INTRODUCTION TO TABLEAU

Authored by Amy D. Such

What is Tableau

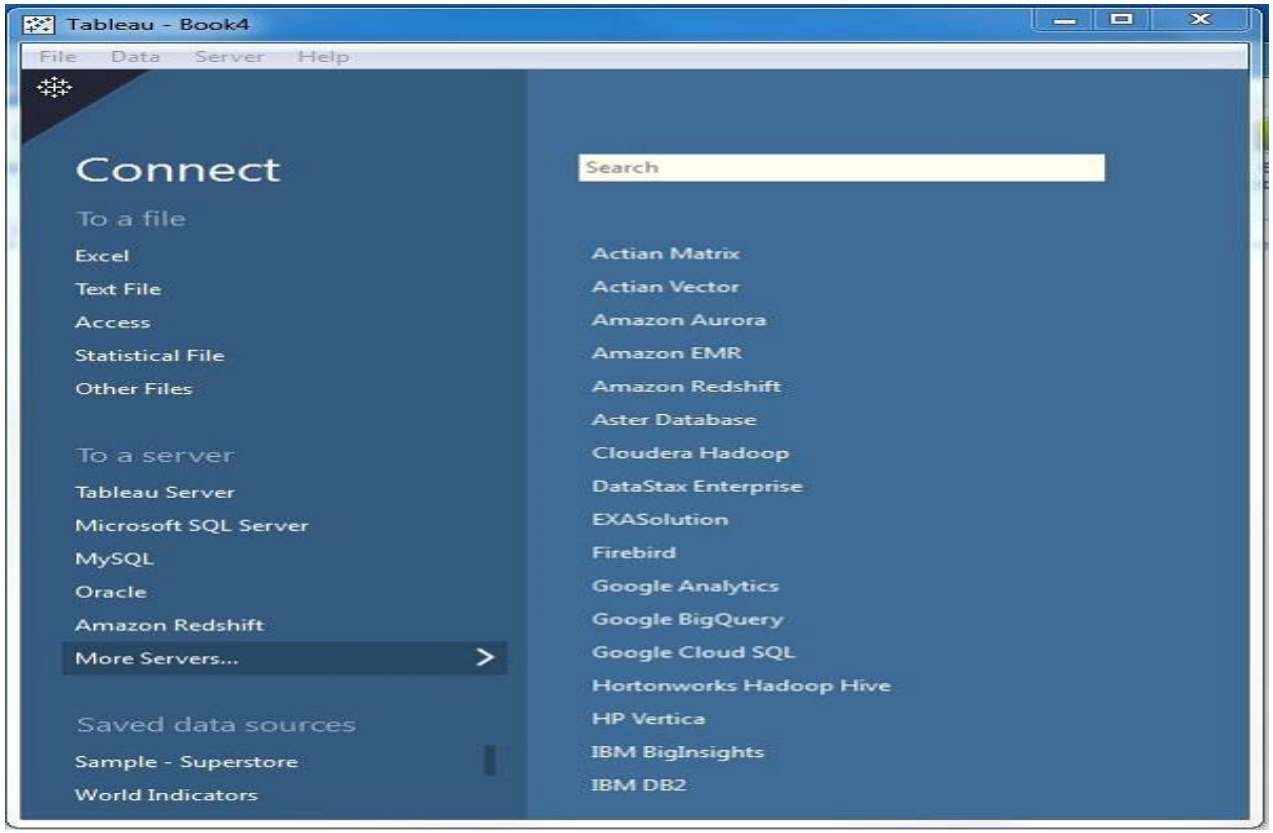
Tableau is a data visualization tool that lets us analyze virtually any type of structured data and produce highly interactive and attractive graphs, dashboards, and reports in minutes.

Using Tableau Effectively

Tableau makes it easier to create powerful, visual information that communicates what is important better than a spreadsheet or text table. Tableau has advanced capabilities for more technical users, but it dramatically lowers the bar for creating dashboards and performing analytical analysis for non-technical analysts and information consumers." – Dan Murray, Interworks

DATA SOURCES

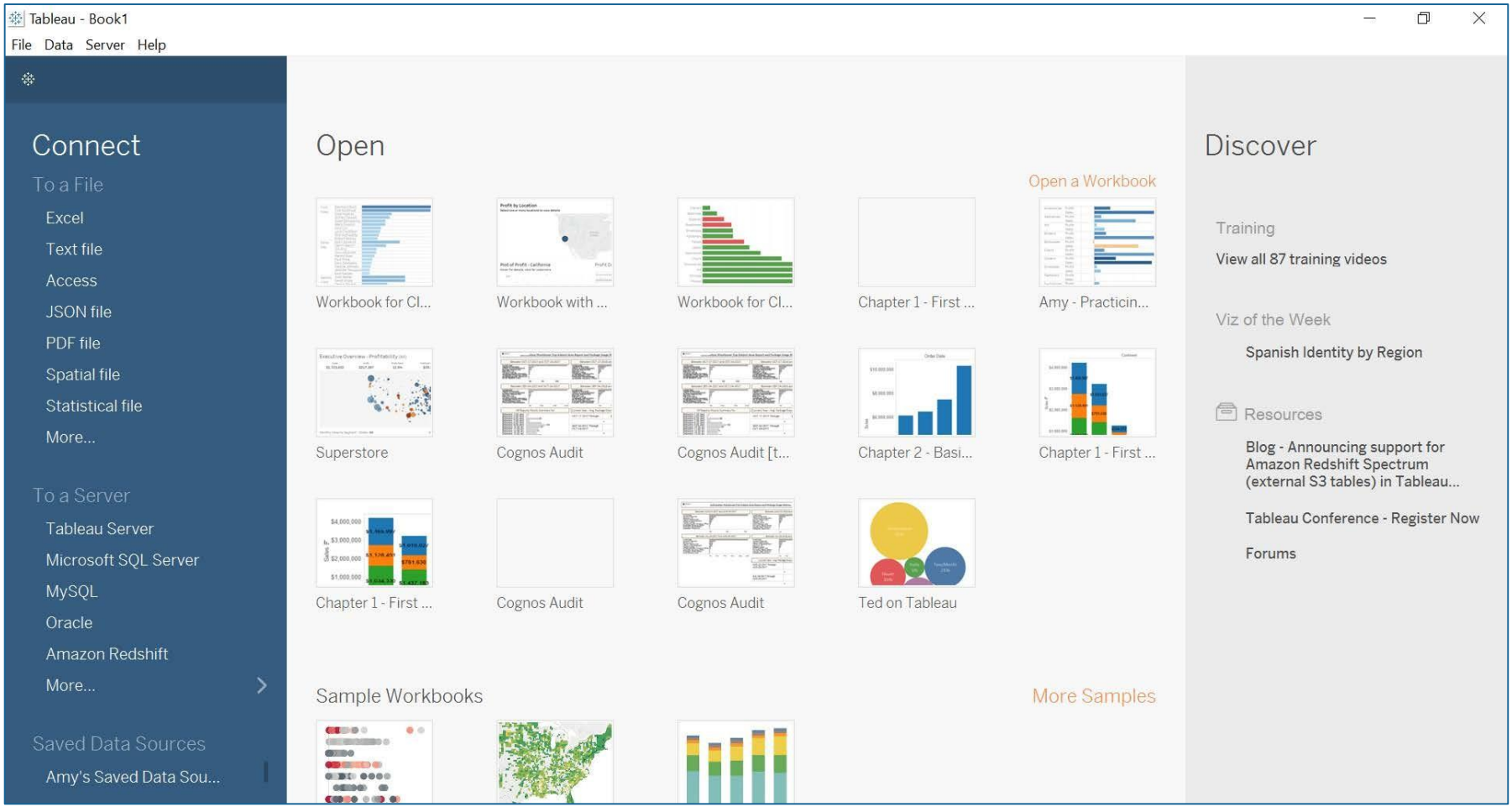
Tableau can connect to all the popular data sources which are widely used.
Tableau's native connectors can connect to the following types of data sources.
File Systems such as CSV, Excel, etc.
Relational Systems such as Oracle, Sql Server, DB2, etc.
Cloud Systems such as Windows Azure, Google BigQuery, etc.
Other Sources using ODBC
The following picture shows most of the data sources available through Tableau's native data connectors.



GETTING YOUR DATA READY

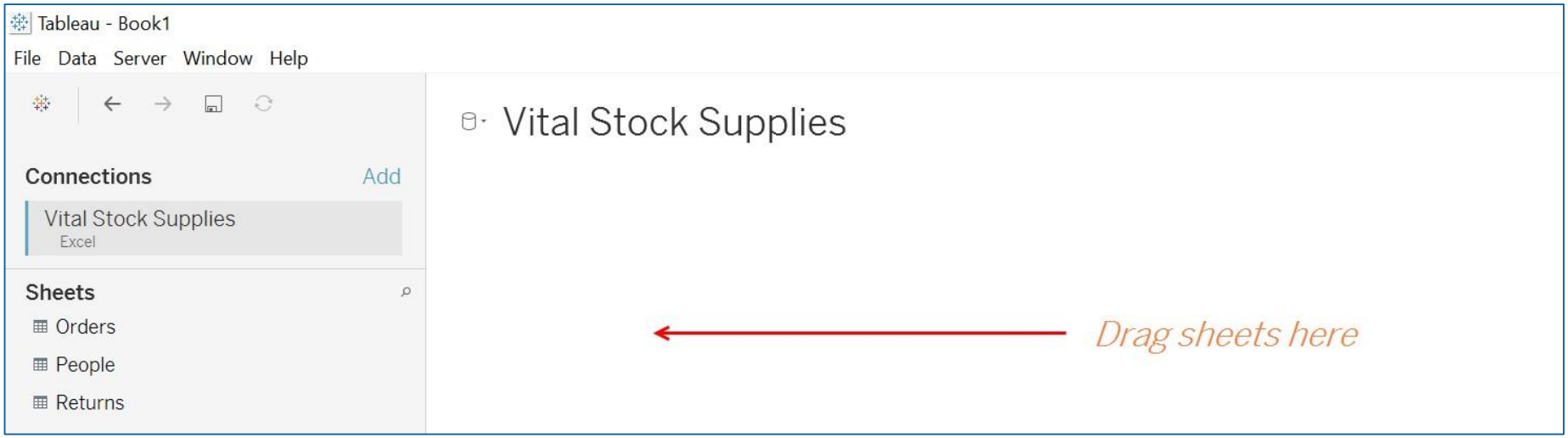
Choose a connection or an existing workbook from the Tableau Welcome Page.

Note: Upon installing Tableau, check your local drive for **My Tableau Repository**



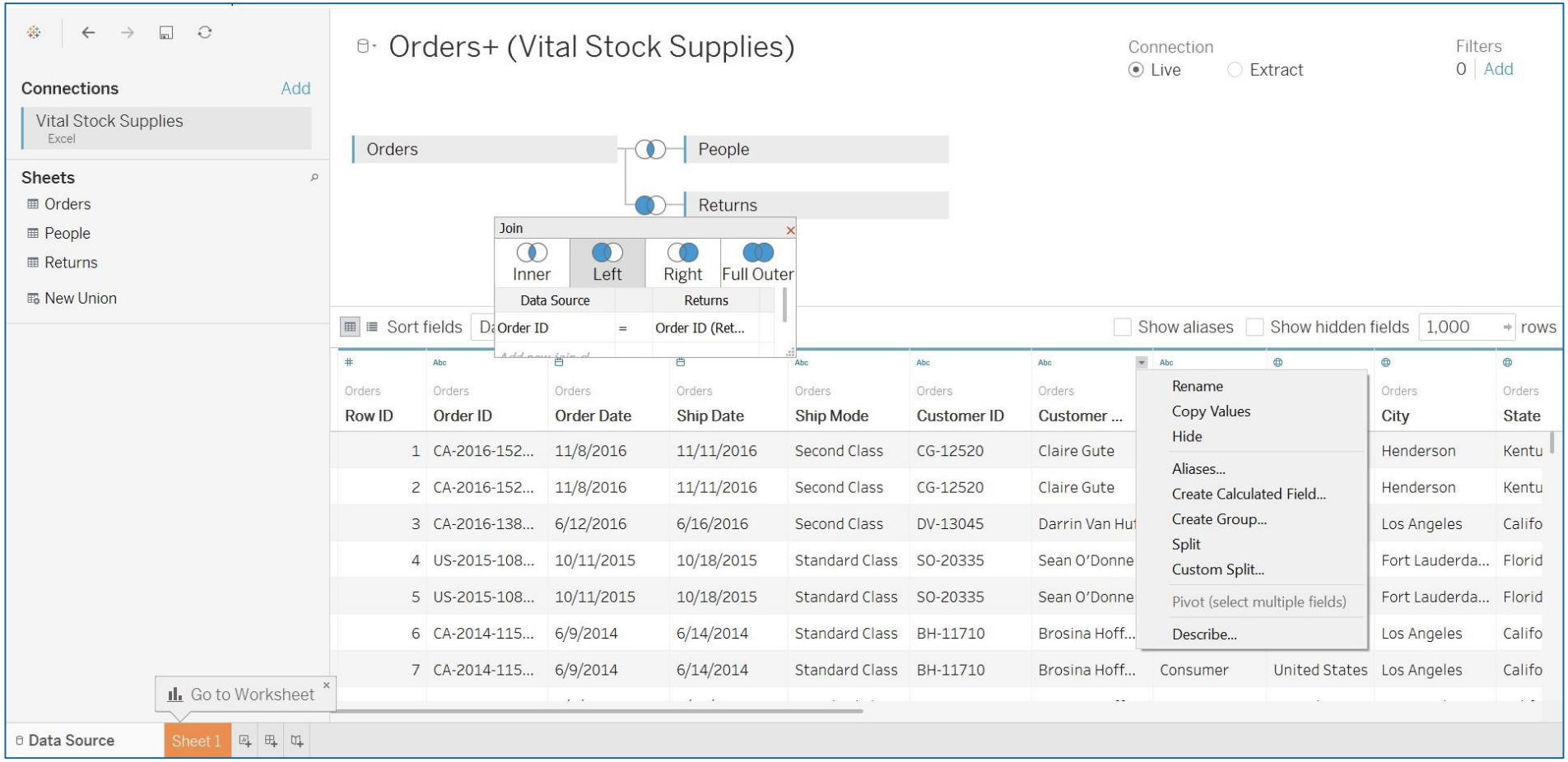
GETTING YOUR DATA READY

Once you choose your data source, you (and it) are brought to the Data Source Page where you c an format your metadata .



Sheets on the left navigation pane behave like tables in a database .

GETTING YOUR DATA READY

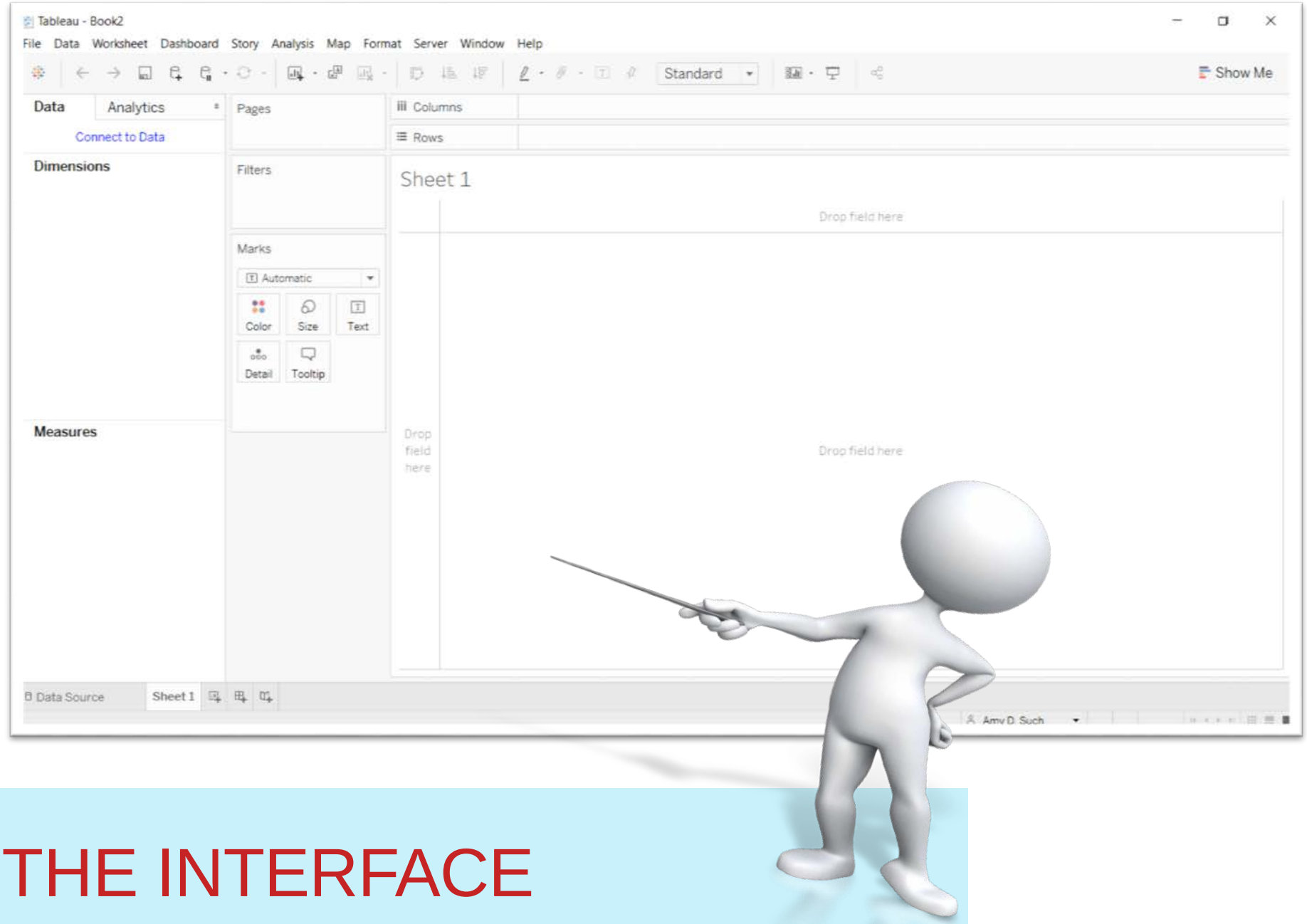


GETTING YOUR DATA READY

- Splitting
  - String fields can be split into multiple fields for easier analysis
  - Automatic or custom split options
- Aliasing
  - Roles (time, ship date/order date)
  - Binning (high/low sales)
- Renaming
- Data types
- Geographic roles
- Calculated fields
- Pivoting
- Data interpreter
  - Helps clean up data

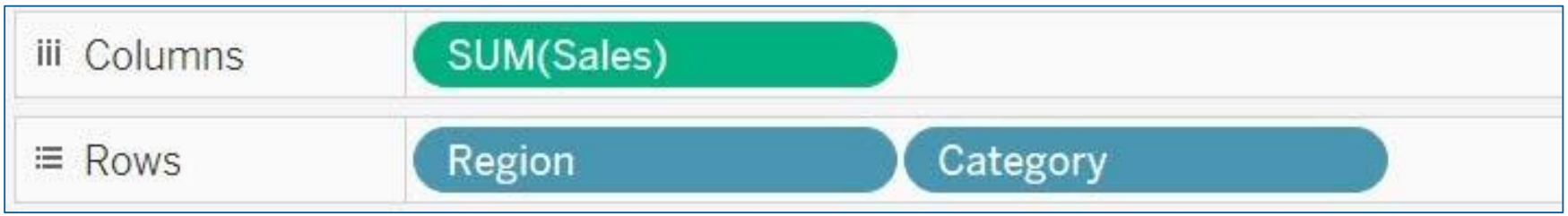
Data type icons in Tableau

Icon	Data type
Abc	Text (string) values
	Date values
	Date & Time values
#	Numerical values
T|F	Boolean values (relational only)
	Geographic values (used with maps)



THE INTERFACE

- We work in workbooks which contain worksheets, dashboards, and stories
- Worksheets are also known as **Views** of your data
- We drag and drop fields from the data source onto **Shelves**
  - Rows, Columns, Filters…
- Items that appear on shelves are called **Pills**
  - If the Pill is **blue**, it is a Dimension
  - If the Pill is **green**, it is a Measure



When you bring your data into Tableau, it automatically places the fields into one of two categories. Dimensions or Measures.

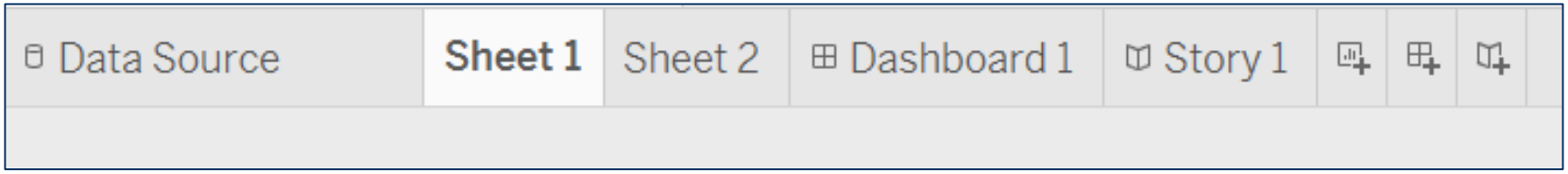
- **Dimensions** are fields that organize your data into categories (or buckets).
  - Individual dimension values are called **Members**
  - Usually Non-numerical and provide critical, contextual meaning to a **Measure**.
    - Who, what, when, where, why

- **Measures** return numeric values for "measuring" different dimensions.
  - Usually aggregated (sum, average, min, max, etc.)
  - Things you can do math on that are generally meaningless without context

Revenue grew $22m during Q2 2017 in the Eastern Region for Technology !

- Change between data source, worksheets, and dashboards using the tabs at the bottom

- You can also add new worksheets, dashboards, or story boards using the “new” buttons.

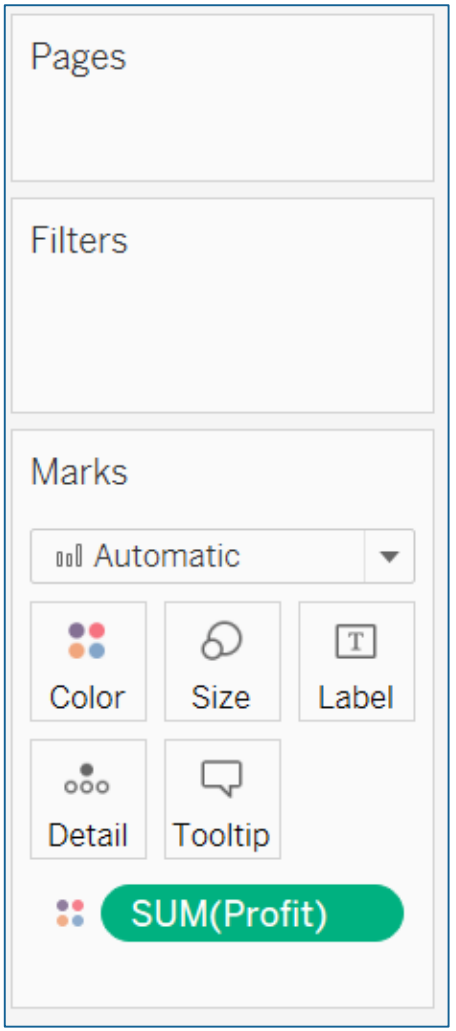


Once you've connected to data and evaluated available dimensions and measures in the Data pane, you'll need to decide where to drag desired dimensions and measures. You may choose to drag directly onto the visualization or onto a particular shelf or card.

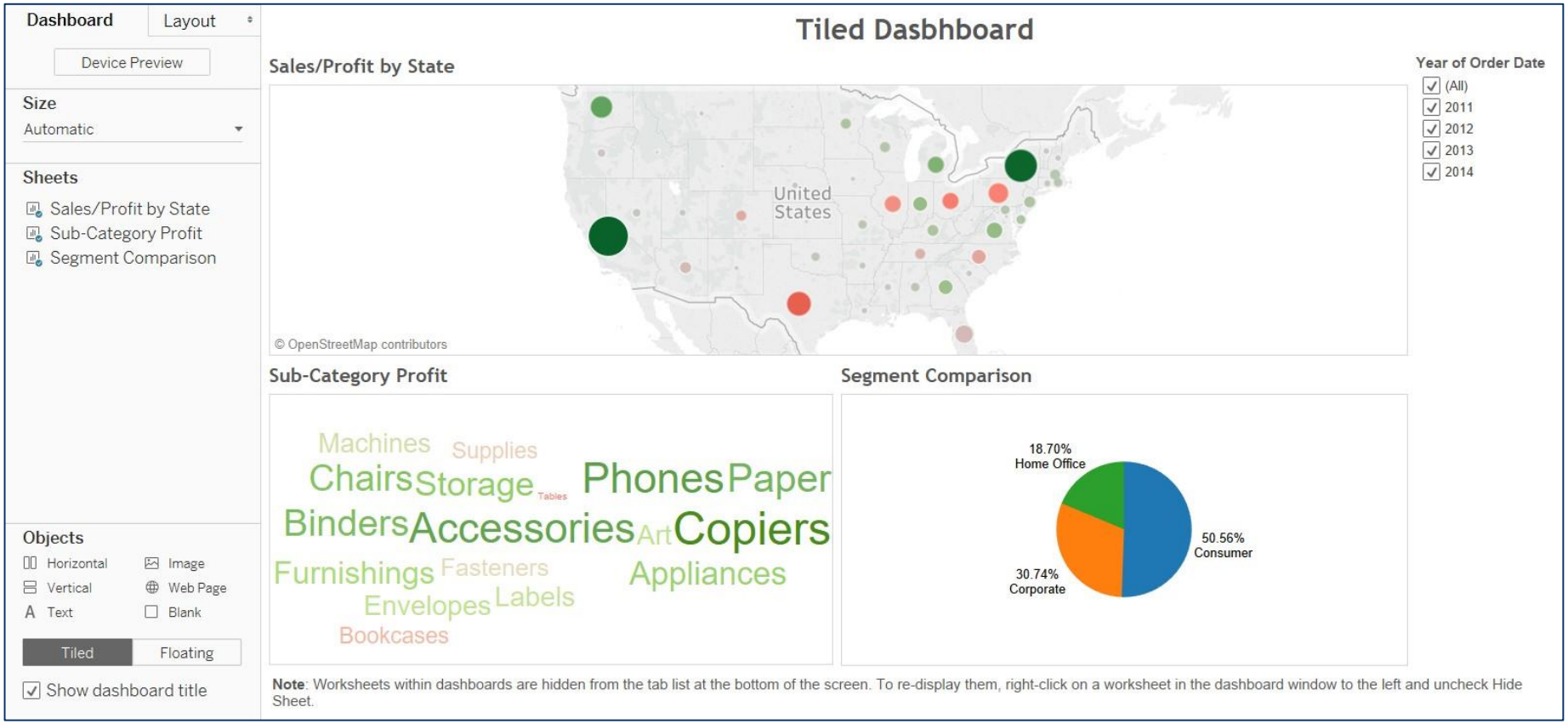
- The most common shelves are the Columns and Rows. Dragging dimensions and measures to these determines the layout of your visualization .

Use cards to navigate pages, filters, and marks on a worksheet .

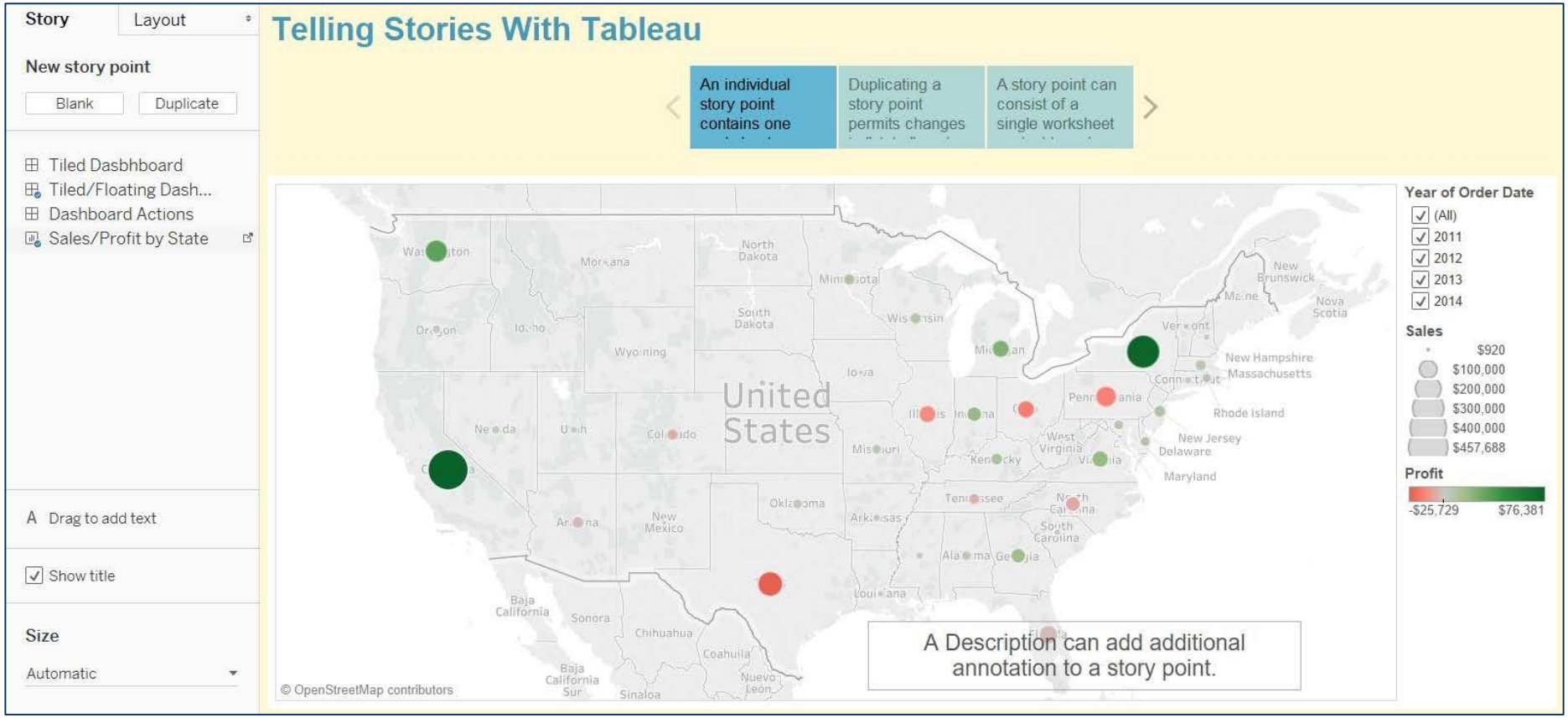
- The **Page** shelf acts as a modified Filters shelf allowing you to page through values .
- The **Filters** shelf allows you to narrow data down.
- The **Marks** card allows you to customize the display of the 'marks' that make up your chart.



Use dashboards to organize and interact with multiple worksheets



Use stories to combine elements and create directed, interactive analysis

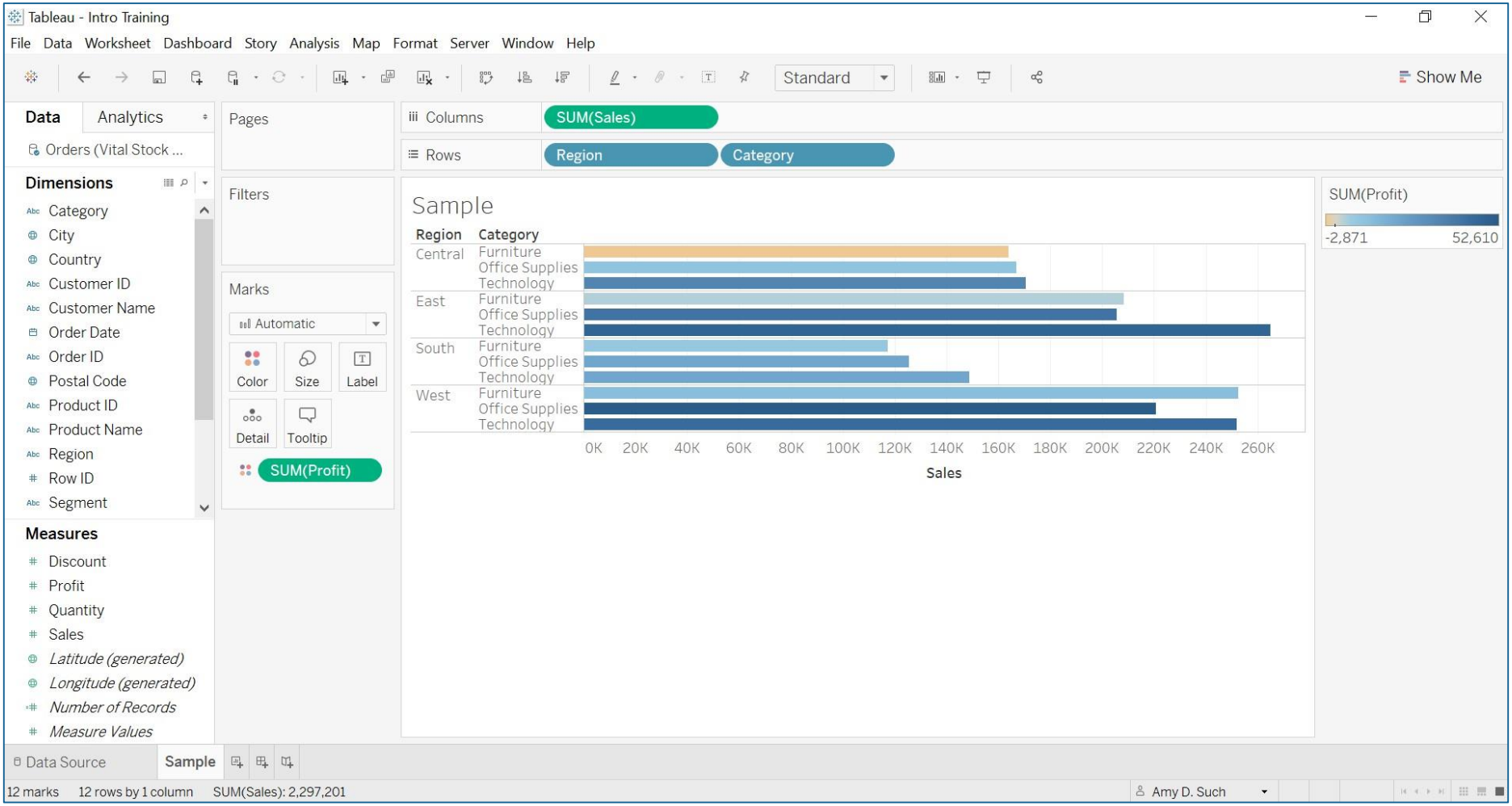


Use **Show Me** to select or change between different visualizations.

- Anytime you're editing a new or existing worksheet, you'll notice the Show Me tab at the upper right of the screen.
- Show Me displays a series of thumbnail images representing the different types of charts you can create with just a few clicks.
- If you've already created a chart but would like to switch to another, clicking a new thumbnail from the Show Me dialog box will change your current visualization.

Note: Only thumbnails for visualizations that are applicable to the dimensions and measures you've selected from your data pane will be active.

Add Dimensions and/or Measures from the Data pane to the Shelves and Cards to create visualizations.



BONUS: AUTOMATIC VIEWS

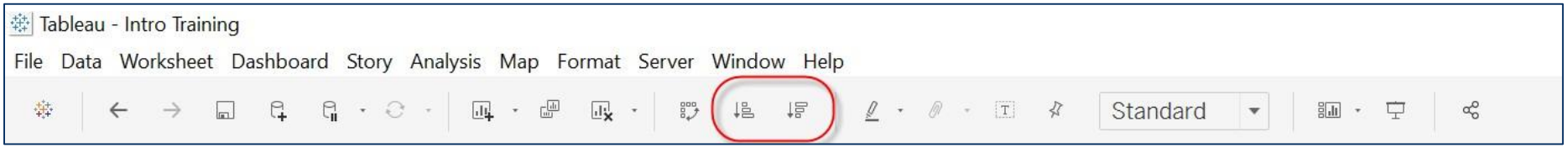
Tableau supports an additional method for automatically generating views of data called Automatic Double-Click. To use this method, double-click fields in the Data pane that you'd like to see in your visualization. Each double-click results in a "best guess" for how you would like the field to be presented in your view.

Note: The order in which you double-click a measure or a dimension is significant.

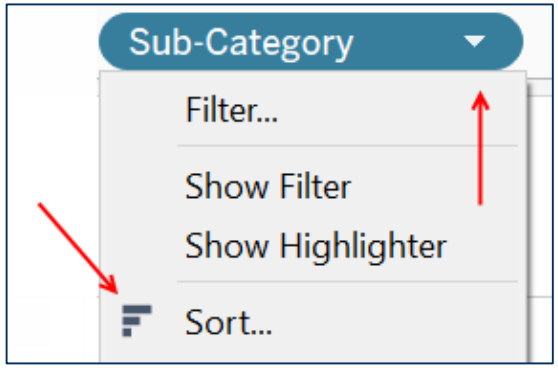
Text Table	Adding a dimension first produces a text table (or cross-tab). All subsequent clicks on fields result in refinement of the text table.
Bars	Adding a measure first and then a dimension produces a bar view. All subsequent clicks result in refinement of the bar view, unless a date dimension is added, at which time the view is changed to a line.
Line	Adding a measure and then a date dimension produces a line view. All subsequent clicks result in refinement of the line view.
Continuous Line	Adding a continuous dimension and then a measure produces a continuous line view. Subsequent dimensions result in refinement of the continuous line view. Subsequent measures add quantitative axes to the view.
Scatter	Adding a measure and then another measure produces a scatter view. Subsequent dimensions result in refinement to the scatter view. Subsequent measures will create a scatter matrix.
Maps	Adding a geographic field produces a map view with latitude and longitude as axes and the geographic field on the Level of Detail shelf. Subsequent dimensions add rows to the view while subsequent measures further refine the map by adding size and color encoding.

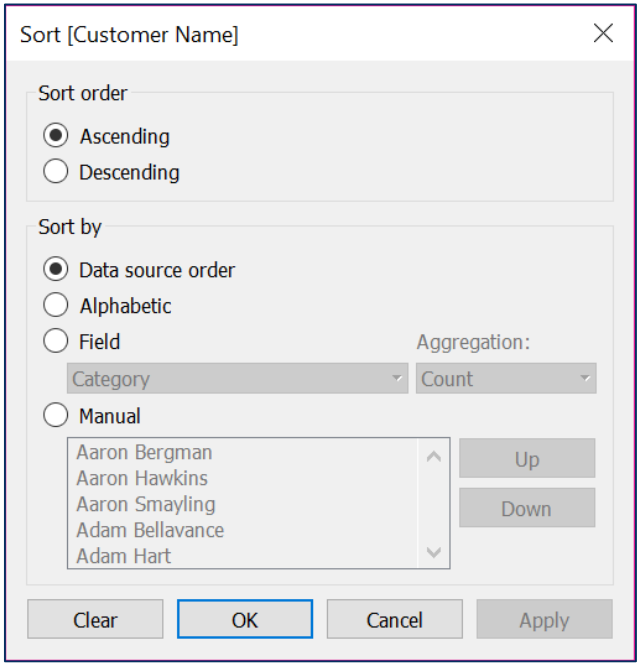
SORTING & FILTERING

- Analysis can be enhanced by sorting results
- Sorting can be done using buttons for simple sorting



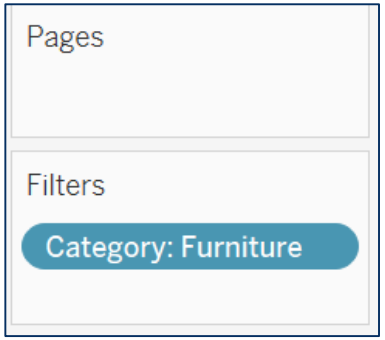
- Or … manually through the advanced sorting dialog which is achieved by right - clicking on a Pill from the Shelf and clicking **Sort**.



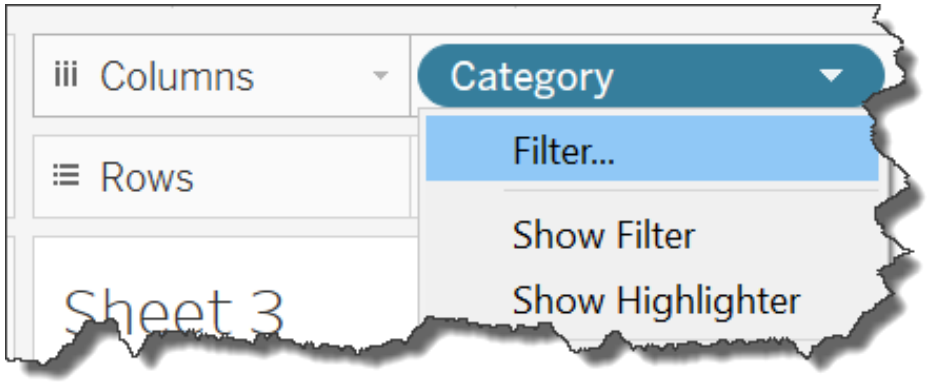


SORTING & FILTERING

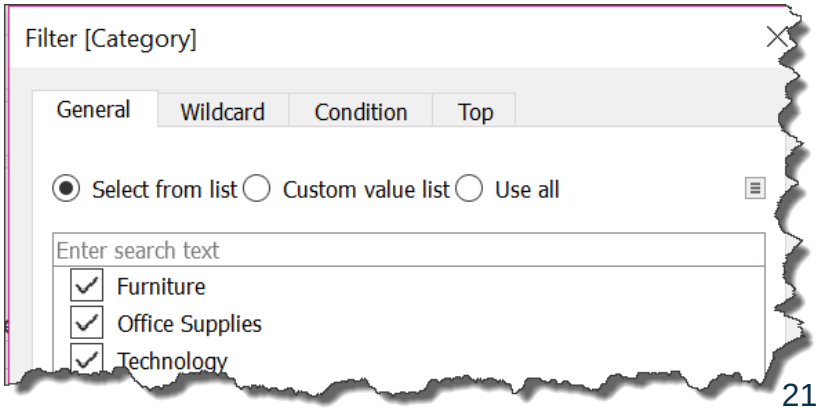
- Filtering results allows you to focus on particular aspects of your data
- Filter by dragging data fields from the Data pane to the Filters card



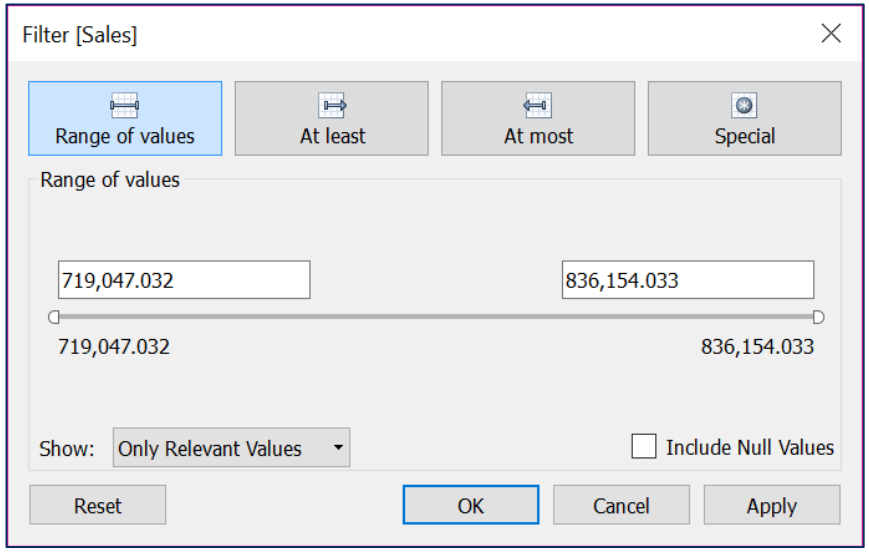
- Filtering can also be created by right-clicking a Pill on a Shelf



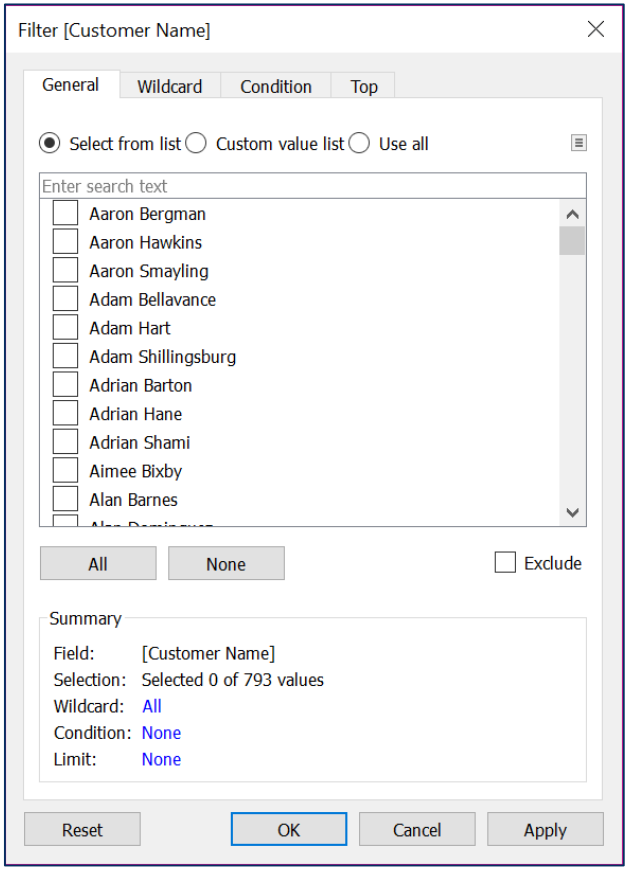
- In either case, a Filter dialog box will reflect the data type chosen and prompt you for additional criteria



Numeric Filter Dialog Box

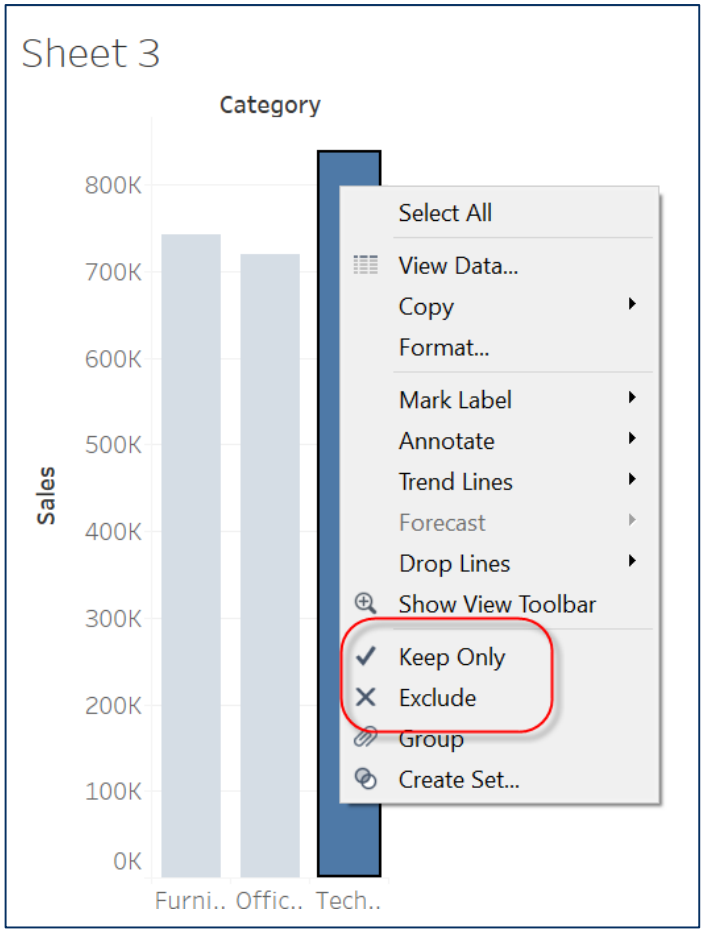


Text Filter Dialog Box



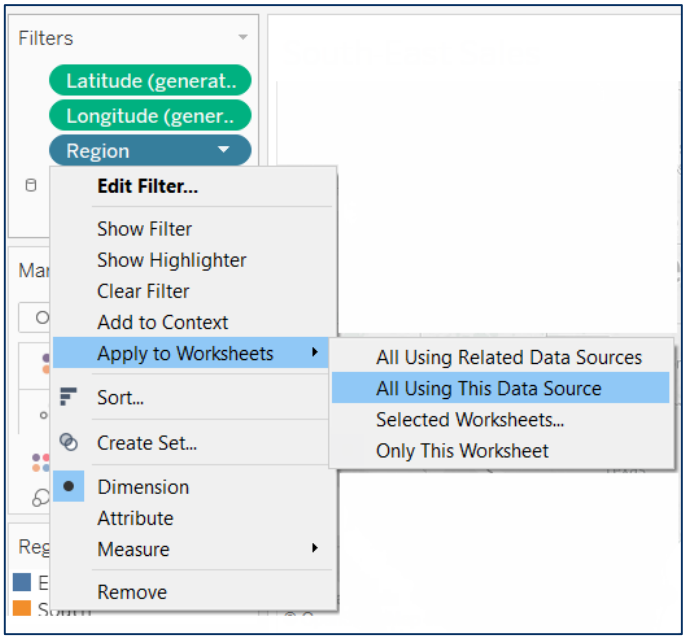
Filtering Data

- Simple filtering can be done by selecting the headers or marks in a View and choosing the **Keep Only** or **Exclude** options
- The Dimension members are removed from the View and the filtered fields are added to the Filters shelf
- When selecting headers that are part of a hierarchy all of the next level members are selected automatically
- Individual marks can be selected on a View, making it easy to focus on or remove outliers



Apply Filters to Multiple Worksheets

- When you had a filter to a worksheet, be default it applies to the current worksheet.
- Sometimes, however, you might want to apply the filter to other worksheets in the workbook.
- You can select specific worksheets to apply the filter to or apply it globally to all worksheets that use the same data source or related data sources.



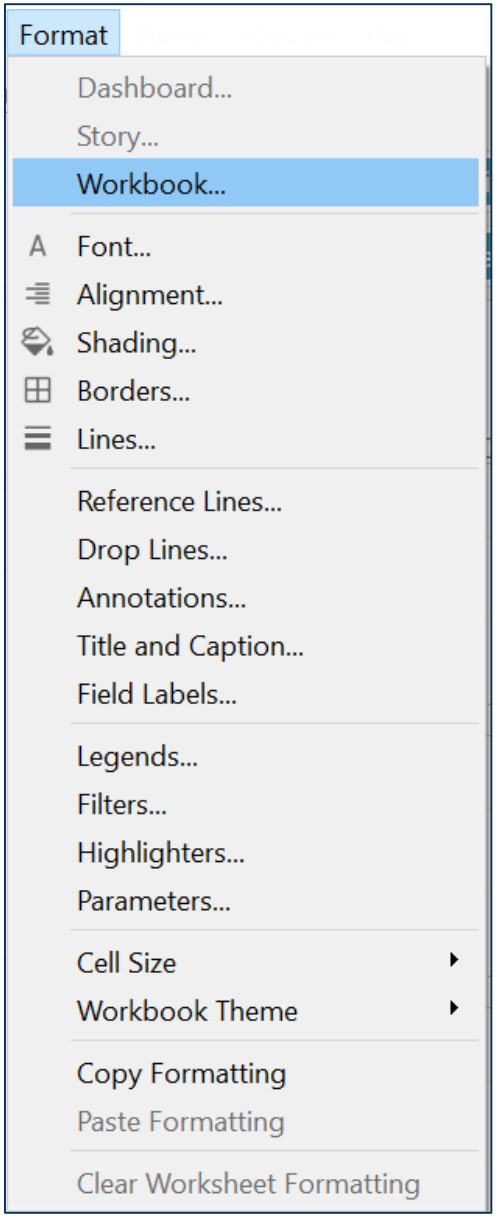
On the Filters shelf, right -click the field and select: **Apply to Worksheets > All Using Related Data Sources**

- Filters that use this option are global across the workbook.

FORMATTING

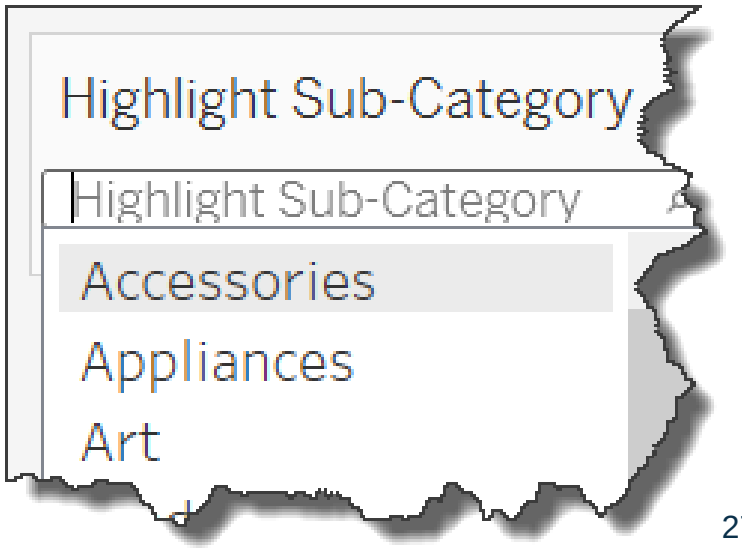
FORMATTING

- Tableau allows for extensive formatting of your views and can be done for an entire worksheet or down to the specific/individual parts.
- There are several ways to initiate formatting in Tableau. You may choose options from the Format drop-down menu or right-click any element and choose **Format** from the context menu.
- Formatting can include changing some or all of the following:
  - Fonts
  - Alignment
  - Shading
  - Borders
  - Lines



Highlighters

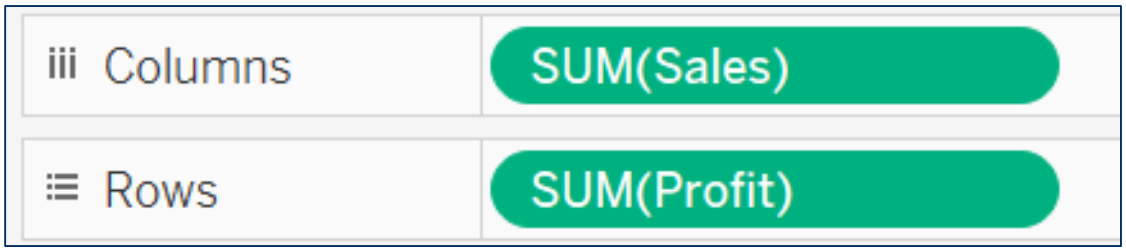
- Highlighters are used to quickly find and highlight data on a visualization
- These come in handy when adding a mark (like color) to a visualization makes it too confusing or hard to read
- Used when there's already a color mark assigned to the visualization
- Previously done through legends and dashbo ard actions
- Can be used with reference bands and lines
- Highlighter legend supports text searches and dropdown
- Multiple highlighters can be added to a single visualization



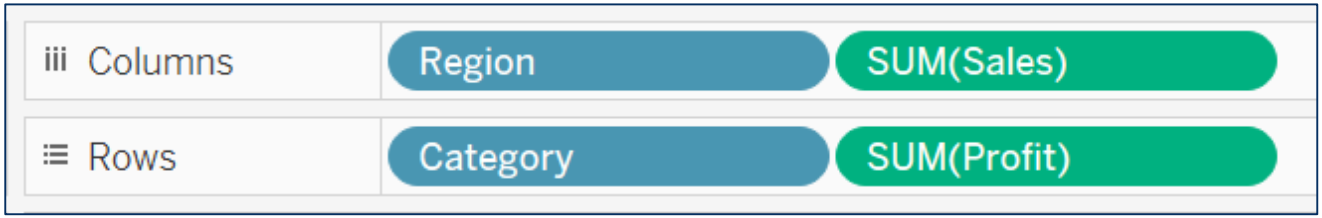
Scatter Plots

- Scatter Plots provide a way to show correlations between numerical values
- Scatter plots require at least two measures
  - One on each of the columns and rows

Creates a simple Scatter Plot



Creates Matrix of Scatter Plots



DUPLICATE AS CROSSTAB

BONUS

A crosstab, sometimes referred to as a Pivot Table, is a table that summarizes data in rows and columns of text .

Sometimes, even in Tableau, you just want to see the data and not a visualization.

You can easily convert your visualization into a crosstab.

Objective: Display the numbers associated with a data view.

1. Right-click any of the visualizations we've created thus far.
2. Select **Duplicate as Crosstab** .

A new sheet will automatically be created beside your visualization, displaying the same information as data in rows and columns of text.

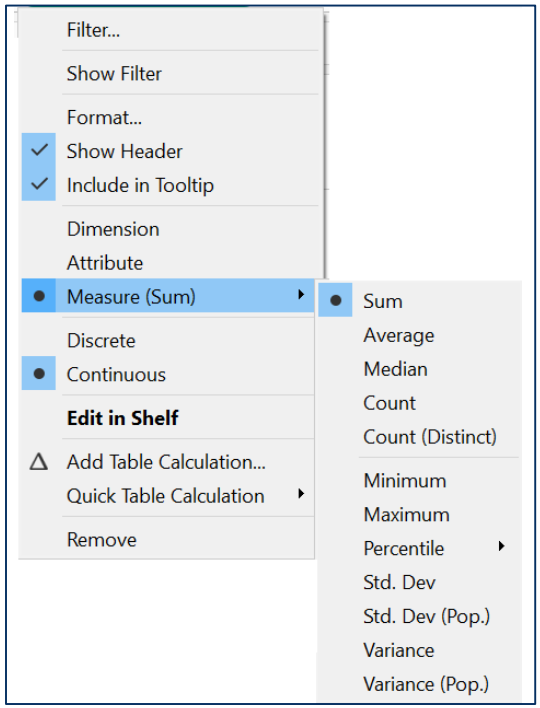
BASIC CALCULATIONS

Basic Calculations

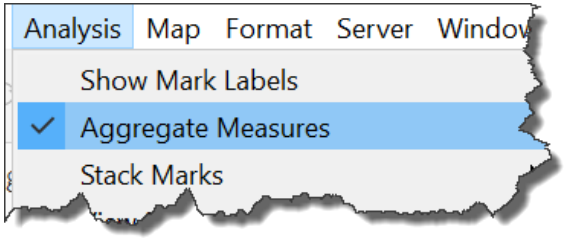
- Calculations are created to enhance the dashboard.
- Tableau takes granular data and aggregates them in the View.
- Calculations help in extending the usage of aggregations.
- If we're looking for a particular dimension or measure that does not seem to exist in our data source, we can form them from calculations.

Aggregating Data

- When a **Measure** is placed on a Shelf, Tableau automatically aggregates the data.
  - Tableau defaults to Sum.
- You can change the aggregation to other types (Average, Max, Min) by clicking on the Pill, navigating to **Measure (Sum)** and selecting from the list of other choices.

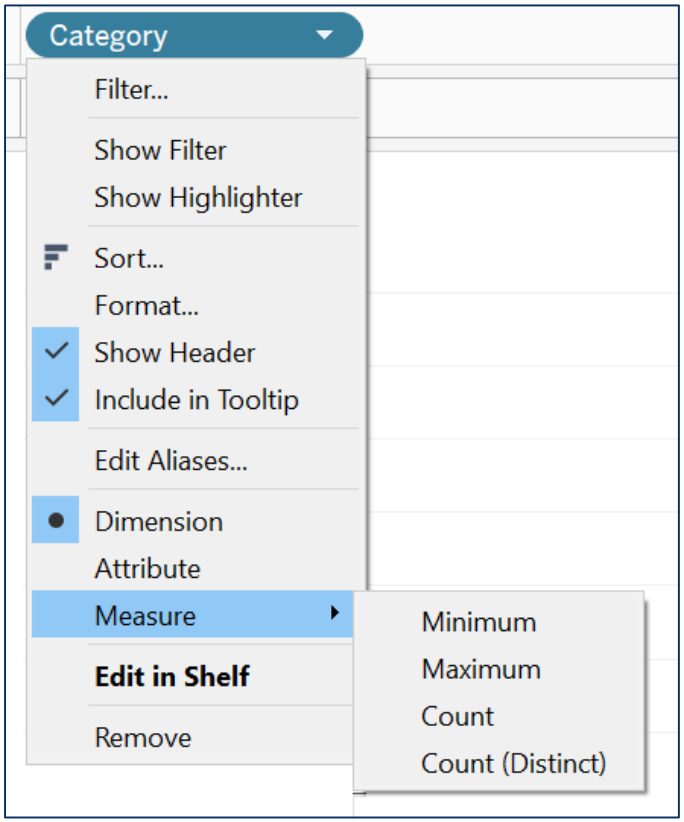


Note: You can disaggregate data by navigating to the **Analysis Menu** and deselecting **Aggregate Measures**.



Aggregating Data

- Dimensions can also be aggregated by right-clicking the Pill, selecting Measure, and choosing from the list of options available:



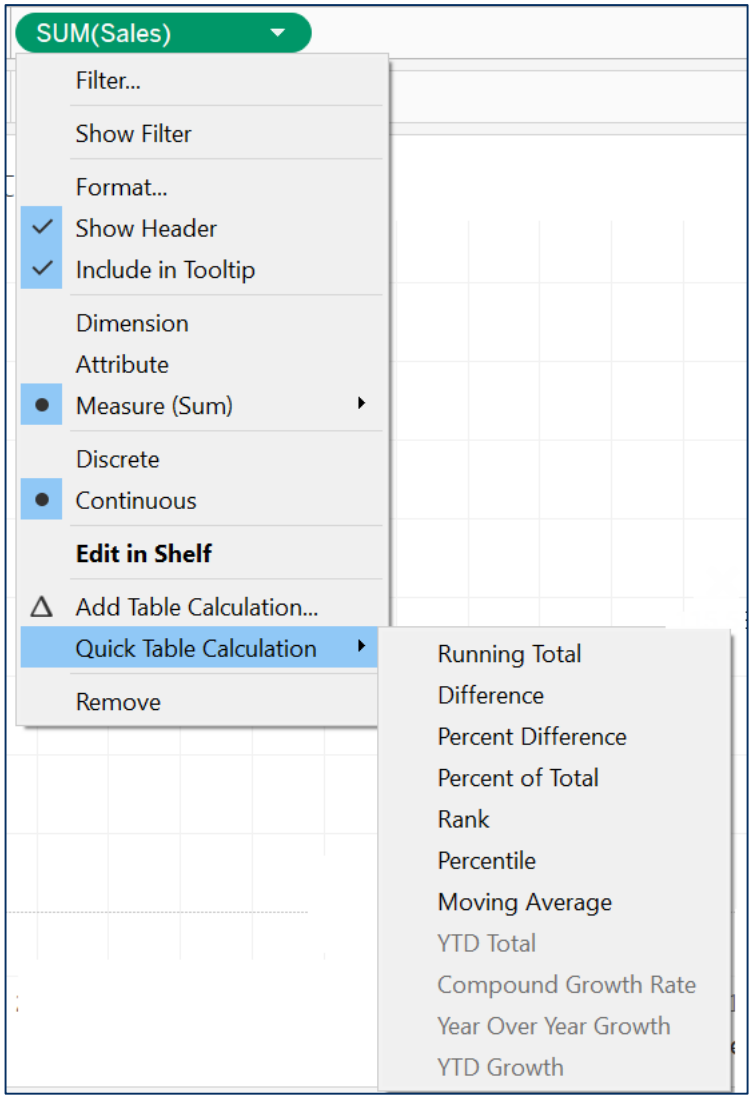
Granularity of Data

- Data granularity refers to the level of detail or depth of data in a View
- Defined by the Dimension fields
- If a Dimension is dragged to Detail, it will change the granularity of the visualization
- Dragging a Measure to Detail will have no effect

Quick Table Calculations

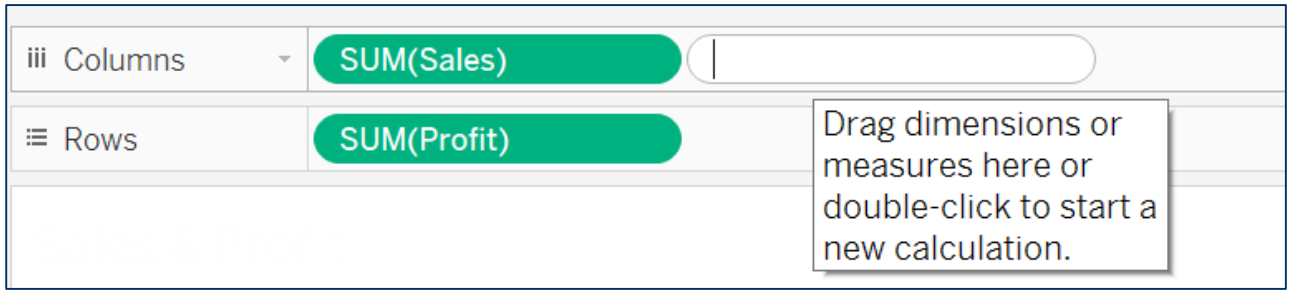
- A quick table calculation is a one-step process where you choose a common, pre-defined c alculation type from a list.
- Tableau automatically applies the most typic al settings for that c alculation type.

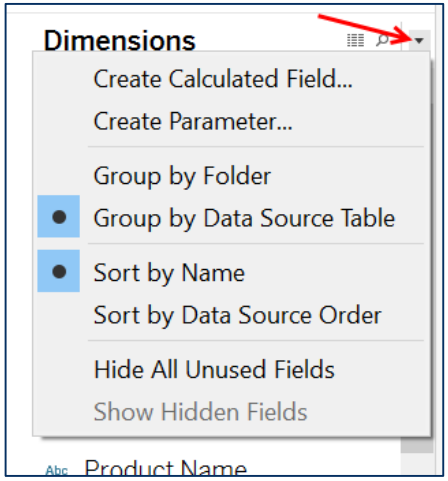
To apply a quick table calculation, simply right -click a measure from a Shelf or Mark and choose **Quick Table Calculation** from the context menu.



Calculated Fields

- When your data source doesn't contain all the fields you need, you can add new ones as calculated fields
- The calculated fields are saved as part of the data source and can be used as often as necessary
- There's several ways to create calculated fields. You can do this easily by using the calculation editor from the drop-down, by double-clicking a Shelf or a Field on a Shelf, or by right-clicking "white space" on the Data pane under Dimensions or Measures
- Calculated fields can be created from existing calculations
- Calculations are based on Formulas and Fields





Formulas Make Up Calculated Fields

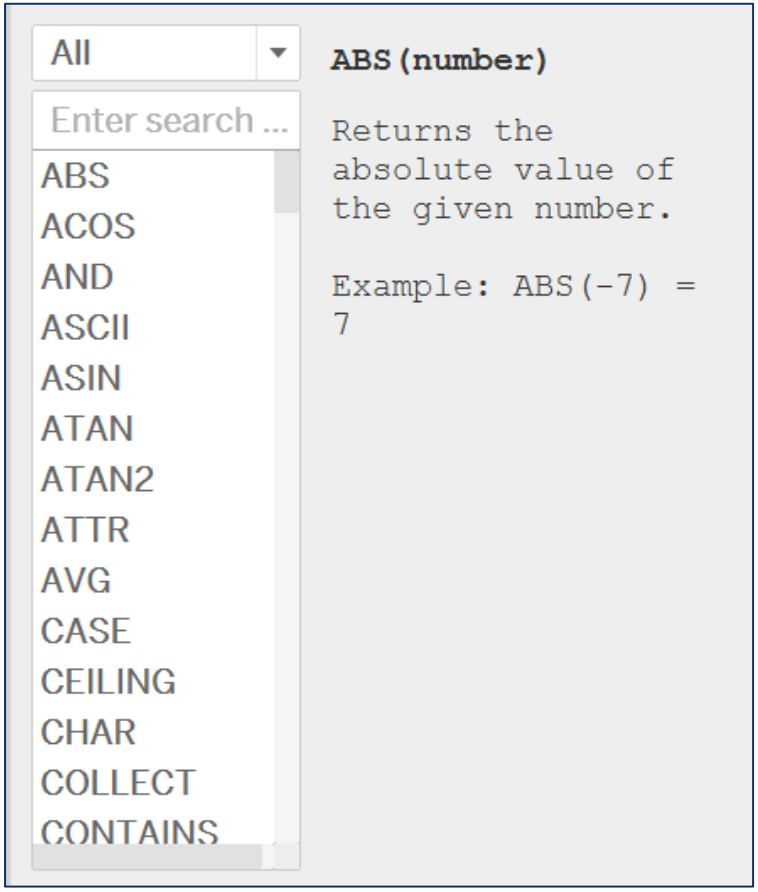
- Formulas can contain functions, fields, operators, parameters, or comments
- The Calculation Editor provides color coding based on the Field added
  - Functions: **Light Blue**
  - Fields: **Orange**
  - Operators: **Black**
  - Parameters: **Purple**
  - Comments: **Grey**

Note: Tableau makes attempts to guess what you're trying to do and provides you with the tools you need for brackets and parenthesis. You should not have to do much typing or guessing.

*Don't "fight" the tool!

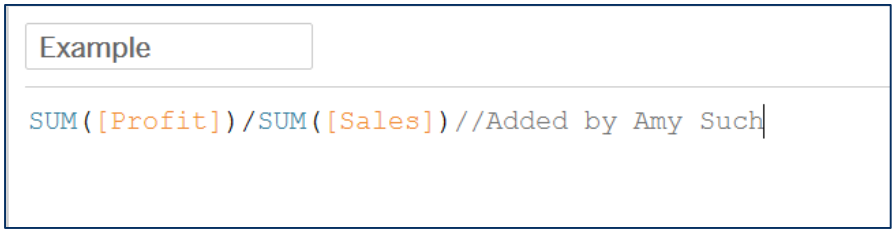
Functions

- Functions can be filtered by category
- Use the **Enter search …** box to find specific functions
- Clicking on a function provides tips and examples to guide you



Comments

- Comments can provide context or detail on calculations
- Define comments using two forward slashes

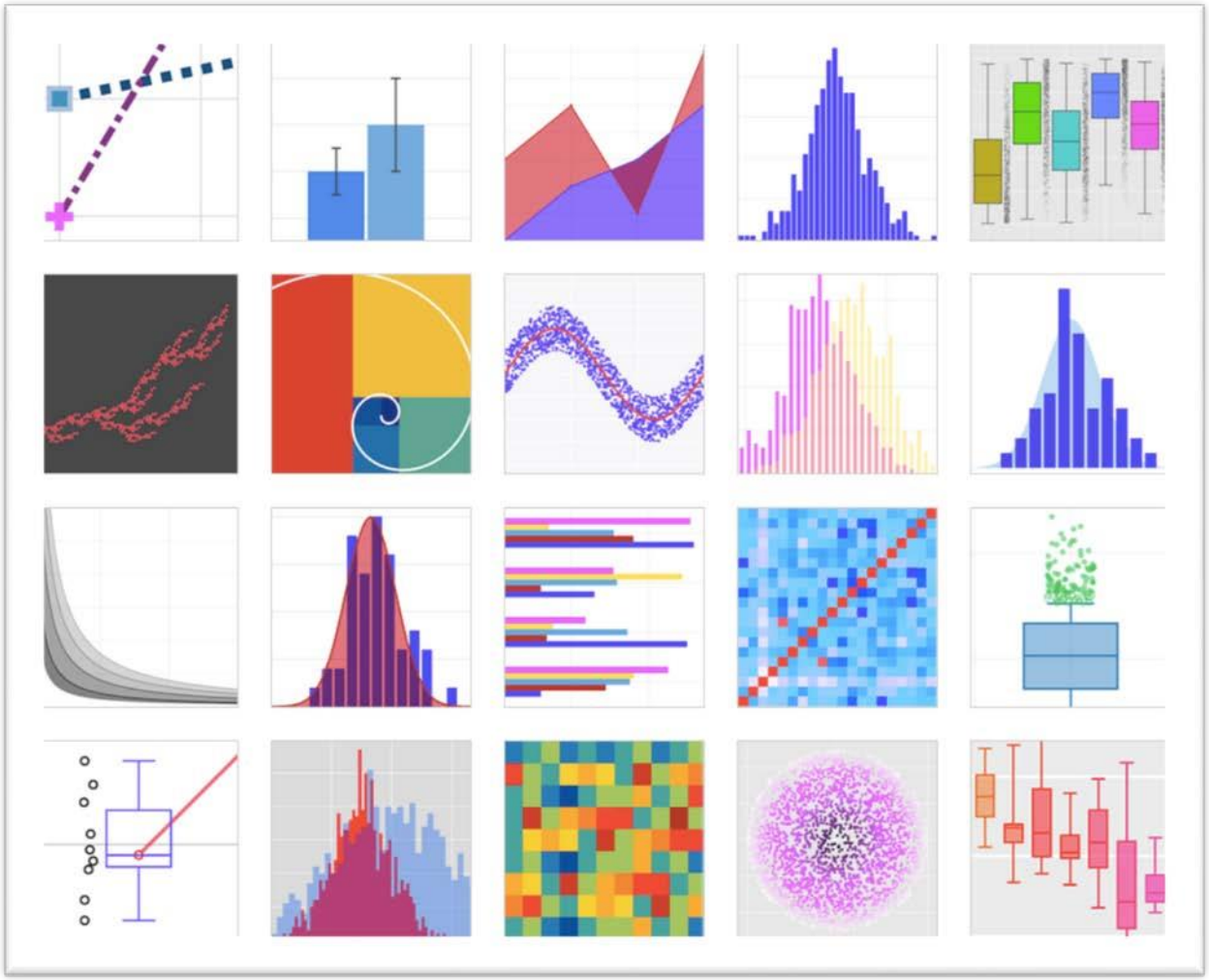


Auto-Completion for Formulas

- Tableau will automatically attempt to fill in a formula
- First attempts to find based on *starts with* then *contains*
- Clicking **Enter** will enter your formula or function into your forming calculation
- If you first type a **[** auto-complete will show fields, parameters, or sets instead of functions
- If you're using multiple data sources, auto-complete will add the fully-qualified name of the Field so you can see exactly where things are being pulled from

Calculations can be used for exception highlighting. In Tableau, this is known as **Spotlighting**.

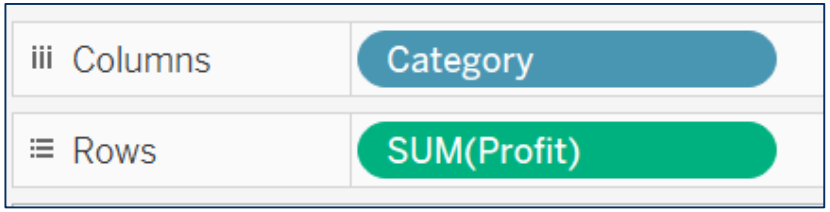
- Spotlighting is based on thresholds set for a selected measure
- It is a technique for showing discrete thresholds based on the values of a measure. For instance, you might want to color-code **Sales** so that those over $10,000 appear green and those below $10,000 appear red.
- A spotlighting calculation is just a special case of a calculation that results in a discrete measure.
  - A discrete measure is a calculation that is a dependent variable (and therefore a measure), but which results in a discrete result (as opposed to a continuous result). Thus the name discrete measure.

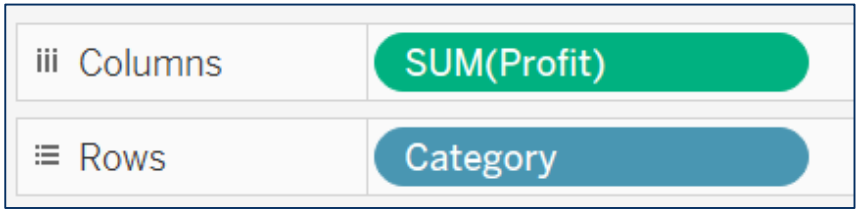


POPULAR VISUALIZATION TYPES

Bar Charts

- Bar Charts are excellent to compare data a cross categories
- To create a bar chart, add a dimension to the Rows shelf and a measure to the Columns shelf or vice versa
- Adding the dimension (blue) to the Columns shelf and the measure (green) to the Rows shelf will result in a Vertical Bar Chart.

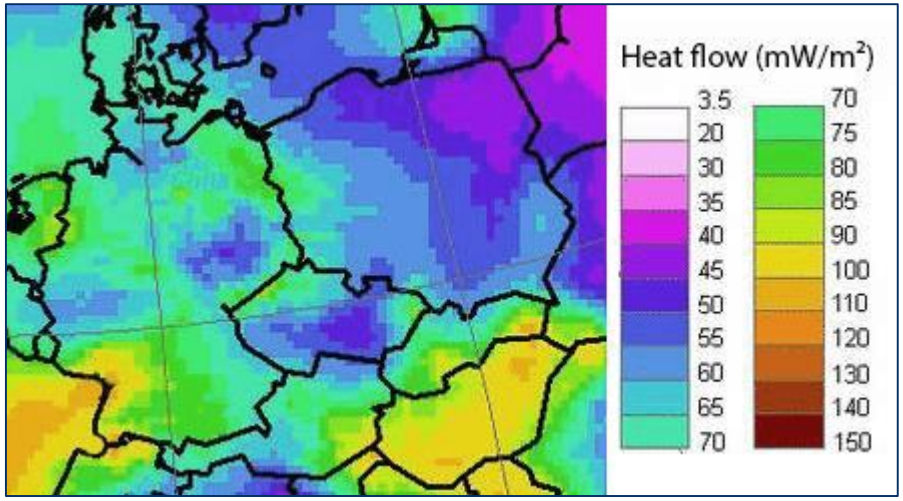




- Adding the measure (green) to the Columns shelf and the dimension (blue) to the Rows shelf will result in a Horizontal Bar Chart.

Heat Maps

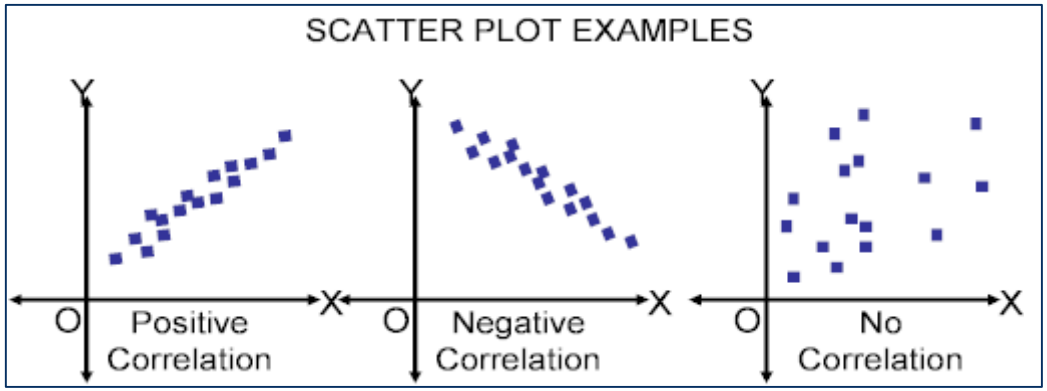
- In a Heat map, data is represented in terms of colors. They provide a quick visual summary of the data.
- Heat maps are best used for showing geographical representation of data where each individual value are shown as color.
- Heat maps enable you to compare data in various categories using different colors.



Example: Heat Map

Scatter Plots

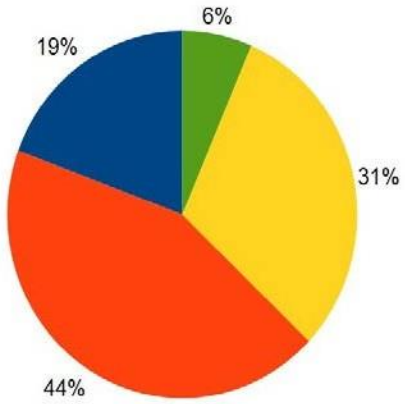
- Scatter Plots can provide a very meaningful visualization of two related numeric measures
- Two different analyses that scatter plots help consider:
  - Comparison/Correlation of the two measures
  - Concentration of data/existence of outliers
- Effective scatter plots include either a small enough number of marks following a general trend to draw quick conclusions about the individual dimensions or a large concentration of marks to draw general conclusions



Example: Scatter Plot

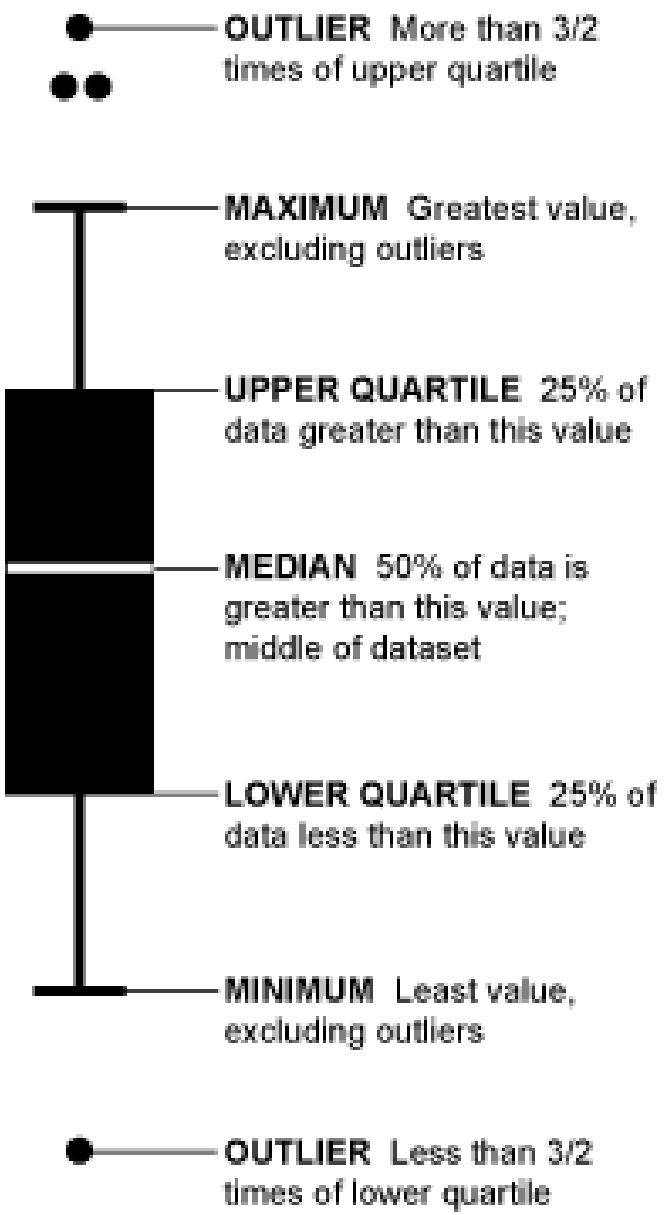
Pie Charts

- Do the parts make up a meaningful whole?
  - If not, use a different chart. Only use a pie chart if you can define the entire set in a way that makes sense to the viewer.
- Are the parts mutually exclusive?
  - If there is overlap between the parts, use a different chart.
- Do you want to compare the parts to each other or the parts to the whole?
  - If the main purpose is to compare between the parts, use a different chart. The main purpose of the pie chart is to show part-whole relationships.
- How many parts do you have?
  - If there are more than five to seven, use a different chart. Pie charts with lots of slices (or slices of very different size) are hard to read.



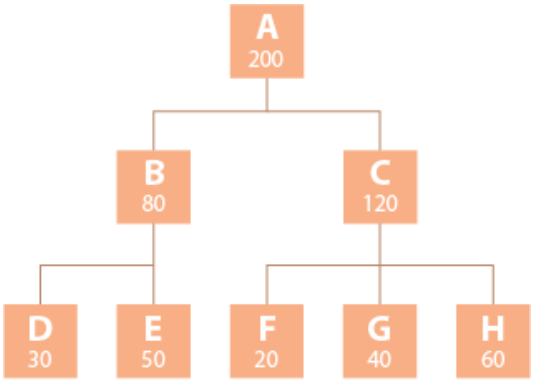
Box Plot

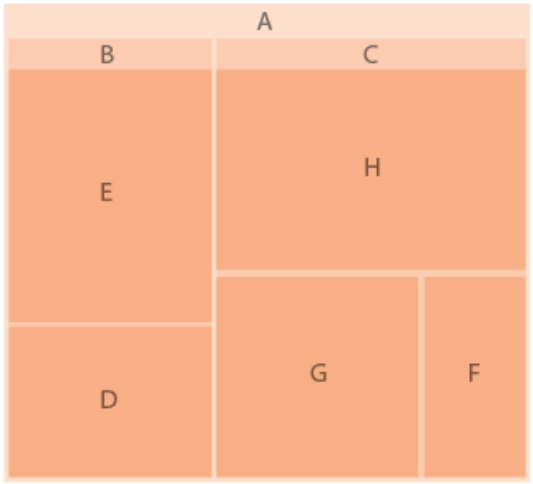
- Also known as Box and Whisker Plot
- Used to show the distribution of data



Tree Maps

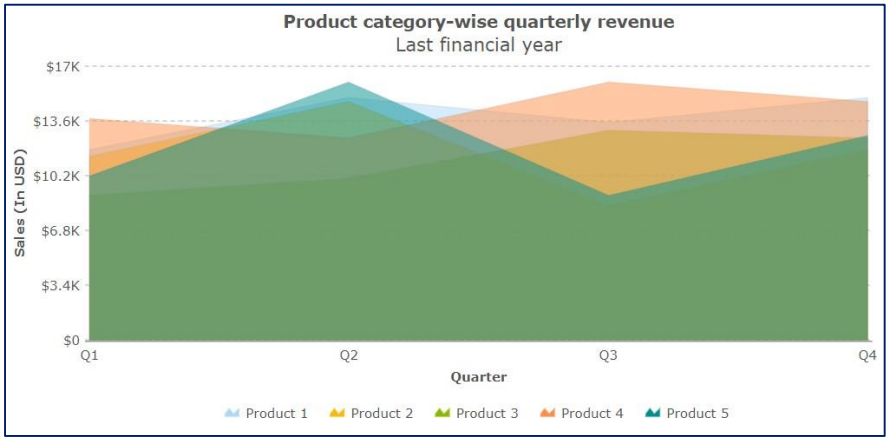
- Treemaps are an alternative way of visualizing the hierarchic al structure of a diagram while also displaying quantities for ea ch category via area size
- Ea ch category is assigned a rectangle area with their sub-category rectangles nested inside of it.
- When a quantity is assigned to a category, its area size is displayed in proportion to that quantity and to the other quantities within the same parent category in a part -to-whole relationship.
- The area size of the parent category is the total of its sub- categories .
- Tree maps start from top left (largest) to bottom right (smallest)





Area Charts

- Area charts are commonly used to show case data that depicts a time -series relationship
- Unlike Line charts, they can also visually represent volume
- Information is graphed on two axes, using data points connected by line segments. The area between the axis and this line is commonly emphasized with color or shading for legibility
- Typically used when you need to:
  - Know the magnitude of a change
  - Compare a small number of categories
  - Compare between trends and not exa ct values



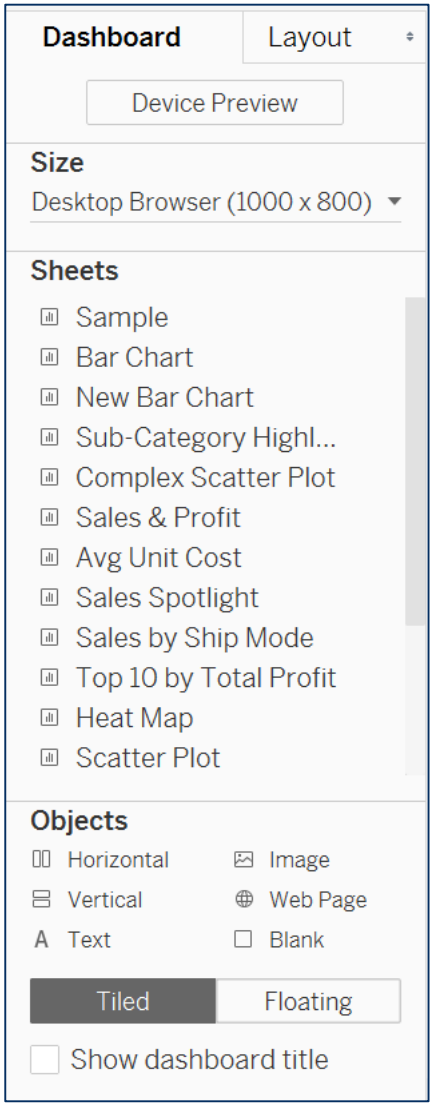
CREATING DASHBOARDS

Dashboards

- Dashboards are collections of worksheets
- They allow you to analyze multiple views at once
- Dashboards are shown at the bottom as tabs like worksheets
- Update automatically when data from the source is updated
- Supports all the same formatting as worksheets

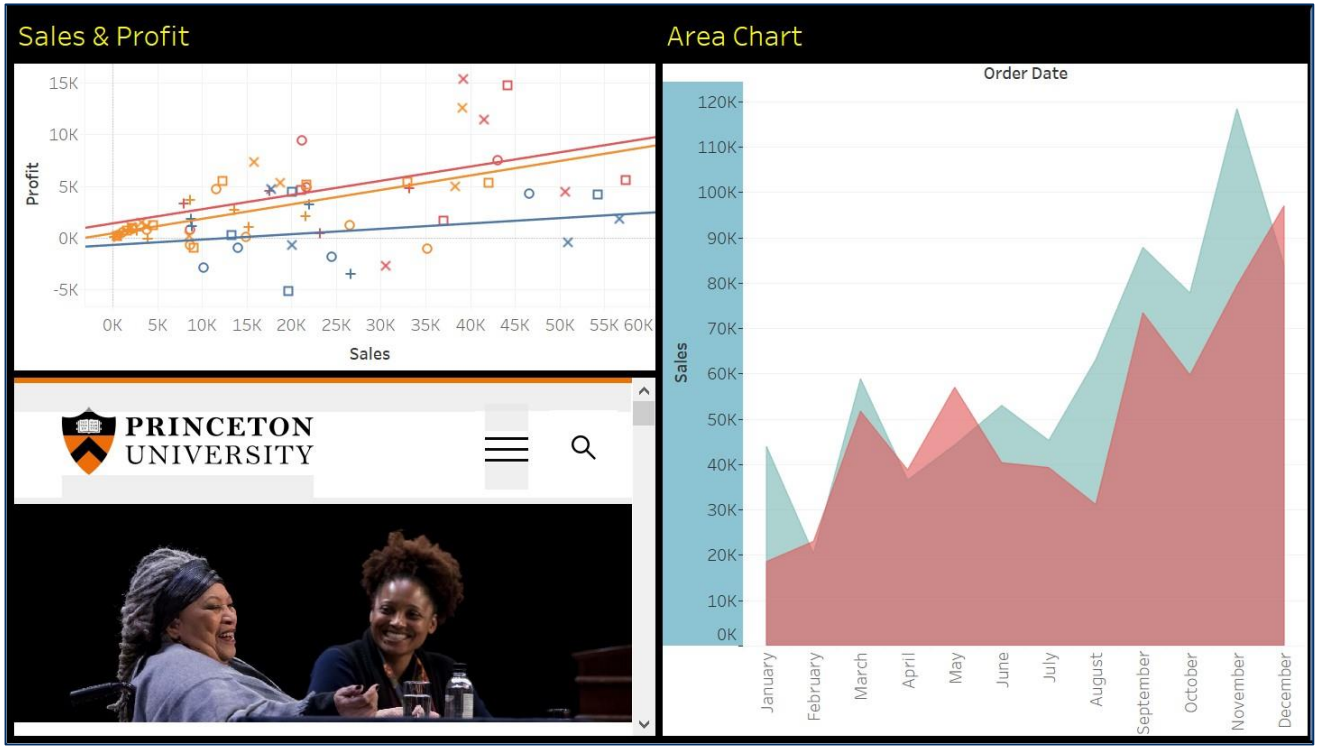
Dashboards

- When creating a new dashboard, the data pane isreplaced with a list of worksheets and objects
- Objects can be added by simply dragging them over to the dashboard
- By default, objects are “tiled” when they are added but can be changed to “floating” to allow overlap



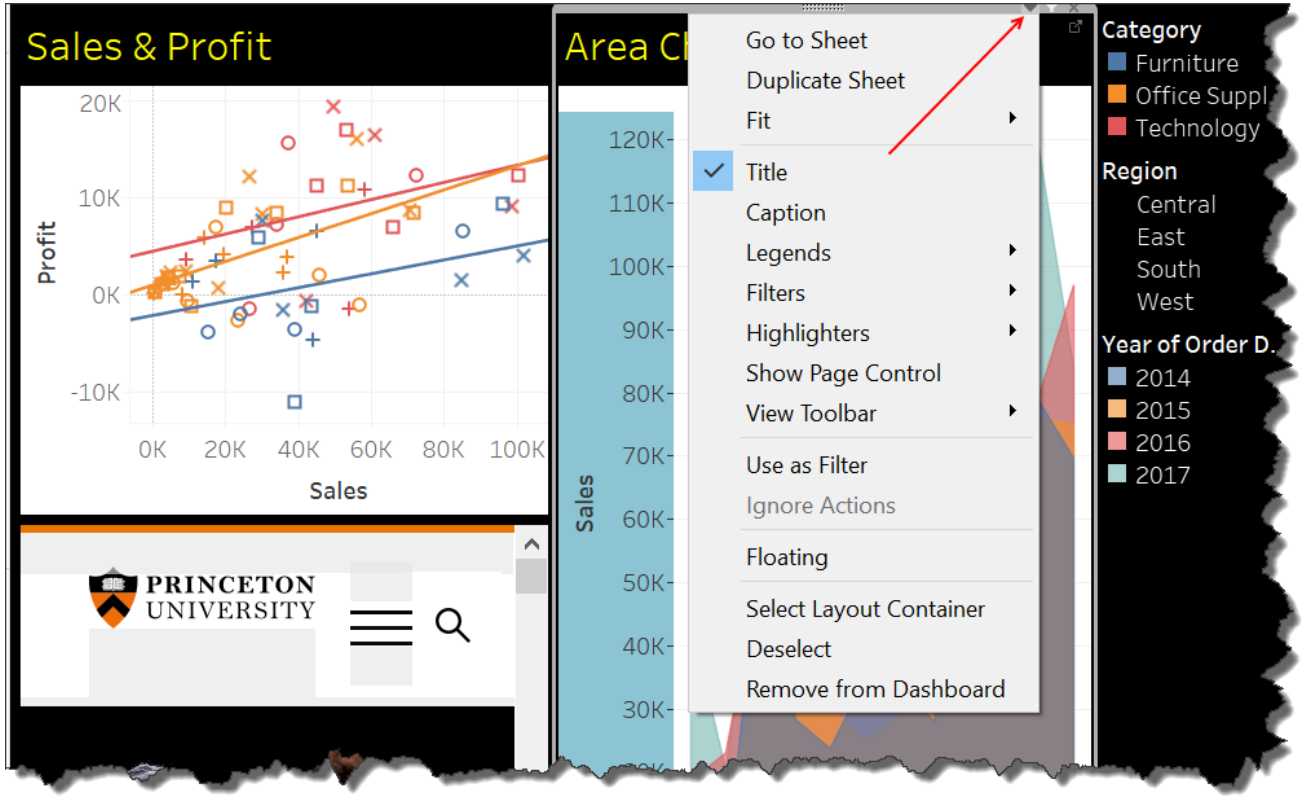
Adding Dashboard Objects

- Dashboard objects allow you to enhance a dashboard with non-worksheet items.
- Objects include:
  - Text boxes
  - Images
  - Web Pages
  - Blank Space



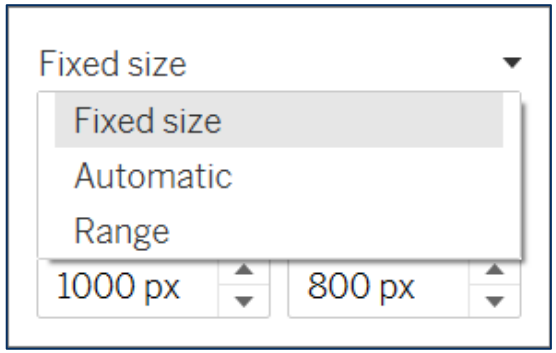
Organizing Dashboards

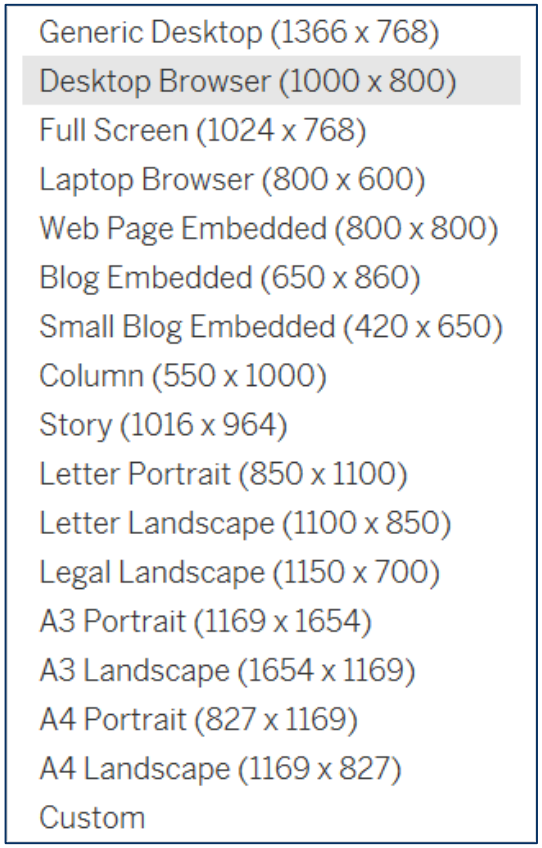
- You have the ability to show and/or hide specific parts of a dashbo ard as desired.
  - Title
  - Caption
  - Legends
  - Filters
  - Parameters

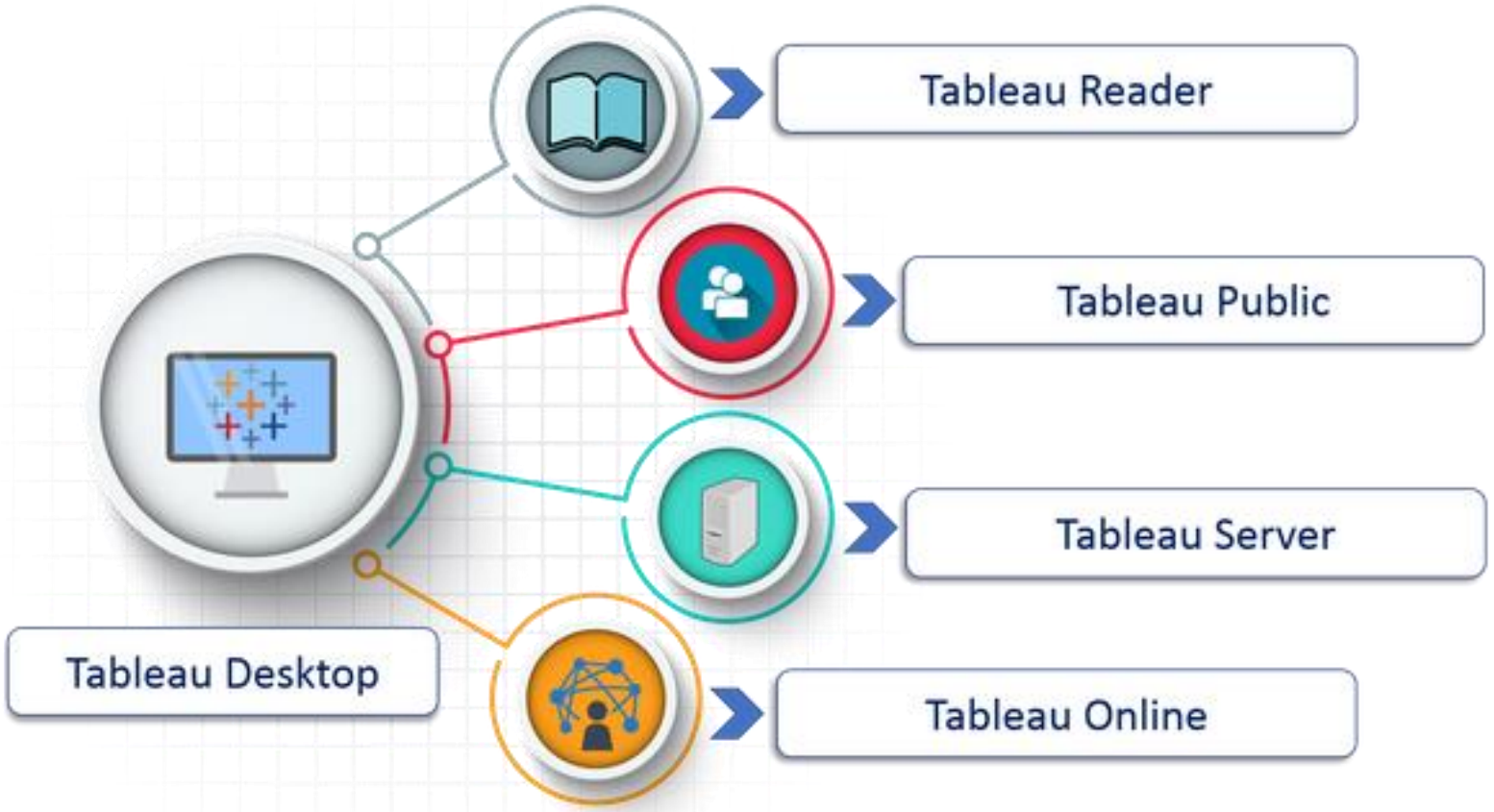


Organizing Dashboards

- Dashboards can be sized based on the following options:
  - Automatic: dashboards adjust as needed
  - Exactly: fixed width, scrollable if larger than window
  - Range: scales between min and max size
  - Presets: select from common sizes







PUBLISHING & SHARING

Publishing & Sharing

- Workbooks can be shared with other Tableau Desktop users, like any other file
- Alternatively, workbooks can be published to private servers
  - Princeton Tableau Server
  - Princeton Tableau Public
  - Tableau Public
- You can publish worksheets, dashboards, or stories

Publishing & Sharing

- Workbooks can be exported and published
- Multiple formats are available to use the output in applications such as PowerPoint, Word, or Excel
- When saving, you can save:
  - Workbooks (.twb)
    - all sheets, dashboards, and stories
  - Packaged workbook (.twbx)
    - All sheets, dashboards, stories, and all reference local data file and images in a single file

THANK YOU

thank you
danke
謝謝
ngiyabonga
شكراً جزيلاً
teşekkür ederim
gracias
merci
obrigado
go raibh maith agat
mochchakkeram
sukriya
kop khun krap
dziękuję
bedankt
спасибо
Баярлалаа
рахмат
dank je
grazie
arigatô
dakujem
terima kasih
감사합니다
ありがとう
мерси
хвала