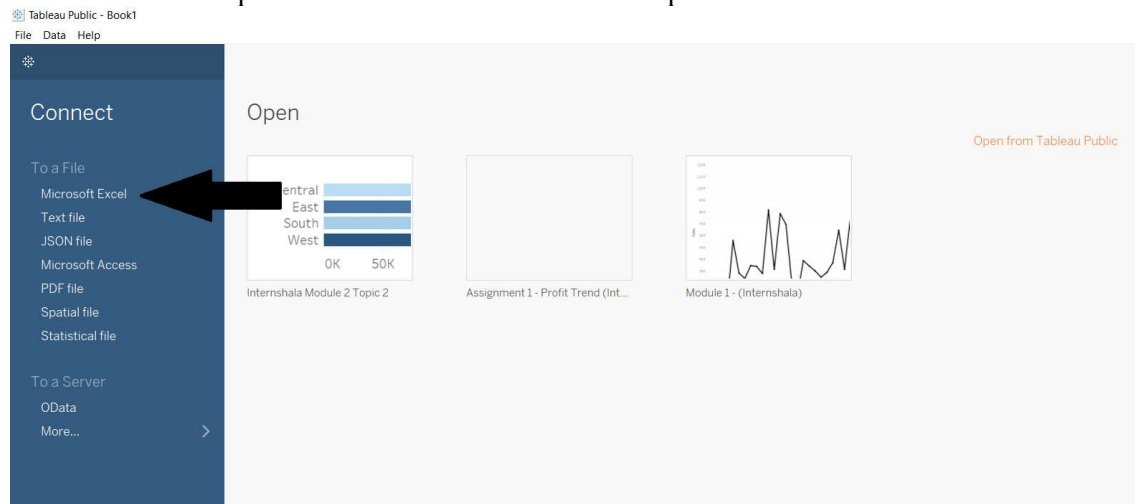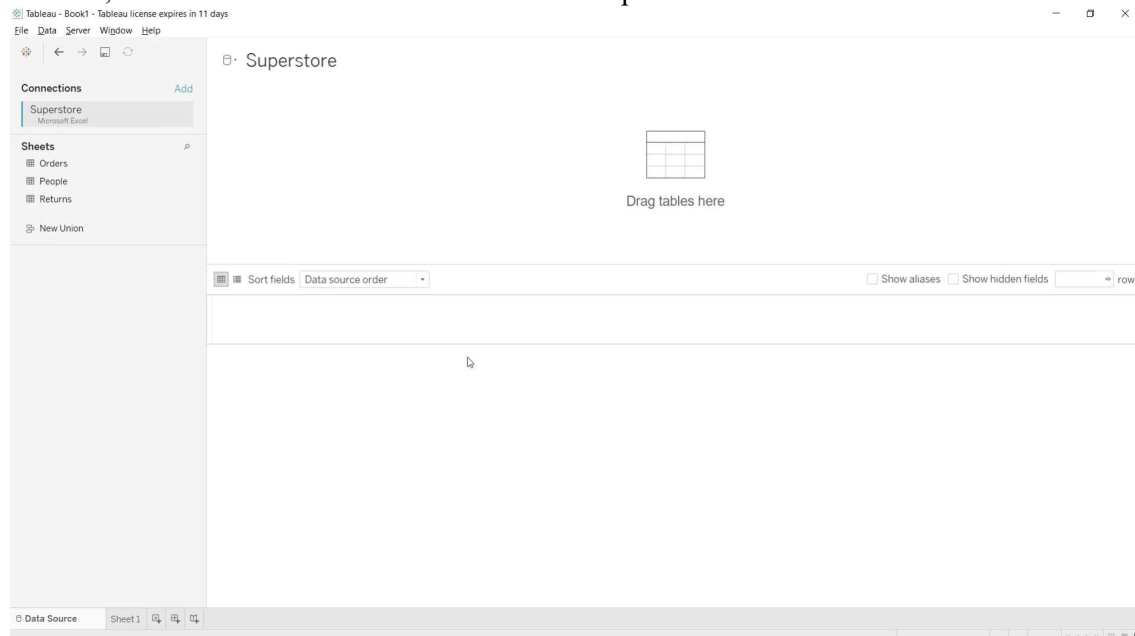# TABLEAU

Introduction to Tableau

Tableau can help anyone see and understand their data. Connect to almost any database, drag and drop to create visualizations, and share with a click. Tableau Software is an American interactive data visualization software company focused on business intelligence. It was founded in 2003 in Mountain View, California, and is currently headquartered in Seattle, Washington.
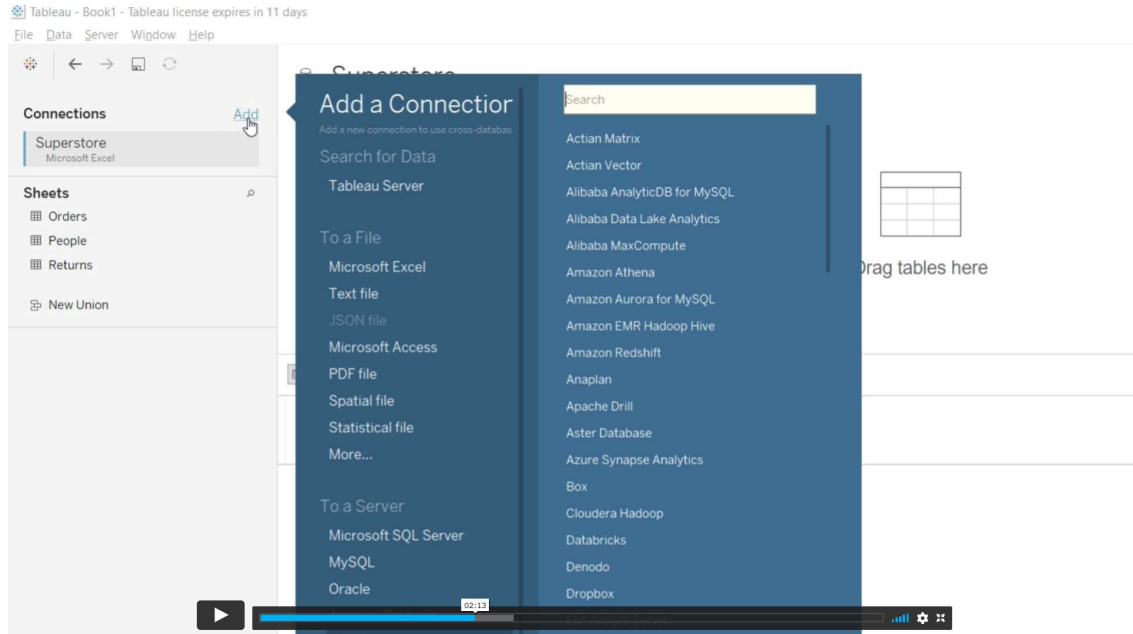
Starting with Tableau

We first import the data set into Tableau workspace.

We have a total of 2 options wherein we can connect a file or a server. Once we select the data, the data is connected to the tableau workspace as seen.

We can even add more data to the added workbook by clicking on the add more option.

Once a file, over here excel is added, we drag and drop the sheet on which we need to analyze using Tableau.
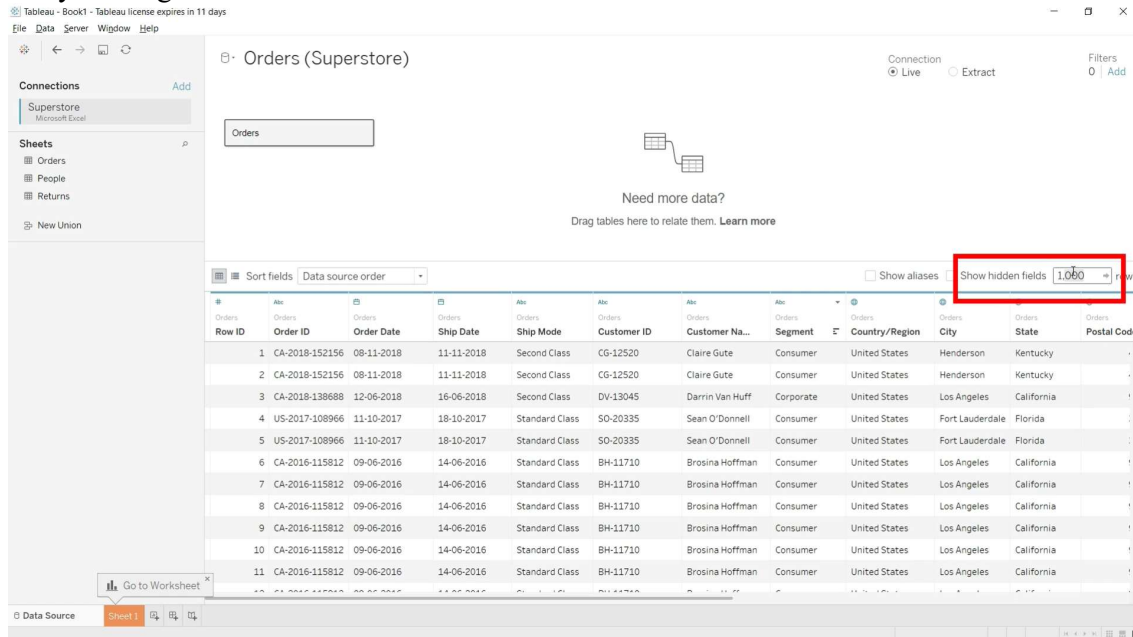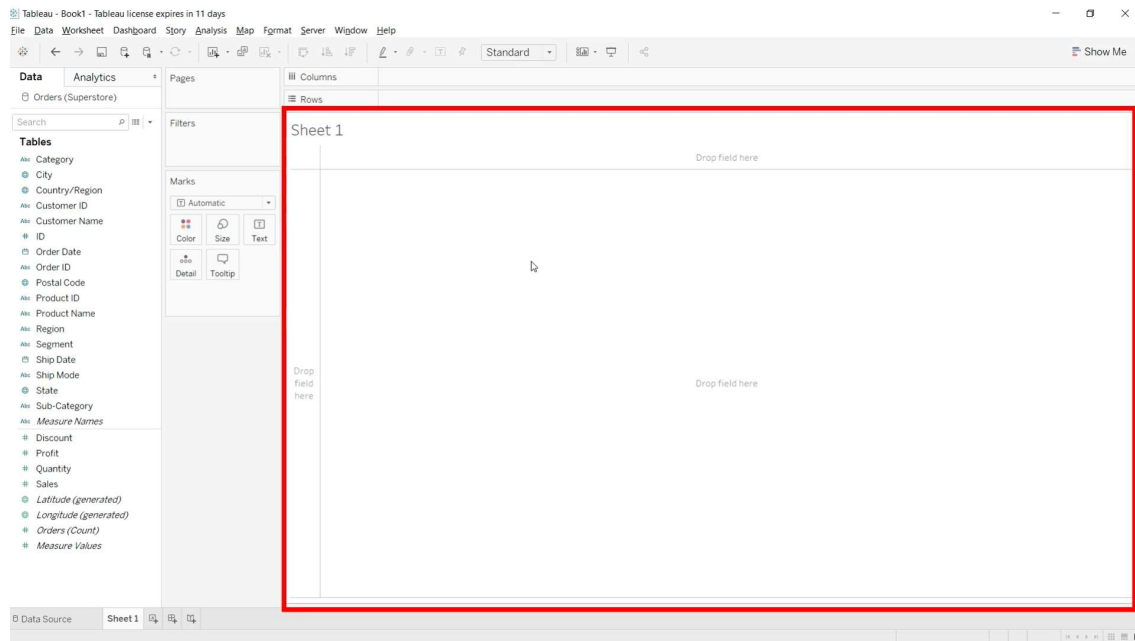
Tableau Interface

The marked area is the workspace where the charts are created.
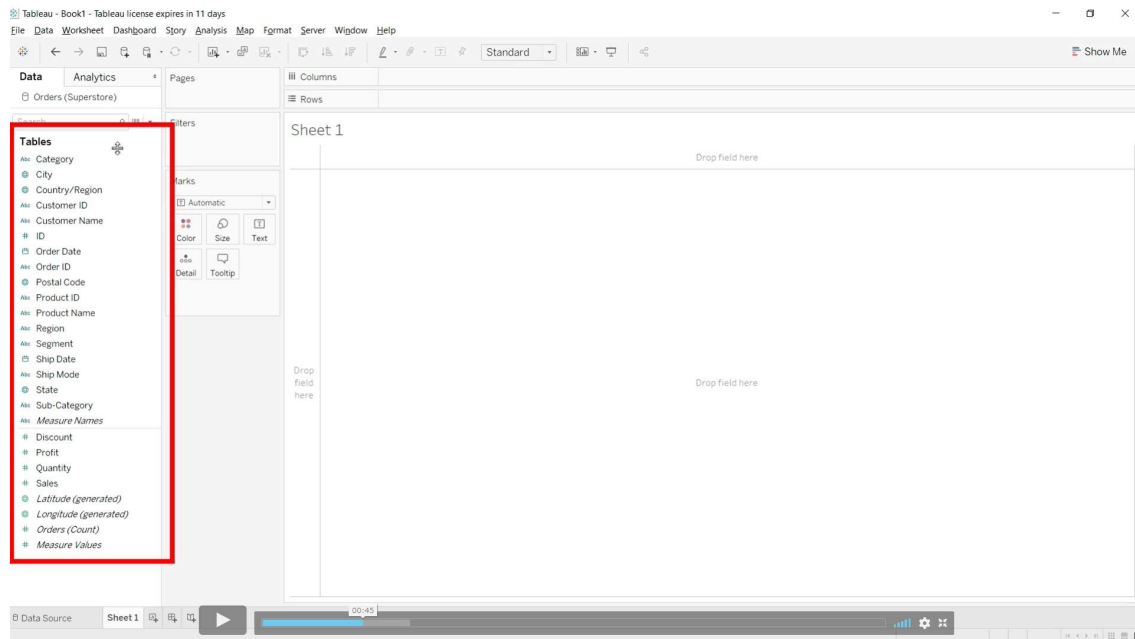
The marked area is the data-pane where we see all the fields from the data set.
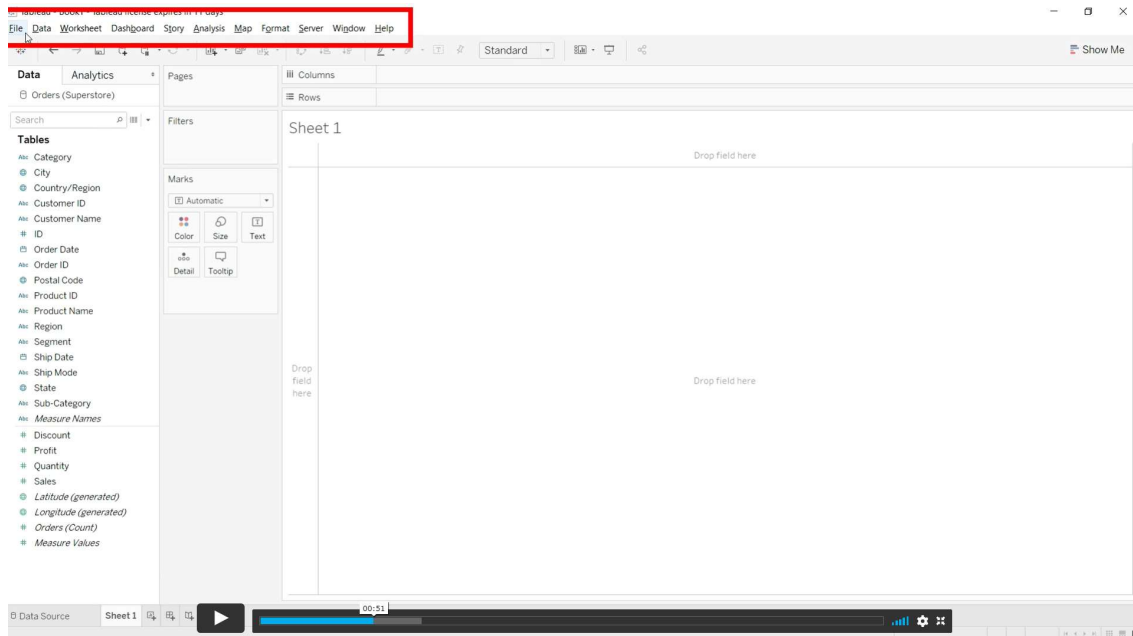
The marked area is the menu.
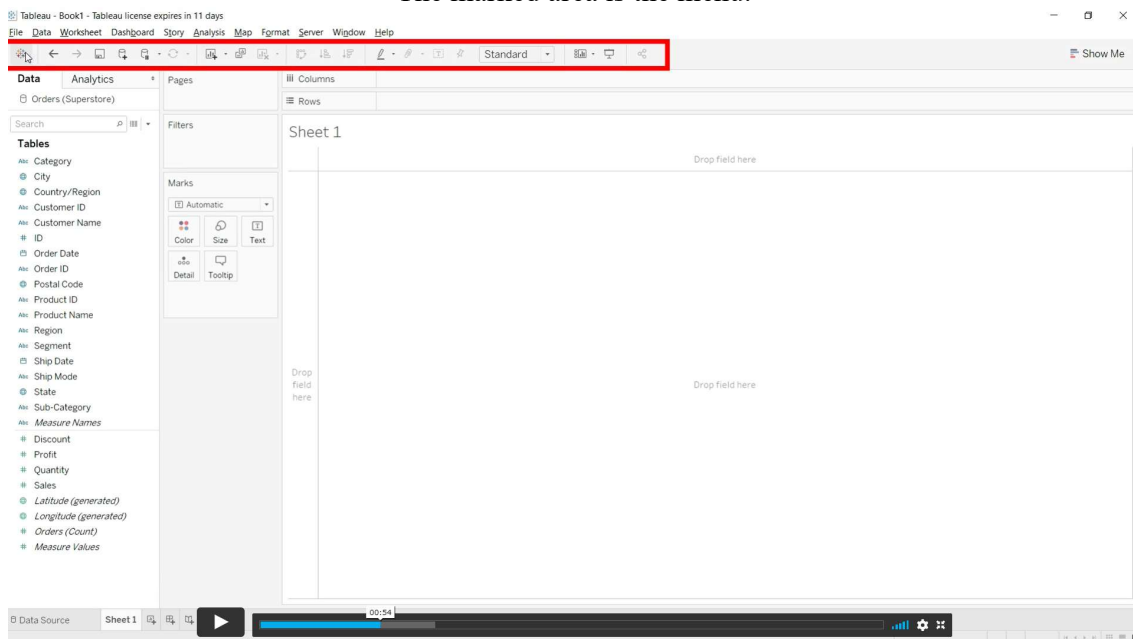
The marked area is the tool bar.
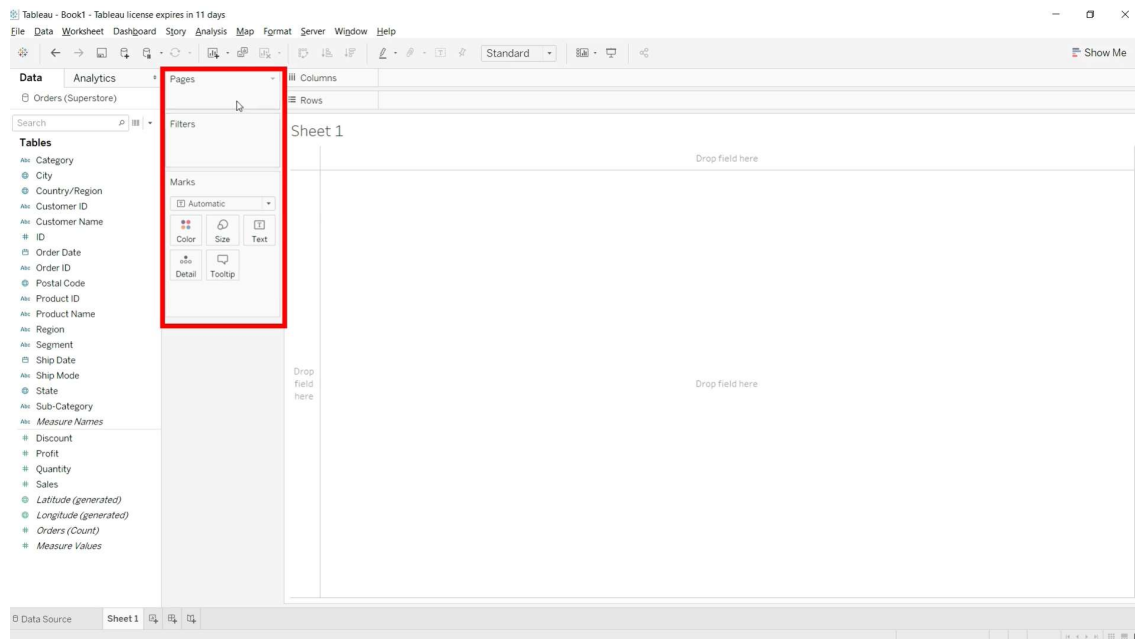
The marked area shows the important cards.
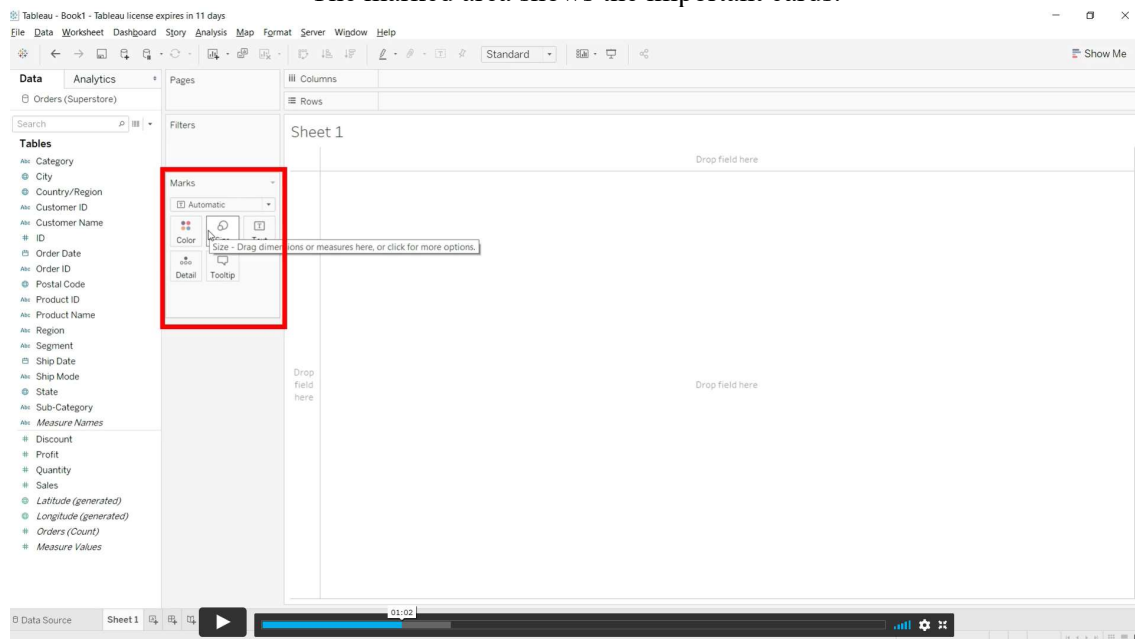
The marked area, Marks allows us to change the way a chart looks.
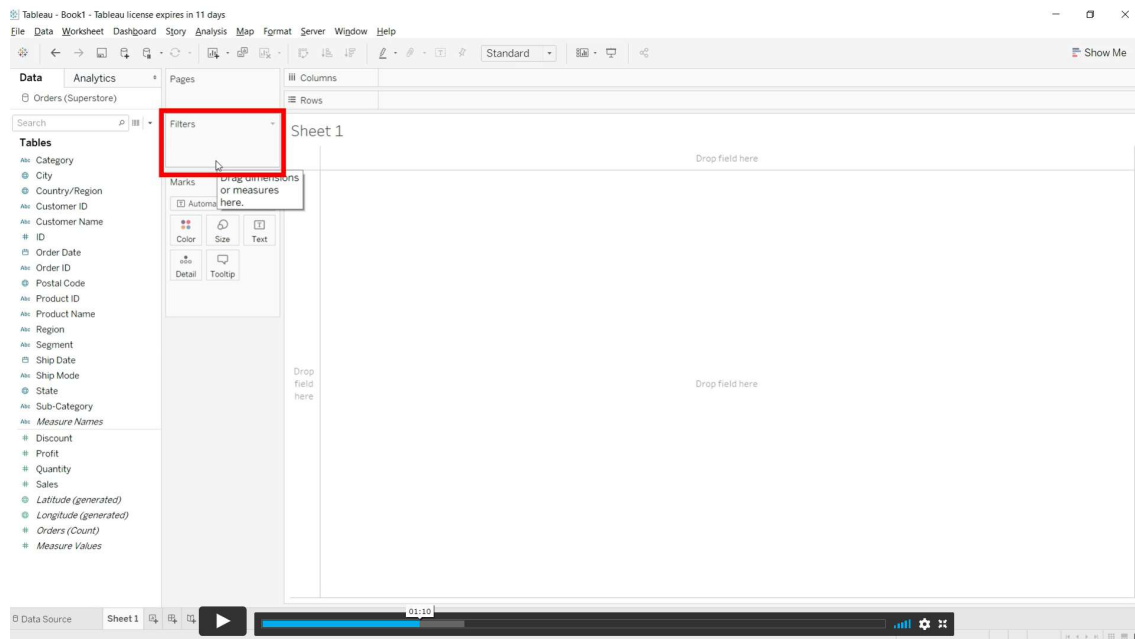
The marked area, filters, allows you to apply and change your filter settings.
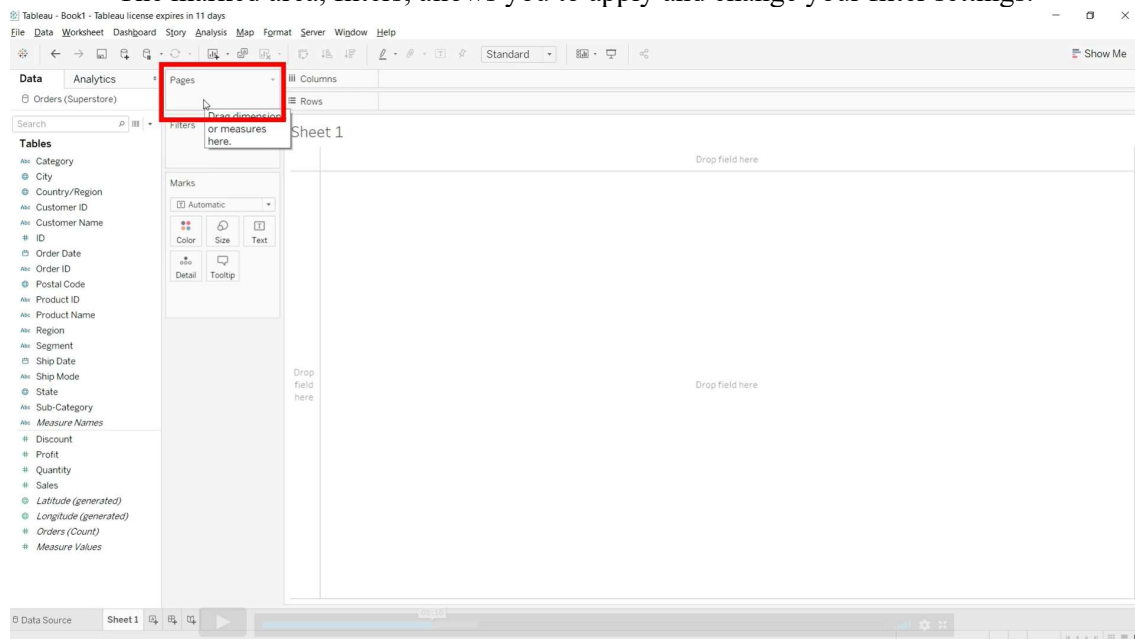
The marked area, Pages, allows you to create multiple sets of pages.
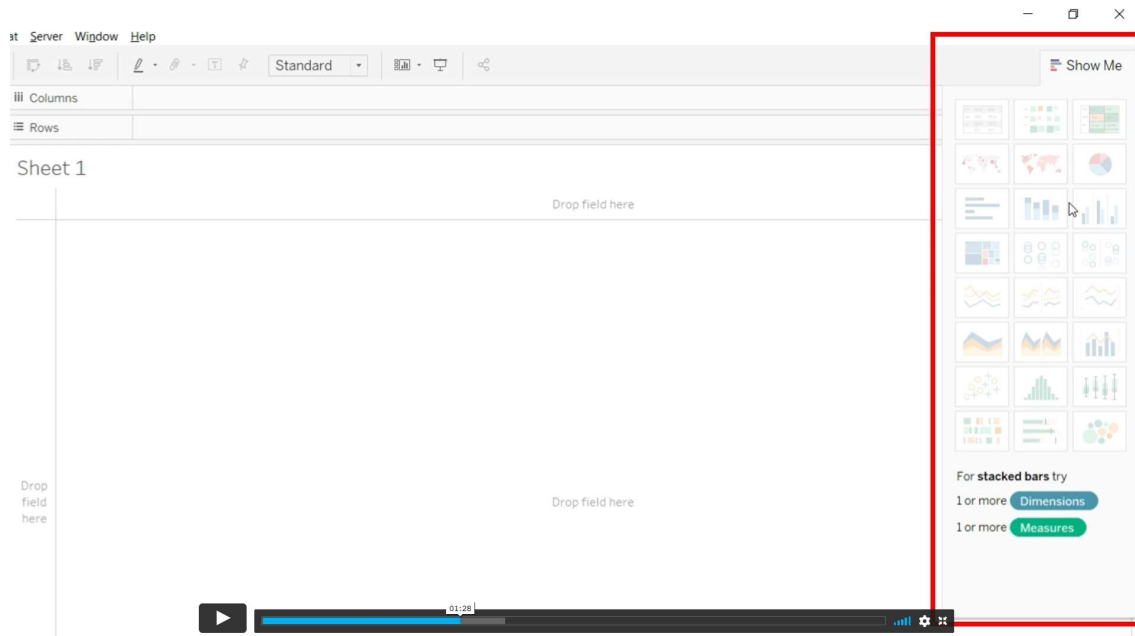
The marked area shows the chart templates.

Data Types in Tableau

Tableau has 7 data types in it. These data types are as follows:
- Number – Any field that has a numerical value in it is called a Number field. It can be a decimal or whole number.
- Date – The date field includes date in the DD/MM/YYYY format.
- Date and time – The date field expands to include time as well.
- String – The string data type stores all the alpha-numeric characters.
- Boolean – Used to store true or false
- Geographic – Unique in Tableau, a country or city name is included here. Tableau can easily procure the latitude and longitude for the given country or city name.
- Mixed – Also called cluster data type which has values of multiple data types.

Tableau also has a concept of Dimensions and Measures.
- Measures – Any quantitative variable is a measure. It is measureable.
- Dimensions – It is a qualitative field. Individually it does not add meaning to the data but when used with Measures brings more information about the data.

Tableau also has a concept of Discrete and Continuous
- **Discrete** –
  - o Fields have values which are finite in nature.
  - o Represented in Blue.
  - o Eg: Category, State, etc.
  - o Most of the time placed in the columns.
- **Continuous** –
  - o Fields have infinite values possible.
  - o Represented in Green.
  - o Eg: Sales, Profit, etc.
  - o Most of the time placed in the rows.

Summary

Data Types
- Numeric, string, date/date & time, Boolean, and Geographic

Dimensions and Measures
- Qualitative Vs Quantitative

Discrete and Continuous
- Finite Vs Infinite

The Marks card

- Color – Allows you to change the color of the chart.

- Size – Allows you to change the size of the contents of the charts.
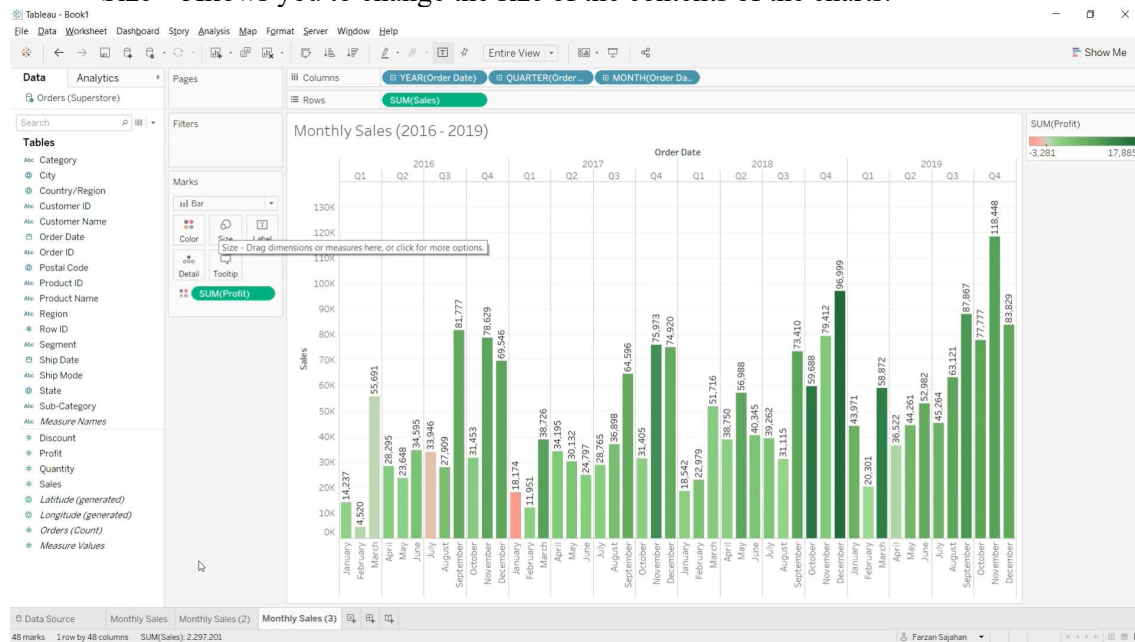
- Label – Allows you to label the chart using the field used.
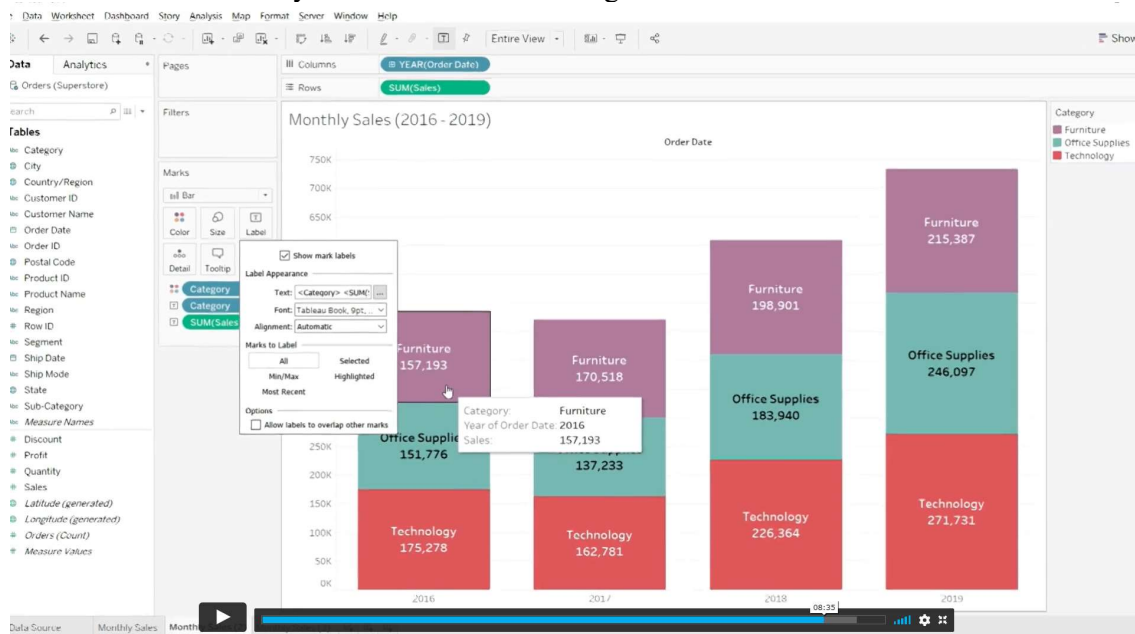
Types of Charts in Tableau

- **Area Charts**
  - An area chart is a line chart where the area between the line and the axis are shaded with a color. These charts are typically used to represent accumulated totals over time and are the conventional way to display stacked lines.

Column Shelf	Dimension Variable
Row Shelf	Measure Variable

- **Bar Charts**
  - Use bar charts to compare data across categories. A bar chart uses the Bar mark type. You create a bar chart by placing a dimension on the Rows shelf and a measure on the Columns shelf, or vice versa.

Column Shelf	Dimension or Measure Variable
Row Shelf	Measure or Dimension Variable

- **Tree Maps**
  - Use treemaps to display data in nested rectangles. You use dimensions to define the structure of the treemap, and measures to define the size or color of the individual rectangles. Treemaps are a relatively simple data visualization that can provide insight in a visually attractive format.

Color	Dimension or Measure Variable
Size	Measure Variable
Label or Detail	Dimension Variable

- **Bubble Charts**
  - o Use packed bubble charts to display data in a cluster of circles. Dimensions define the individual bubbles, and measures define the size and color of the individual circles.

Detail	Dimension
Size	Measure
Color	Dimension or Measure
Label	Dimension or Measure

- **Pie Chart**
  - o Use pie charts to show proportions of a whole.

Color	Dimension
Angle	Measure

- **Histograms**
  - o A histogram is a chart that displays the shape of a distribution. A histogram looks like a bar chart but groups values for a continuous measure into ranges, or bins.

Row Shelf	Continuous Measure
Column Shelf	Bins (Continuous or Discrete)

- **Box Plots**
  - o Use box plots, also known as box-and-whisker plots, to show the distribution of values along an axis.

Row Shelf	Measure
Column Shelf	Dimension
Detail	Dimension
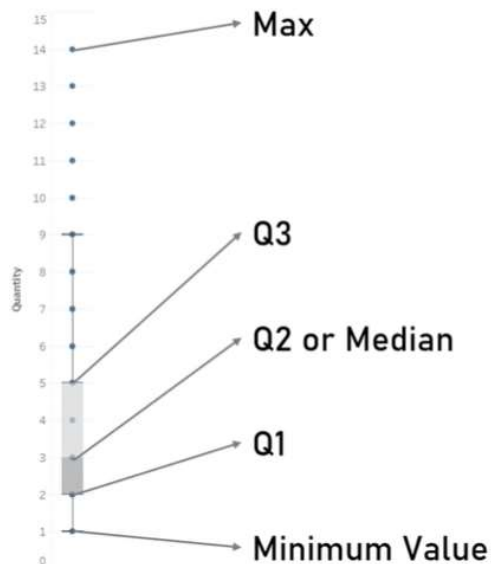
- **Scatter Plots**
  - Use scatter plots to visualize relationships between numerical variables. In Tableau, you create a scatter plot by placing at least one measure on the Columns shelf and at least one measure on the Rows shelf.

- **Fill Maps**
  - For Filled Maps, you require a Geographical variable. Along with which you could use a Measure

- **Symbol Maps**
  - Proportional symbol maps are great for showing quantitative values for individual locations. They can show one or two quantitative values per location.

Little bit more about Maps

- Tableau gives fields geographical variable tag on the basis of its name.
- Tableau also gives the fields its own geographical role.
- Tableau generates the Latitude and Longitude values based on the geographical fields.
- Latitude and Longitude values if need to be used can be used only with geographical fields as it doesn't exist in isolation.

Formatting the Charts

- We can remove/hide the axes by right clicking and switching off the "Show header" option, or by clicking on Edit Axis.
- We can hide the legend card by clicking on the tiny arrow mark on the top right of the card and clicking on Hide card option.
- We can use the Edit axes option by right clicking on the axes and choosing the Edit axes button. We can do the following:
  - Edit range of the axes.
  - Edit axis titles.
  - Edit scale into normal, logarithmic or reversed.
  - Edit how tick marks appear.
- We can change the looks of the chart by using the Marks card on the left. We can do the following:
  - Change the color of the chart using the color option.
- We can change the looks of the chart by right clicking on the chart and selecting format. We can do:
  - Click on shading option on the left pop up and change the background color of the chart by clicking on the sheet tab and select the desired color corresponding to the worksheet tab.
  - Click on the shading option of the left pop up and change the color of the axes by clicking on the Rows/Columns tab and select the desired color corresponding to the Header tab.
- Basically explore the format option when u right click anywhere on the chart.

Combined Axis Charts

- To make combined axes chart, just bring another measure variable and drop on the left axis.
- If the order of magnitude of the variables used is same then no need for any changes, else create a Dual Axis Charts.

Dual Axis Charts

- To make Dual Axes Charts we start of by dropping one of the measures into the chart's left side axis.
- Right click/Click on the drop down menu and select the dual axis mode.
- In case of differing magnitudes of axes, we can right click on the right side axis and select synchronize axis.

Editing the interactivity – Tooltips

- Click on the Tooltips option in the marks card and let loose on however u want to see it.
- U can even insert charts in the tooltips by using the insert option a clicking on the sheets and the chart you have created before to insert.

Filters in charts

- To apply filter, drag and drop the filter variable (the variable being used to filter the chart) in the filter card above the Marks card.
- Normally the filter option is not shown to the viewers, in case you want to do that, right click on the filter card and click on show filter. This allows u to maneuver the filter options using the right hand side pop up.
- Filters are basically of 4 types, namely:
    - Simple filters
        - To apply simple filters, one just needs to drag and drop the filter variable in the filter card.
        - You have 3 options in the simple filters:
            - Select from list wherein all unique values under the filter variable is seen and can be chosen.
            - Custom value list wherein you type in the unique value which would be shown as a filter.
            - Use all, wherein tableau will use every single unique item in the data field and apply it in the filter.
    - Wild Card Filters
        - To apply wild card filters, one just needs to drag and drop the filter variable in the filter card.
        - You have 4 options in the wild card filter:
            - Filter and display the fields which "contains" the keyword entered.
            - Filter and display the fields which "starts with" the keyword entered.

- Filter and display the fields which "ends with" the keyword entered.
- Filter and display the field which is "exactly the same" as the keyword entered.
  - You can also choose to exclude the fields entered.
- Conditional Filters
  - To apply conditional filters, one just needs to drag and drop the filter variable in the filter card.
  - You have 2 options in the Conditional Filter:
    - Filter by field where you can select your choice of filter variable from the given list of fields, their respective option (such as sum, average, count, etc.) and give a value to it.
    - Filter by formula where you can input a formula using the given fields too.
- Top/Bottom Filters
  - To apply Top/Bottom filters, one just needs to drag and drop the filter variable in the filter card.
  - You have 2 options in the Top/Bottom Filter:
    - Filter Top/Bottom "number you input" by field and their respective option (such as sum, average, count, etc.)
    - Filter Top/Bottom "number you input" by the formula you enter using the fields given.

Groups

- Grouping in Tableau is grouping multiple members in a SINGLE dimension into a higher category.
- To group members of a certain variable we right click on the variable, click on create, and then groups.
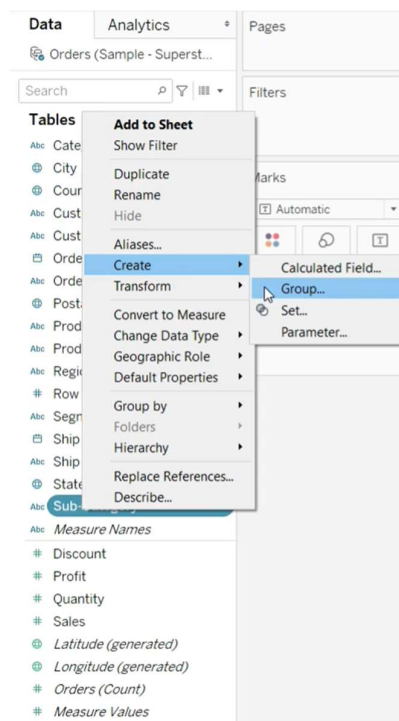
Sets

- Sets are custom fields that are created within Tableau Desktop based on dimensions from your data source.
- To create a set of a certain variable we right click on the variable, click on create, and then sets.
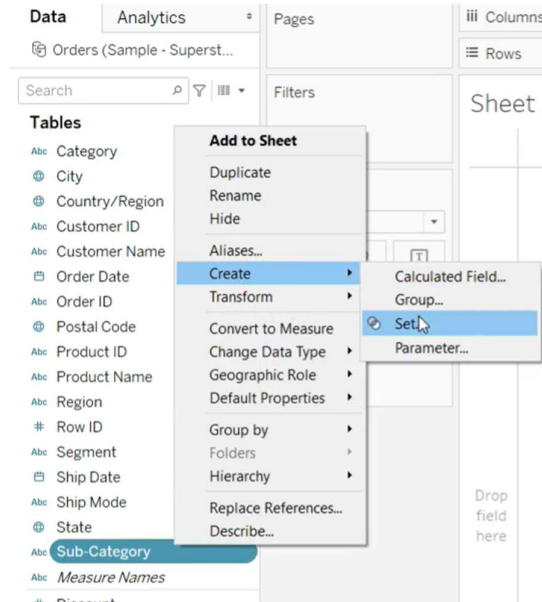
Dashboards

- A dashboard is a consolidated display of many worksheets and related information in a single place. It is used to compare and monitor a variety of data simultaneously. The different data views are displayed all at once.
- There are 2 ways of creating a dashboard:
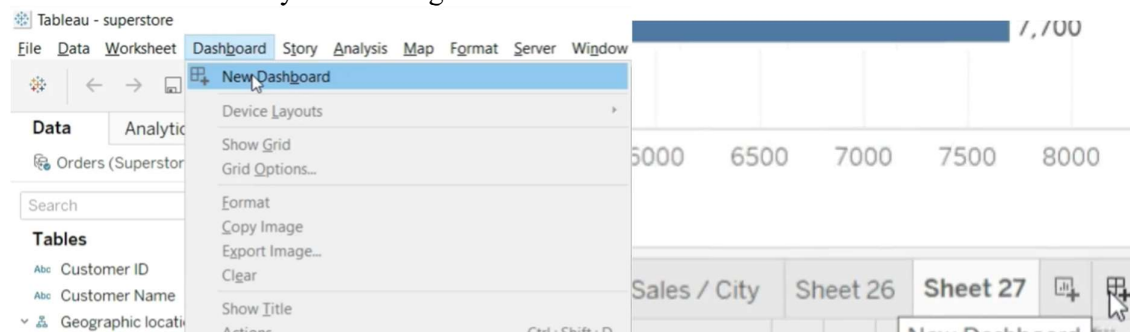
- To add charts, one just need to drag and drop the charts in the same workbook into the dashboard workspace.
- One can add containers in the dashboard within which we can add the charts too.
- We can build interactivity in the dashboard by adding objects such as containers (vertical and horizontal), navigation buttons, text, extensions, web pages, etc.
- We can delete charts or their legends by simple clicking on the cross which appears when hovered above the respective chart/legend.
- We can build interactivity in the charts too:
  o We can add filters to the chart and show the filter along with the chart in the dashboard.

o   We can make the chart a filter of another chart.
o   We can also use action filters which give an enhanced feel of using filters in the charts.
o   We can also highlight stuff.

Numerical Calculated Fields

- We can create new fields in the data using the create Numerical Field option in Tableau.
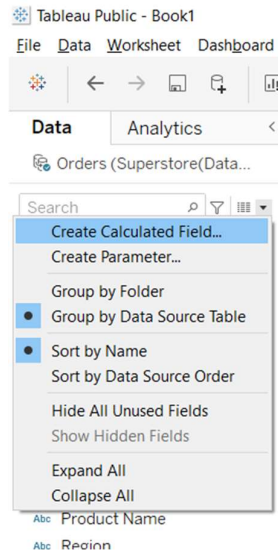
- While creating the calculated field, we would require to input a formula to create one. Thus signifying, a new field can be created only using measure variables by Calculated Field.
- We have multiple options of stuff in the calculated field menu too. A few categories are as follows:
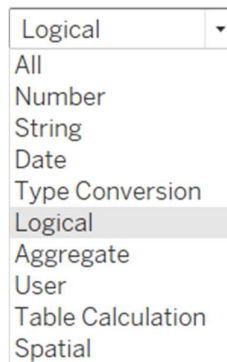
- Each of these categories also have various sub categories which can be used when formulating/coding for the new field.

Parameters in Filters

- A parameter is a workbook variable such as a number, date, or string that can replace a constant value in a calculation, filter, or reference line.

- You can make your parameters more dynamic and interactive by using them in Parameter Actions. Parameter actions let your audience change a parameter value through direct interaction with a viz, such as clicking or selecting a mark.
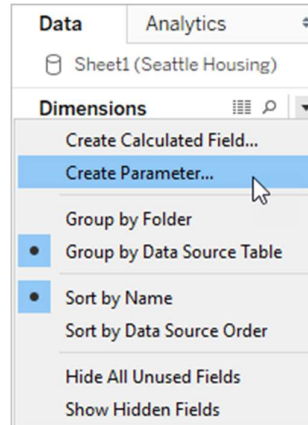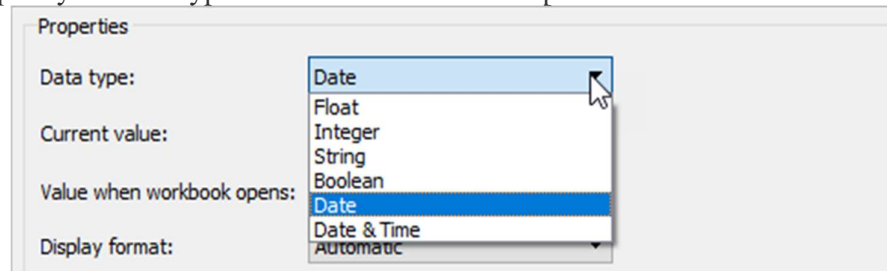- To create a parameter, we follow the following steps:
  - In the Data pane, click the drop-down arrow in the upper right corner and select Create Parameter.

  - In the Create Parameter dialog box, give the field a Name.

  - Specify the data type for the values it will accept:

  - Specify a current value. This is the default value for the parameter. In this case, let's leave the field as is because we'll be using the latest data, which we'll configure in the next step.
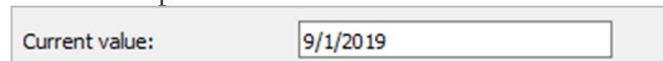
  - Specify a value when the workbook opens.
  - Specify the display format to use in the parameter control/
  - Specify how the parameter will accept values. The availability of these options is determined by the data type. You can select from the following options:
    - All - The parameter control is a simple text field.
    - List - The parameter control provides a list of possible values for you to select from. If you select List, you must specify the list of values. Click in the left column to type your list of values, or you can add members of a field by selecting Add values from.
    - Range - The parameter control lets you select values within a specified range. If you select Range, you must specify a minimum, maximum, and step size.

- When finished, click OK. The parameter is now listed in the Parameters section at the bottom of the Data pane.

Table Calculations

- Table Calculations are special type of calculations which work specially for what's shown on the worksheet.
- Right click and click on table calculations.
- You can combine 2 table calculations by clicking on the edit table calculation option.

Pareto Chart

- A Pareto chart is a type of chart that contains both bars and a line graph, where individual values are represented in descending order by bars, and the cumulative total is represented by the line.
- Can be made easily by creating a dual axis chart by combining Bar chart and a Line chart.

Level of Detail (LOD) Expressions

- Level of Detail expressions (also known as LOD expressions) allow you to compute values at the data source level and the visualization level according to our formulation.
- There are 3 types of LOD Expressions, namely:
  - INCLUDE
    - Includes the dimension in visualization as well as the dimension used in the calculation.
  - EXCLUDE
    - Includes the dimension in visualization but excludes the dimension in the calculation.
  - FIXED
    - Use only the dimension in the calculation irrespective of anything else.

Stories

- A story is an extension of a Dashboard.
- A story is a sequence of visualization, charts or dashboards that work together to convey information.
- It is very similar to Power Point Presentation or Keynotes.
- We can create a story by clicking on the story option in the menu bar.
- We can add animations too by clicking on format in menu bar and clicking on animations.

Data Preparation

- Data comes in from multiple sources and is almost always messy.
- Data preparation involves several custom steps such as removing unnecessary records and fields, converting data types, splitting/combining of data fields, etc.
- To remove a data field, in tableau we can just hide the field in the table.
- To rename a data field, in tableau we can just rename the field in the table.

- To change the data type in tableau, we can just click on the data type icon on the field name in the table and change it.
- We can split a field in the data in tableau by right click and clicking on split. It will split automatically with the delimiter given.
- We can filter the data too at the source level by clicking on filters and adding the necessary filter on the top right of the window.
- We can group the data in the source itself by right clicking on the field and the selecting group.
- We can create aliases, that is, name the field something else in the visualization.
- We can also combine data by using any of the following 4 methods:
  - Unions
    - We stack the data vertically. For example, data of 2017 is stacked over data of 2016.
    - 2016 + 2017 = 2016 & 2017.
    - To create a union, just drag a sheet to the workspace and then another sheet right below the previous one such that union word is seen.
    - Another way to create union is by using wildcard union wherein workbooks are used into a union directly then the sheets.
  - Join
    - We join the data horizontally using one or more common fields.
    - Order ID, Sales, Profit + Order ID, Returned Order = Order ID, Sales, Profit, Order ID.
    - To join 2 tables we just bring the main sheet in the workspace and double click it. Then we bring the other sheet next to the previous sheet. This creates a inner join by default. We can change the type of join by clicking on the symbol of join.
    - There are 4 types of join, namely:
      - Left Join

Order ID	Sales	Profit		Order ID	Returned Y/N?		Order ID	Sales	Profit	Returned Y/N?
001	100	30		002	N		001	100	30	-
002	150	45		003	N		002	150	45	N
003	120	40		005	Y		003	120	40	N
004	180	100					004	180	100	-

      - Right Join

Order ID	Sales	Profit		Order ID	Returned Y/N?		Order ID	Returned Y/N?	Sales	Profit
001	100	30		002	N		002	N	150	45
002	150	45		003	N		003	N	120	40
003	120	40		005	Y		005	Y	-	-
004	180	100								

- Inner Join

Order ID	Sales	Profit		Order ID	Returned Y/N?		Order ID	Sales	Profit	Returned Y/N?
001	100	30		002	N		002	150	45	N
002	150	45		003	N		003	120	40	N
003	120	40		005	Y					
004	180	100								

- Outer Join

Order ID	Sales	Profit		Order ID	Returned Y/N?		Order ID	Sales	Profit	Returned Y/N?
001	100	30		002	N		001	100	30	–
002	150	45		003	N		002	150	45	N
003	120	40		005	Y		003	120	40	N
004	180	100					004	180	100	–
							005	–	–	Y

- Blending
  - It is by default a left join.
  - Similar to join but doesn't combine tables.
  - We basically are querying 2 tables at the level of visualization.
  - Useful when working with multiple large data-sources.
  - Its faster than join and union.
  - To blend 2 data sources, we import the required sheet and open the chart section. In there we click on the new data source icon (or Control + D) and import the other data source. Then we link the tables.
- Relationships & Data Modelling
  - Works almost same as blend but much more intuitive than blend and join.
  - All advantages of blend are applied to this.
  - Combines multiple tables together using related or common fields.
  - For this method we simply bring the worksheets/workbooks directly onto the workspace and tableau creates an automatic link between them.