Chapter 3: Analyzing Data Using SAS® Visual Analytics

Objectives

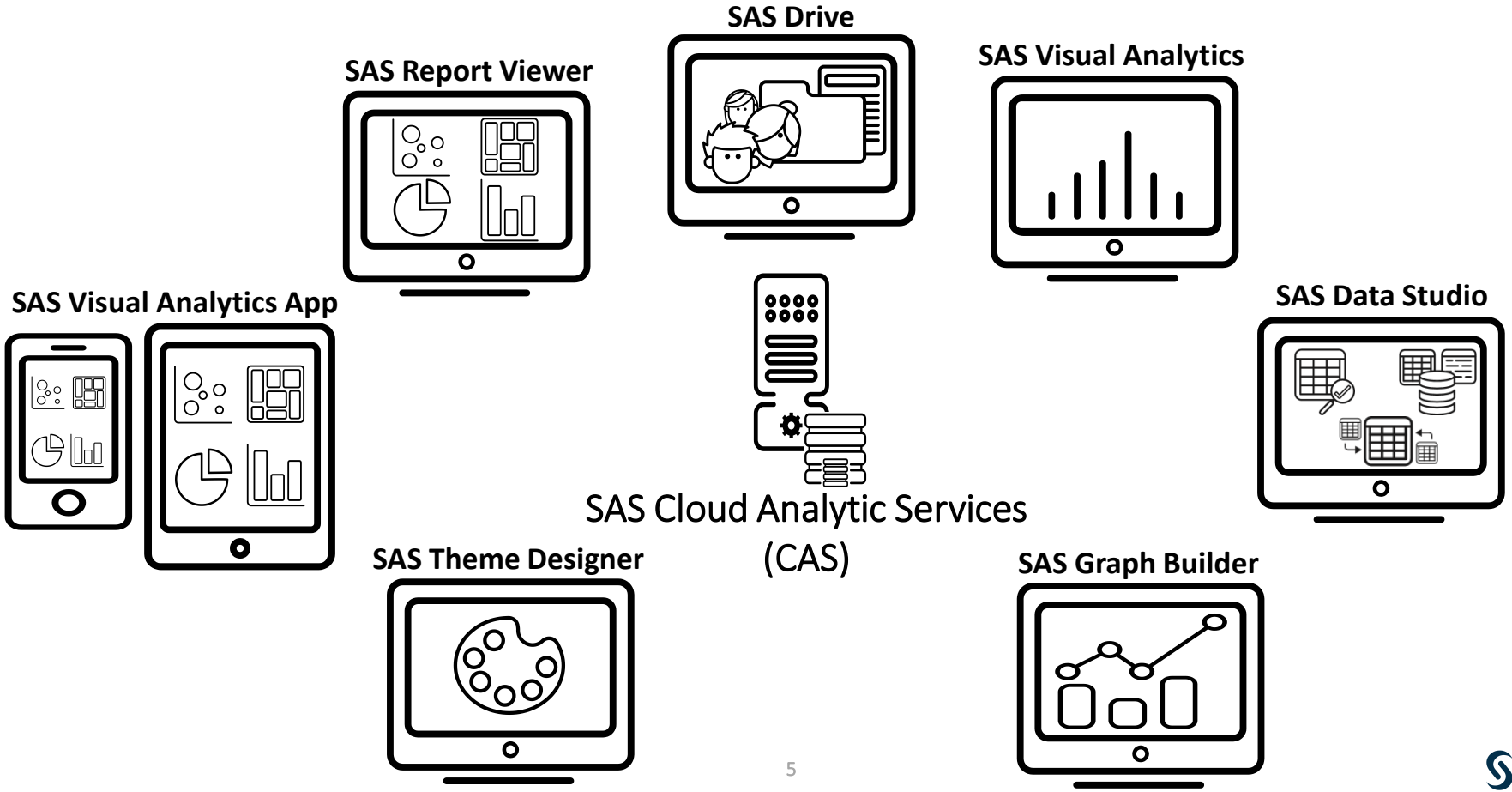
- Discuss the Analyze phase of the SAS Visual Analytics methodology.
- Change data items (modify formats, modify aggregations, modify classifications, rename data items) in Visual Analytics for the analysis.

Visual Analytics Methodology: Analyze

In the *Analyze* phase, you can evaluate the data by doing the following:

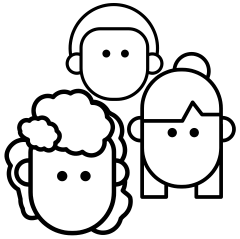
- modifying data item properties
- creating new calculated items needed for analysis
- applying any necessary filters for the analysis
- exploring relationships between data items using charts and graphs
- discovering trends and patterns between data items
- creating, testing, and comparing models based on patterns discovered*

SAS Viya Applications



Business Scenario: Customers

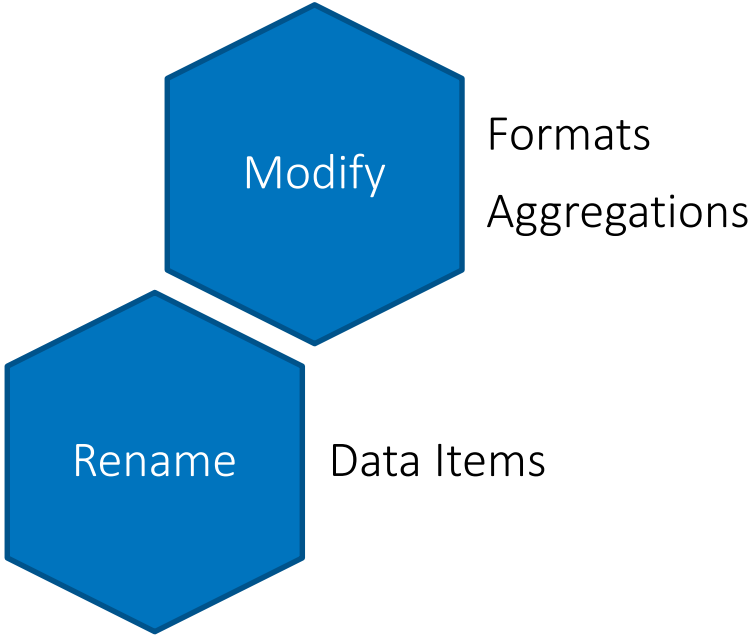
Based on the investigation of the data and the assignment (analyze profits for the Marketing team and analyze delivery times for the Shipping team), you need to make some changes to data items in the **CUSTOMERS** table.



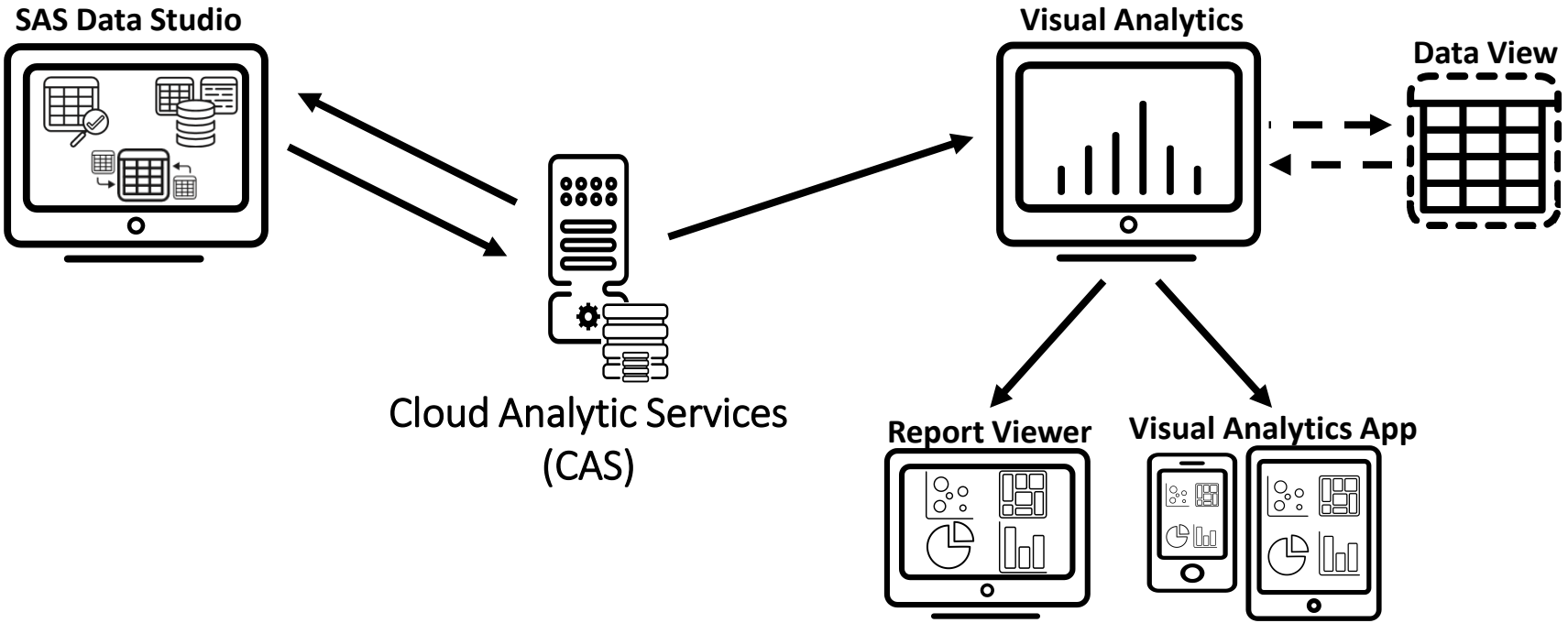
68,300 customers

747,953 orders

You can make more changes as you perform the analysis.



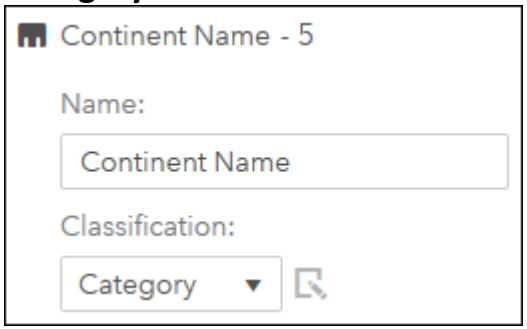
SAS Data Studio versus Visual Analytics



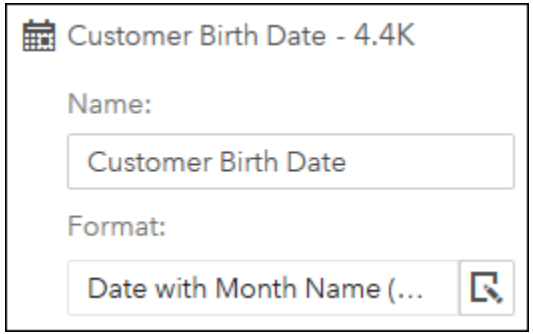
Data Item Properties

In the Data pane, properties can be modified for each data item to aid in your analysis.

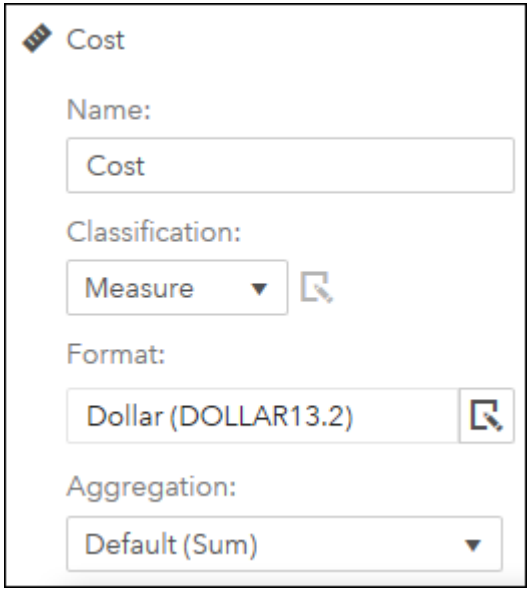
Category



Datetime



Measure



Working with Data Items

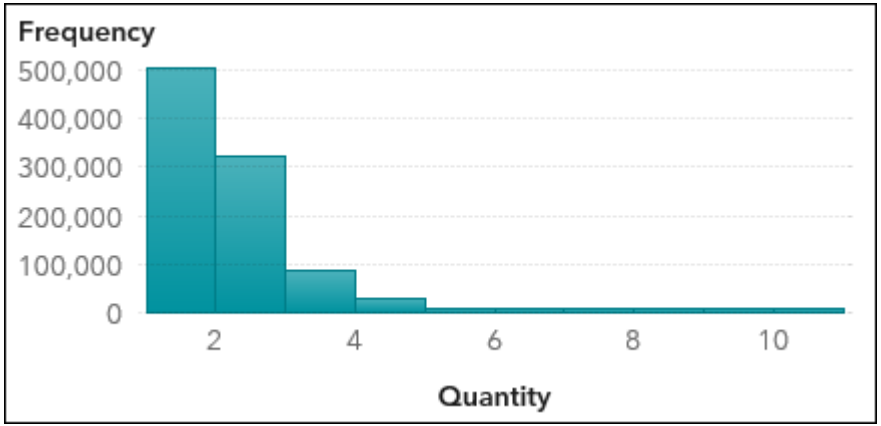
This demonstration illustrates how to modify data item properties (name, format, aggregation) in Visual Analytics.

Chapter 3: Analyzing Data Using SAS® Visual Analytics

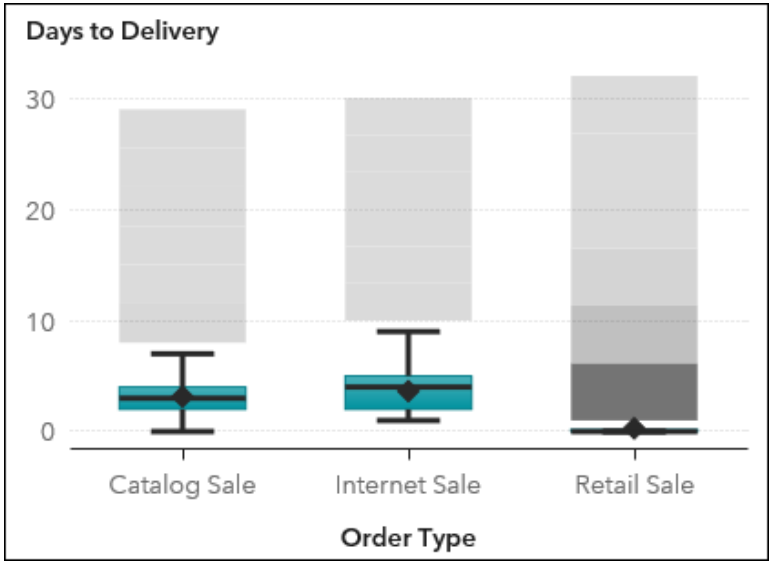
Objectives

- Discuss when to use descriptive graphs (histogram, box plot, bar chart) in Visual Analytics.
- Maximize graphs objects to view details.
- Modify roles and options for graph objects.

Objects: Graphs (Descriptive)



Use a *histogram* to view the distribution of a single measure.

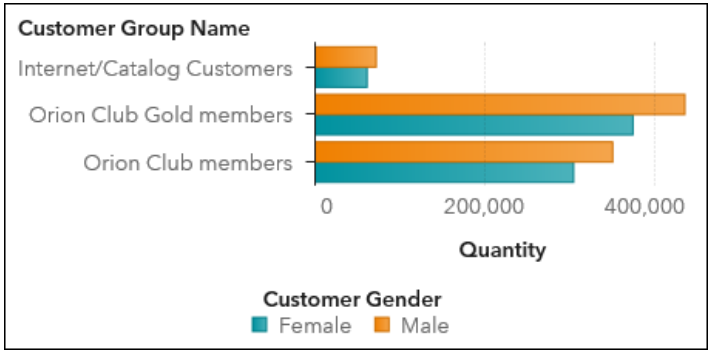


Use a *box plot* to view information about the variability of the data and extreme values.

Objects: Graphs (Descriptive)

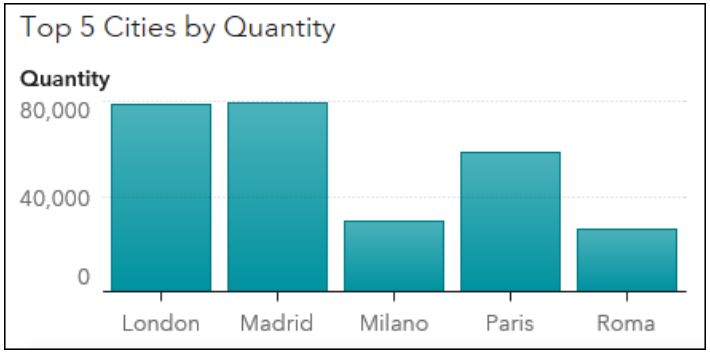
Use a *bar chart* to compare summarized data for the following:

Nominal values

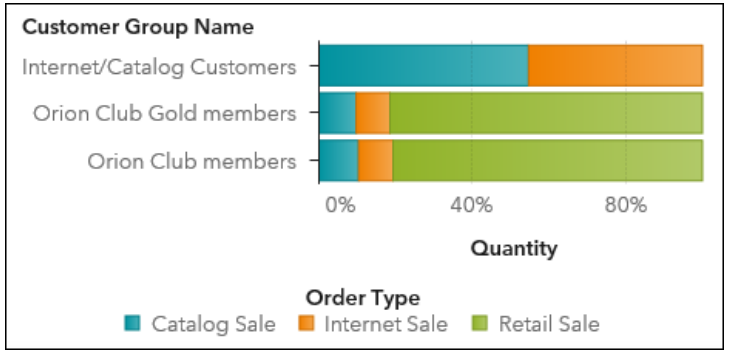


Time series data

Rankings



Parts of a whole



Chapter 3: Analyzing Data Using SAS® Visual Analytics

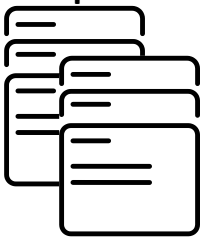
Objectives

- Describe the types of data items that can be created in Visual Analytics.
- Discuss the difference between calculated items and aggregated measures.
- Describe the various ways that data can be filtered in Visual Analytics.
- Discuss when to use geographic maps in Visual Analytics.

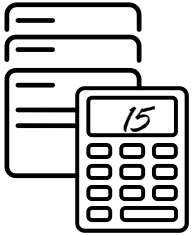
Creating Data Items

The following data items can be created in Visual Analytics:

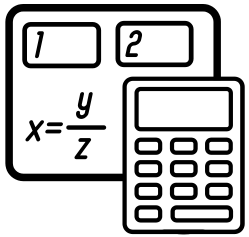
Duplicate



Distinct count

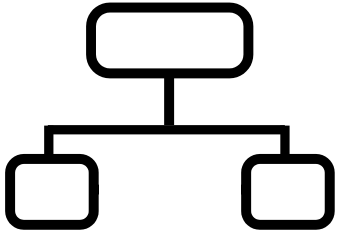


Calculated

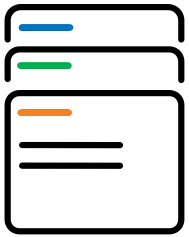


Geography

Hierarchy

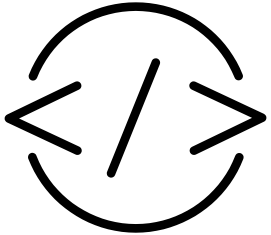


Custom category



Derived items

Parameters

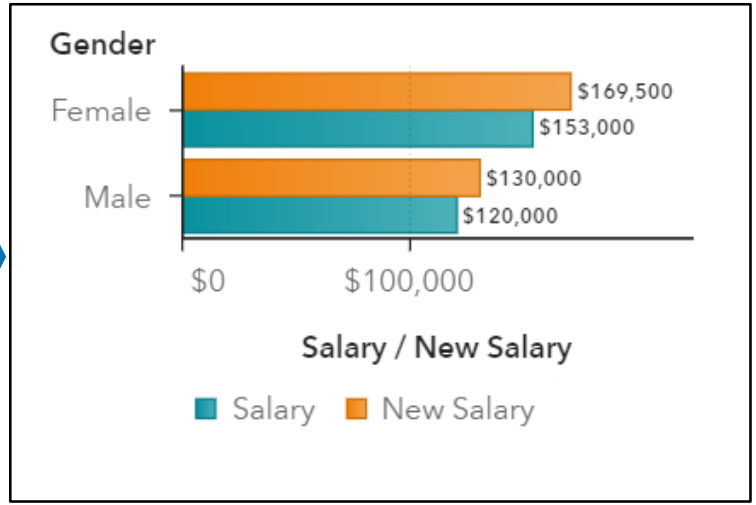


Calculated Item: Example

Calculated items are created by performing operations on unaggregated data.

(Salary * Increase)

Gender	Salary	Increase	New Salary
Male	40,000	1.05	42,000
Female	65,000	1.10	71,500
Female	32,000	1.05	33,600
Male	80,000	1.10	88,000
Female	56,000	1.15	64,400



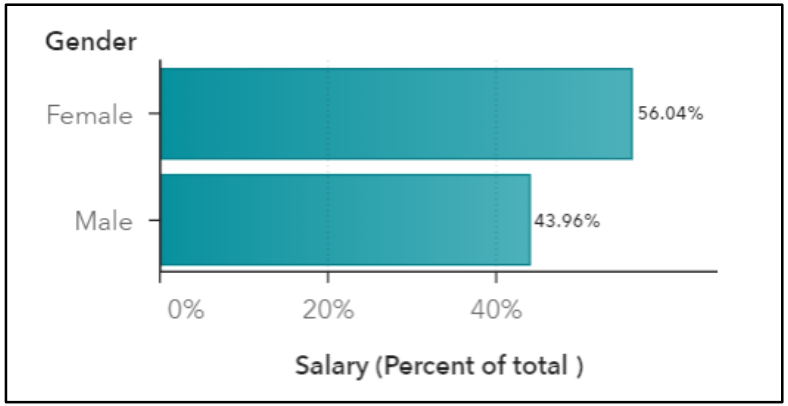
Aggregated Measure: Example

Aggregated measures are created by aggregating and then performing the operation.

(Sum _ByGroup_ (Salary) / Sum _ForAll_ (Salary))

Gender	Salary
Male	40,000
Female	65,000
Female	32,000
Male	80,000
Female	56,000

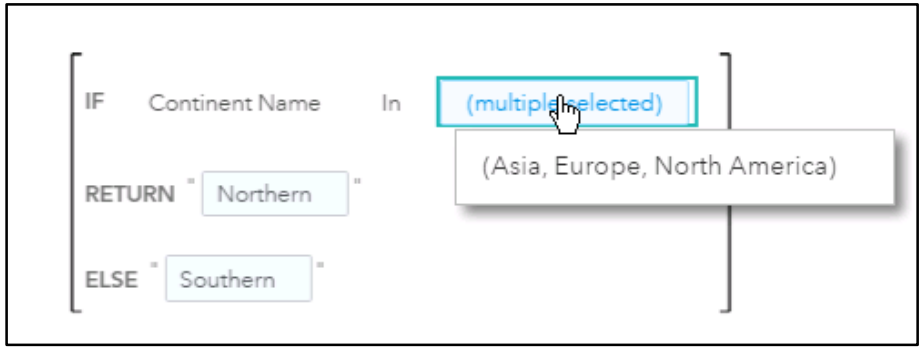
Gender	Salary
Male	120,000
Female	153,000
TOTAL	273,000



Custom Category: Example

Custom categories create labels for groups of category or measure data items.

Calculated item



Custom category

This calculated item and custom category produce equivalent results.

Calculated Columns: Customer Age

The *Now* operator creates a datetime value using the current date and time, where the current date and time is evaluated every time you view the report.

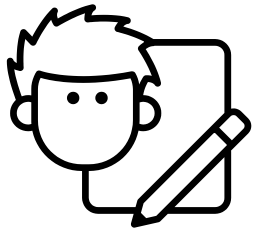
The *DatePart* operator converts a datetime value to a date value.

The *TreatAs* operator enables a numeric, or datetime, value to be used as a different data type within other operators.

The *Floor* operator rounds the number down to the nearest integer.

Filtering Data

Many different types of filters can be created to subset data in Visual Analytics:

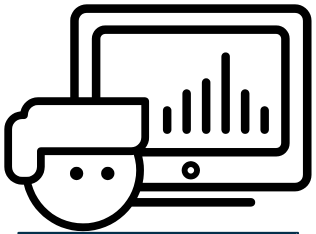


Report Designer

Detail report filters

- Data source
- Basic
- Advanced

Post-aggregate report filters



Report Viewer

Prompts

- Report
- Page

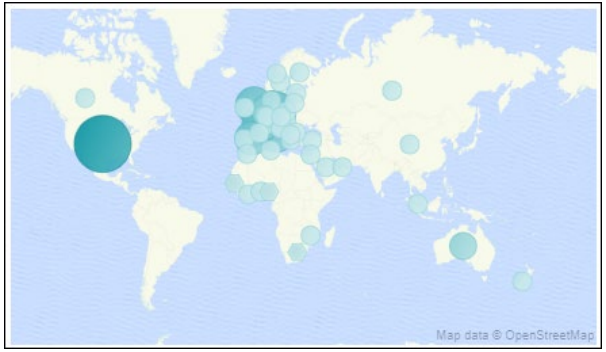
Actions

- Filter
- Links

Objects: Graphs (Geography)

Use a *geo map* when location is a critical component of the analysis.

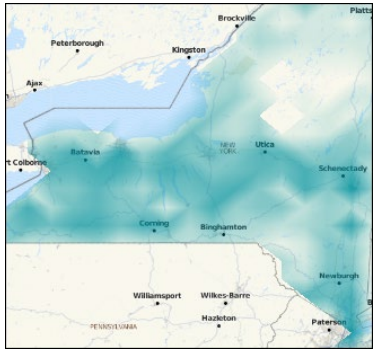
Bubbles



Coordinates

Use a *geo region map* or *geo coordinate map* only when there is an even distribution of values within each region.

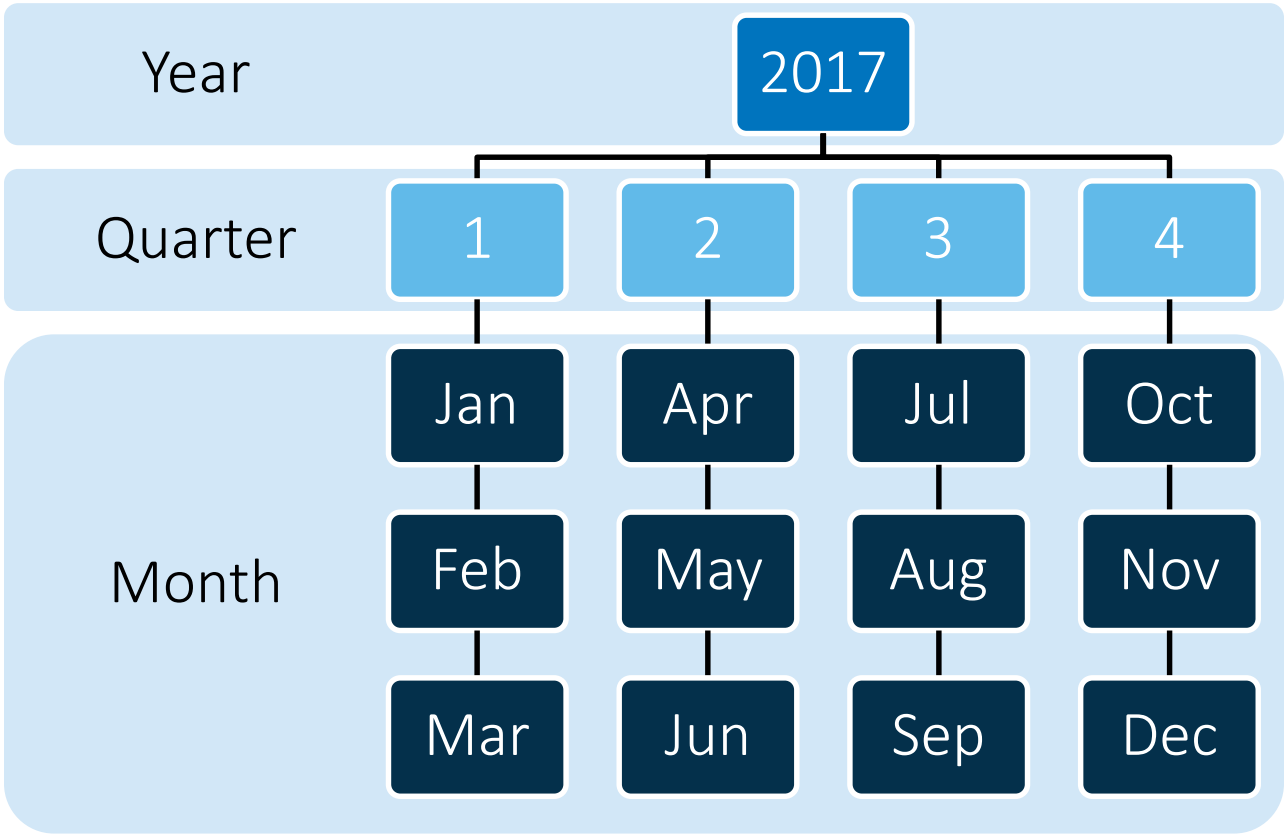
Contour



Regions

Use a *geo contour map* to show very dense data.

What Is a Hierarchy?

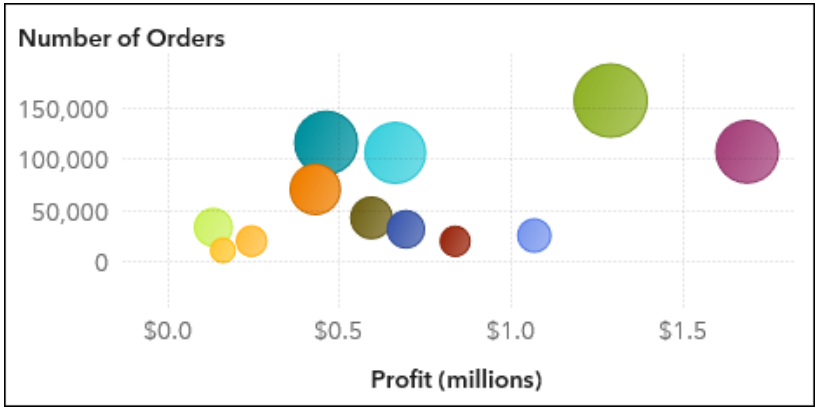


Chapter 3: Analyzing Data Using SAS® Visual Analytics

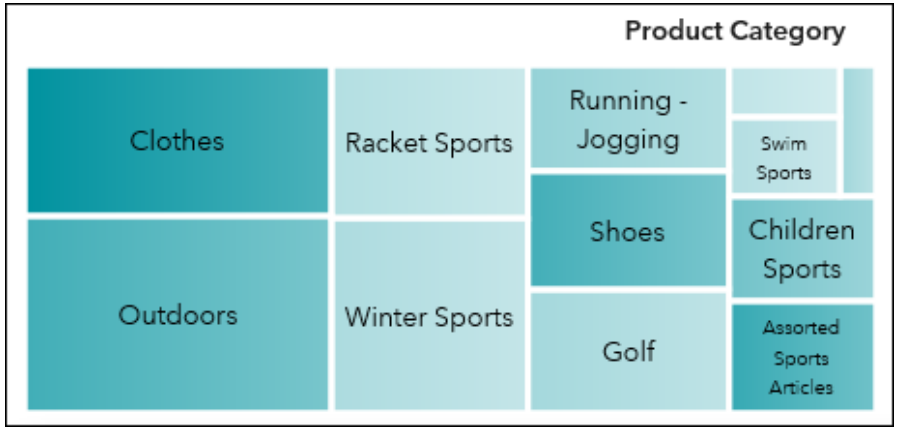
Objectives

- Discuss when to use analysis graphs in Visual Analytics.
- Describe the types of fit lines that can be added to analysis graphs.
- Describe the forecasting capabilities available in Visual Analytics.

Objects: Graphs (Analysis)



Use a *bubble plot* to display three dimensions of data (horizontal location, vertical location, size of bubble) for some group of category values.

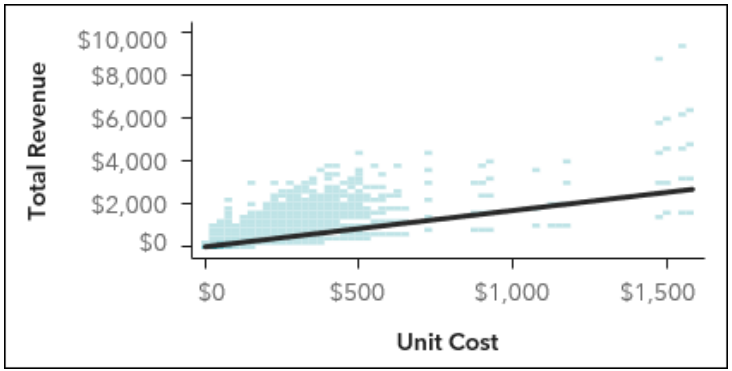


Use a *treemap* to display lots of information in a small amount of space. Use size and color to draw attention to specific areas of interest.

Objects: Graphs (Analysis)

Use a *correlation matrix* to evaluate the linear relationship between measures.

Use a *scatter plot* to evaluate the relationship between two measures.

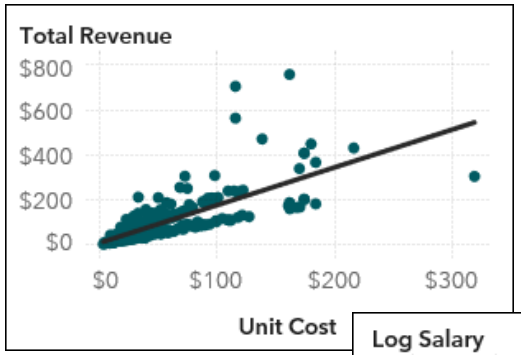


Use a *heat map* to evaluate the relationship between two high-cardinality measures, between two categories, or between a category and a measure.

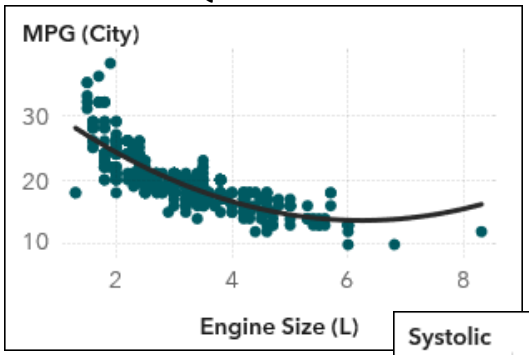
Fit Lines

Fit lines can be added to scatter plots and heat maps to plot the relationship between variables.

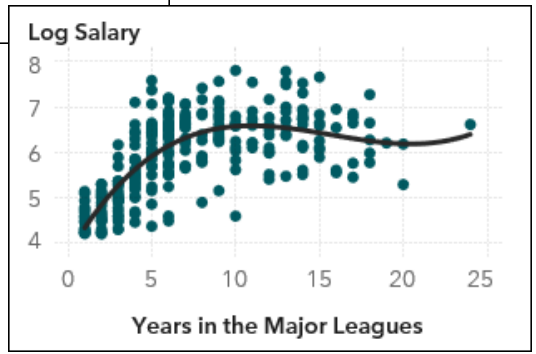
Linear



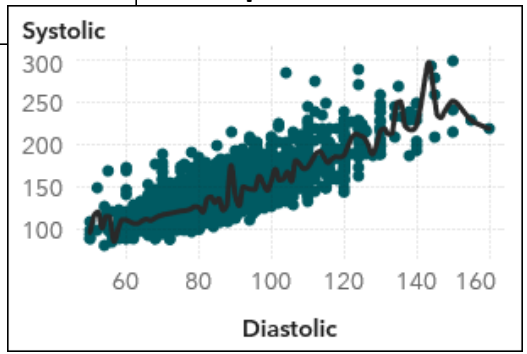
Quadratic



Cubic

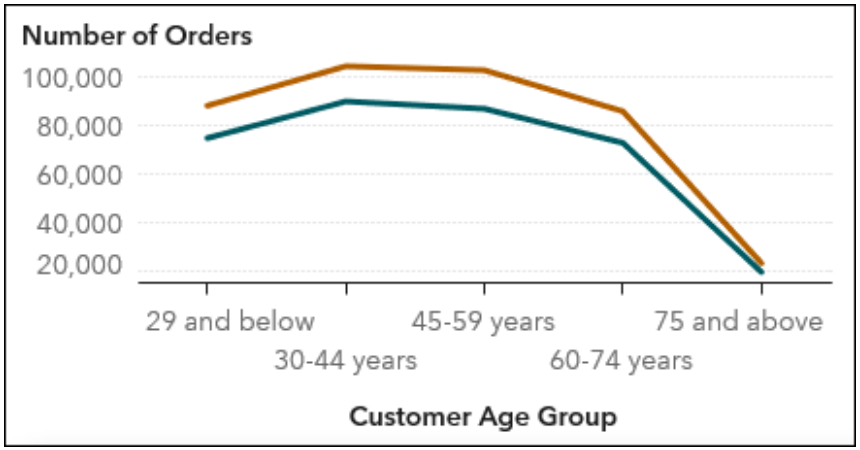


PSpline



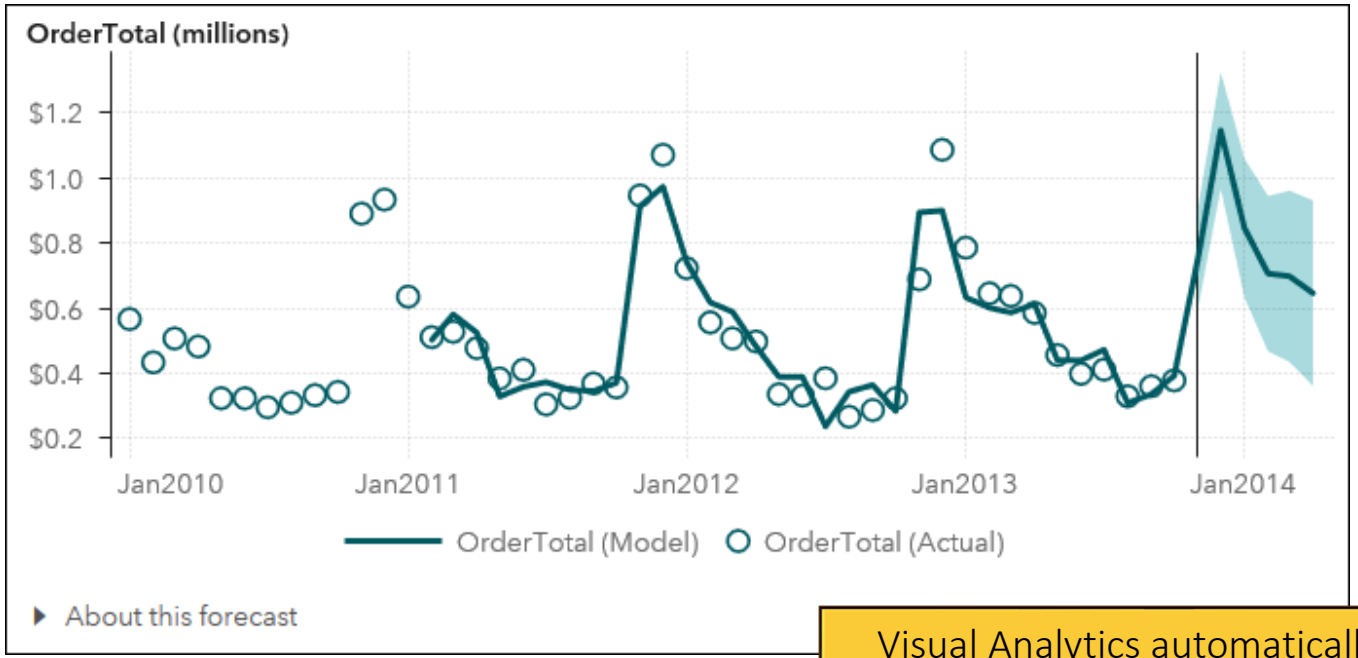
Objects: Graphs (Analysis)

Use a *time series plot* to show trends of measures over time.



Use a *line chart* to show trends over some ordinal variable (time, age group).

Objects: Analytics (Forecasting)



Use a *forecasting* object to show estimates of future values based on historical trends in the data.

Visual Analytics automatically selects the best forecasting model for your data.