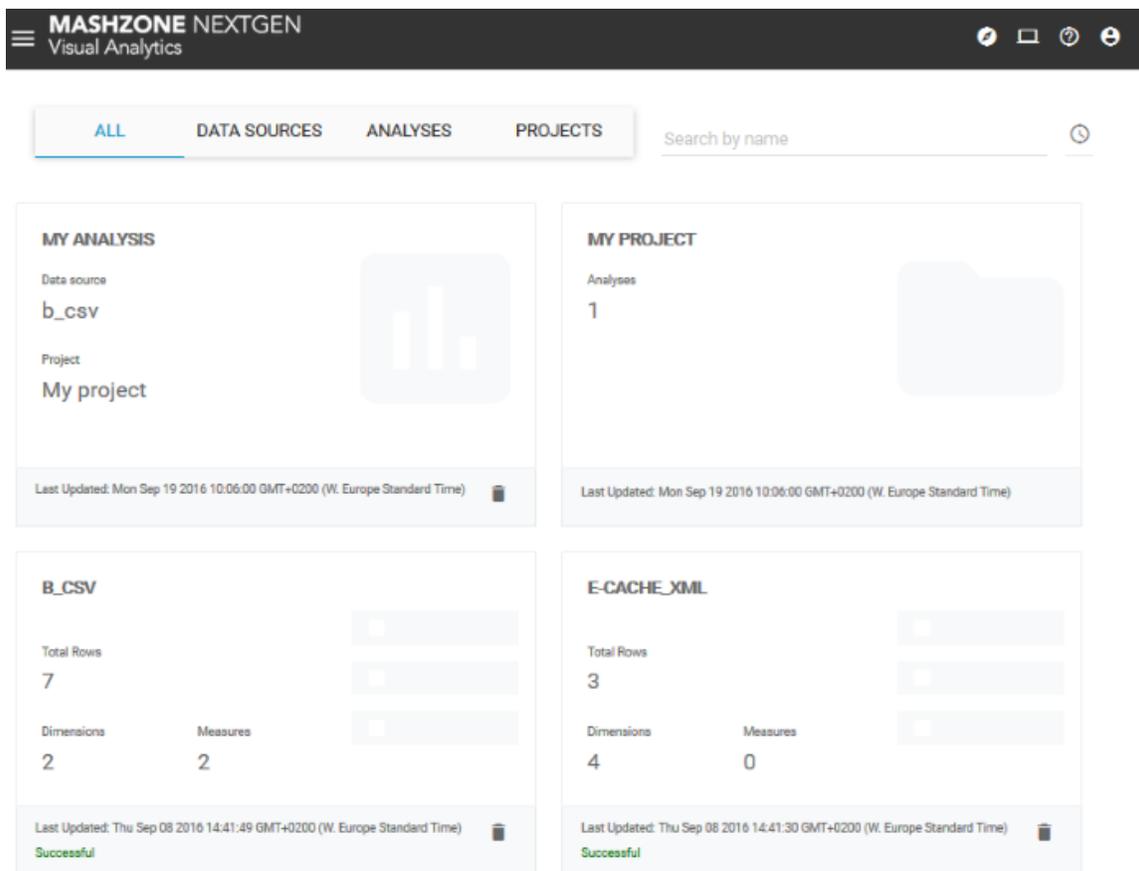
Only the data sources, analyses, and projects are available for which you have the corresponding privileges.

The search box **Search by name** helps you to find quickly your data sources, analyses, or projects desired.

You can choose the **Sort by name** (AZ) or **Sort by last modified** (🕒) options for the elements displayed on the several views. Click the AZ or 🕒 Sort button and select the sort option desired.

Example: Visual Analytics Explorer - All view

The following screenshot shows the **ALL** view of the Visual Analytics Explorer containing **My analysis** (assigned to My project), **My project**, and the two data sources **B-CSV** and **E-CACHE_XML**.



Additionally, Visual Analytics Explorer provides several views with all contents and functions you need for your analysis creation. Using the various configuration bars, you are able to set the criteria (dimensions and measures) you want to analyze and the values you want to display. Furthermore you can select filters to narrow down the query results and configure the analysis visualization.

- The **INSIGHTS** view displays details of the dimensions and measures contained in the data source. It gives you an overview of what is insight the data in the forms of several criteria values, e.g., count, min or max values, and appropriate distribution charts.
- The **VISUALIZATIONS** view enables you to set the visualization of your analysis, e.g., selecting chart types or grouping criteria. Additionally, the view enables you to

view a table with the raw data of the data source corresponding to your analysis configuration.

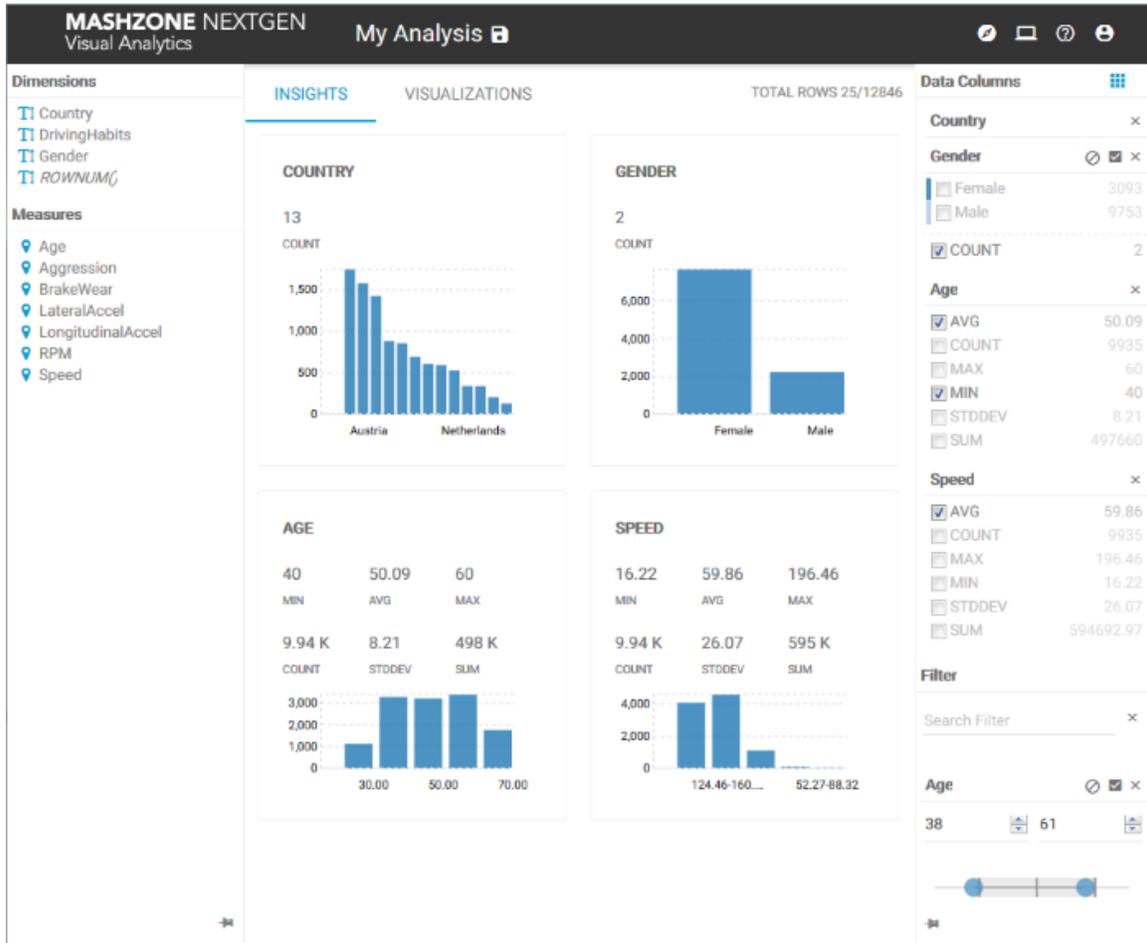
Visual Analytics Explorer provides several configuration panels that supports you to configure your analyses.

- On the left-hand side of a view the **Dimensions** and **Measures** bars are placed. The criteria listed there are provided by the data source selected. Whereby each criterion represents a separate data column of the data source. The criteria names match the column names of the data source. In the **VISUALIZATIONS** view the left configuration panel provides additional visualization options.
- On the right-hand side of a view the criteria selected are listed in the **DATA COLUMNS** bar. Here you can choose the aggregated values of the columns which you want to display in the analysis and you can set additional filters.

In the analysis area in the middle, you are able to configure your analysis, explore your source data, and visualize the results of your queries using charts or tables.

Example: Visual Analytics Explorer - Insights view

The following example shows the **INSIGHTS** view of the analysis explorer displaying two dimensions and two measures selected. The **Age** measure is set as filter criterion.



Visual Analytics Console

Visual Analytics Console provides a technical API to Visual Analytics Server.

You can open the Visual Analytics Console by clicking the  **Console** icon in the program bar.

Example: Visual Analytics Console

The screenshot displays the MashZone NextGen Visual Analytics console interface. On the left is a navigation menu with a 'Registry' section containing various endpoints, with 'searchProjects' selected. The main area is titled 'SearchProjects' and features an 'Invoke' button. It shows a REST client interaction with a 'Payload' and a 'Response'.

Registry

- searchDataSources
- searchByCriteria
- searchProjects**
- loadInitialDataSet
- createDataSource
- getDataSource
- updateDataSource
- deleteDataSource
- searchAnalyses
- createAnalysis
- getAnalysis
- updateAnalysis
- deleteAnalysis
- getLastUpdatedEntities

Request Broker

- Statistics Provider
- License Management Service
- User Service

SearchProjects Invoke

Payload

```
{
  "pattern": "",
  "partial": true
}
```

Response

```
[
  {
    "name": "Contour",
    "analyses": 1,
    "timestamp": 1470236824267
  },
  {
    "name": "Olympics",
    "analyses": 2,
    "timestamp": 1474382582685
  },
  {
    "name": "Plant",
    "analyses": 2,
    "timestamp": 1470234596315
  },
  {
    "name": "Telematics",
    "analyses": 1,
    "timestamp": 1472667292910
  }
]
```

2 Using Visual Analytics

■ Import data source	14
■ Create analysis	14
■ Set visualization	20
■ Share analyses	26
■ Create projects	26
■ Using the configuration panels	27
■ Which aggregated values are available?	28

Import data source

Visual Analytics enables you to discover and analyse historical data and real-time data sources provided in file-based or cache-based data sources.

Before you are able to explore and analyze a data source you have to deploy and register the data source required in Visual Analytics. For importing data sources Visual Analytics supports you with an auto-deploy tool.

For details see:

- ["Manage data sources" on page 29](#)
- ["How does Visual Analytics handle data sources?" on page 37](#)

Create analysis

You can easily configure the content of your analyses by using the configuration bars and analysis area available in the relevant Visual Analytics Explorer views.

See also ["Visual Analytics Explorer" on page 7](#) for details.

A analysis defines a query that is set for only a single data source. You configure a data analysis by selecting a data source and the dimensions and measures that you want to analyze. For the selected criteria you then set the values you want to display in the analysis. Optionally, you can specify filters to narrow down the query result. After configuring your query you are able to analyze interactively your source data using different kinds of visualizations.

To create a data analysis perform the following steps.

Select data source

You can select a data source registered to create your analysis.

Procedure

1. In Visual Analytics Explorer (🔍) click **DATA SOURCES** to open the view containing all registered data sources.
2. Click a data source.

By using the search box you are able to find quickly your data source desired.

The data source selected is opened in the **INSIGHTS** view of Visual Analytics Explorer.

Select analysis criteria

Select the dimensions and measures you want to analyse.

Procedure

1. ["Select data source" on page 14.](#)
2. Drop one or more dimensions into the analysis area by dragging them from the **DIMENSIONS** bar on the left configuration panel. See ["Using the configuration panels" on page 27.](#)

If you select more than one dimension the last dimension selected is used for partitioning the data. Each dimension is visually highlighted by a unique color. See ["Partitioning data" on page 15.](#)

3. Drop one or more measures into the analysis area by dragging them from the **MEASURES** bar on the left configuration panel. See ["Using the configuration panels" on page 27.](#)

The criteria selected are displayed in the analysis area with detail views and they are listed in the **DATA COLUMNS** bar on the right configuration panel.

By moving your mouse pointer over the criteria in the chart further detailed information is displayed.

Partitioning data

You can partition data evaluated in your analysis.

If you select more than one dimension the last dimension selected is used for partitioning the data. Partitioning essentially groups data on the selected dimension and the dimension values are visually highlighted by a unique color. The corresponding values of the partitioning dimension are displayed in the **Data Columns** bar with the colors used for the grouped data (comparing to a chart legend).

Procedure

1. ["Select the analysis criteria" on page 15](#) required.

You need to select at least two dimensions.

2. Select partition values to be displayed in the analysis from the right configuration panel under the **Data Columns** bar. See ["Using the configuration panels" on page 27.](#)

The data is grouped by the partitioning dimension values and is visually highlighted by a unique color for each dimension value.

Change criteria type

You can change a measure to dimension and a dimension of type **INTEGER**, **DOUBLE**, and **LONG** to measure.

Procedure

1. ["Open the left configuration panel." on page 27](#)
2. To change a dimension to measure drag a dimensions from the **DIMENSIONS** bar and drop it into the **MEASURE** bar.
3. To change a measure to dimension drag a measure from the **MEASURE** bar and drop it into the **DIMENSIONS** bar.

The type of the criteria selected is changed.

Select criteria values

Select the dimension and measure values you want to display in your analysis.

See also: ["Which aggregated values are available?" on page 28](#)

Procedure

1. ["Open the right configuration panel." on page 27](#)
2. Click the  **Group, Ungroup** button and select **Group**.
Now the criteria values are selectable.
3. Activate the check boxes of the relevant criteria values in the **DATA COLUMNS** bar, e.g., **MIN** or **MAX**.
4. Click a criteria name to collapse or to expand the value list.

The relevant criteria values are selected and will be displayed in the analysis.

Specify query filters

Optionally, you can define filters for your analysis to narrow down the query result. Each filter describes a specific value range of a dimension or measure.

Procedure

1. ["Open the left and right configuration panels." on page 27](#)
2. Drag one or more dimensions or measures from the criteria bar of the left configuration panel and drop them into the **Filter** bar of the right configuration panel.
3. Set the filter value.

Depending on the dimension or measure type, you can select particular values or expression, or specify periods by which you want to filter.

- a. For **numerical** filter criteria you can use up and down buttons (▲▼) or alternatively a slider to set your filter. To set a data range move the start or the end of the slider at the positions desired.
- b. The values of **text** filter criteria are displayed as a list. In the list, click the values by which you want to filter.

The slider shows the minimum and maximum criterion values indicated by the blue slider ends. And additionally, it shows the standard deviation and average value. See the example below.

To set the minimum or maximum data value to the lower or upper value of the standard deviation click on the corresponding marker on the slider.

- c. For **date** filter criteria you can use a slider to easily set your filter. To set a date range click the start or the end of the slider and place it at the desired positions.

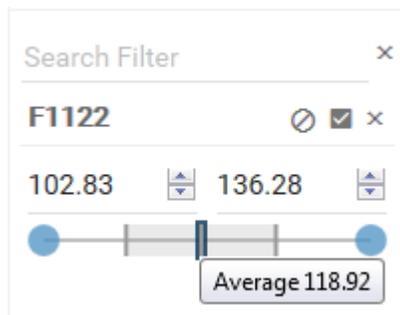
You can set the date scaling of the filter by clicking the 📅 **Change Datetime Scale** icon and selecting the date format (e.g., Year, Month, or Day). The default scaling of the data column is displayed underlined.

4. You can include (default) or exclude the selected filter values in the query result. Click the /☐ **Include/Exclude** icon and select the required option.
5. Click the **Restore** icon to set the default values.
6. Click a criteria name to collapse or to expand the filter settings.

Your filter settings are applied.

Example with numerical filter criterion

The example shows the measure **F1122** used as filter with **106.3** as maximum criterion value and **132** as minimum criterion value. The grey area between shows the standard deviation range with the average value in the middle.



Delete criteria

You can delete criteria selected.

Procedure

1. "Open the right configuration panel." on page 27
2. Click the  **Delete** icon of a criteria.

The corresponding criteria is removed from your analysis.

Refresh criteria data

You can refresh the data displayed of cache based data sources.

If a historical, cache based data source has been updated you can refresh the criteria values displayed in your current analysis, e.g, MIN, MAX or COUNT.

Procedure

1. Move the mouse cursor over the analysis area.
 - The  **Settings** button is displayed.
2. Click the  **Settings** button.
3. Click  **Refresh**.

All source data are updated and the criteria values are refreshed accordingly.

Calculate full result

You are able to calculate the full result of your analysis query.

When you select a dimension you might only retrieve partial result for that dimension if the value exceeds a predefined threshold value. This threshold value is set to 1000 by default. In this case an exclamation icon () is displayed on the detail view of the relevant dimension. You can calculate the full result for this dimension.

The option is not available for measures.

Note: Once you have triggered a calculation of full result for a dimension, you will always get full result for that dimension until the server is restarted.

Procedure

1. Click the  exclamation icon on the detail view of the dimension.
2. Click **Calculate full result**.

The full result is calculated for the dimension selected.

If one or more dimensions have partial results, then you can calculate the full result for all of them at once. For this, move the mouse cursor over the analysis area and click the  **Settings** button on the analysis area. Then click **Calculate full results** in the drop-down menu.

Save analysis

You can save your analysis settings under any name.

Note: The name given has to be unique.

Procedure

1. Click the **Save** icon in the program bar.
2. Click the **Name** input box and enter a name for your analysis.
The preset name corresponds to the data source file of your analysis.
3. You can assign your analysis to a new or an already existing project. Click the **Project** input box and enter a new project name or select a project name available in the pop-up menu.
4. Additionally, you can enter a description for your analysis. Click the **Description** input box and enter a text.
5. Click the **Save** button.

Your settings are saved. Your analysis is available on the **ANALYSIS** tab and assigned to a project if applicable.

To open an analysis assigned to a project, you can also open the **PROJECTS** view in the Visual Analytics Explorer and select the relevant project. Then select the analysis desired.

Edit analysis properties

You can edit the properties of an existing analysis, e.g., name, description, or project assignment.

Procedure

1. ["Open an analysis." on page 20](#)
2. Click the analysis name in the program bar.
3. You can assign your analysis to a new or an already existing project. Click the **Project** input box and enter a new project name or select a project name available in the pop-up menu.
4. Additionally, you can enter a description for your analysis. Click the **Description** input box and enter a text.
5. Click the **Save** button.

Your settings are saved and your analysis is assigned to a project if applicable.

Set visualization

Visual Analytics provides you with various visualization possibilities to analyze interactively your data.

With the **VISUALIZATIONS** view in the analysis editor you have an own view where you can set your data visualization and analyze the correlation between the selected criteria. You are able to change the chart type, to group or ungroup the criteria, to split the analysis view, and to change the criteria axis assignment, etc.

In the **VISUALIZATIONS** view you have additionally the same options to configure your analysis as described in ["Create analysis" on page 14](#).

See the following chapters for details.

- ["Select analysis criteria" on page 15](#)
- ["Select criteria values" on page 16](#)
- ["Delete criteria" on page 17](#)
- ["Save analysis" on page 19](#)
- ["Specify query filters" on page 16](#)

Open an analysis

You can open an existing analysis to edit your analysis settings.

Procedure

1. In Visual Analytics Explorer click **ANALYSIS** to open the view containing all created analyses.

Using the search box you can find quickly your analysis desired.

2. Click **VISUALIZATIONS** to open the corresponding view.

The **VISUALIZATIONS** view is opened displaying a preset visualization according to your analysis settings.

To open an analysis assigned to a project, you can also open the **PROJECTS** view in Visual Analytics Explorer and select the relevant project. Then select the analysis desired.

Select chart type

You are able to change the chart type displayed in your analysis view.

Visual Analytics provides several chart types that are suitable for different analysis requirements.

Procedure

1. Click the **Chart Type** selection field in the **CONFIGURATION** bar.
2. Select a chart type in the drop-down menu.

Depending on the chart type selected the **CONFIGURATION** bar provides additional options to configure the analysis view. See the chart option list below.

The chart type selected is displayed in your analysis. Your chart settings are applied.

Chart options

The following specific chart options are available.

Chart type	Options
Bar chart	<p>Stacked / Grouped</p> <p>Displays the criteria values stacked or grouped.</p> <p>Switch the slider to the option desired.</p>
Line chart	<p>Interpolation</p> <p>Provides a number of interpolation methods. Interpolation is used to provide smooth or angled curves to connect the data points on the chart.</p> <p>Select an option in the Interpolation drop-down menu.</p> <ul style="list-style-type: none"> ■ None: Default interpolation used. ■ Basis: A basis-spline, with control point duplication on the ends. ■ Linear: Piecewise linear segments, as in a polyline. ■ Monotone: Cubic interpolation that preserves monotonicity in y-values. ■ Step: Alternate between horizontal and vertical segments, as in a step function. ■ Step-After: Alternate between horizontal and vertical segments, as in a step function.

Chart type	Options
	<ul style="list-style-type: none"> ■ Step-Before: Alternate between vertical and horizontal segments, as in a step function.
Area chart	<p>Interpolation</p> <p>Provides a number of interpolation methods. Interpolation is used to provide smooth or angled curves to connect the data points on the chart.</p> <p>Select an option in the Interpolation drop-down menu.</p> <ul style="list-style-type: none"> ■ None: Default interpolation used. ■ Basis: A basis-spline, with control point duplication on the ends. ■ Linear: Piecewise linear segments, as in a polyline. ■ Monotone: Cubic interpolation that preserves monotonicity in y-values. ■ Step: Alternate between horizontal and vertical segments, as in a step function. ■ Step-After: Alternate between horizontal and vertical segments, as in a step function. ■ Step-Before: Alternate between vertical and horizontal segments, as in a step function.
Scatter chart	<p>Size by</p> <p>Displays the bubble size depending on the criterion selected in the drop-down menu.</p> <p>Select an option in the Size by drop-down menu.</p> <p>Aggregation</p> <p>Displays the bubble size depending on the criterion value selected.</p>

Chart type	Options
	<p>Select an option in the Aggregation drop-down menu.</p>
Contour chart	<p>Intensity</p> <p>Displays the color intensity depending on the criterion selected in the drop-down menu.</p> <p>Select an option in the Intensity drop-down menu.</p> <p>Aggregation</p> <p>Displays the color intensity depending on the criterion value selected.</p> <p>Select an option in the Aggregation drop-down menu.</p> <p>Radius</p> <p>Controls the size of the circles or points on the chart.</p> <p>Default: 6</p> <p>Select a value in the Radius drop-down menu.</p>

Change axes assignment

You can change the assignment of the criteria to the different axes.

The criteria axes assignment correlates with the criteria list order in the **DATA COLUMNS** bar. If you have selected only two criteria each criterion is assigned to one axis. If you have selected dimensions and measures then the dimensions are assigned to the x-axis and are placed on the top of the criteria list. You can change the axes between two dimensions or two measures but you can not change the axis of a dimension with the axis of a measure.

Procedure

1. Drag a criteria in the **DATA COLUMNS** bar.
2. Drop it on the criteria for that you want to change the axes assignment.

Your changes are applied in the analysis view.

Show trend line

You can show a trend line in your line, bar, area, and scatter chart.

Pre-requisites:

You display a line, bar, area, and scatter chart with one dimension or measure assigned to the x-axis and one measure assigned to the y-axis.

Procedure

1. Click the **Chart Type** selection field in the **CONFIGURATION** bar.
2. Select one of the above-mentioned chart types in the **Chart Type** drop-down menu.
3. Click the **Show trend line** switch.

A trend line is displayed in your chart.

Split the analysis view

You can split the chart displayed in various charts depending on the criteria used in the analysis.

By default, all criteria used in an analysis are displayed in one chart. For example, in a bar chart several measures are displayed stacked or several dimensions are displayed grouped. If you want to display one chart per dimension or per measure you can split the chart in various charts according to the number of criteria used. You have the following options.

- No Split: Displays the analysis criteria in one chart (default).
- Split & Group by column: The number of charts displayed depends on the number of criteria (columns) selected.
- Split & Group by aggregate: The number of charts displayed depends on the number of criteria values selected.
- Split in groups of 3: Displays always three criteria per chart. Additionally, if you have four or more measures, you get multiple charts with three and subsequent measures used as radius for scatter or contour charts.
- Split: Displays an own chart for each measure selected.

Procedure

Select the split option in the **CONFIGURATION** bar.

Your changes are applied.

Set data limit

For reasons of performance you can set the maximum number of rows analysed.

The numbers of the rows analysed and the total number of rows available in the data source is shown above the chart.

Procedure

1. Click the **Limit** input box in the **Configuration** bar.
2. Enter a value for the maximum row number.

The default value is 3000.

Your settings are applied.

Display data table

The **VISUALIZATIONS** view enables you to take a look at the data driving your analysis.

In addition to your chart, you can choose to display a table with the data of your analysis, including dynamically calculated aggregations, filtered by current criteria values of your analysis configuration.

Procedure

1. To display a data table click the  **Show/Hide Data Table** icon in the program bar and click **Show**.
2. To hide a data table displayed click the  **Show/Hide Data Table** icon in the program bar and click **Hide**.

The data table of your analysis is displayed.

Save analysis

You can save your analysis settings under any name.

Note: The name given has to be unique.

Procedure

1. Click the  **Properties** icon in the program bar.
2. Click the **Name** input box and enter a name for your analysis.

The preset name corresponds to the data source file of your analysis.

3. You can assign your analysis to a new or an already existing project. Click the **Project** input box and enter a new project name or select a project name available in the pop-up menu.

4. Additionally, you can enter a description for your analysis. Click the **Description** input box and enter a text.
5. Click the **Save** button.

Your settings are saved. Your analysis is available on the **ANALYSIS** tab and assigned to a project if applicable.

To open an analysis assigned to a project, you can also open the **PROJECTS** view in the Visual Analytics Explorer and select the relevant project. Then select the analysis desired.

Freeze real-time data visualization

You can freeze the visualization of real-time data.

When you analyse real-time data, new source data will be constantly reloaded and the visualization is updated continuously. To freeze the changing visualization you can pause and restart the data loading process.

Procedure

1. Open the **VISUALIZATIONS** view of an analysis based on a real-time data source.
2. Click the  **Toggle Live Pause** button to freeze the visualization.
3. Click the  **Toggle Live Pause** button to restart the visualization.

The data loading process is stopped and the corresponding visualization is frozen.

Share analyses

You can share your analyses with any user.

Procedure

1. ["Open an analysis." on page 20](#)
2. Click the analysis name in the program bar.
3. Click  **Copy link** in the properties bar.

The URL of your analysis is copied to the clipboard.

For example, you can past the link copied into an email and share it with any user.

Create projects

You are able to group a number of analyses as own project.

You can assign your analyses to a new or an already existing project. All your projects are listed on the **PROJECTS** view of the Visual Analytics Explorer.

Procedure

1. "Open an analysis" on page 20.
2. Click the  **Properties** icon in the program bar.
3. Click the **Project** input box and enter a new project name or select a project name available in the pop-up menu.
4. Click the **Save** button.

Your settings are saved and the analysis selected is assigned to a project. Where appropriated a new project is created on the **PROJECTS** view.

To open an analysis assigned to a project, you can open the **PROJECTS** view in the Visual Analytics Explorer and select the relevant project. Then select the analysis desired.

To remove an analysis from a project, open a project and click the  **Delete** icon of an analysis.

Using the configuration panels

By means of the configuration panels you are able to configure easily your analyses.

The configuration panels are available in the **INSIGHTS** and **VISUALIZATIONS** views of Visual Analytics Explorer. If you open one of these views using a lower screen resolution or a smaller screen, either only the right panel or both panels are hidden automatically.

Procedure

1. To display or to hide the left configuration panel click the  **Toggle left panel** icon.
You can only hide an unfixed panel.
2. To display or to hide the right configuration panel click the  **Toggle right panel** icon.
You can only hide an unfixed panel.
3. To fix or to unfix the configuration panel click the  **Pin** or  **Unpin** icon.

Your setting are applied.

- If you drag a criteria from the left configuration panel the right panel is automatically displayed.
- Alternatively, you can hide the panels displayed by clicking in the analysis area or by using **Escape** on the keyboard. You can only hide unfixed panels.

Visual Analytics user interface

Which aggregated values are available?

You can select the dimensions and measures values you want to display in your analysis.

See "[Select criteria values](#)" on page 16 for details.

Depending on the criteria data type different aggregated values are available for each criteria.

- Average value (**AVG**)
- Count (**COUNT**)
- Maximum value (**MAX**)
- Minimum value (**MIN**)
- Standard deviation (**STDDEV**)
- Sum (**SUM**)

3 Manage data sources

■ Manage historical data sources	30
■ Manage real-time data sources	34
■ Define date and time formatting pattern	36
■ How does Visual Analytics handle data sources?	37

Visual Analytics enables you to discover and analyse historical data and real-time data streams provided in file-based or cache-based data sources.

Manage historical data sources

You can deploy and register historical datasets, repositories of time-stamped data, or BigMemory data sources. Visual Analytics supports different file-based data source formats incl. CSV, XML and JSON files, or Terracotta BigMemory caches.

- ["Manage file-based data sources" on page 30](#)
- ["Manage cache-based data sources" on page 32](#)

Manage real-time data sources

You can also deploy and register real-time data sources for your data analyses in Visual Analytics. Visual Analytics consumes streaming data provided by Apama and stored in separate data stores (caches). The caches are managed by so-called cache managers and configured in a cache configuration file. Visual Analytics is able to read the cache configuration file and use its contents to access the data in the external Apama cache.

- ["Manage real-time data sources" on page 34](#)

Prerequisite: You have administrator privileges.

See also: ["How does Visual Analytics handle data sources?" on page 37](#)

Manage historical data sources

You can deploy and register historical data sets, repositories of time-stamped data, or BigMemory data sources. Visual Analytics supports different file-based data source formats incl. CSV, XML and JSON files, or Terracotta BigMemory caches.

The following chapters describe how to create, update, and delete file-based and cache-based data sources.

Manage file-based data sources

You can create, update, and delete file-based data sources for your exploration and analyses of historical source data.

Prerequisite: You have administrator privileges.

See also: How does Visual Analytics manage ["Historical file-based data sources" on page 37?](#)

For deploying your historical file-based data sources Visual Analytics provides a specific auto-deploy folder.

<Visual Analytics installation> /data/auto-deploy/file

Create file-based data source

You can create file-based data sources for your data explorations. Visual Analytics supports source data of CSV, XML and JSON data types.

Note: Visual Analytics only accepts JSON source files with self-contained JSON documents, i.e., source files that comply with JSON Lines format. See <http://jsonlines.org/> for a detailed documentation.

Note: Slow copy processes on MS Windows systems (e.g., caused by copying large files or slow server connection) can lead to unexpected behavior while deploying the data source file. Therefore, we recommend that you first copy the source file in a local folder and then move it into the relevant auto-deploy folder.

Procedure

1. Copy your data source file in the auto-deploy/file folder.
2. Check the .vsd data source description file in the auto-deploy folder.

You can open the .vsd file with a default text editor.

Your data source is deployed and registered in Visual Analytics. A corresponding `<data source name>`.vsd data source description file is written into the auto-deploy folder.

Delete existing file-based data source

You can delete an existing data source.

Procedure

1. Identify the corresponding .vsd data source description file in the auto-deploy/file folder.

If you do not delete the data source file as well, the description file will be created again with the next Visual Analytics server restart.

2. Delete the file manually.

The Visual Analytics server detects automatically the deletion of the data source description file. The corresponding data source is deregistered and the cache containing the data is removed from the cache manager.

Update data of existing file-based data source

You can replace the existing data source by a new data source containing unchanged meta data.

Procedure

- Replace the existing data file in the auto-deploy/file folder with the new one.

Both files must have the same name.

The Visual Analytics server detects that the data file has been updated. The new file will be parsed and data in the cache will be replaced by new data.

Update meta data of existing file-based data source

You can update the meta data of an existing data source. The meta data information of a data source covers inter alia the data source name, cache name, schema of imported data, result of the last import run, etc. For example, it is possible that the data type of a column is not guessed correctly by the auto-deployer, because some columns contain integer values and others double values or similar.

The meta data are stored in the `<data source name>.vsd` data source description file.

Procedure

1. Edit the appropriate `.vsd` file in the `auto-deploy/file` folder.

You can edit the `.vsd` file with a default text editor.

2. Save your changes.

Saving the file triggers the Visual Analytics server to update the data source using meta data from the edited `.vsd` file.

The meta data and the corresponding data source are updated.

Manage cache-based data sources

You can create, update and delete cache-based data sources for your exploration and analyses of historical source data.

Prerequisite: You have administrator privileges.

See also: How does Visual Analytics handle ["Historical cache-based data sources" on page 37?](#)

For deploying your historical cache-based data sources Visual Analytics provides a specific `auto-deploy` folder.

`<Visual Analytics installation>/data/auto-deploy/cache`

Create cache-based data source

You can create cache-based data sources for your data explorations. Visual Analytics supports BigMemory caches as source data.

Procedure

1. Copy your cache configuration file in the `auto-deploy/cache` folder.

A corresponding `<data source name>.vsd` data source description file will be written into the `auto-deploy` folder.

2. Check the .vsd data source description file in the auto-deploy folder.

You can open the .vsd file with a default text editor.

Your data source is deployed and registered in Visual Analytics. A corresponding *<data source name>*.vsd data source description file is written into the auto-deploy folder.

Delete existing cache-based data source

You have administrator privileges.

You can delete an existing cache-based data source.

Procedure

1. Identify the corresponding .vsd data source description file in the auto-deploy/cache folder.

If you do not delete the cache configuration file as well, the description file will be created again with the next Visual Analytics server restart.

2. Delete the file manually

The Visual Analytics server detects automatically the deletion of the data source description file and the corresponding data source is deregistered in Visual Analytics.

Note: The external cache containing the data will not be deleted.

Update cache-base data source

You have administrator privileges.

You can update the cache data source by redeploying your cache configuration file.

Visual Analytics server will detect that the cache configuration file has been updated. The new file will be parsed and data in the cache will be replaced by new data.

Note: Please note that changes made in the configuration file are lost in this case.

Procedure

1. Remove the relevant cache configuration file and the corresponding *<data source name>*.vsd data source description file from the auto-deploy/cache folder.
2. Copy your new cache configuration file in the auto-deploy/cache folder.

A corresponding *<data source name>*.vsd data source description file will be written into the auto-deploy folder.

3. Check the .vsd data source description file in the auto-deploy folder.

You can open the .vsd file with a default text editor.

The data source is updated.

Manage real-time data sources

With Visual Analytics you are able to discover and analyse real-time data sources. Visual Analytics uses Apama data stores (real-time caches) for processing analytic requests on streaming data.

Prerequisite: You have administrator privileges.

See also: How does Visual Analytics handle ["Real-time data sources" on page 38?](#)

The following chapters describe how to create, update, and delete real-time data sources.

Create real-time data source

You have administrator privileges.

You can create real-time data sources for your real-time data explorations.

Procedure

1. Copy your Apama cache configuration file in the auto-deploy/stream folder.
A corresponding *<data source name>* .vsd data source description file is written into the auto-deploy folder.
2. Check the .vsd data source description file in the auto-deploy folder.
You can open the .vsd file with a default text editor.

Your data source is deployed and registered in Visual Analytics.

Delete existing real-time data source

You have administrator privileges.

You can delete an existing real-time data source.

Procedure

- Identify the corresponding .vsd data source description file in the auto-deploy/stream folder and delete it manually.
If you do not delete the Apama cache configuration file as well, the description file will be created again with the next Visual Analytics server restart.

The Visual Analytics server detects automatically the deletion of the data source description file and the corresponding data source is deregistered in Visual Analytics.

Note: The external cache containing the data will not be deleted.

Update real-time data source

You have administrator privileges.

You can update the real-time data source.

Visual Analytics server will detect that the data file has been updated. The new file will be parsed and data in the cache will be replaced by new data.

Note: Please note that changes made in the configuration file are lost in this case.

Procedure

1. Remove the relevant cache configuration file and the corresponding *<data source name>.vsd* data source description file from the auto-deploy/stream folder.
2. Copy your new cache configuration file in the auto-deploy/stream folder.
A corresponding *<data source name>.vsd* data source description file will be written into the auto-deploy folder.
3. Check the *.vsd* data source description file in the auto-deploy folder.

You can open the *.vsd* file with a default text editor.

The data source is updated.

Real-time data source registration

Visual Analytics uses Apama data stores (real-time caches) for processing analytic requests on real-time data.

With means of a cache configuration file provided by Apama Visual Analytics generates a configuration file and the configured cache manager is instantiated. For each real-time cache a new data source is registered in Visual Analytics server.

See also "[Manage real-time data sources](#)" on page 34 and "[How does Visual Analytics handle data sources?](#)" on page 37 for more details.

Visual Analytics uses BigMemory's cache event listener framework for receiving notifications on cache changes. The Ehcache configuration file specifying the real-time cache must meet the following requirements:

1. It must provide a default cache configuration.
2. It must contain a configuration for the cache **__apama_internal_table_schemas** holding Apama's internal schema information. The configuration should prevent eviction of schema information (`<pinning store="incache" />`) and must define a distributed cache (`<terracotta />`). This will enable Visual Analytics to access and interpret the schema information provided by Apama.
3. It must contain a configuration for each real-time cache hosted by the cache manager that should be analyzed using Visual Analytics. The configuration

must enable the replication of cache events (`<cacheEventListenerFactory class="net.sf.ehcache.event.TerracottaCacheEventReplicationFactory" />`) and must define a distributed cache (`<terracotta />`). This will enable Visual Analytics to be notified on cache changes.

Since these requirements cannot be fulfilled completely using the Apama MemoryStore API configuration, the according Spring file has to be modified in a way that an existing Ehcache configuration file is used (`<property name="ehcacheConfigFile" value="PATH_OF_EHCACHE_CONFIG_FILE" />`).

Define date and time formatting pattern

You can specify your own formatting pattern for data source values of **Date** and **Time** type.

Prerequisite: You have administrator privileges.

By default, Visual Analytics supports several standard date and time patterns for auto-deployed data sources. For non-matching date and time patterns you can specify your custom formatting patterns.

If the data source file contains columns with values of **Date** or **Time** type in the standard format, these columns are detected as **DATE** type columns. The corresponding .vsd file contains an additional entry **datePattern**, specifying the detected date and time pattern.

If the values contained do not match the standard formatting patterns, the corresponding columns are detected as **STRING** type. To specify your own formatting pattern edit manually the relevant .vsd file.

Procedure

1. Deploy your data source file in the relevant auto-deploy folder.

See "[Manage data sources](#)" on page 29 for details.

2. Edit the appropriate `<data source name>` .vsd data source description file.

You can edit the .vsd file with a default text editor.

3. Change the data type of all relevant columns from **STRING** to **DATE** type within the **schema** attribute.
4. Specify your date and time format by adding the **datePattern** attribute.

The syntax for specifying custom date and time patterns is described here: <https://docs.oracle.com/javase/8/docs/api/java/text/SimpleDateFormat.html>.

5. Save your changes.

The data source is redeployed automatically and all columns that were specified as **DATE** type columns are parsed using the custom **datePattern** attribute.

How does Visual Analytics handle data sources?

Visual Analytics enables you to discover and analyse historical data and real-time data sources provided in file-based or cache-based data sources.

Visual Analytics detects automatically a new data source file, analyses the content, and recognize the data source structure. Based on the data structure Visual Analytics determines the data source configuration.

- Is it a static (historical) or a dynamic (real-time) data source?
- What are the dimensions and measures?
- What are the data types?
- etc.

The data source meta data determined are written in a separate description file. By editing the description file you are able to reconfigure the data source configuration, e.g., changing dimensions names or measures data types.

Historical file-based data sources

Visual Analytics supports different file-based data sources for your historical data analysis. These include source data of CSV, XML, and JSON data types. For deploying your file-based data sources Visual Analytics provides a specific auto-deploy folder.

<Visual Analytics installation> /data/auto-deploy/file

New and updated data sources are automatically deployed and configured for your data exploration by the Visual Analytics auto-deployer. As soon as a source file is dropped into the Visual Analytics auto-deploy folder, the file is parsed, repeating elements are identified and the schema of these elements is determined. The data is stored in a Visual Analytics internal data cache and the new data source is registered in the Visual Analytics registry to make it accessible for other Visual Analytics components.

After the automatic parsing process a corresponding *<data source name>* .vsd data source description file is written into the auto-deploy/file folder. This file contains all meta data information of the data source, incl. name of the data source, cache names, dimensions, measures, schema of imported data, result of the last import run, etc..

See also: "[Manage file-based data sources](#)" on page 30

Historical cache-based data sources

Visual Analytics supports Terracotta BigMemory caches as data sources for your historical data analyses. For deploying your cache-based data sources Visual Analytics provides a specific auto-deploy folder.

<Visual Analytics installation> /data/auto-deploy/cache

New and updated data sources are automatically deployed and configured for your data exploration by the Visual Analytics auto-deployer. As soon as a cache configuration file is dropped into the Visual Analytics auto-deploy folder, the configuration file is parsed, the configured cache manager is instantiated, and for each data cache a new data source is registered in Visual Analytics server. The new data source is registered in the Visual Analytics registry to make it accessible for other Visual Analytics components.

Please contact the system administrator in charge for further information on cache administration for your Terracotta BigMemory connection.

After the automatic parsing process a corresponding *<data source name>* .vsd data source description file is written into the auto-deploy/cache folder. This file contains all meta data information of the data source, incl. name of the data source, schema of data contained in the external cache and information how to connect to the cache (Cache manager name, cache name), dimensions, measures, etc.

The *<data source name>* is composed by the name of the cache manager and the name of the cache. E.g., a data source pointing to a cache "StockData2015" managed by a cache manager with name "StockData" would be registered with the name "StockData-StockData2015".

See also: ["Manage cache-based data sources" on page 32](#)

Real-time data sources

With Visual Analytics you are able to discover and analyse real-time data sources. Visual Analytics uses Apama data stores (real-time caches) for processing analytic requests on streaming data.

For deploying your real-time data sources Visual Analytics provides a specific auto-deploy folder.

<Visual Analytics installation> /data/auto-deploy/stream

Please contact the system administrator in charge for further information on cache administration for your Apama data store connection.

New and updated data sources are automatically deployed and configured for your data exploration by the Visual Analytics auto-deployer. As soon as a cache configuration file is dropped into the corresponding Visual Analytics auto-deploy folder, the configuration file is parsed, the configured cache manager is instantiated and for each real-time cache a new data source is registered in Visual Analytics server. The new data source is registered in the Visual Analytics registry to make it accessible for other Visual Analytics components.

After the automatic parsing process a corresponding *<data source name>* .vsd data source description file is written into the auto-deploy/stream folder. This file contains all meta data information of the data source, incl. name of the data source, schema of data

contained in the external cache and information how to connect to the cache (Cache manager name, cache name), dimensions, measures, etc.

The *<data source name>* is composed by the name of the cache manager and the name of the cache. E.g., a data source pointing to a cache "StockData2015" managed by a cache manager with name "StockData" would be registered with the name "StockData-StockData2015".

See also: ["Manage real-time data sources" on page 34](#)

4 Visual Analytics Administration

- Manage user privileges 42
- Create user password 44
- Set SSL configuration 45
- Apache Spark Integration 47

Manage user privileges

Visual Analytics enables you to create new users, change their passwords and roles, and delete existing users.

Depending on the user role assigned the user has different access permissions for data sources and analysis.

By default, the following user roles with the corresponding permissions are available.

- administrator:
 - Full control of all data sources and analysis.
- analyst:
 - Full Create, Read, Update and Delete permissions for data sources and analysis owned by the user or if users has an **EDIT** permission.
 - Read only control on data sources and analysis if **VIEW** permission is present.
- guest:
 - Read only permissions for analysis if **VIEW** permission is present.

The Visual Analytics users available and roles assigned are listed in the auth.properties file stored in the <Visual Analytics installation> \config folder.

The users are listed with username, password and user role in the form of:

```
user.<username> = <password encrypted>,<user role>,va
```

Example

```
user.Administrator = $shiro1$SHA-256$500000$BK02RfNycT5X9+SYFOw==$WVnmao8JwY5+KlOfodA=,administrator,va
```

Create user

You can create new Visual Analytics users and assign the relevant permissions.

To create a new Visual Analytics user you have to specify the user name, user password, and the user role. Using the Visual Analytics Console you are able to create the required user password. See "[Create user password](#)" on page 44 for details.

Procedure

1. Open the auth.properties file in an appropriate text editor.
2. Add a user to the user list in the following form.

```
user.<username> = <password encrypted>,<user role>,va
```

- Replace the <username> variable by a unique user name.

- Replace the <password encrypted> variable by the password generated (without quotes).
 - Replace the <user role> variable by the user role with the access permissions required.
3. Save your changes.
 4. Restart the Visual Analytics server.

The user is created and can log in to Visual Analytics using the appropriate user name and user password.

See also ["Starting Visual Analytics" on page 6](#).

Delete user

You can delete an already existing Visual Analytics user.

Procedure

1. Open the auth.properties file in an appropriate text editor.
2. In the user list, delete the row with the relevant user name.
3. Save your changes.
4. Restart the Visual Analytics server.

The user is deleted and can not longer log in to Visual Analytics.

Change user password

You can change the password of an existing Visual Analytics user.

Using the Visual Analytics Console you are able to create the new user password. See ["Create user password" on page 44](#) for details.

Procedure

1. Open the auth.properties file in an appropriate text editor.
2. In the user list, replace the existing password of an user with the new one.

The password is encrypted, starts with \$, and ends with =.

3. Save your changes.
4. Restart the Visual Analytics server.

The password is saved and the user can log in to Visual Analytics using the appropriate user name and user password.

See also ["Starting Visual Analytics" on page 6](#).

Change user role

You can change the user role assigned to an existing Visual Analytics user.

Procedure

1. Open the `auth.properties` file in an appropriate text editor.
2. In the user list, replace the user role assigned to an user.

You can assign the **administrator**, **analyst** and **guest** user role. See ["Manage user privileges" on page 42](#).

3. Save your changes.
4. Restart the Visual Analytics server.

Your changes are applied and the user can log in to Visual Analytics with the new user role assigned.

Create user password

Using the Visual Analytics Console you are able to create an encrypted password.

You can use the password randomly generated for the Visual Analytics user authentication.

Procedure

1. Click the  **Console** icon in the program bar.
2. Click **User Service** in the **Command** bar.

The subordinated commands are expanded.

3. Click the **hashPassword** command.
4. Enter your password into the **Payload** field.
5. Click the **Invoke** button.
6. Copy the password without the double quotes to clipboard.

The encrypted password is generated and displayed in the **Response** field. The password starts with **\$** and ends with **=** and is set in double quotes (**"**).

Example

```
"$shiro1$SHA-256$500000$BK02RfNycT5X9+SYFOw==$WVnmao8JwY5+KlOfodA="
```

You can copy the password (without the double quotes) to clipboard and use it for, e.g, ["creating a new Visual Analytics user" on page 42](#).

Set SSL configuration

Visual Analytics enables you to setup a secure communication connection between Visual Analytics Client and Visual Analytics Server.

In most cases, Visual Analytics Server will not run on the same machine as Visual Analytics Client. To secure the data transfer between client and server, the communication between the client and server should be encrypted.

The default setup after installation of Visual Analytics will run the Visual Analytics Server without support for encrypted communication (HTTPS / TLS / SSL), i.e., with Transport Layer Security (TLS) disabled. Thus, you cannot connect via https to the server.

Enabling TLS

Visual Analytics allows to enable TLS via settings in the config.json configuration file.

The config.json configuration file is located in the following folder:

<Visual Analytics installation> /config.

Procedure

1. Open the config.json configuration file in a text editor.
2. Enter the following code line and save your changes.

```
"httpServer.tls.enabled" : true,
```

TLS is enabled.

Note: Visual Analytics ships with a self-signed certificate. Thus, when you enable TLS without replacing the key material by your own material "[Setup of certificates and key material](#)" on page 45, Web browsers will issue a security warning when connecting to Visual Analytics Server.

Setup of certificates and key material

Visual Analytics supports three different kinds of key material / certificates:

- JKS (Java KeyStore)
- PEM (Privacy Enhanced Mail)
- PKCS#12 (from RSA, Inc., based on Microsoft's PFX)

Enabling TLS with JKS-based key material

Visual Analytics allows to use public/private key material stored in a Java Key-Store (JKS) to be used as certificate for TLS communication.

Procedure

1. Open the config.json configuration file in a text editor.
2. Enter the following code lines and save your changes.

```
"httpServer.tls.certType" : "JKS",
"httpServer.tls.keyPath"   : "<key path>",
"httpServer.tls.keyPassword" : "<key password>",
```

- The **keyPath** parameter must point to the key file (usually ends with **.jks**). The path may either be relative (to Visual Analytics' root folder) or absolute.
- The **keyPassword** parameter is the password required to open the key store.

TLS with JKS is enabled.

Enabling TLS with PEM-based key material

Procedure

1. Open the config.json configuration file in a text editor.
2. Enter the following code lines and save your changes.

```
"httpServer.tls.certType" : "PEM",
"httpServer.tls.keyPath"   : "<key path>",
"httpServer.tls.certPath"  : "<key password>",
```

- The **keyPath** parameter must point to the key file (usually ends with **.pem**). The path may either be relative (to Visual Analytics' root folder) or absolute.
- The **certPath** parameter must point to the public key file (usually ends with **.pem**). The path may either be relative (to Visual Analytics' root folder) or absolute.

TLS with PEM is enabled.

Enabling TLS with PKCS#2-based key material

Procedure

1. Open the config.json configuration file in a text editor.
2. Enter the following code lines and save your changes.

```
"httpServer.tls.certType" : "PKCS12",
"httpServer.tls.keyPath"   : "<key path>",
"httpServer.tls.keyPassword" : "<key password>",
```

- The **keyPath** parameter must point to the key file (usually ends with **.p12**). The path may either be relative (to Visual Analytics' root folder) or absolute.
- The **keyPassword** parameter is the password required to open the **PKCS#12** file.

Apache Spark Integration

You can use Apache Spark to process large CSV and JSON source files in Visual Analytics.

Apache Spark is a fast and general-purpose cluster computing system. Visual Analytics version 10.0 ships with Apache Spark and integrates it to process large CSV and JSON source files and execute queries on those data sources. XML source files are currently not supported. When dealing with bigger source files, the query execution time is lower compared to Visual Analytics default engine. Thus, the overall performance of Visual Analytics will benefit in those cases. As a guidance, data sets with 500,000 lines and above should be processed using the now-integrated Spark engine.

By default, the Visual Analytics installation runs the Visual Analytics Server without Apache Spark integration. To enable Apache Spark you have to perform the following steps.

Procedure

1. Open the config.json configuration file in a text editor.
2. Add the following code line and save your changes.

```
"spark.enabled" : true,
```

3. To run Visual Analytics with Spark open a command line program and run `<Visual Analytics installation>\bin\vaserver --embeddedSpark`

The default installation path is `C:\SoftwareAG\MashzoneNG\VisualAnalytics`.

Visual Analytics runs with Apache Spark.

Advanced Configuration

By default, Apache Spark engine utilizes up to four CPU cores and 1 GB of RAM per core. If more resources are available on the machine running Visual Analytics, you can configure the maximal number of CPU cores and RAM to be used by Spark.

Procedure

1. Open the config.json configuration file in a text editor.
2. Add the following code line and save your changes.

```
"spark.cores.max" : <number of cores>,  
"spark.executor.memory" : "<amount of memory>",
```

Your settings are applied.

Thereby, `<number of cores>` is a numeric value, such as 3, 7, 15. It is recommended to not assign all cores of the machine to Spark, i.e., when `N` is the number of cores of the machine, then use `N-1` cores at max for Spark.

`<amount of memory>` is a numeric value followed by the unit, e.g., 1g for 1 GiByte, 512m for 512 MiB etc.

Note: The amount of memory is defined per CPU core, i.e., 1g with 16 cores equals an allowed total memory consumption of 16 GiB of RAM.

Allowed units are

```
1b (bytes)
1k or 1kb (kibibytes = 1024 bytes)
1m or 1mb (mebibytes = 1024 kibibytes)
1g or 1gb (gibibytes = 1024 mebibytes)
1t or 1tb (tebibytes = 1024 gibibytes)
1p or 1pb (pebibytes = 1024 tebibytes)
```

Deployment of Spark on a separate machine

Visual Analytics integrates Apache Spark via Spark JobServer. This provides the benefit that Visual Analytics and Spark do not necessarily have to be deployed on the same machine, i.e., you can install Visual Analytics on server A and Spark JobServer on server B (whereby B could be a more capable machine than A).

Procedure

1. Copy the folder <Visual Analytics installation>/va-sjs to the target machine where you want to run Spark JobServer.
2. Start Spark JobServer on the target machine via
 running va-sjs\bin\va-sjs.bat (Windows)
 running va-sjs/bin/va-sjs (Linux / Unix-derivates)
3. On the Visual Analytics machine, open the config.json configuration file in a text editor.
4. Add the following code lines and save your changes.

```
"spark.enabled"      : true,
"spark.jobserver.url" : "https://<target machine IP>:8090",
```

Whereby <target machine IP> is the IPv4-address of the machine running Spark JobServer, e.g., "spark.jobserver.url" : "https://172.16.0.1:8090", if the server has the IP 172.16.0.1.

Your settings are applied.

Importing larger data sources

To deploy data files larger than 1GB, if Spark is deployed on a separate machine, you need to increase the relevant parameters.

Edit the following parameters in the Spark JobServer config file va.conf.

The file is located in the <Visual Analytics installation>/va-sjs/config/ directory.

- Set request-chunk-aggregation-limit = 2000m in section spray.can/server.
- Set max-chunk-size = 2000m in section spray.can/server/parsing.

- Set `short-timeout = 10 s` in section `spark/jobserver`.

Edit the following parameters in the Spark JobServer startup scripts `va-sjs.bat` for MS Windows and `va-sjs` for Linux.

The files are located in the `<Visual Analytics installation>/va-sjs/bin/` directory.

- In `va-sjs.bat`: Change line `set DEFAULT_JVM_OPTS=""` to `set DEFAULT_JVM_OPTS="-Xmx8192m"`.
- In `va-sjs`: Change line `DEFAULT_JVM_OPTS=""` to `DEFAULT_JVM_OPTS="-Xmx8192m"`.