Tableau to get acquainted with its interface. There are three basic steps involved in creating any Tableau data analysis report.

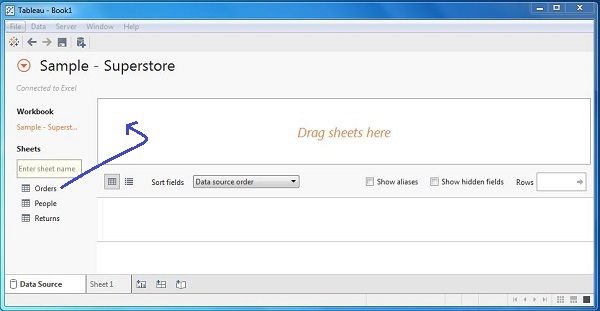
These three steps are −

* **Connect to a data source** − It involves locating the data and using an appropriate type of connection to read the data.
* **Choose dimensions and measures** − This involves selecting the required columns from the source data for analysis.
* **Apply visualization technique** − This involves applying required visualization methods, such as a specific chart or graph type to the data being analyzed.

For convenience, let’s use the sample data set that comes with Tableau installation named sample – superstore.xls. Locate the installation folder of Tableau and go to **My Tableau Repository**. Under it, you will find the above file at **Datasources\9.2\en\_US-US**.

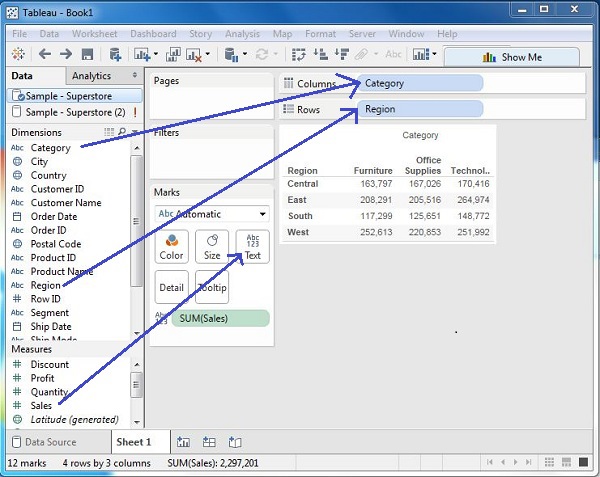
Connect to a Data Source

On opening Tableau, you will get the start page showing various data sources. Under the header **“Connect”**, you have options to choose a file or server or saved data source. Under Files, choose excel. Then navigate to the file **“Sample – Superstore.xls”** as mentioned above. The excel file has three sheets named Orders, People and Returns. Choose **Orders**.



Choose the Dimensions and Measures

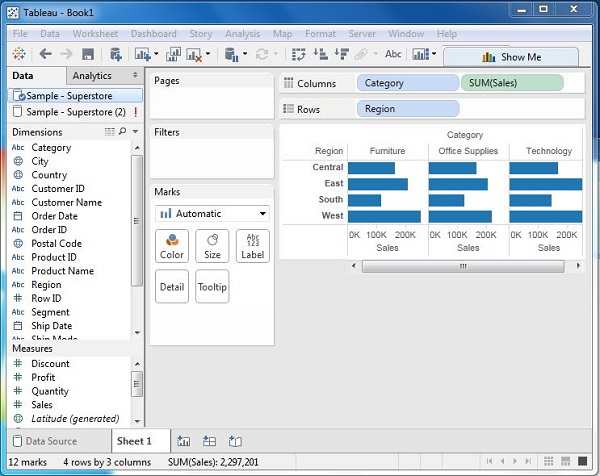
Next, choose the data to be analyzed by deciding on the dimensions and measures. Dimensions are the descriptive data while measures are numeric data. When put together, they help visualize the performance of the dimensional data with respect to the data which are measures. Choose **Category** and **Region** as the dimensions and **Sales** as the measure. Drag and drop them as shown in the following screenshot. The result shows the total sales in each category for each region.



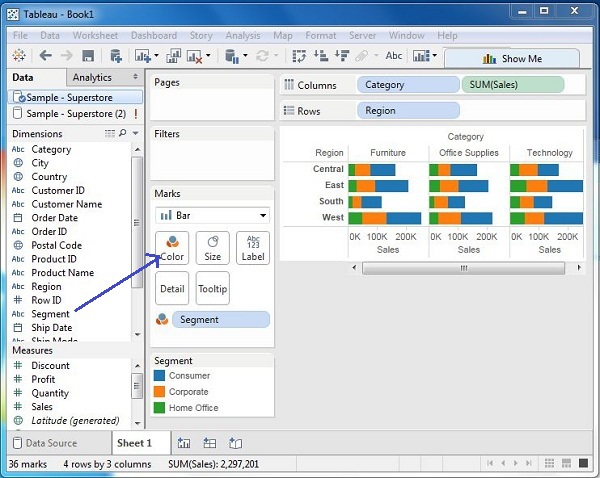
Apply Visualization Technique

In the previous step, you can see that the data is available only as numbers. You have to read and calculate each of the values to judge the performance. However, you can see them as graphs or charts with different colors to make a quicker judgment.

We drag and drop the sum (sales) column from the Marks tab to the Columns shelf. The table showing the numeric values of sales now turns into a bar chart automatically.

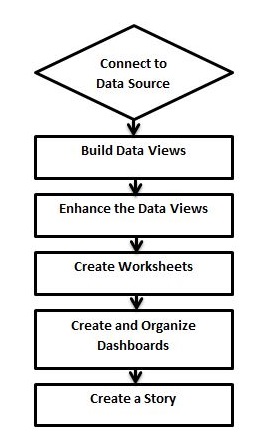


You can apply a technique of adding another dimension to the existing data. This will add more colors to the existing bar chart as shown in the following screenshot.



As Tableau helps in analyzing lots of data over diverse time periods, dimensions, and measures, it needs a very meticulous planning to create a good dashboard or story. Hence, it is important to know the approach to design a good dashboard. Like any other field of human endeavor, there are many best practices to be followed to create good worksheets and dashboards.

Though the final outcome expected from a Tableau project is ideally a dashboard with story, there are many intermediate steps which needs to be completed to reach this goal. Following is a flow diagram of design steps that should be ideally followed to create effective dashboards.



## Connect to Data Source

Tableau connects to all popular data sources. It has inbuilt connectors which take care of establishing the connection, once the connection parameters are supplied. Be it simple text files, relational sources, SQL sources or cloud data bases, Tableau connects to nearly every data source.

## Build Data Views

After connecting to a data source, you get all the column and data available in the Tableau environment. You classify them as dimensions and measures, and create any hierarchy required. Using these you build views, which are traditionally known as Reports. Tableau provides easy drag and drop feature to build views.

## Enhance the Views

The views created above needs to be enhanced further by the use of filters, aggregations, labeling of axes, formatting of colors and borders, etc.

## Create Worksheets

Create different worksheets to create different views on the same or different data.

## Create and Organize Dashboards

Dashboards contain multiple worksheets which are linked. Hence, the action in any of the worksheet can change the result in the dashboard accordingly.

## Create a Story

A story is a sheet that contains a sequence of worksheets or dashboards that work together to convey information. You can create stories to show how facts are connected, provide context, demonstrate how decisions relate to outcomes, or simply make a compelling case.

As a data analysis tool, Tableau classifies every piece of data into one of the four categories namely - String, Number, Boolean and datetime. Once data is loaded from the source, Tableau automatically assigns the data types. Contrarily, you can also change some of the data types if it satisfies the data conversion rule. The user has to specify the data type for calculated fields.

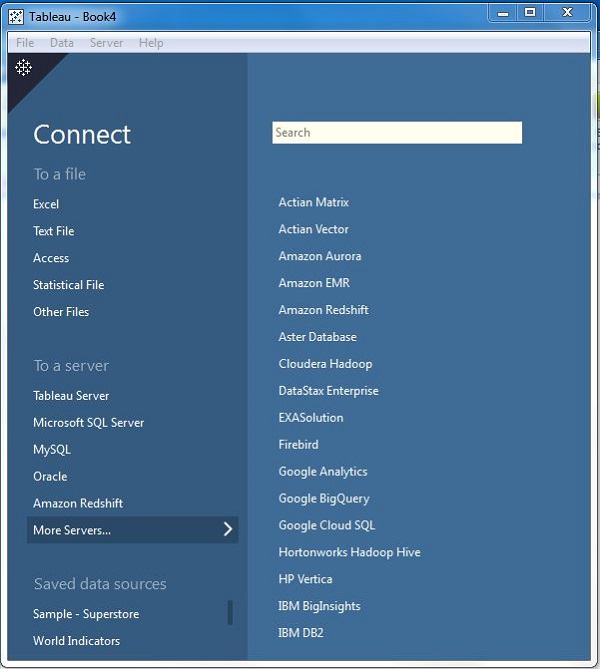
Following table lists the description of data types supported by Tableau.

|  |  |  |
| --- | --- | --- |
| **Data Type** | **Description** | **Example** |
| **STRING** | Any sequence of zero or more characters. They are enclosed within single quotes. The quote itself can be included in a string by writing it twice. | 'Hello' 'Quoted' 'quote' |
| **NUMBER** | These are either integers or floating points. It is advised to round the floating point numbers while using them in calculations. | 3  142.58 |
| **BOOLEAN** | They are logical values. | TRUE  FALSE |
| **DATE & DATETIME** | Tableau recognizes dates in almost all formats. But in case we need to force Tableau to recognize a string as date, then we put a # sign before the data. | "02/01/2015"  "#3 March 1982" |

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.



Connect Live

The Connect Live feature is used for real-time data analysis. In this case, Tableau connects to real-time data source and keeps reading the data. Thus, the result of the analysis is up to the second, and the latest changes are reflected in the result. However, on the downside, it burdens the source system as it has to keep sending the data to Tableau.

In-Memory

Tableau can also process data in-memory by caching them in memory and not being connected to the source anymore while analyzing the data. Of course, there will be a limit to the amount of data cached depending on the availability of memory.

Combine Data Sources

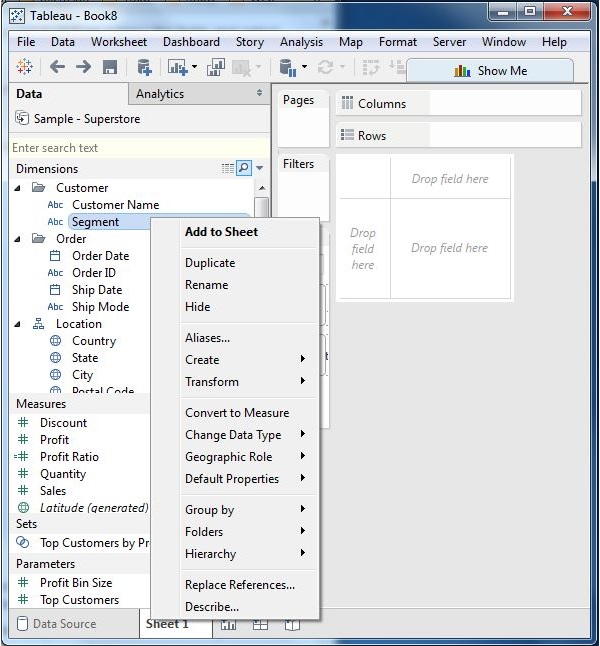
Tableau can connect to different data sources at the same time. For example, in a single workbook you can connect to a flat file and a relational source by defining multiple connections. This is used in data blending, which is a very unique feature in Tableau.

Tableau has many features to manipulate the fields present in Tableau data pane. You can rename the fields or combine two fields to create one field. Such operations help in better organization of the dimensions and measures, as well as accommodate two or more fields with the same name for better data analysis.

Following are the important examples of such Field Operations.

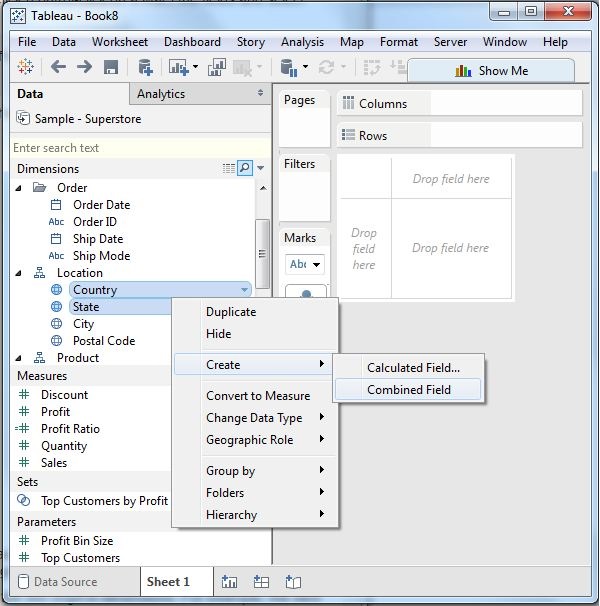
## Adding Fields to Worksheet

You can add any field to the worksheet by right-clicking and choosing the option Add to Sheet. You can also drag and drop the fields into different shelves present in the worksheet, like Columns shelf, Rows shelf, Filters shelf, and many other shelves under the Marks card. The following diagram shows the right-click option.



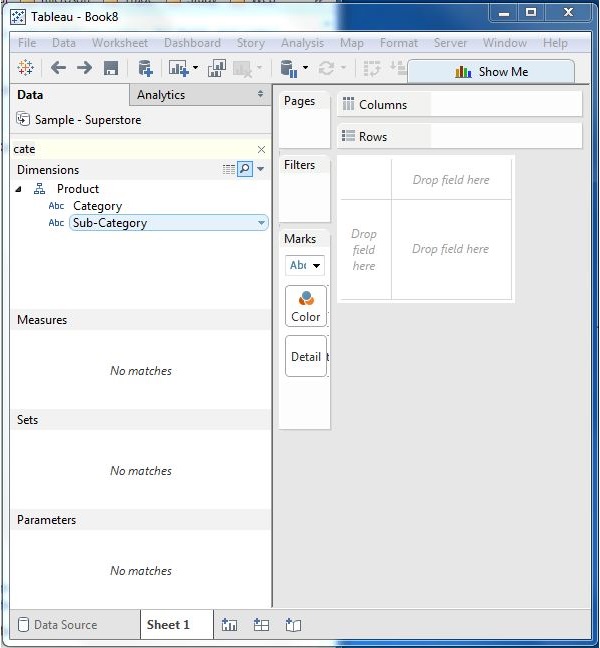
## Combining Two Fields

You can combine two dimension fields to create one field. This combined field has a name which is a combination of the individual fields. The values in the dimension get combined to a single value by joining the two strings into one string separated by a comma. However, this default name can be changed by using the rename field operation. The following diagram shows the step to combine two fields.



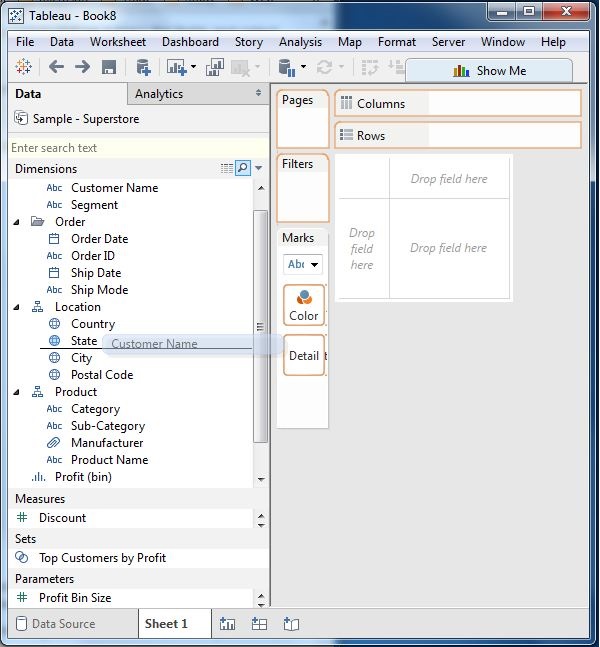
## Searching Fields

You can search for names of fields by using the search box option. Writing first three or more letters of the field name brings out the result showing only the fields whose name contains these letters.



## Reordering Fields

You can change the position of fields by simply dragging them up and down. In the following example, we drag the field customer name to the place between state and city. This is usually done to bring similar fields together which are frequently used for analysis.

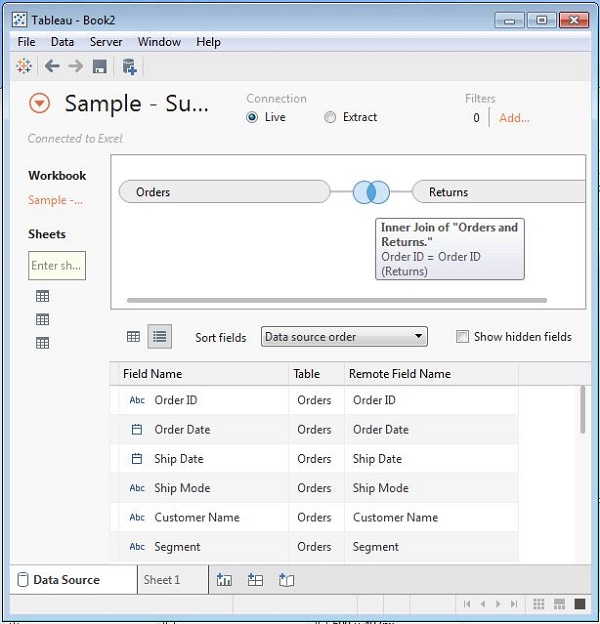


Data joining is a very common requirement in any data analysis. You may need to join data from multiple sources or join data from different tables in a single source. Tableau provides the feature to join the table by using the data pane available under Edit Data Source in the Data menu.

## Creating a Join

Consider the data source ‘Sample superstore’ to create a join between Orders and Returns table. For this, go to the Data menu and choose the option Edit Data Source. Next, drag the two tables, Orders and Returns to the data pane. Depending on the field name and datatype, Tableau will automatically create a join which can be changed later.

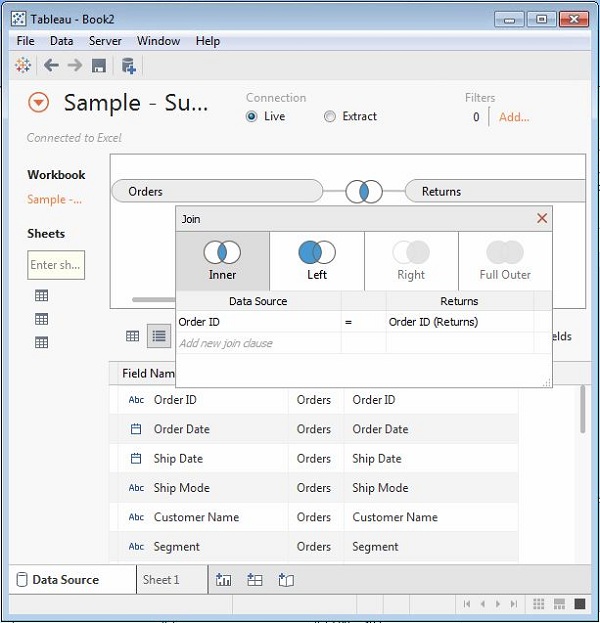
The following screenshot shows the creation of an inner join between Orders and Returns using the Field Order ID.



## Editing a Join Type

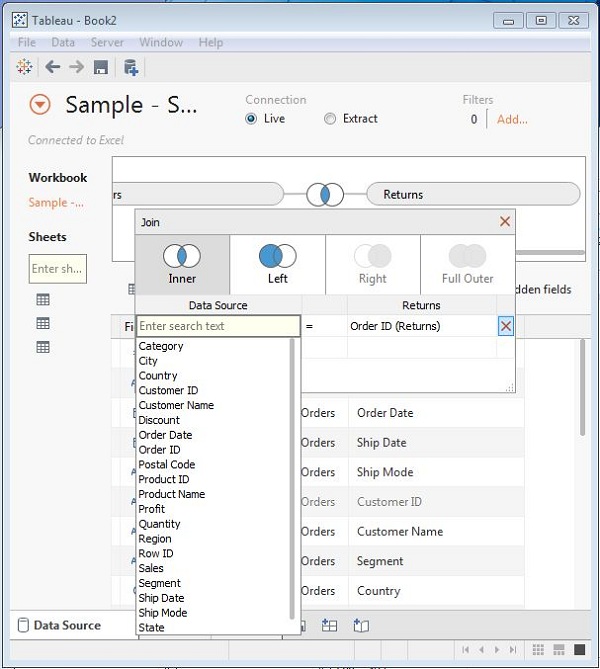
The type of join which the table creates automatically can be changed manually. For this, click the middle of the two circles showing the join. A popup window appears below which shows the four types of joins available. Also Tableau automatically greys out some types of joins, which it finds irrelevant on the basis of data present in the data source.

In the following screenshot, you can see the inner and left outer join as the available joins.



## Editing Join Fields

You can also change the fields forming the join condition by clicking the Data Source option available in the join popup window. While selecting the field, you can also search for the field you are looking for using a search text box.

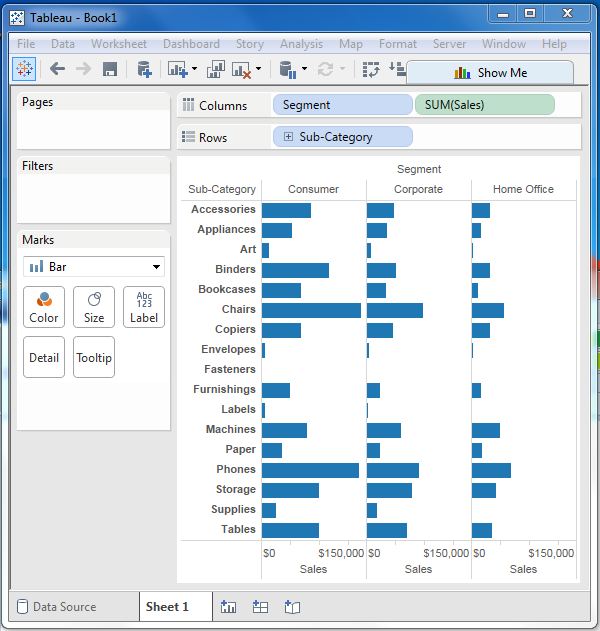


One of the important filtering options in Tableau is to apply some conditions to already existing filters. These conditions can be very simple like finding only those sales which are higher than a certain amount or it can be a complex one based on a certain formula. The conditions can also be applied to create a range filter.

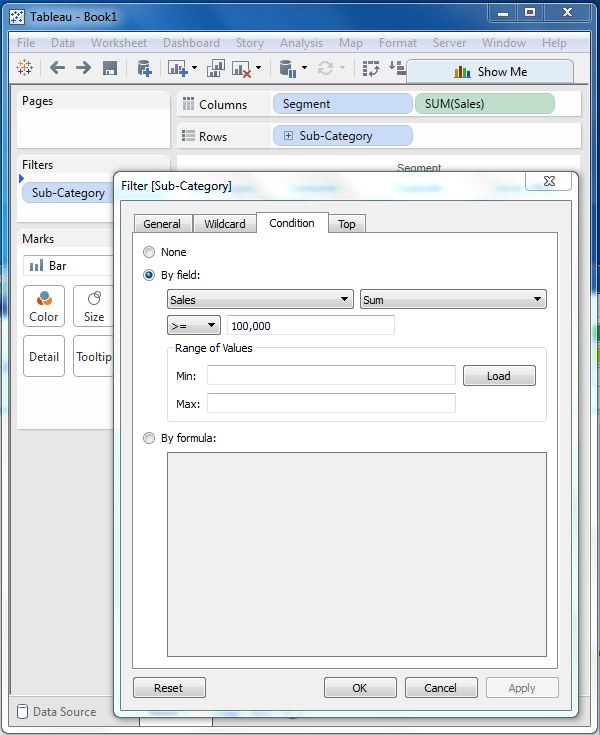
## Creating a Condition Filter

Using the Sample-superstore, let's find that sub-category of products across all segments whose sales exceed one million. To achieve this objective, following are the steps.

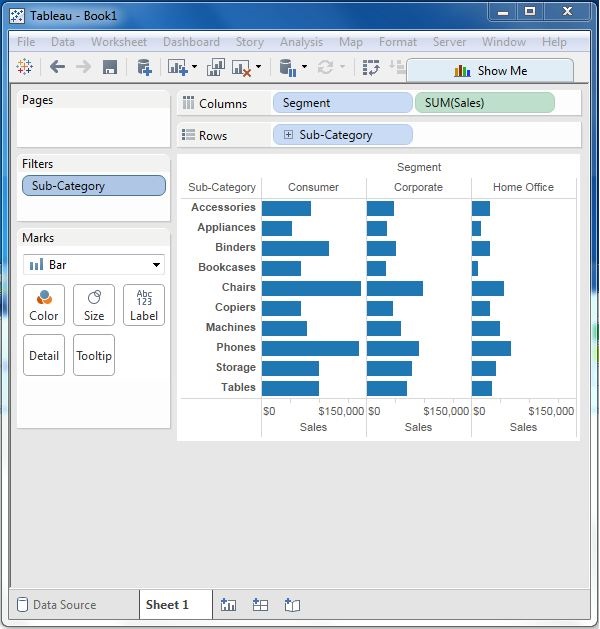
**Step 1** − Drag the dimension segment and the measure Sales to the Column shelf. Next, drag the dimension Sub-Category to the Rows shelf. Choose the horizontal bar chart option. You will get the following chart.



**Step 2** − Drag the dimension Sub-Category to the Filters Shelf. Right-click to edit and go to the tab Condition. Here, choose the radio option by field. From the drop-down, select Sales, Sum and greater than equal to symbol specifying the value 100000.



On completion of the above two steps, we get a chart which shows only those subcategory of products, which have the required amount of sale. Also this is shown for all the available segments where the condition is met.

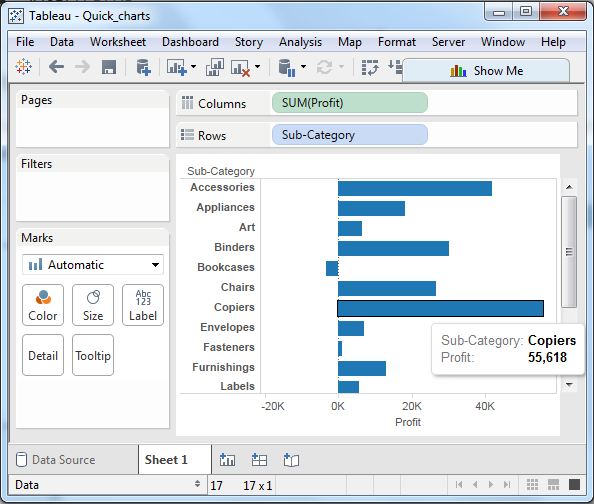


A bar chart represents data in rectangular bars with the length of the bar proportional to the value of the variable. Tableau automatically produces a bar chart when you drag a dimension to the Row shelf and measure to the Column shelf. We can also use the bar chart option present in the Show Me button. If the data is not appropriate for bar chart, then this option will be automatically greyed out.

In Tableau, various types of bar charts can be created by using a dimension and a measure.

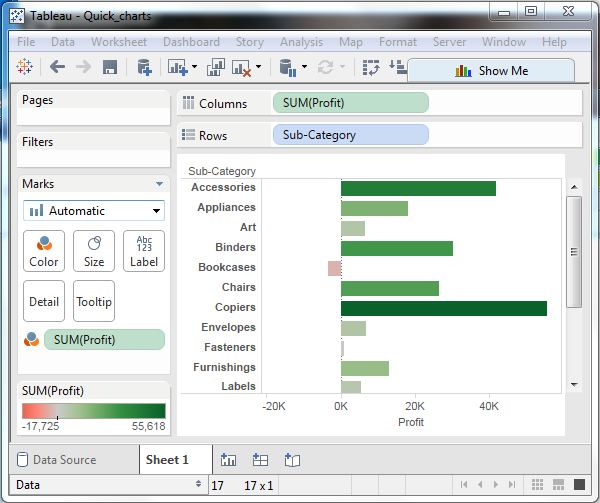
## Simple Bar Chart

From the Sample-Superstore, choose the dimension, take profit to the columns shelf and Sub-Category to the rows shelf. It automatically produces a horizontal bar chart as shown in the following screenshot. In case, it does not, you can choose the chart type from the Show Me tool to get the following result.



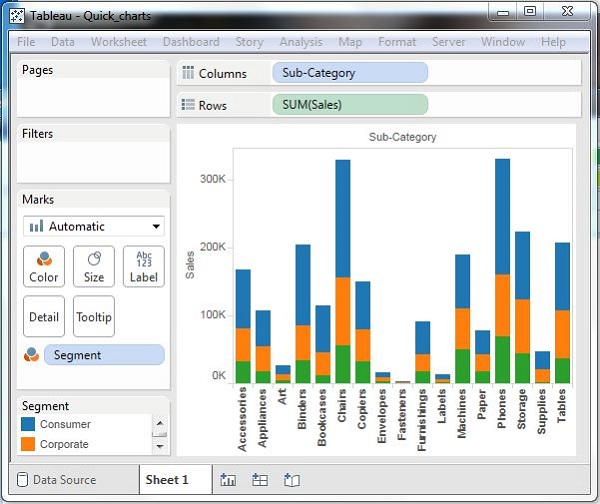
## Bar Chart with Color Range

You can apply colors to the bars based on their ranges. The longer bars get darker shades and the smaller bars get the lighter shades. To do this, drag the profit field to the color palette under the Marks Pane. Also note that, it produces a different color for negative bars.



## Stacked Bar Chart

You can add another dimension to the above bar chart to produce a stacked bar chart, which shows different colors in each bar. Drag the dimension field named segment to the Marks pane and drop it in colors. The following chart appears which shows the distribution of each segment in each bar.



**Level of Detail (LOD) expressions** are used to run complex queries involving many dimensions at the data source level instead of bringing all the data to Tableau interface. A simple example is adding dimension to an already calculated aggregate value.

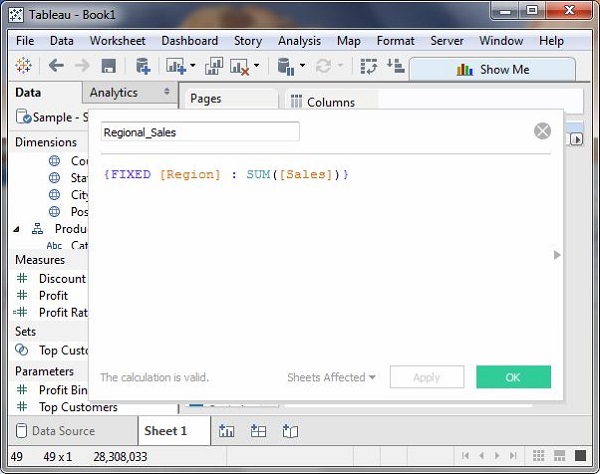
Types of LOD

There are three main types of LOD expressions.

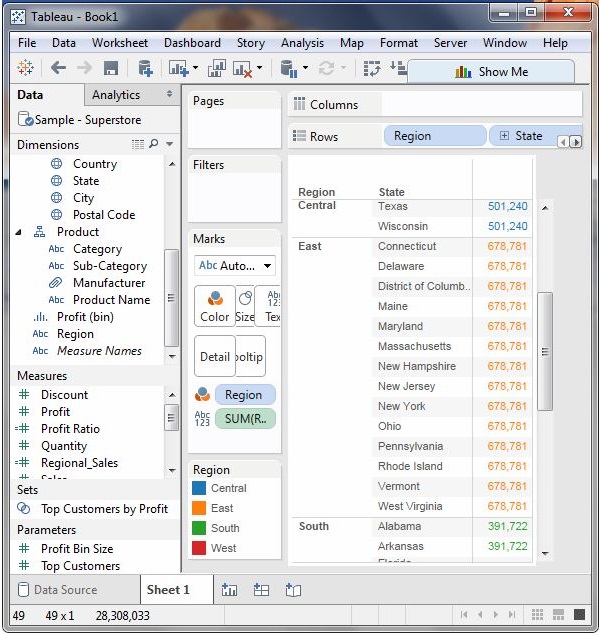
* **FIXED LOD** This expression computes values using the specified dimensions without reference to any other dimensions in the view.
* **INCLUDE LOD** This level of detail expressions compute values using the specified dimensions in addition to whatever dimensions are in the view.
* **EXCLUDE LOD** These levels of detail expressions subtract dimensions from the view level of detail.

FIXED Level of Detail Expressions

Find the amount of Sales for each state in each region. Here, first create the formula field named Regional Sales using the formula as shown in the following screenshot.



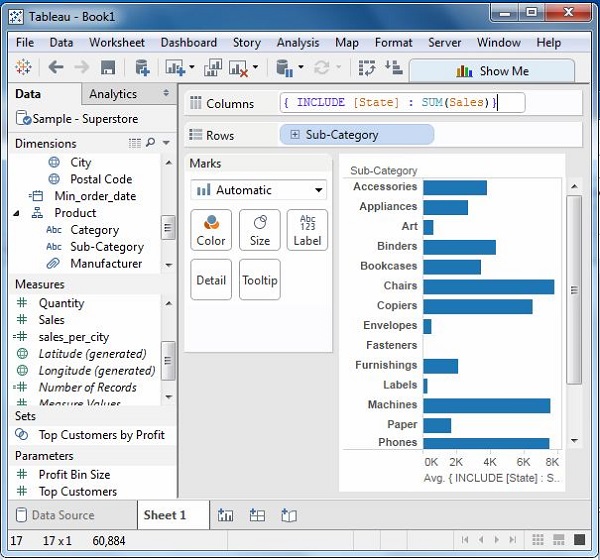
Next, drag the Region and State field to the Rows shelf and the calculated field to the Text shelf under the Marks card. Also drag the Region field to the Color shelf. This produces the following view, which shows a fixed value for different states. That is because we have fixed the dimension as region for the calculation of Sales value.



INCLUDE Level of Detail Expressions

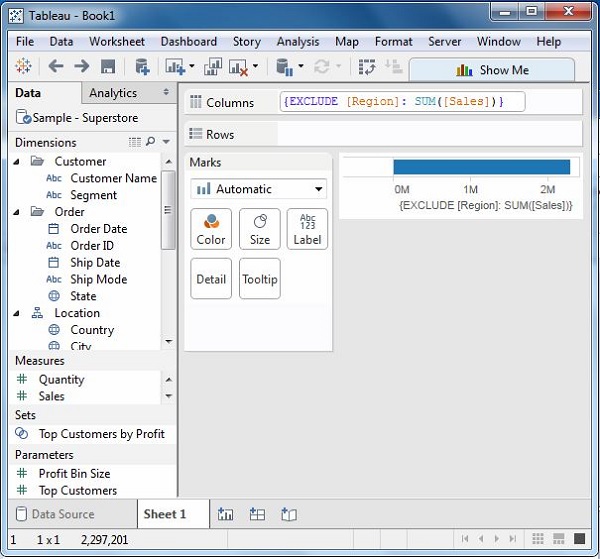
INCLUDE level of detail expressions compute values using the specified dimensions in addition to whatever dimensions are in the view. Calculate the sum of sales per state for each sub-category of products.

For this, drag the Sub-Category field to the Rows shelf. Then, write the expression in the Columns shelf as shown in the following screenshot. It produces the following view which includes both the dimensions in the calculations.



EXCLUDE Level of Detail Expressions

EXCLUDE level of detail expressions specify dimensions to exclude from the view level of detail. Exclude Region from Sales figure calculated for every month. Create the formula as shown in the following screenshot.



On dragging the relevant fields to the respective shelves, you will get the final view for the EXCLUDE LOD as shown in the following screenshot.

