



Mu Sigma

Tableau

Training document

Do The Math

Chicago, IL
Bangalore, India
www.mu-sigma.com

Aug 2013

Proprietary Information

"This document and its attachments are confidential. Any unauthorized copying, disclosure or distribution of the material is strictly forbidden"

the
**business
intelligence**
renaissance

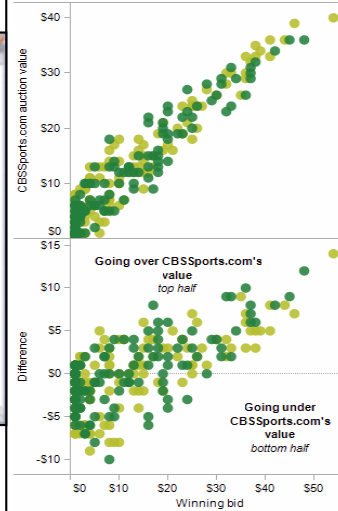


Tableau - MuCel
Mu Sigma



Mixed League

CBSports.com Mixed League analyst auction



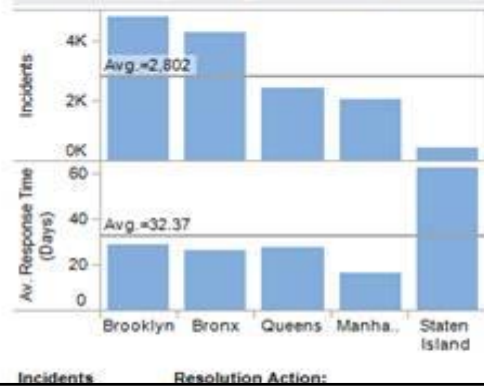
- Select Position:**
- (All)
 - C
 - 1B
 - 2B
 - 3B
 - SS
 - OF
 - DH
 - P
- Select winning bidder:**
- (All)
 - Adam Aizer
 - Al Melchior
 - Corey Guerrero
 - Dave Richard
 - Eric Mack
 - Fazal Inam
 - Jamey Eisenberg
 - Kevin Cruz
 - Peter Madden
 - Ross Devonport
 - Scott White
 - Steven Spierer
- Highlight League:**
- American League
 - National League

Winning bids

Player	Position	Value
Pujols, Albert	1B	\$54
Ramirez, Hanley	SS	\$46
Braun, Ryan J.	OF	\$44
Mauer, Joe	C	\$45
Rodriguez, Alex	3B	\$48
Utley, Chase	2B	\$41
Kemp, Matt	OF	\$38
Fielder, Prince	1B	\$39
Teixeira, Mark	1B	\$42
Cabrera, Miguel	1B	\$38
Howard, Ryan	1B	\$36
Lincecum, Tim	P	\$41
Halladay, Roy	P	\$37
Longoria, Evan	3B	\$38
Kinsler, Ian		
Sabathia, CC		
Ellsbury, Jacoby		
Holliday, Matt		
Reynolds, Mark		
Tulowitzki, Troy		
Crawford, Carl		
Greinke, Zack		

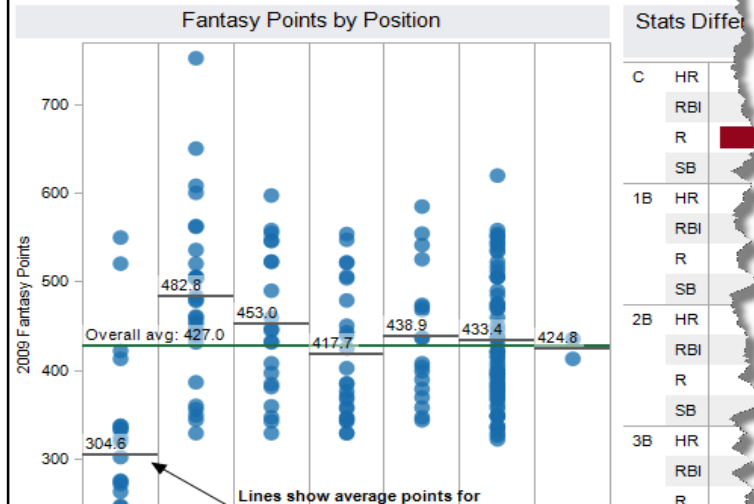
Graffiti in NYC

Graffiti by Borough: Select to Filter



Setting the Position Standards

Select a point in the chart below for player details - hold Ctrl to multi-select



5 reasons why dashboards matter now

1 Right Information to the Right Person

2 Distributed Decision-Making

3 Drill-up, Drill-down and Drill-through

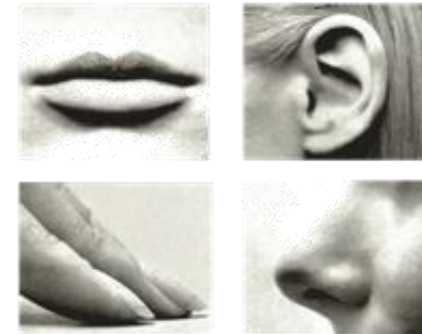
4 Domain Experts Build Analytics

5 Visual, Easy to Understand

Vision: Most **Powerful** Perception Sense



70%



30%

Total Sense Receptors



How many 9s in the grid below?

3	3	0	3	0	1	8	7	6	8	2	1	4	0	3	8	3	7	7	2	0	5	2	3	2	7	0	2	0
7	1	4	6	0	2	1	3	2	7	6	0	2	5	6	3	2	5	7	6	3	3	0	2	0	3	0	7	2
8	7	5	7	2	8	3	8	7	7	8	2	0	7	7	5	2	3	1	1	5	6	3	8	4	7	8	2	0
0	5	0	5	1	6	1	7	5	6	8	0	4	4	6	7	4	7	1	4	0	0	8	4	4	3	0	3	2
2	4	3	1	3	5	4	9	5	0	7	6	0	7	4	3	1	8	2	7	3	4	6	0	2	4	8	2	3
8	6	2	2	6	5	4	6	7	0	7	6	0	0	3	9	0	2	4	7	1	7	2	3	3	5	8	7	0
0	8	4	5	1	3	1	7	6	4	5	4	1	2	4	5	3	3	5	4	9	6	7	7	6	3	4	2	5
4	7	7	0	2	2	0	1	1	7	7	7	0	2	6	6	4	7	5	8	6	1	4	3	7	8	5	4	6
4	3	6	6	4	6	6	2	8	4	8	5	3	7	8	8	1	3	8	5	4	5	7	4	0	3	2	8	4
5	5	0	3	5	3	5	3	8	3	2	3	8	2	3	1	6	2	7	2	4	6	3	6	4	4	3	2	5
4	4	0	2	1	7	2	4	4	7	4	1	9	2	4	5	2	5	0	4	0	0	5	3	6	3	3	6	7
7	4	6	6	8	7	5	7	9	2	0	2	8	8	8	8	3	2	4	2	6	4	0	4	6	3	7	2	1
0	1	7	1	5	9	1	4	2	8	7	3	7	1	4	5	1	8	7	8	0	5	1	7	0	5	8	8	1
2	8	5	2	1	2	8	7	7	6	2	5	6	2	6	4	1	5	1	6	1	2	1	1	0	5	6	4	0
2	1	1	7	7	2	0	0	1	8	7	0	2	9	0	2	8	5	7	8	4	6	0	6	5	0	7	1	2
0	5	2	4	1	5	3	3	1	5	5	1	4	0	1	6	4	3	3	9	8	8	3	4	6	8	4	8	6
7	3	7	5	2	4	0	2	7	6	3	8	5	5	4	5	8	8	7	5	5	6	5	6	7	9	7	7	4
0	3	2	8	1	4	4	6	0	8	2	3	0	1	3	4	6	2	0	5	7	7	3	6	1	8	7	3	5
4	4	8	3	3	3	5	0	1	0	3	8	6	3	2	0	5	0	6	1	3	3	4	3	6	1	5	8	6
1	0	2	2	7	6	3	3	0	8	8	0	3	1	8	8	1	2	1	7	5	2	9	3	5	8	3	2	5



Is it **easier** to count now?

3 3 0 3 0 1 8 7 6 8 2 1 4 0 3 8 3 7 7 2 0 5 2 3 2 7 0 2 0
7 1 4 6 0 2 1 3 2 7 6 0 2 5 6 3 2 5 7 6 3 3 0 2 0 3 0 7 2
8 7 5 7 2 8 3 8 7 7 8 2 0 7 7 5 2 3 1 1 5 6 3 8 4 7 8 2 0
0 5 0 5 1 6 1 7 5 6 8 0 4 4 6 7 4 7 1 4 0 0 8 4 4 3 0 3 2
2 4 3 1 3 5 4 **9** 5 0 7 6 0 7 4 3 1 8 2 7 3 4 6 0 2 4 8 2 3
8 6 2 2 6 5 4 6 7 0 7 6 0 0 3 **9** 0 2 4 7 1 7 2 3 3 5 8 7 0
0 8 4 5 1 3 1 7 6 4 5 4 1 2 4 5 3 3 5 4 **9** 6 7 7 6 3 4 2 5
4 7 7 0 2 2 0 1 1 7 7 7 0 2 6 6 4 7 5 8 6 1 4 3 7 8 5 4 6
4 3 6 6 4 6 6 2 8 4 8 5 3 7 8 8 1 3 8 5 4 5 7 4 0 3 2 8 4
5 5 0 3 5 3 5 3 8 3 2 3 8 2 3 1 6 2 7 2 4 6 3 6 4 4 3 2 5
4 4 0 2 1 7 2 4 4 7 4 1 **9** 2 4 5 2 5 0 4 0 0 5 3 6 3 3 6 7
7 4 6 6 8 7 5 7 **9** 2 0 2 8 8 8 8 3 2 4 2 6 4 0 4 6 3 7 2 1
0 1 7 1 5 **9** 1 4 2 8 7 3 7 1 4 5 1 8 7 8 0 5 1 7 0 5 8 8 1
2 8 5 2 1 2 8 7 7 6 2 5 6 2 6 4 1 5 1 6 1 2 1 1 0 5 6 4 0
2 1 1 7 7 2 0 0 1 8 7 0 2 **9** 0 2 8 5 7 8 4 6 0 6 5 0 7 1 2
0 5 2 4 1 5 3 3 1 5 5 1 4 0 1 6 4 3 3 **9** 8 8 3 4 6 8 4 8 6
7 3 7 5 2 4 0 2 7 6 3 8 5 5 4 5 8 8 7 5 5 6 5 6 7 **9** 7 7 4
0 3 2 8 1 4 4 6 0 8 2 3 0 1 3 4 6 2 0 5 7 7 3 6 1 8 7 3 5
4 4 8 3 3 3 5 0 1 0 3 8 6 3 2 0 5 0 6 1 3 3 4 3 6 1 5 8 6
1 0 2 2 7 6 3 3 0 8 8 0 3 1 8 8 1 2 1 7 5 2 **9** 3 5 8 3 2 5



Tableau is a new-age self service BI tool, that aids in data analysis and creation of business dashboards

About the company

- ▶ Tableau Software, Inc. was founded in 2003 and is headquartered in Seattle, U.S.
- ▶ It is a recognized leader in providing data visualization products, focused on business intelligence

Offerings

- ▶ **Tableau Desktop** to create and publish workbooks
- ▶ **Tableau Server** to distribute the published workbooks with enterprise level scalability
- ▶ **Tableau Public** to create and quickly publish interactive visuals on the web
- ▶ **Tableau Reader** to view workbooks and dashboards only

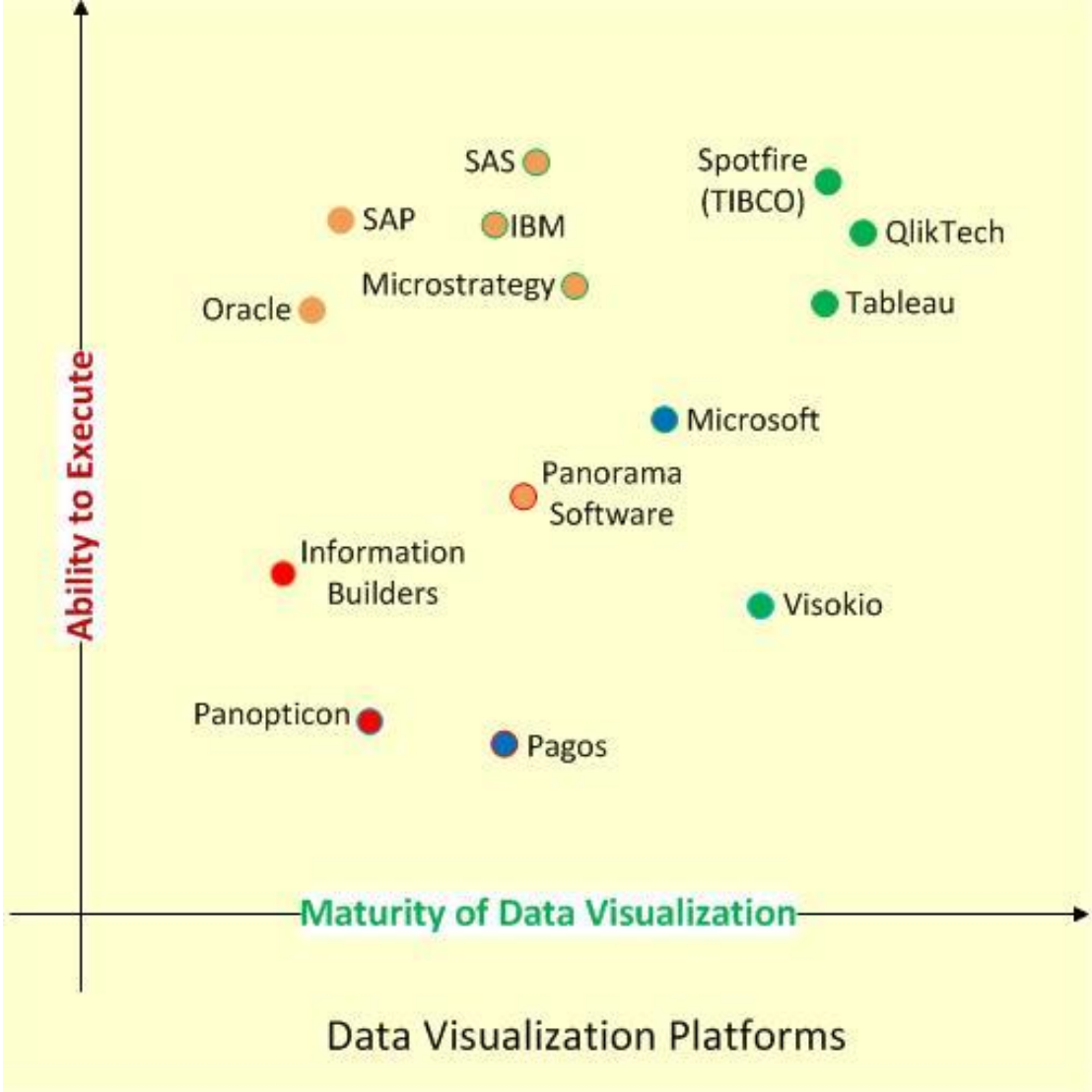
Customers

- ▶ Some of Tableau’s customers across various verticals include:
 - **BFSI** - Allstate, Bank of America, Barclays, Citigroup
 - **Retail** - Barnes & Noble, Starbucks Coffee, Wal-Mart
 - **Pharmaceutical** - Merck, Pfizer, Sanofi Aventis
 - **Technology** – eBay, EMC, Facebook

Key Differentiators

- ▶ Tableau stands out because of the ease of usage and control it allows the users to manipulate the data
- ▶ Tableau can easily connect to a huge variety of data sources, ranging from excel sheets to delimited files, oracle databases and several more
- ▶ Interactive analysis can be shared with a report consumer equipped with a Web browser

Tableau provides enhanced visualization capabilities, all the while maintaining its ease of usage





Excerpts from Gartner B.I. platform survey on Tableau

Strengths

- ▶ Overwhelmingly positive customer satisfaction in terms of ease of use, functionality, product quality, performance, support
- ▶ Tableau delivers strong interactive visualization for analysis, dashboards, information delivery and managed analytic applications
- ▶ Self contained BI platform which integrates with a large variety of data sources and drag-n-drop feature enable user engagement and customization

Mu Sigma POV

- ▶ Tableau enables the end user to move away from just a consumer to a creator of analytics

Cautions

- ▶ Tableau lacks broader BI platform capabilities, such as production reporting and predictive analytics
- ▶ Tableau's products often fill an unmet need in organizations that already have a BI standard, and are frequently deployed as a complementary capability to an existing BI platform
- ▶ Above average license price per user

Mu Sigma POV

- ▶ Individual license cost might be higher, but the cost of development is much lower

Tableau allows user to do an in-depth data analysis & visualization and create interactive dashboards with minimum effort

- ▶ Tableau has three products to suit varying requirements

Tableau Desktop

- Allows the user to connect to data , analyze them, then visualize and create interactive dashboards
- The Software requires license which costs ~ \$2000 (with life time free updates)

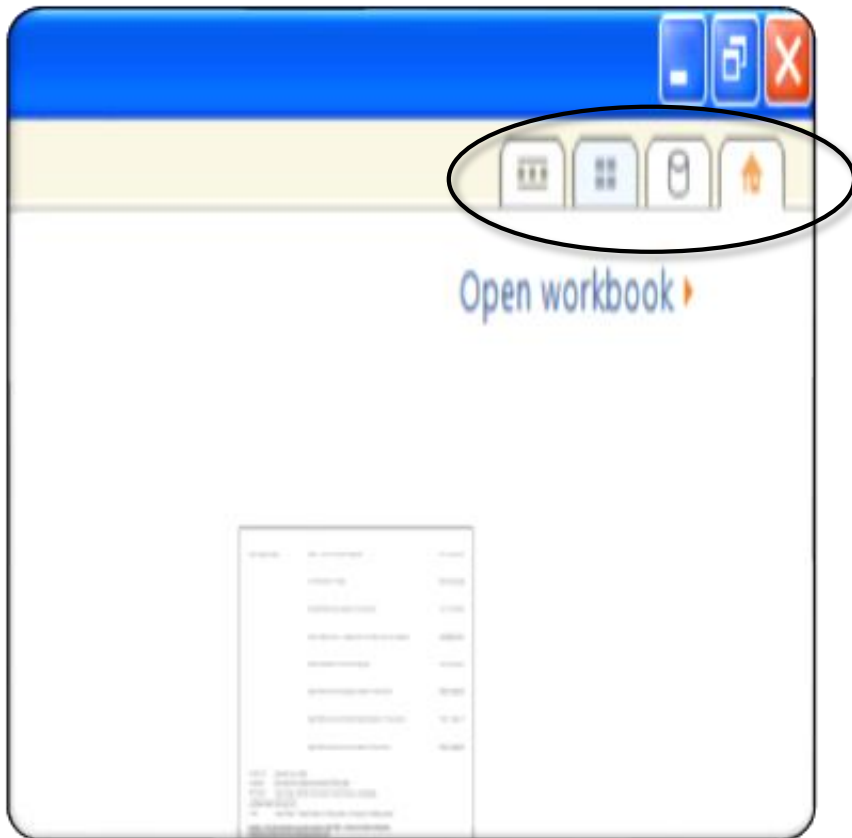
Tableau Reader

- Allows the user to view at the data analysis, visualization and business intelligence information from Tableau packaged workbook
- The Software does require any license and is free of cost

Tableau Server

- Allows the user to upload the tableau workbook & data connections and share the link
- Allows scheduling of periodic automatic refresh of the workbook with the latest data

Tableau has four Tabs, icons of which can be found on the left top corner of Home Tab



Home : Has the set of recently opened workbooks, some samples workbooks and some saved data sources



Data Connection Tab : Provides a list of data connections that can be made, that includes connections from flat files as well connections to different servers



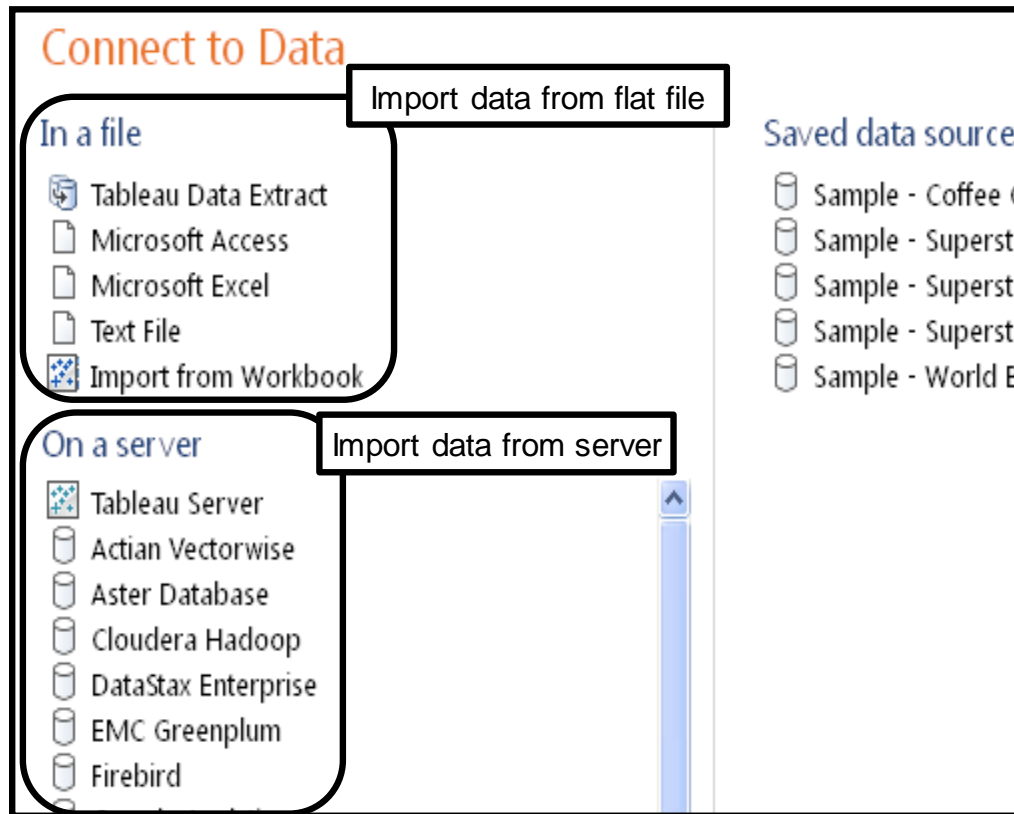
Thumbnail view : Provides the thumbnails of all the dashboards/sheets in the workbook



Sheet Tab : Shows a sheet/dashboard for working on it

Data Connections in Tableau

Tableau provides a list of data connections, in the data connection tab, clicking from which data can be imported either from flat files or from server.



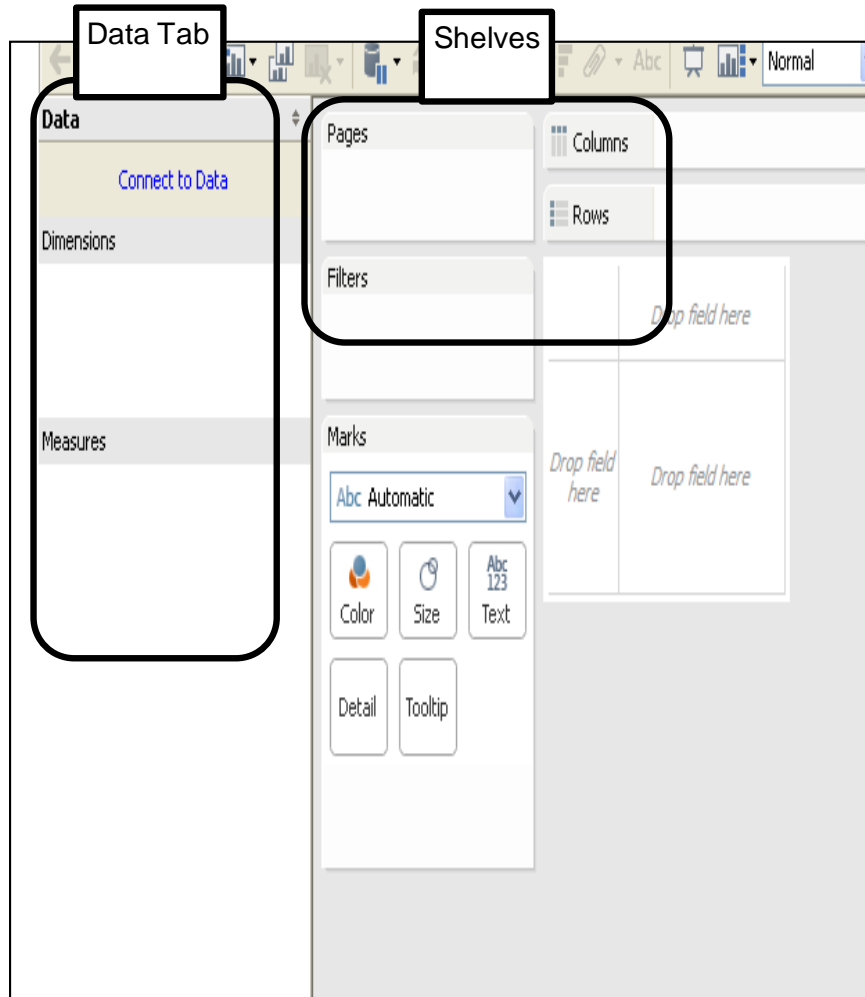
Different ways of connecting to the data

The screenshot shows the 'Data Connection' dialog box with three options:

- Connect live**: Represented by a double-headed green arrow icon. A callout box explains: "Directly connects to the data source and refers to it every time any field is used (generally used when the data is small)". Below the option, it says: "Connect directly to your data. The speed of your data source will determine performance."
- Import all data**: Represented by a green arrow pointing to a spreadsheet icon. A callout box explains: "Creates an extract of the data, which moves along with the packaged workbook (used when data is large)". Below the option, it says: "Import all of your data into Tableau's fast data engine."
- Import some data**: Represented by a green arrow pointing to a spreadsheet icon. A callout box explains: "Enables filtering of data over different fields before extracting it". Below the option, it says: "Select a subset of your data to import into Tableau's fast data engine."

At the bottom of the dialog box, there is a checkbox labeled "Always do this for Microsoft Excel" and a "Learn More" link.

Basic view of a worksheet



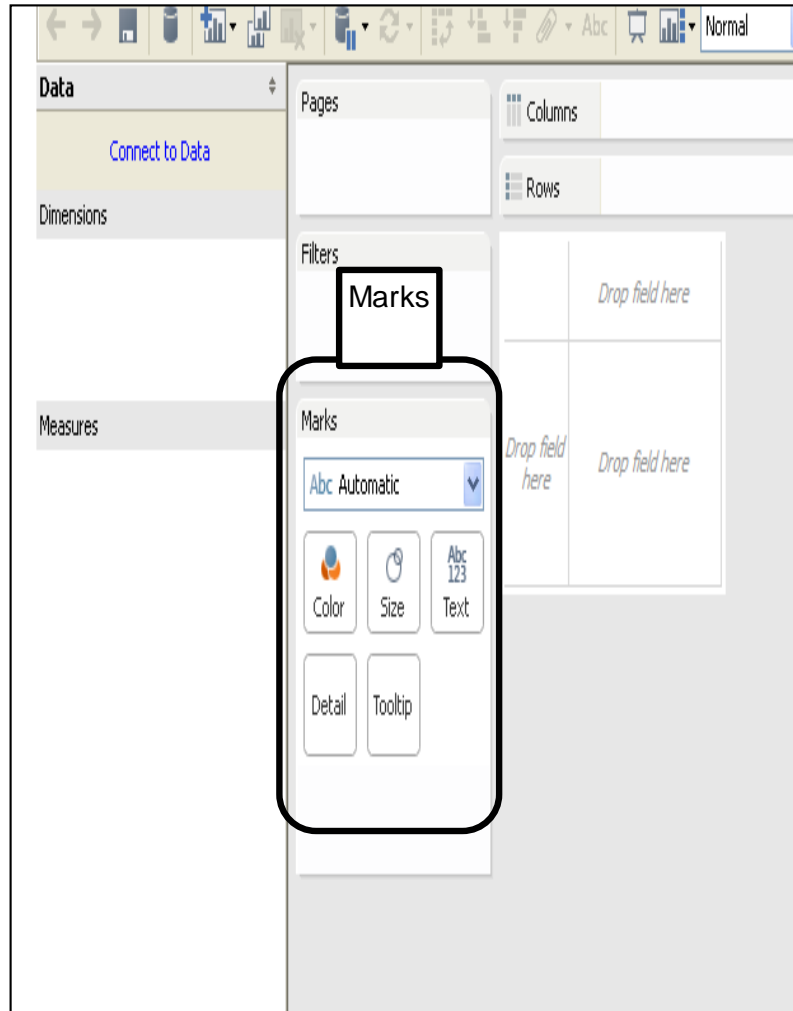
Data Tab

- ▶ Data consists of “Dimensions” and “Measures”.
 - Dimensions – All string and date variables fall under this category
 - Measures – All numeric fields on which calculations are formed, fall under this category.
- ▶ Tableau automatically identifies the data types and groups them in dimensions and measures

Shelves

- ▶ The different shelves present are
 - Rows – Contains the field that has to appear in rows
 - Columns – Contains the field that has to appear in columns
 - Filters – Contains the fields on which data has to be filtered. (Filters can be applied only for the current sheet/all sheets/specific sheets based on the requirement)
 - Pages – Fields in this shelf are not filtered but are restricted with respect to the view

Basic view of a worksheet



Marks

- ▶ The different options present under marks are
 - Drop down – This allows to choose the type of graph to be presented in the view
 - Color – The colors in the view will be based on the field pulled into this box
 - Size – The size of the representations in the view (bar, line etc.) can be adjusted based on the field pulled into this box
 - Text/Label – Any field pulled into this box, gives the label to the graph
 - Detail – The field pulled into this indicates the level of data for the sheet
 - Tooltip – Fields pulled into this, appears as tooltip (values appear when hovered over the graphs)

Hierarchy – This is used to group fields of decreasing levels in the form of a hierarchy

The screenshot shows the Tableau interface with the 'Data' pane on the left. A context menu is open over the 'Country' field, with 'Create Hierarchy...' selected. The menu includes options like 'Copy', 'Paste', 'Duplicate', 'Rename...', 'Hide', 'Create Calculated Field...', 'Create Group...', 'Create Set...', 'Create Parameter...', 'Create Hierarchy...', 'Convert to Measure', 'Change Data Type', 'Geographic Role', 'Default Properties', 'Replace References...', and 'Describe...'. The 'Dimensions' pane lists fields such as 'Abc Brand', 'Abc Channel', 'Abc Corporation', 'Abc Corporation Lipid Lowering Market', and 'Country'.

Once the hierarchy is created various fields can be pulled into it in the decreasing order of the hierarchy

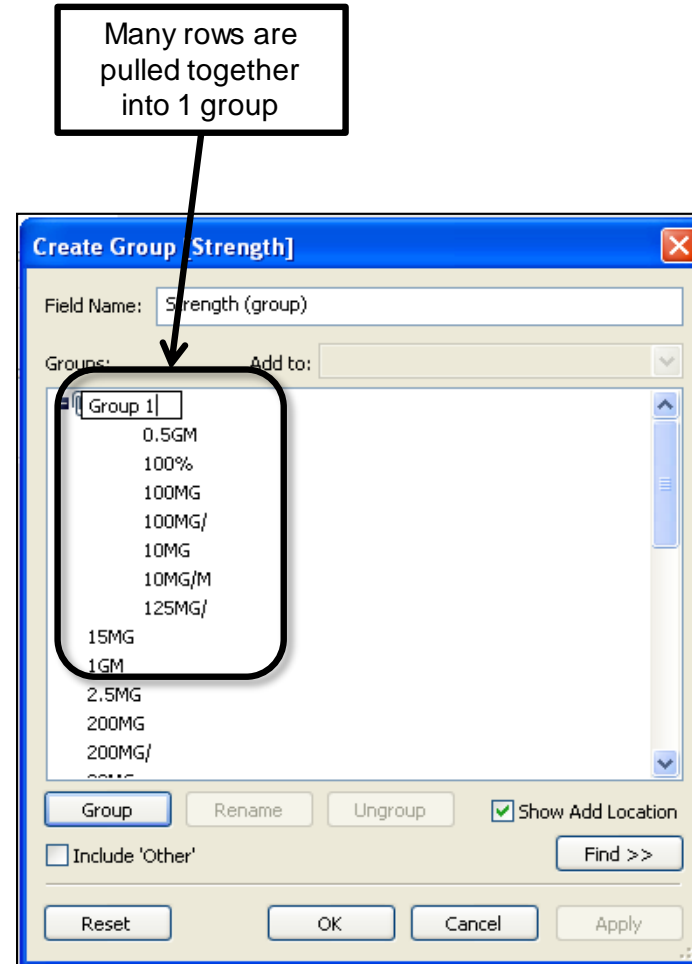
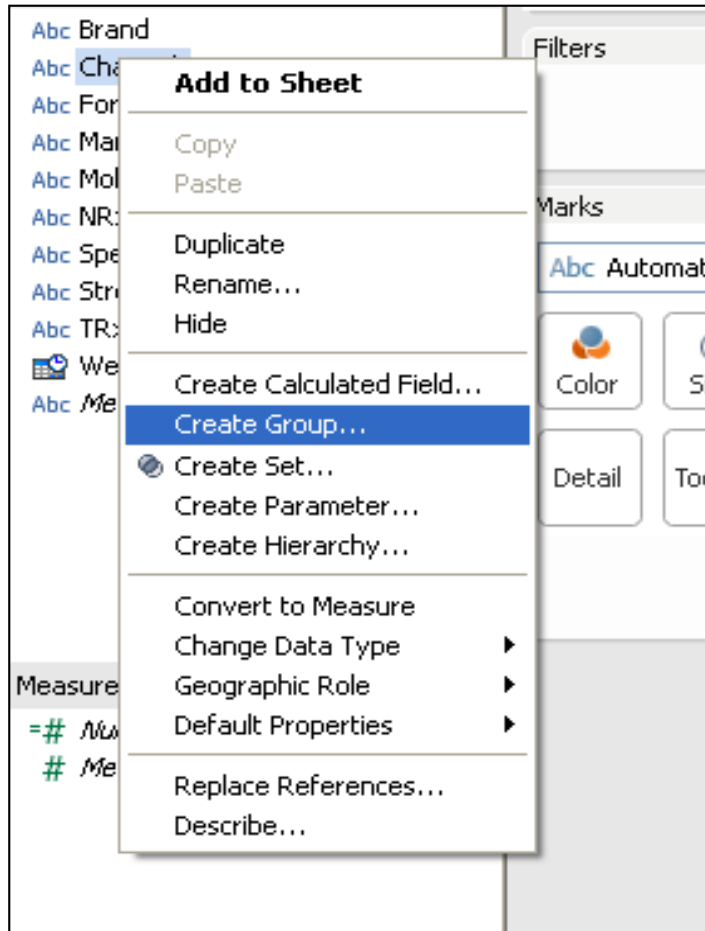
Clicking on this “+” sign, helps to drill down to lowers levels of the hierarchy



The screenshot shows the Tableau interface with a hierarchy of 'Geography' fields in the 'Dimensions' pane. The hierarchy includes 'Abc Worldwide', 'Country', 'Abc Super_Region', 'Abc Region', and 'Abc Sub_Region'. The 'Columns' shelf contains 'Worldwide' and 'Country'. The 'Rows' shelf contains 'Worldwide' and 'Country'. A table view is displayed with columns for 'Worldwide' and 'Country'. A '+' sign is highlighted in a box, indicating the drill-down functionality.

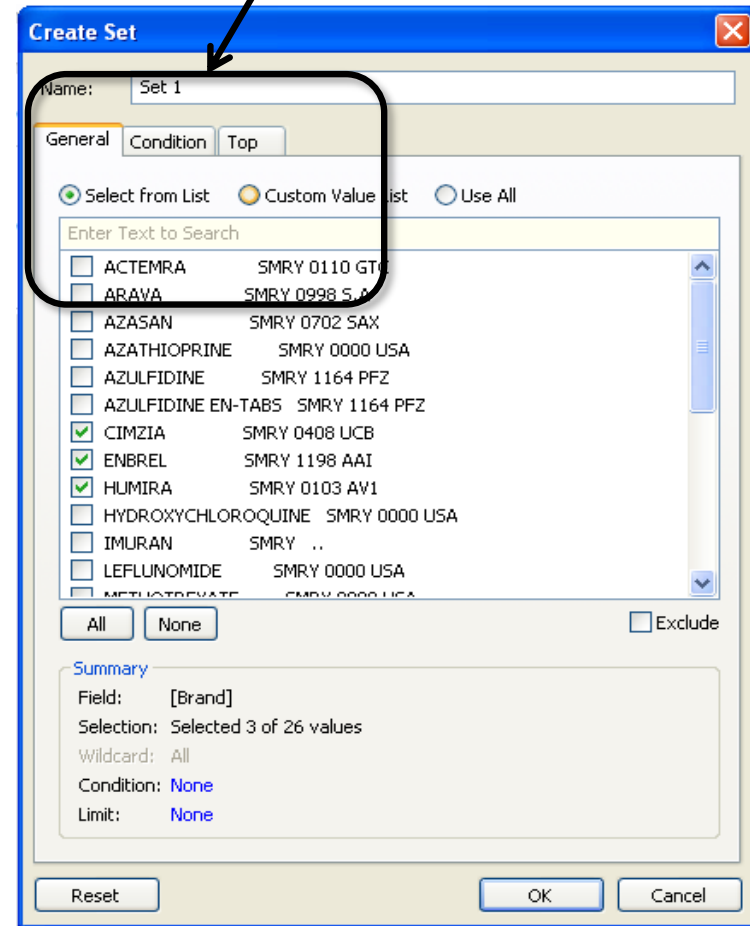
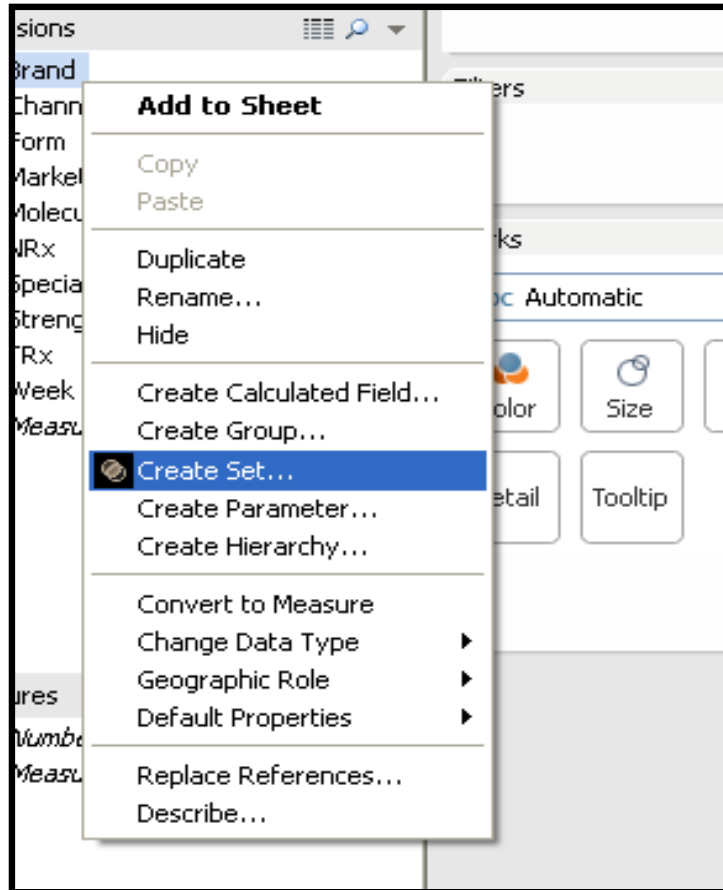
Worldwide	Country	
WorldWide	Algeria	Abc
	Argentina	Abc
	Australia	Abc
	Austria	Abc
	Belgium & Luxembourg	Abc
	Brazil	Abc
	Bulgaria	Abc
	Canada	Abc
	Central America	Abc

Groups – These are used to create groups among different rows of one field.



Sets – An easy way of adding a filter. Rows assigned to a set, filter the data accordingly when pulled into the sheet.

Sets can be defined, by manually selecting the required fields or by some condition or based on the top few rows



Filters

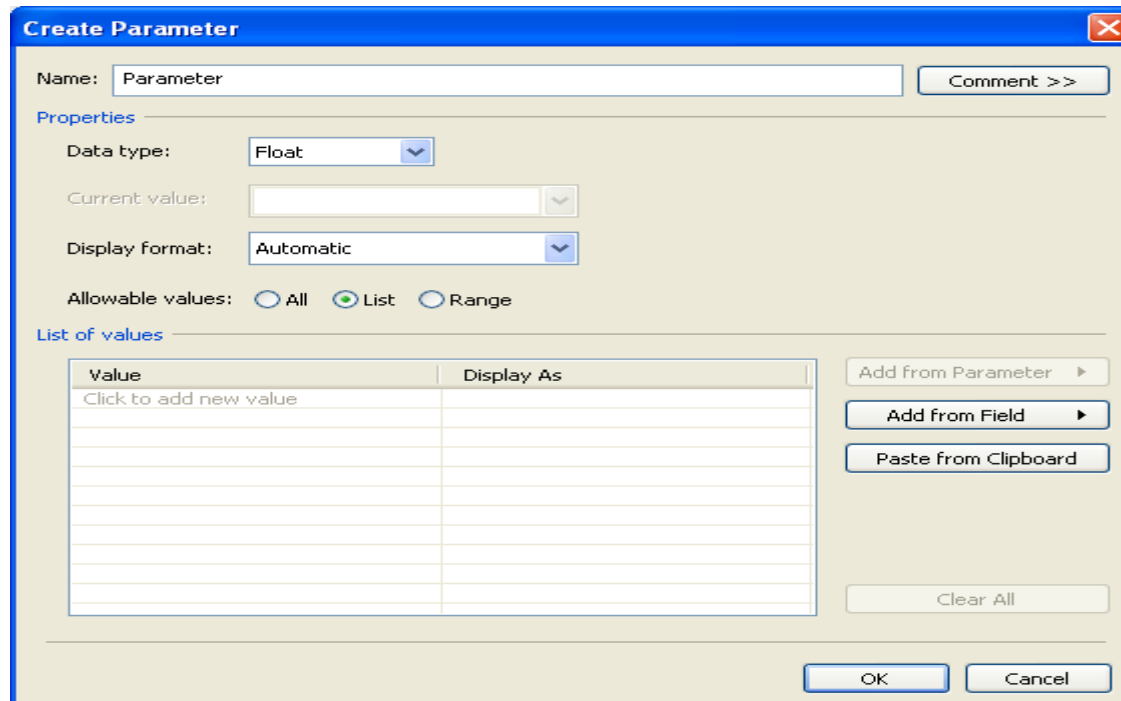
- ▶ Filters can be used for limiting the data for analysis/visualization and can be applied on
 - Members of a dimension field
 - Range of values of a measure field
 - Relative dates / Range of dates of a Date field
- ▶ By default, all filters are computed independently i.e., each filter accesses all rows the data source without regard to other filters.
- ▶ A filter can be set as a context filter with add to context option. With setting a field as a context any other filters that are used act as dependent filters i.e., they process only the data that passes through the context filter.
- ▶ Advantages of context filter :
 - Improve performance – If you set a lot of filters or have a large data source, the queries can be slow. You can set one or more context filters to improve performance.
 - Create a dependent numerical or top N filter – You can set a context filter to include only the data of interest, and then set a numerical or a top N filter

Calculated Fields and Aggregated Calculations

- ▶ Tableau allows the user to create a new calculated field by defining a formula based on data source fields and other calculated fields, and using standard functions and operators
- ▶ The new calculated field displays in either the Dimensions area or the Measures area of the Data window depending on the data type returned by the calculation
- ▶ Calculated fields using aggregated functions like sum, max, min, avg, etc., are called aggregated calculations
- ▶ An aggregated value cannot be combined with a disaggregated value. For example, $\text{SUM}(\text{Price}) * [\text{Items}]$ is not valid while $\text{SUM}(\text{Price} * \text{Items})$ and $\text{SUM}(\text{Price}) * \text{SUM}(\text{Items})$ are both valid.
- ▶ An aggregate calculation is always a measure and are always aggregated at “level of detail” which is defined by the fields placed in rows and detail shelf

Parameters

- ▶ Parameters are user defined variables that can be used to create a dynamic visualization
- ▶ Since they are user defined, they wouldn't have any connection with the data unless used in a calculated field, filter or reference lines.
- ▶ Multiple value selection is not possible with parameters and the selections works globally across all sheets



Create Parameter

Name:

Properties

Data type:

Current value:

Display format:

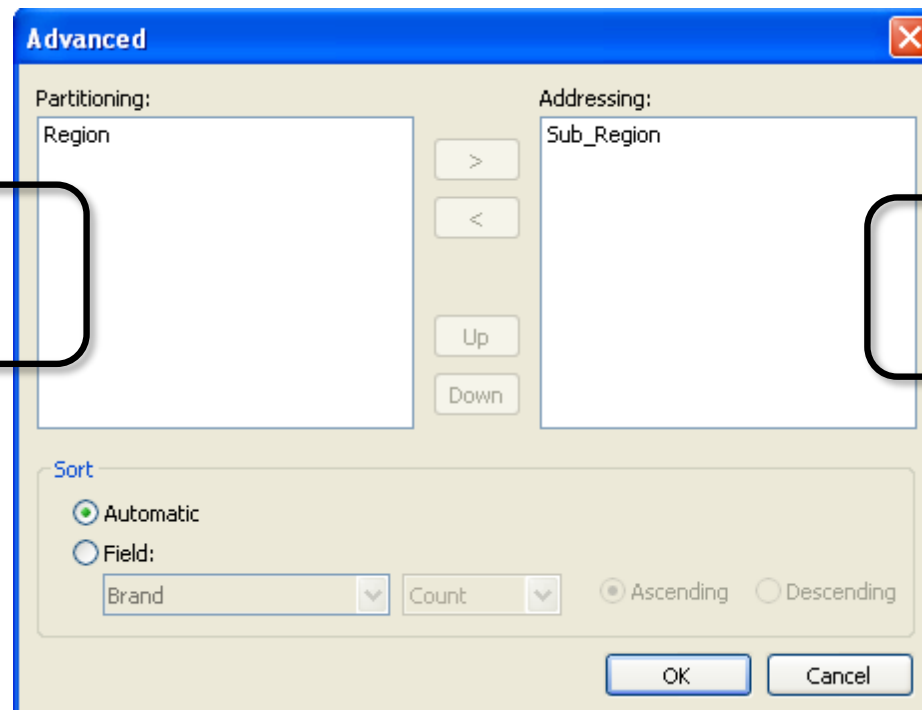
Allowable values: All List Range

List of values

Value	Display As
Click to add new value	

Table Calculations

- ▶ Table Calculations allows the user to select the level of aggregation irrespective of fields in rows and detail shelves
- ▶ Quick table calculations option allows the user to use few common table calculations like Running Total, Percent of Total, Moving Average etc., without having to define the calculation
- ▶ The level of aggregation can be selected using the option – edit table calculation option

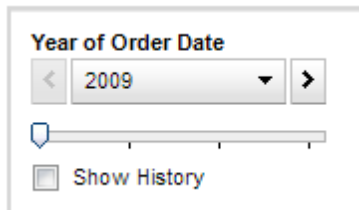


The calculations will be aggregated at the level of all fields included in 'Partitioning' box

All the fields included in the 'Addressing' box will be ignored will computing aggregations

Pages

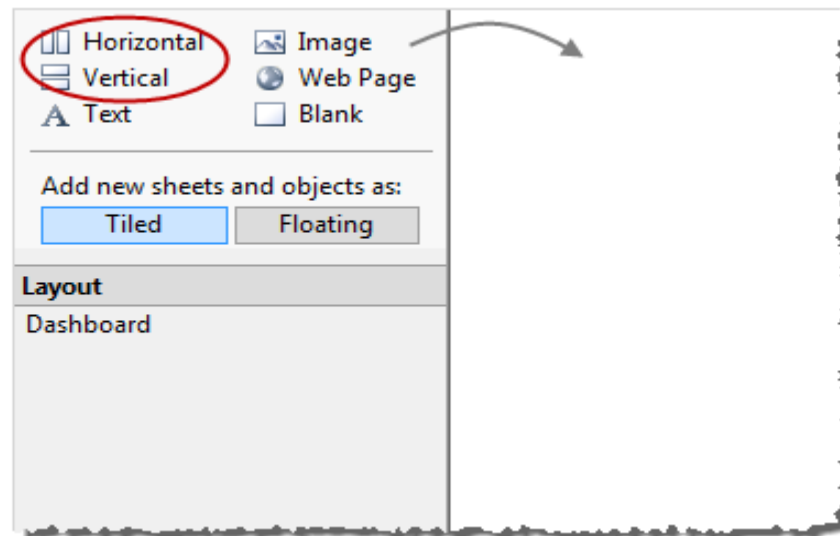
- ▶ The Pages shelf allows to break a view into a series of pages
- ▶ When a dimension is placed on the Page shelf, a new row is added for each member in the dimension.
- ▶ When a measure is placed on the Pages shelf, the measure is converted into a discrete measure



- ▶ User can manually advance through the sequence of pages in any of the following ways:
 - Use the drop-down menu to select a value.
 - Use the forward and back buttons on either side of the drop-down list to navigate through the pages one at a time.
 - Use the Page slider to quickly scroll forward and backward in the sequence of pages.
- ▶ Selecting the 'Show History' option allows to show marks from previous pages in addition to marks for the current page

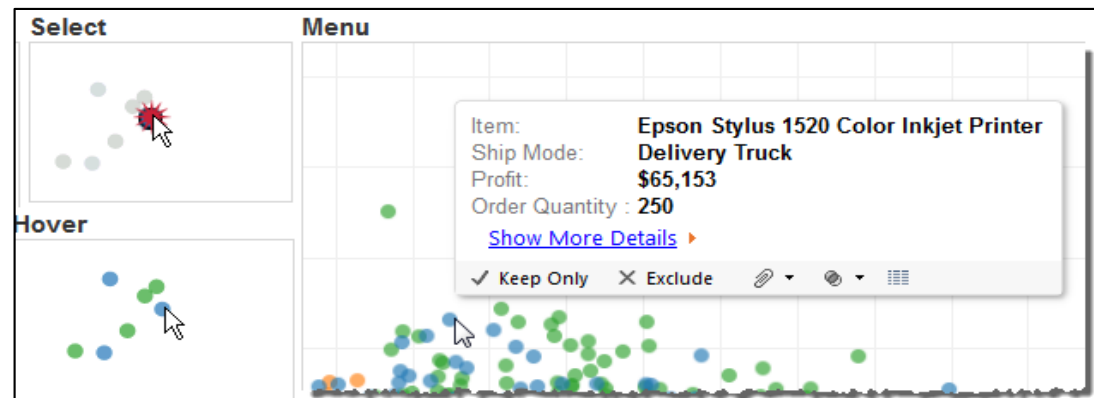
Dashboard

- ▶ A dashboard is a collection of several worksheets and supporting information shown in a single place
- ▶ A variety of supporting objects such as text areas, web pages, and images can also be added in a dashboard
- ▶ All sheets and objects placed in the dashboard can be organized / arranged using containers
 - **Horizontal Containers:** Allows the user to place the sheets one beside the other
 - **Vertical Containers:** Allows the user to place the sheets one below the other



Actions

- ▶ Tableau allows the user to add context and interactivity to your data using actions
- ▶ There are three kinds of actions in Tableau:
 - **Filter** : Allows the user to use the data in one view to filter data in another
 - **Highlight** : Allows the user to use the data in one view to highlight the data in another
 - **URL actions** : It's a hyperlink that points to a Web page, file, or other web-based resource outside of Tableau
- ▶ Actions can be initiated by three following ways:
 - **Hover** - rest the pointer over a mark in the view to run the action. This option works well for highlight and filter actions within a dashboard.
 - **Select** - click on a mark in the view to run the action. This option works well for all types of actions.
 - **Menu** - right-click a selected mark in the view and then select an option on a the context menu. This option works well for filter and URL actions.



Best Practices : Structuring the dashboard

- ▶ Place the most important view in the top or top left. Your eye is first drawn to upper left corner.
- ▶ Make sure your most important data is on rows or columns
- ▶ Have the views which are meant to be interacted with on the top
- ▶ Use at most three views in a dashboard
- ▶ Make sure that the legends and filters grouped or placed in an intuitive place. They are best viewed if positioned on the bottom or right hand side (Hint: top left is not usually ideal)
- ▶ Avoid using multiple color schemes per dashboard. If you have to, then Multiple color schemes could be Purple-gray, Blue-red and Green-orange
- ▶ In case of scatter, your most important data should be on rows or columns. Less important data should be in color, size, shape etc.
- ▶ Always switch on color legend highlighting, so when someone clicks on a value in legend, the graph gets highlighted

Best Practices : Formatting

- ▶ Axes/labels should have proper understandable names. (NOT “Value” or Sum(Apples))
- ▶ Axes/labels should have a reasonable number of decimals (NOT 1.000000)
- ▶ Make sure the mark labels make sense without an explanation
- ▶ The Quick Filters is preferable with titles as commands (“Select a Team:”)
- ▶ All sheets and documents must be named appropriately (NOT Sheet1, etc)
- ▶ The title color should not clash with your color scheme used in the visualizations
- ▶ As a general rule: axes and labels should be dark grey

A close-up, green-tinted image of a person's face wearing glasses. The text "what about specifics?" is overlaid on the image. The word "what" is in red, "about" is in red, "specifics" is in yellow-green, and the question mark is in red.

**what about
specifics?**