ACS College of Engineering

(Affiliated Visvesvaraya Technological University, Belagavi , Approved by AICTE, UGC & GoK, Accredited by ISO 9001-2015 Certified Institution)

Sponsored by: MOOGAMBIGAI CHARITABLE AND EDUCATIONAL TRUST

Department of Computer Science and Engineering

LABORATORY MANUAL

DATA VISUALIZATION LABORATORY
(BAIL504)

V SEMESTER

DATA VISUALIZATION LABORATORY

Sub Code: BCS358C   **IA Marks: 50**
Total Hours: 25   **Hours/Week: 02**

DATA VISUALIZATION LAB		Semester	V
Course Code	**BAIL504**	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	0:0:2:0	SEE Marks	50
Credits	01	Exam Hours	100
Examination type (SEE)	Practical		

Course objectives:
- Understand the Importance of data Visualization for business intelligence and decision making.
- Learn different approaches to understand the importance of visual perception.
- Learn different data visualization techniques and tools.
- Gain knowledge of effective data visuals to solve workplace problems.

Sl.NO	Experiments
1	**Getting Started -** Tableau Workspace, Tableau terminologies, basic functionalities.
2	**Connecting to Data Source –** Connecting to Database, Different types of Tableau Joins.
3	**Creating a View -** formatting charts, adding filters, creating calculated fields and defining parameters.
4	**Dashboard Design and Storytelling –** Components of Dashboard, Understanding how to place worksheets in Containers, Action filters and its types.
5	**Introducing Power BI –**Components and the flow of work. Power BI Desktop Interface-The Report has five main areas.
6	**Querying Data from CSV -** Query Editor, Connecting the data from the Excel Source, Clean, Transform the data.
7	**Creating Reports & Visualizations -** Different types of charts, Formatting charts with Title, Colors.
8	**Dashboards -** Filters in Power BI, Formatting dashboards.
9	Analysis of revenue in sales dataset: i) Create a choropleth map (fill the map) to spot the special trends to show the state which has the highest revenue. ii) Create a line chart to show the revenue based on the month of the year. iii) Create a bin of size 10 for the age measure to create a new dimension to show the revenue. iv) Create a donut chart view to show the percentage of revenue per region by creating zero access in the calculated field. v) Create a butterfly chart by reversing the bar chart to compare female & male revenue based on product category. vi) Create a calculated field to show the average revenue per state & display profitable & non-profitable state. vii) Build a dashboard.

10	Analysis of GDP dataset:
	   i) Visualize the countries data given in the dataset with respect to latitude and longitude along with country name using symbol maps.
	   ii) Create a bar graph to compare GDP of Belgium between 2006 – 2026.
	   iii) Using pie chart, visualize the GDP of India, Nepal, Romania, South Asia, Singapore by the year 2010.
	   iv) Visualize the countries Bhutan & Costa Rica competing in terms of GDP.
	   v) Create a scatter plot or circle views of GDP of Mexico, Algeria, Fiji, Estonia from 2004 to 2006.
	   vi) Build an interactive dashboard

11	Analysis of HR Dataset:
	   i) Create KPI to show employee count, attrition count, attrition rate, attrition count, active employees, and average age.
	   ii) Create a Lollipop Chart to show the attrition rate based on gender category.
	   iii) Create a pie chart to show the attrition percentage based on Department Category- Drag department into colours and change automatic to pie. Entire view, Drag attrition count to angle. Label attrition count, change to percent, add total also, edit label.
	   iv) Create a bar chart to display the number of employees by Age group,
	   v) Create a highlight table to show the Job Satisfaction Rating for each job role based on employee count.
	   vi) Create a horizontal bar chart to show the attrition count for each Education field Education field wise attrition – drag education field to rows, sum attrition count to col,
	   vii) Create multiple donut chart to show the Attrition Rate by Gender for different Age group.

12	Analysis of Amazon Prime Dataset:
	   i) Create a Donut chart to show the percentage of movie and tv shows
	   ii) Create a area chart to shows by release year and type
	   iii) Create a horizontal bar chart to show Top 10 genre
	   iv) Create a map to display total shows by country
	   v) Create a text sheet to show the description of any movie/movies.
	   vi) Build an interactive Dashboard.

	Course outcomes (Course Skill Set):
	At the end of the course the student will be able to:
	  1. Design the experiment to create basic charts and graphs using Tableau and Power BI.
	  2. Develop the solution for the given real world problem.
	  3. Analyze the results and produce substantial written documentation.

Introduction to Various Data Visualization tools

Data visualization is the practice of translating information into a visual context, such as a map or graph, to make data easier for the human brain to understand and pull insights from. It is the representation of information and data through use of common graphics, such as charts, plots, infographics, and animations. Data visualization is a powerful way for people, especially data professionals, to display data so that it can be interpreted easily.

Data Visualization enables decision-makers of any enterprise or industry to look into analytical reports and understand concepts that might otherwise be difficult to grasp.

Benefits of Data Visualization:

1. It is easy to understand the information with graphics

2. It made data to be represented in attractive way

3. Shows complex relationships

4. Helps to process large datasets

5. Useful for identifying trends

6. Minimizes ambiguity

Data visualization tools provide the ability to see and understand data trends, outliers, and patterns in an easy, intuitive way. There are various data visualization tools available. One must choose the tool based on various factors such as its ease of use, types of graphical representations the tool can produce, size of the dataset the tool can handle etc. some of Data Visualization tools are Tableau, Power BI, Google Charts, Jupyter, Grafana etc.

The following are some common types of data visualizations:

1. **Table:** A table is data displayed in rows and columns, which can be easily created in a Word document or Excel spreadsheet.

2. **Chart or graph:** Information is presented in tabular form with data displayed along an x and y axis, usually with bars, points, or lines, to represent data in comparison.

3. **Geospatial visualization:** Data is depicted in map form with shapes and colours that illustrate the relationship between specific locations, such as a choropleth or heat map.

4. **Dashboard:** Data and visualizations are displayed, usually for business purposes, to help analysts understand and present data.

Introduction to Tableau and Installation

Tableau is a data visualization tool that provides pictorial and graphical representations of data. It is used for data analytics and business intelligence. Tableau provides limitless data exploration without interrupting flow of analysis.

With an intuitive drag and drop interface, user can uncover hidden insights in data and make smarter decisions faster.

Tableau is a Business Intelligence tool for visually analyzing the data. Users can create and distribute an interactive and shareable dashboard, which depict the trends, variations, and density of the data in the form of graphs and charts. Tableau can connect to files, relational and Big Data sources to acquire and process data. The software allows data blending and real-time collaboration, which makes it very unique. It is used by businesses, academic researchers, and many government organizations for visual data analysis. It is also positioned as a leader Business Intelligence and Analytics Platform in Gartner Magic Quadrant.

As a leading data visualization tool, Tableau has many desirable and unique features. Its powerful data discovery and exploration application allows you to answer important questions in seconds. You can use Tableau's drag and drop interface to visualize any data, explore different views, and even combine multiple databases easily. It does not require any complex scripting. Anyone who understands the business problems can address it with a visualization of the relevant data. After analysis, sharing with others is as easy as publishing to Tableau Server.

Tableau Features

- **Speed of Analysis** − As it does not require high level of programming expertise, any user with access to data can start using it to derive value from the data.

- **Self-Reliant** − Tableau does not need a complex software setup. The desktop version which is used by most users is easily installed and contains all the features needed to start and complete data analysis.

- **Visual Discovery** − The user explores and analyzes the data by using visual tools like colors, trend lines, charts, and graphs. There is very little script to be written as nearly everything is done by drag and drop.

- **Blend Diverse Data Sets** − Tableau allows you to blend different relational, semi structured and raw data sources in real time, without expensive up-front integration costs. The users don't need to know the details of how data is stored.

- **Architecture Agnostic** − Tableau works in all kinds of devices where data flows. Hence, the user need not worry about specific hardware or software requirements to use Tableau.

- **Real-Time Collaboration** − Tableau can filter, sort, and discuss data on the fly and embed a live dashboard in portals like SharePoint site or Salesforce. You can save your view of data and allow colleagues to subscribe to your interactive dashboards so they see the very latest data just by refreshing their web browser.

- **Centralized Data** − Tableau server provides a centralized location to manage all of the organization's published data sources. You can delete, change permissions, add tags, and manage schedules in one convenient location. It's easy to schedule extract refreshes and manage them in the data server. Administrators can centrally define a schedule for extracts on the server for both incremental and full refreshes.

Working on Tableau ---Some important screen shots and steps of sheets (Not Related to Programs)
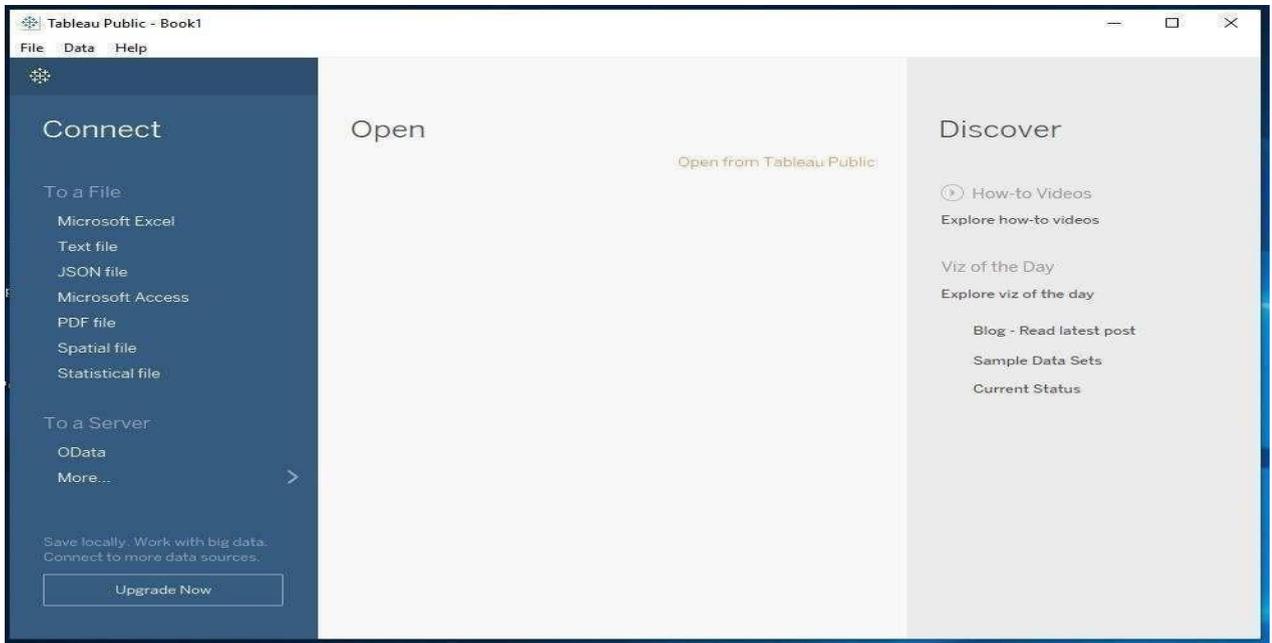

Connecting to Data and preparing data for visualization in Tableau

Tableau supports connecting to a wide variety of data, stored in a variety of places. For example, data might be stored on computer in a spread sheet or a text file, or in a big data, relational, or cube (multidimensional) database on a server in enterprise or the data can be from a public domain available on the web.

Data can be imported in Tableau Public from Connect panel on left side. For example, an Excel sample dataset was loaded into Tableau as follows:
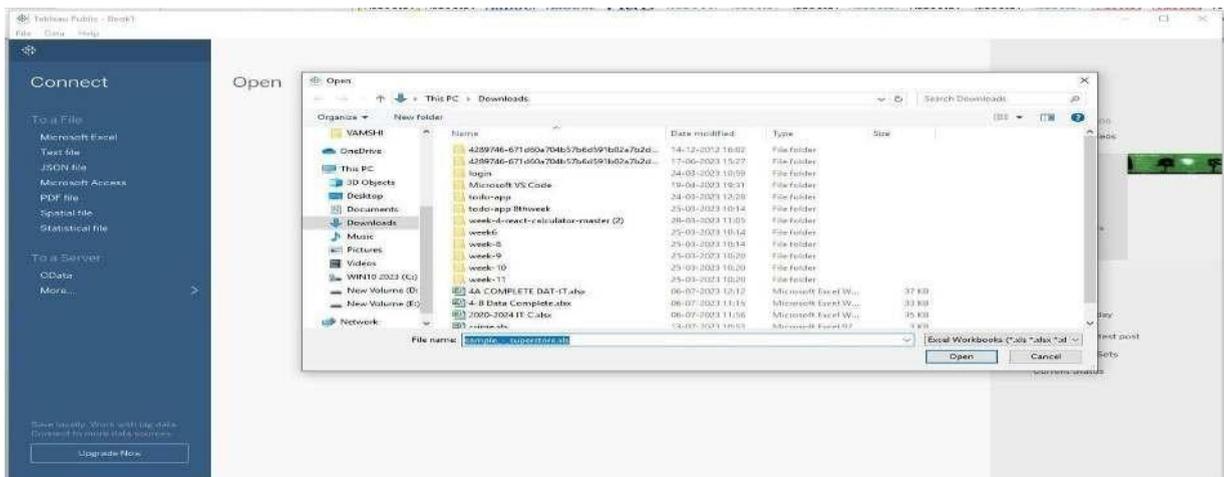
After clicking on open, screen is as follows:
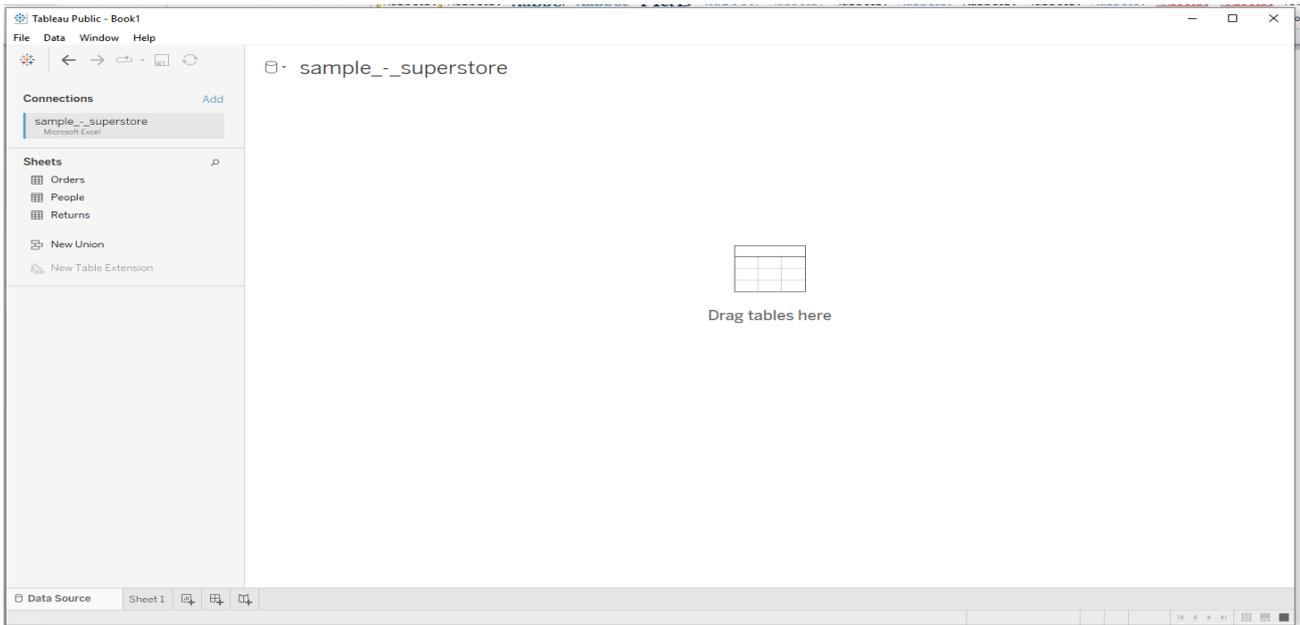
The data store page appears as above. The left pan shows that above dataset consists of 3 worksheets. If we drag orders table, screen appears as follows: Tableau automatically identifies the data type of each column.
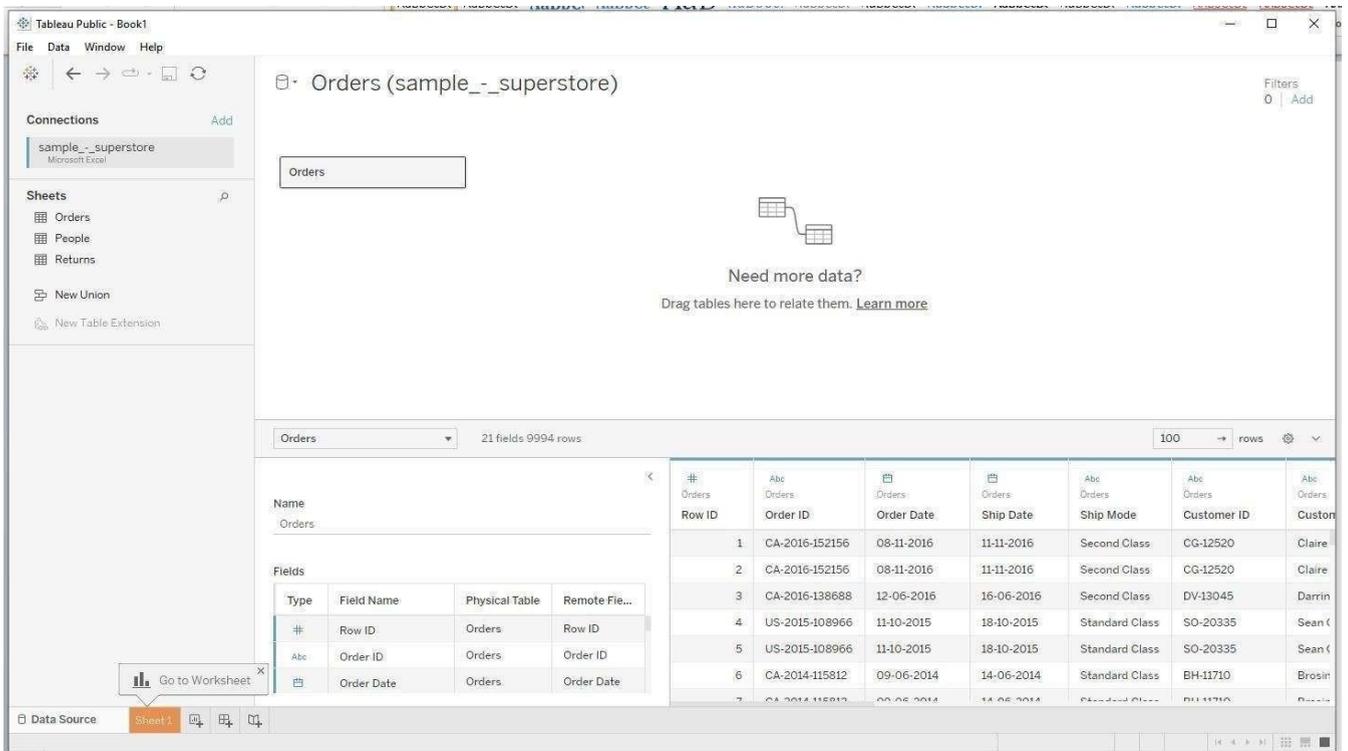
Now drag Returns table onto the Canvas to the right of Orders table. This shows the relation between the two tables Orders and Returns.

If we click on the link between Orders and Returns table names at the top gives the

summary of the relationship between the tables. Now rename the data store and click on Sheet1 at the bottom left to proceed. This step creates a data extract which improves query performance
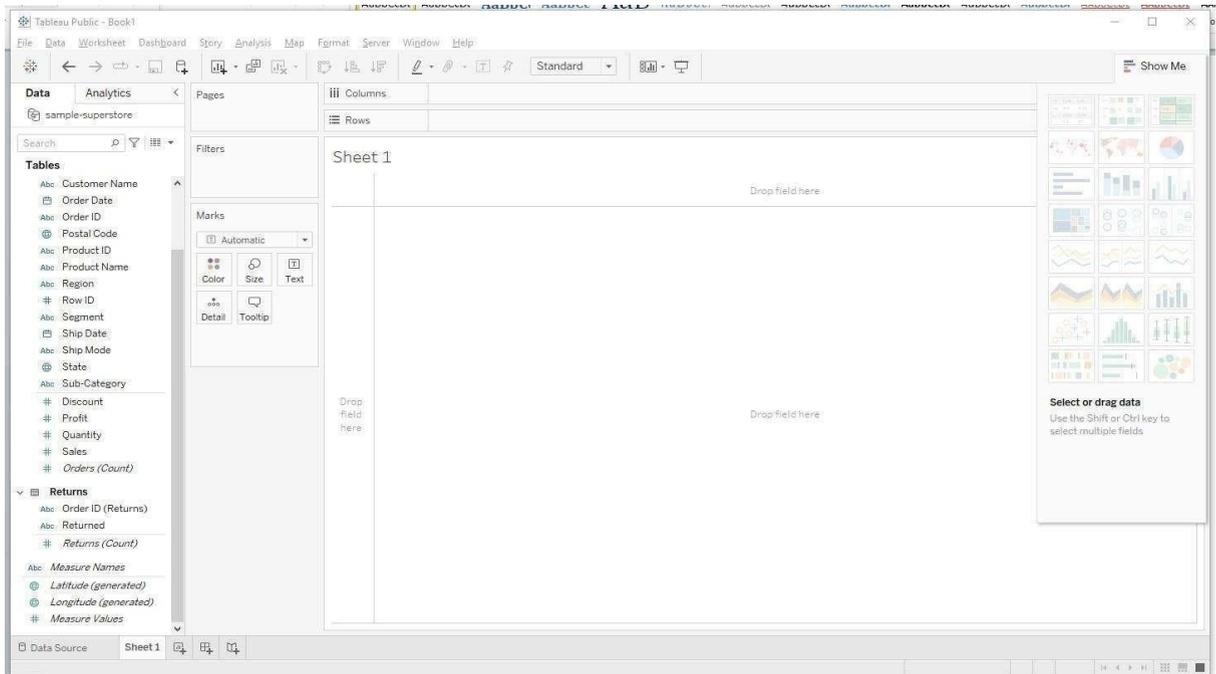
Data aggregation and statistical functions

We can apply various aggregation and statistical functions on data such as count, minimum, maximum, standard deviation, variance etc. This is shown below. This can be done by right clicking on the required field of dataset, click on Default properties and click on aggregation.
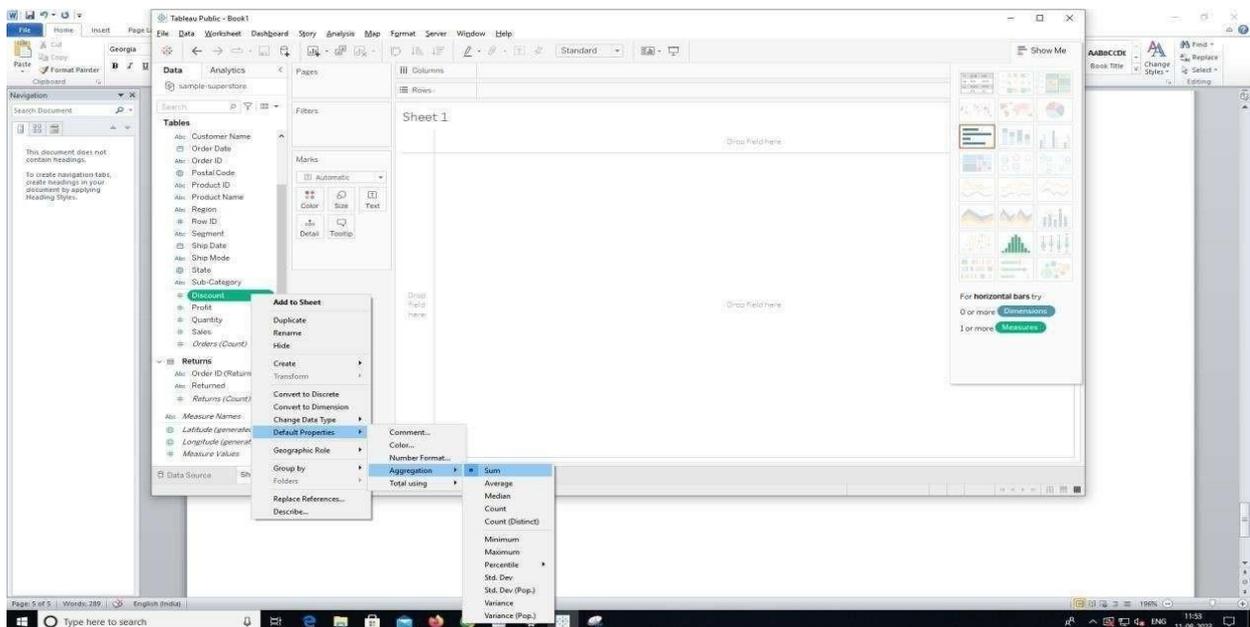
Or the above operation can be done by creating a calculated field as shown below. To create a calculated field, click on the down arrow button beside search tab above Tables panel, drag a field to that calculated field window.
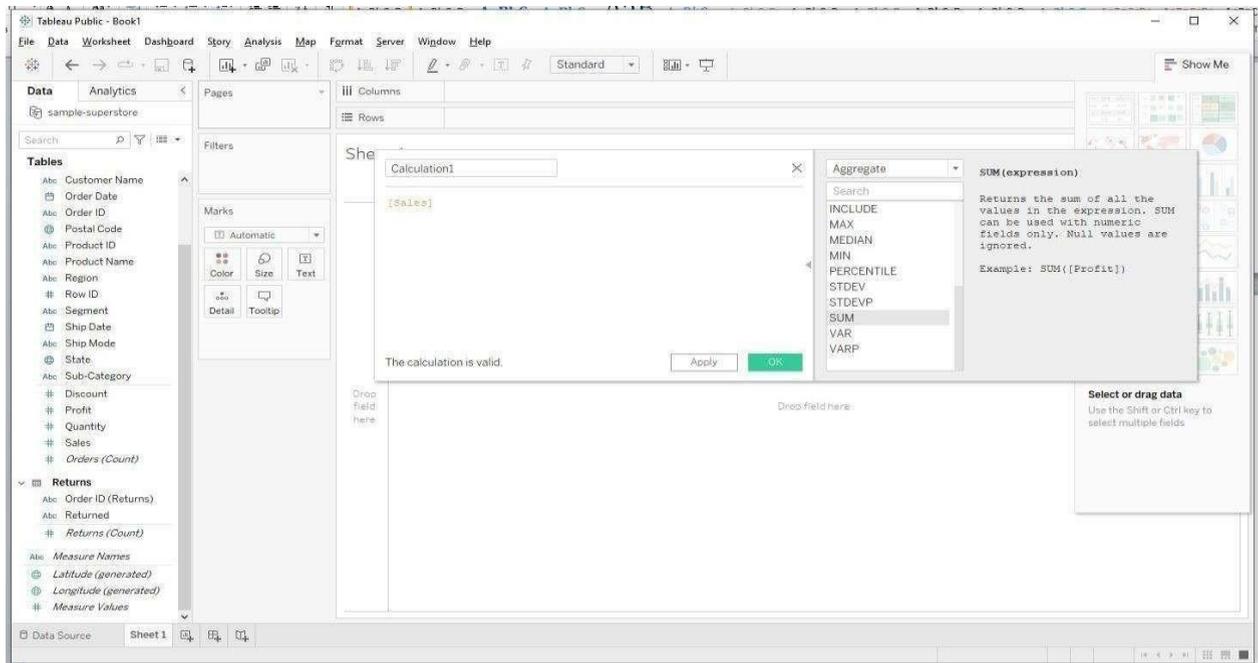
Then click on apply and results are shown below:

Dept. of CS & E (Data Science)
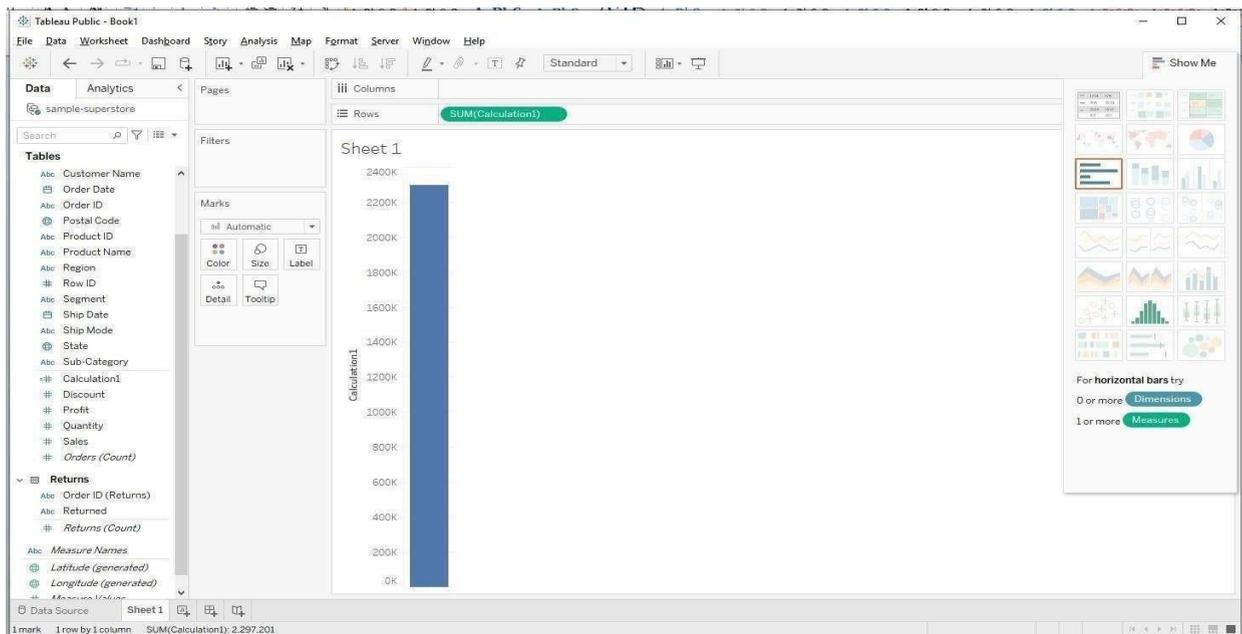
In the same way we can apply any aggregate or statistical function on data with the help of calculated fields.

Data Visualization

we can perform various visualization operations on data in Tableau. Some of them are bar cart, histogram, bubble chart, gantt chart, scatter plot, heat map etc.

Bar chart:

Bar charts can be created in 3 variations in Tableau: Horizontal bars, stacked bars, side-by-side bars. Horizontal bars can be created by selecting that type of chart from Show Me menu on right hand side of Canvas. The type of chart in box on right hand side represents horizontal bar graph.
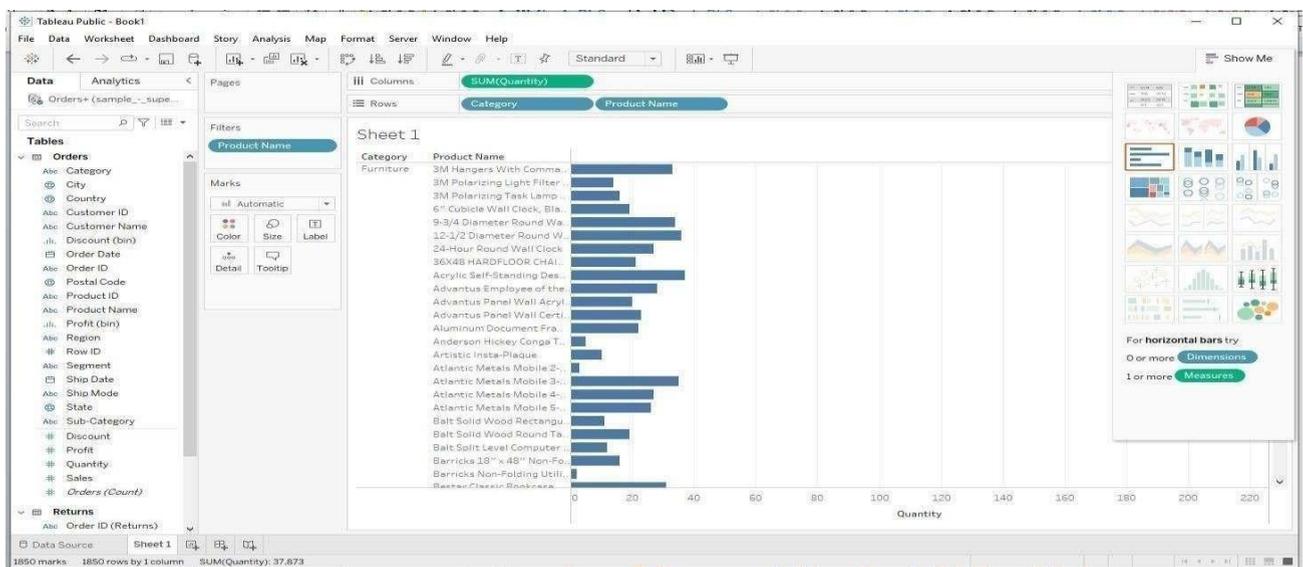
Side-by-side bar chart can be created in following way.

Dept. of CS & E (Data Science)
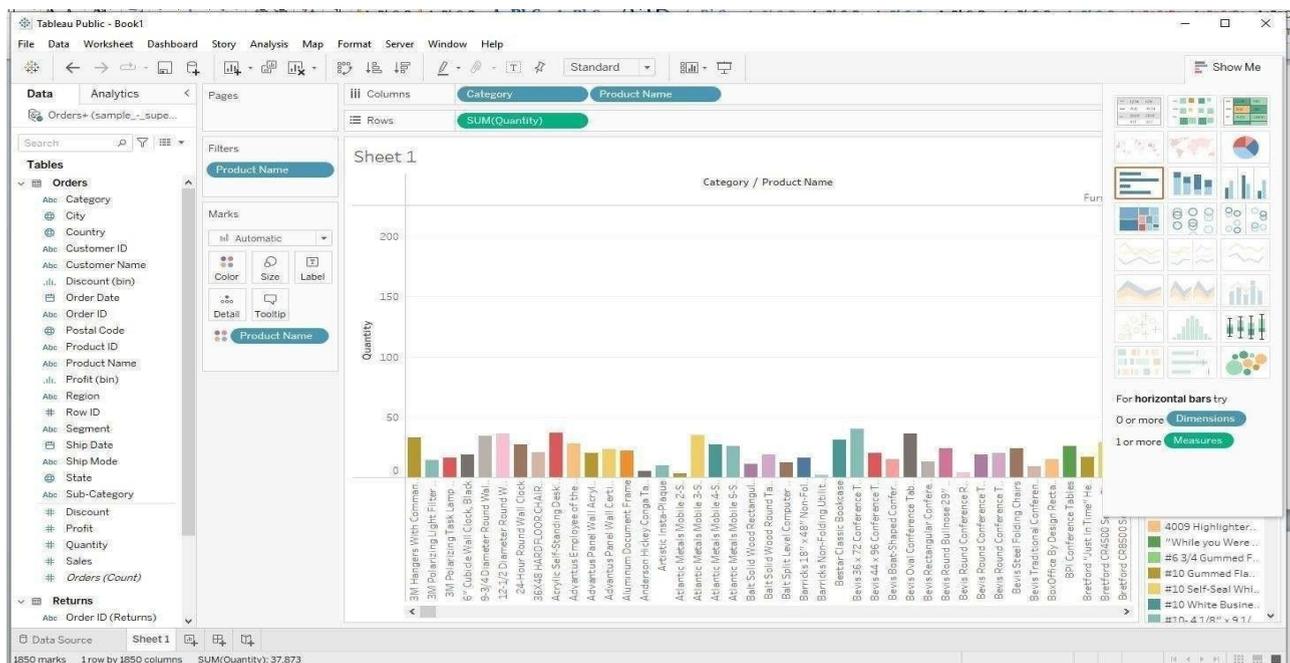
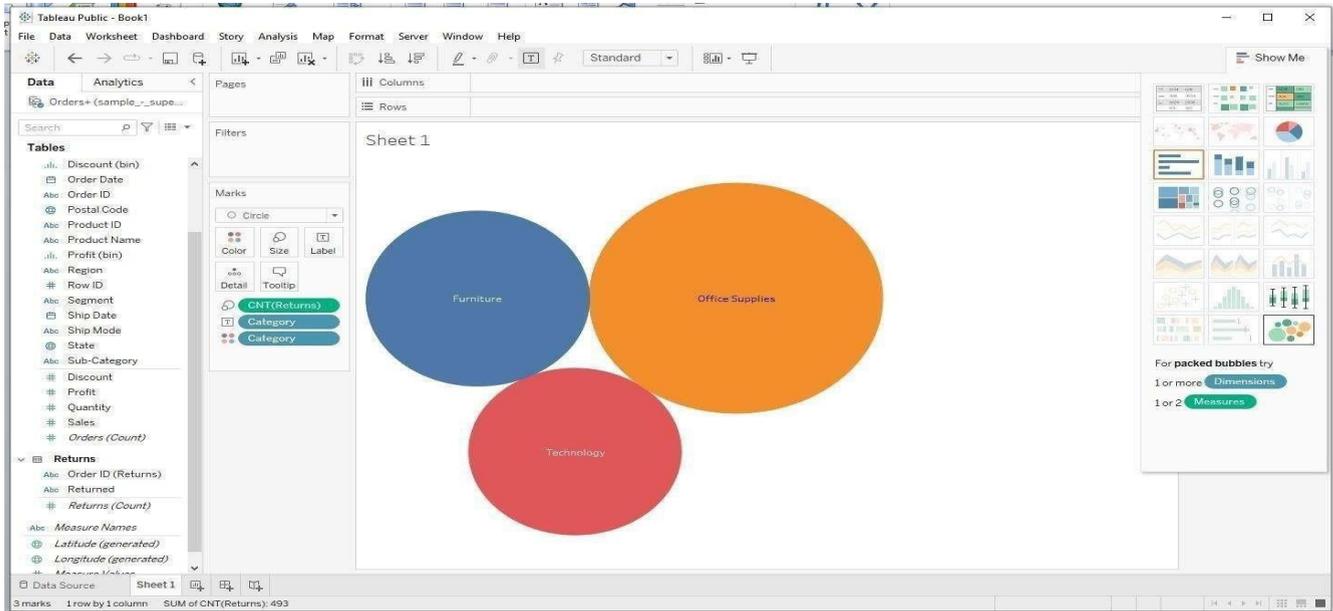**Line graph:** Line graph can be continuous or discrete
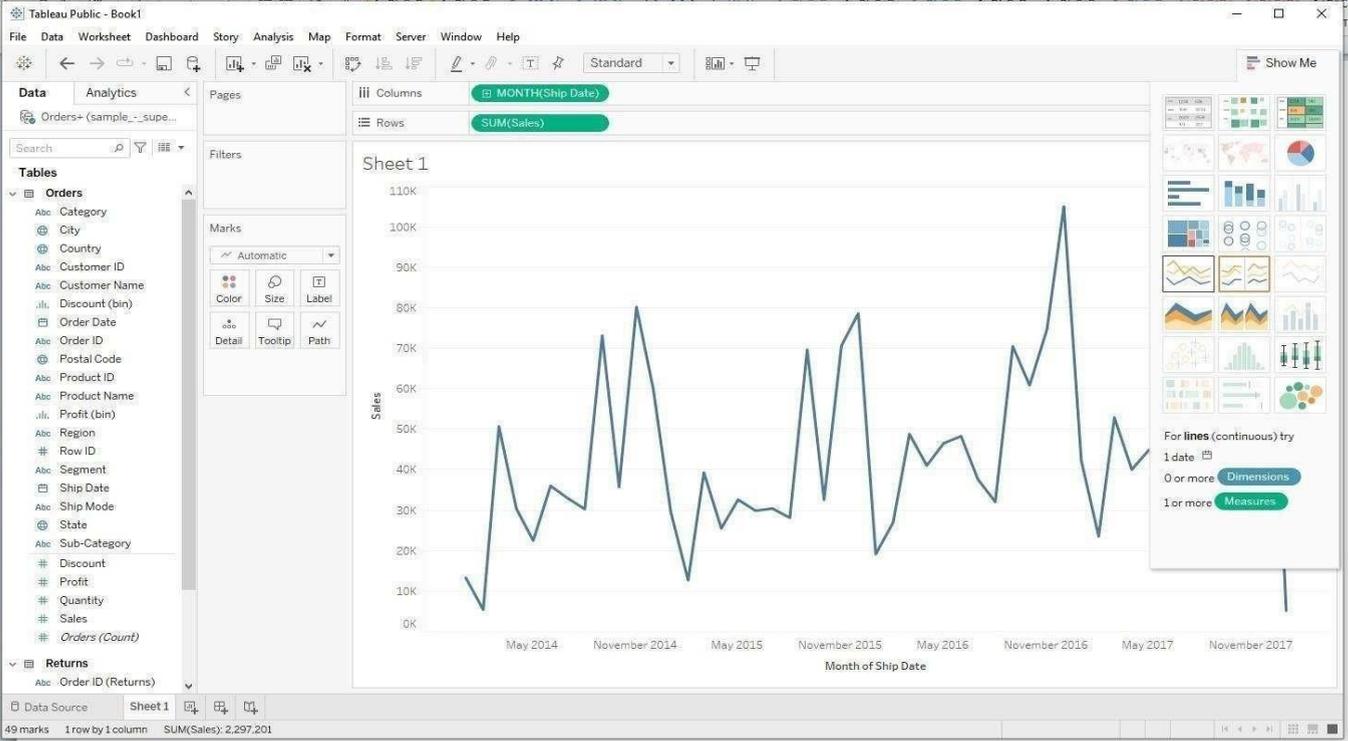
Discrete line graph is shown below:
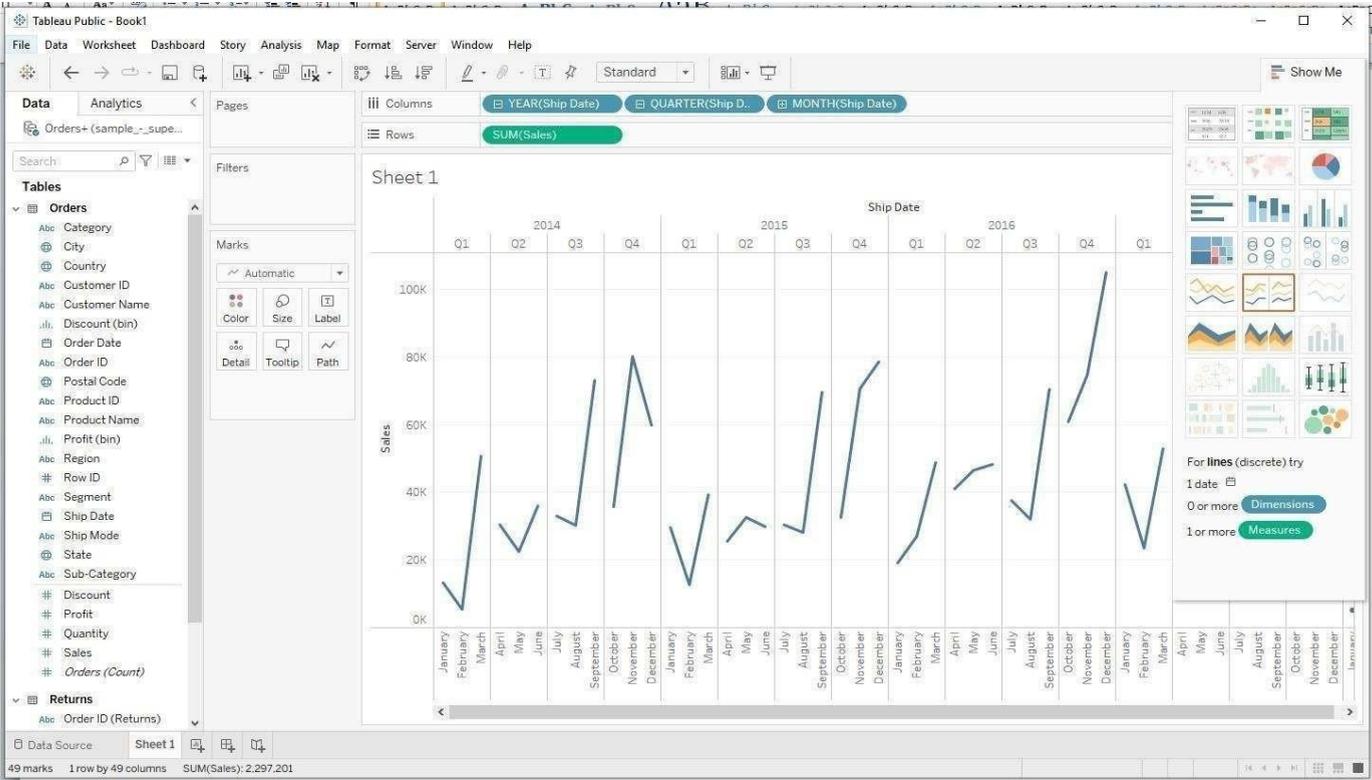
Pie chart:
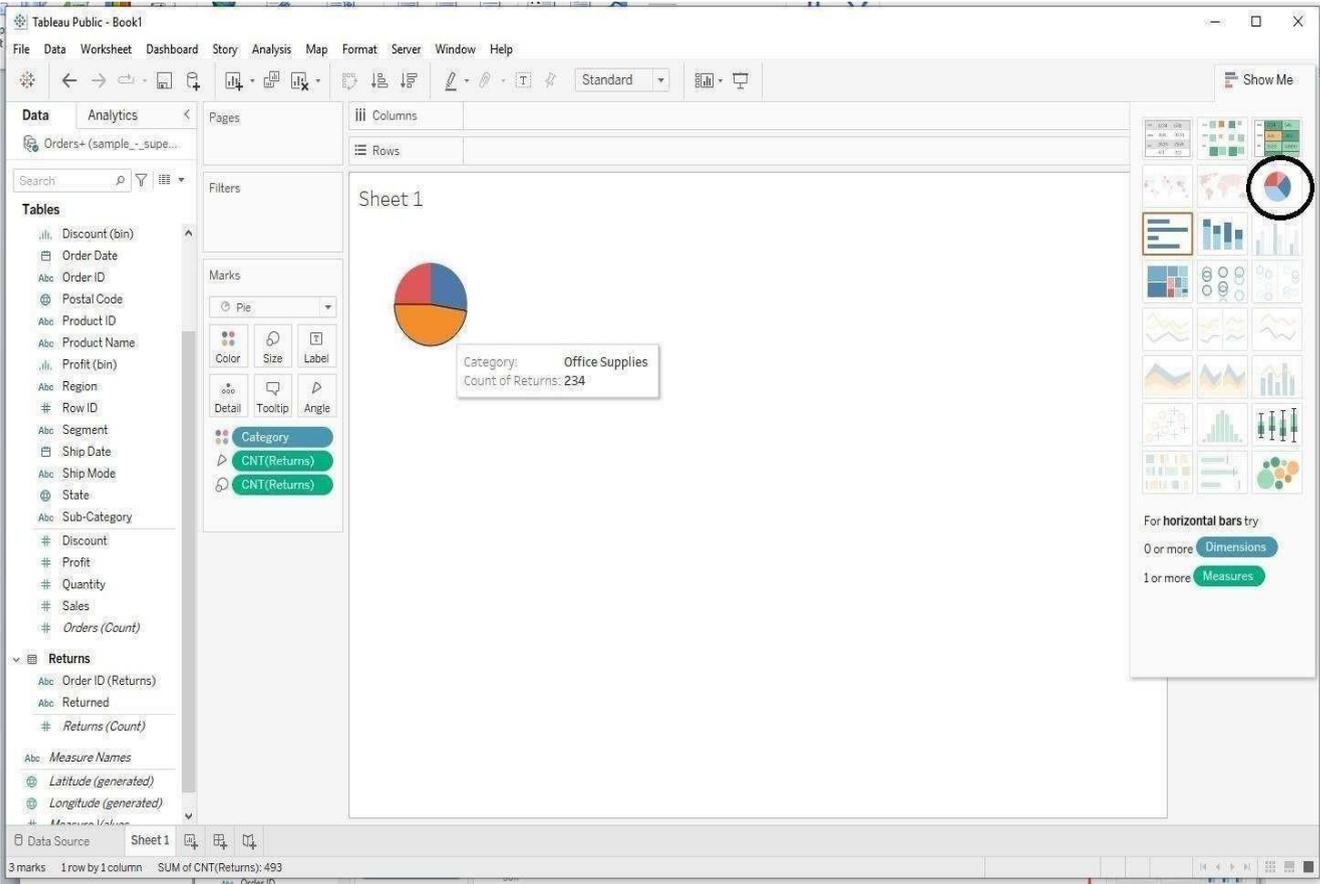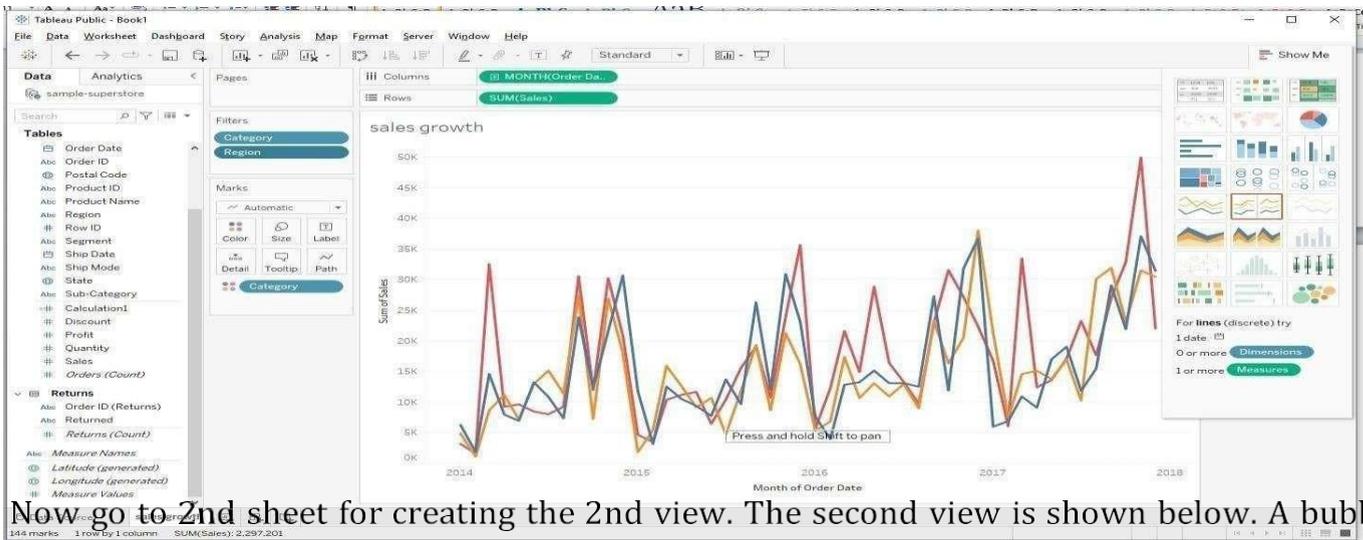
Bubble chart:

Dashboards

Dashboard is a way of displaying various types of visual data in one place. Usually, a dashboard is intended to convey different, but related information in an easy-to-digest form. And oftentimes, this includes things like key performance indicators (KPI)s or other important business metrics that stakeholders need to see andunderstand at a glance.

Dashboards are useful across different industries and verticals because they're highly customizable. They can include data of all sorts with varying date ranges to help you understand: what happened, why it happened, what may happen, and what action should be taken.

For example, category of sales across months in a year, region is the field added. The first view is shown below. This can be renamed at the bottom of the screen.
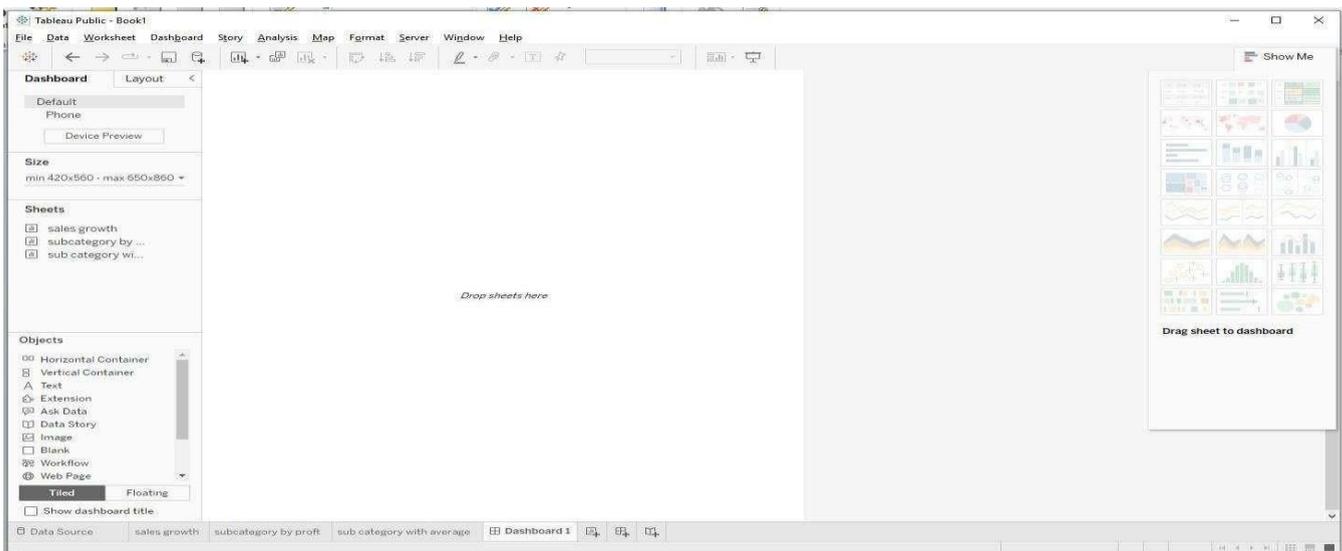

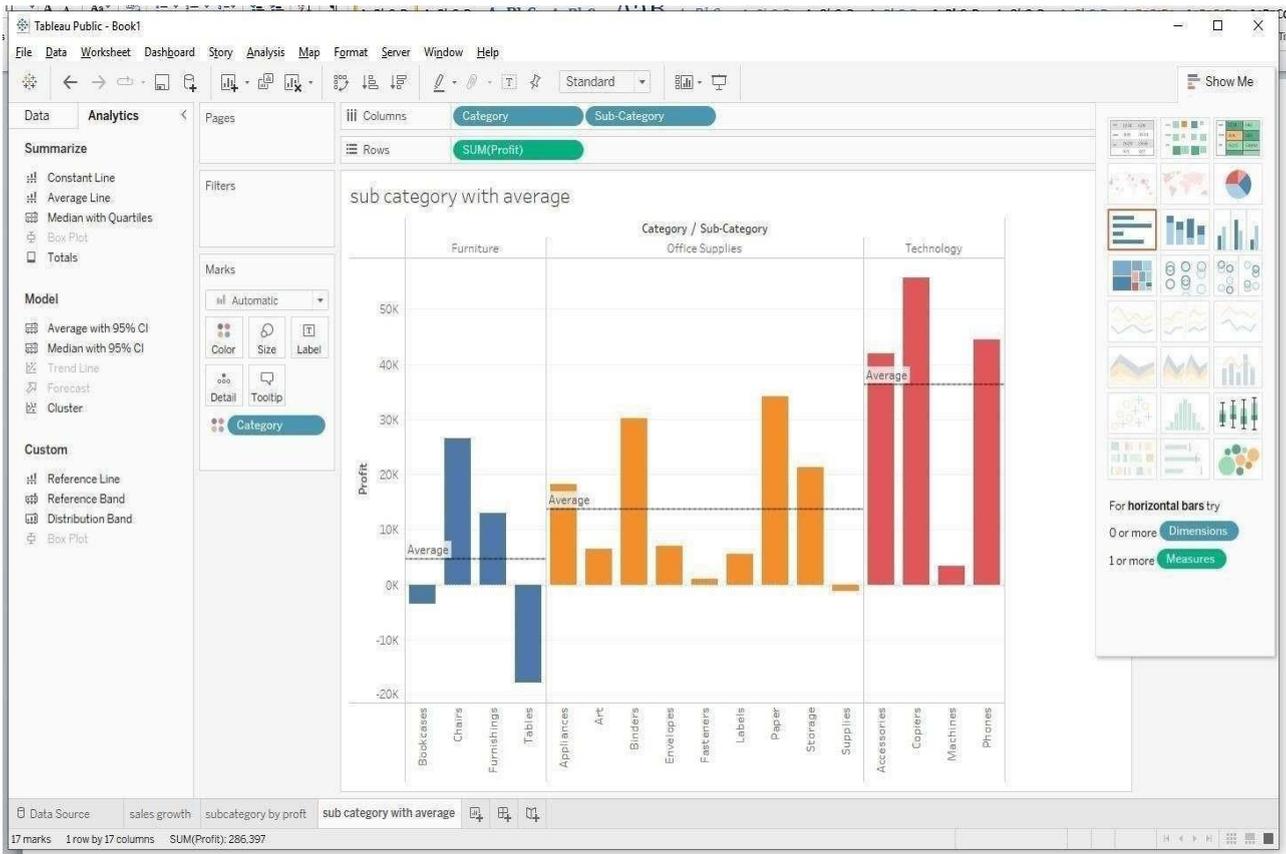Now go to 2nd sheet for creating the 2nd view. The second view is shown below. A bubble

chart was drawn between profit and subcategory. Then rename the sheet.
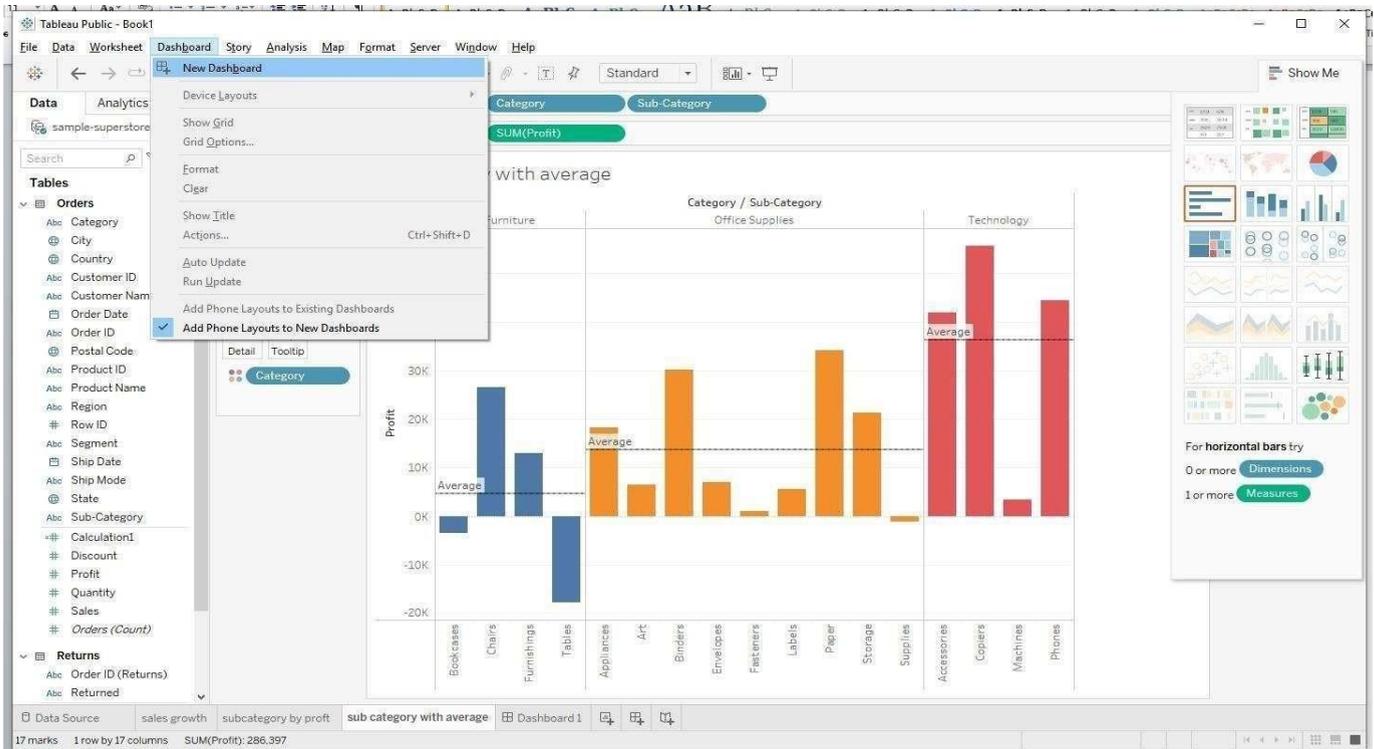
Next 3rd view is created as follows for profit for each subcategory in the category with averages.


After creating individual views, now a Dashboard can be created by clicking on create dashboard at the toolbar.

After clicking on new dashboard option, the screen is shown below.

Now the sheets or views which are created earlier can be drag and dropped on this dashboard. The above three created views are placed in the dashboard as follows. One can follow their own way of importing sheets on the dashboard. After creating dashboard, title can be given to the dashboard from Dashboard tab. Dashboard can be customized in terms of its appearance by the user if required. Dashboard once created can be saved on user's system and can be retrieved whenever required.
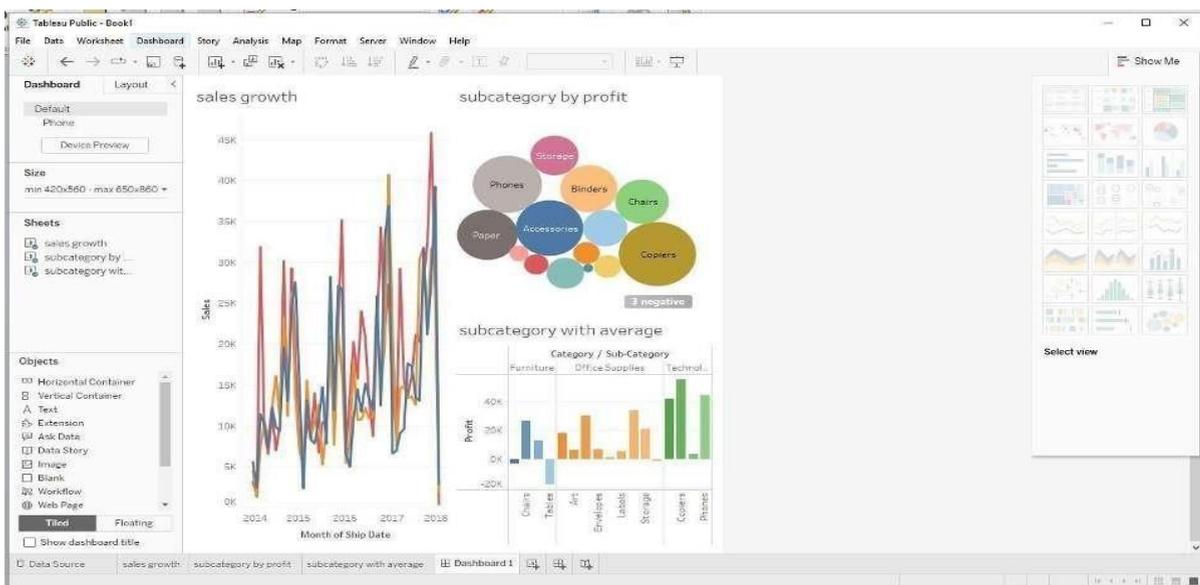
POWER BI

What is Power BI?

Power BI is a business intelligence tool that allows you to connect to various data sources, visualize the data in reports and dashboards, and then share them with anyone you want.

Power BI is a Data Visualization and Business Intelligence tool that converts data from different data sources to interactive dashboards and BI reports.

What is Power BI Used For

Power BI is a tool in the category of Business Intelligence (BI). The purpose of BI is to track Key Performance Indicators (KPIs) and uncover insights in business data so as to better inform decision-making across the organization.

Power BI is used in different ways depending on the role of the individual, from developers, analysts, managers, and directors, to everyone in between.

How Does Power BI Compare to Other Tools Like Tableau and Excel?

Power BI and Tableau are both business intelligence tools and have a lot of overlap in terms of their capabilities. There are 2 key differences between Power BI and Tableau:

1. Power BI only works on Windows, whereas Tableau supports both Windows and MacOS.
2. Pricing options differ between Power BI and Tableau. However, Tableau is generally the more expensive option.

Why Power BI?

"DATA "Analysis and Decision Making

Organizations need a tool that can help them understand the large amount of data that they are collecting. It is a powerful data visualization and analysis tool that allows **businesses to turn raw data into actionable insights and reports.**

Microsoft Power BI comes with a **free** or **paid version**. The free version only provides Power BI tools like **Power BI Desktop** and **Power Q&A** to dashboards. Whereas, in the Pro version they provide services like **live report sharing**, **Power View**, and more Power BI apps.

Key Differences Between Power BI and Tableau

Power BI	Tableau
Power BI uses DAX for measuring and calculating columns.	Tableau deploys MDX for dimensions and measures.
Power BI is best for a limited volume of data.	Tableau can handle huge columns of data and still offer better performance.
Power BI offers many data points for data visualization.	Tableau has better data visualization.

PROGRAM 1: Getting Started - Tableau Workspace, Tableau terminologies, Basic functionalities.

Working with Tableau that focuses on understanding the Tableau Workspace, Tableau terminologies, and basic functionalities.

Dataset used: vgsales.csv
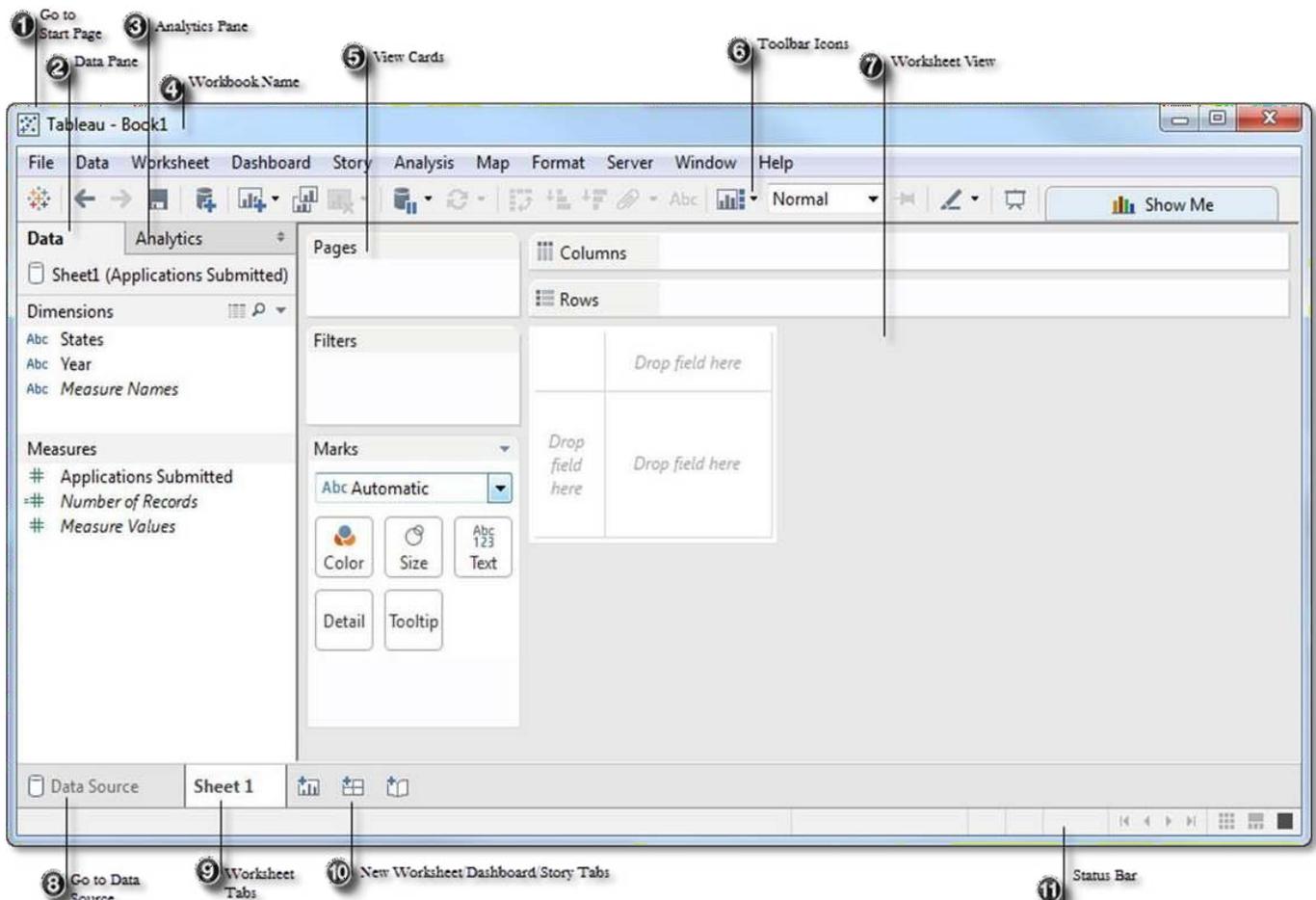
1. Go to Start Page: Toggle between the active sheet and the Desktop Start Page.
2. Data Pane: Includes dimensions and measures, populated from your selected data source. May also include calculated fields, parameters, or sets.
3. Analytics Pane: Includes options you can use to apply reference lines, forecasts, trend lines, to add totals to crosstabs, and to build boxplots.
4. Workbook Name: The file name of our workbook.
5. View Cards: Used for modifying the worksheet.
6. Toolbar Icons: Icons are available for quick access to popular features.
7. Worksheet/View: Workspace for building your visualizations.
8. Go to Data Source: Returns you to the data source specification page.
9. Worksheet Tabs: Click to view a specific worksheet, dashboard, or story
10. New Worksheet, Dashboard, and Story abs: Click to create a new Worksheet, Dashboard, or Story.
11. Status Bar: Displays data about the fields and marks included in the view.

Steps:

Tableau Workspace Setup:

- Connect to Data:
  - Open Tableau, and on the "Start Page," select Connect -> To a File -> Text File.
  - Browse to the location of vgsales.csv and open it.
  - Data Preview:
  - After loading, Tableau will show a preview of the data. You can rename columns if necessary.
  - Click on the "Sheet 1" tab at the bottom to go to your first worksheet.


Tableau Terminologies:

- **Dimensions:** These are qualitative fields. In vgsales.csv, examples include Platform, Genre and Publisher.
- **Measures:** These are quantitative fields used for calculations. Examples are Global_Sales, NA_Sales, and Year.
- **Rows and Columns Shelf**: Drag dimensions and measures to the Rows or Columns shelves to build the structure of your visualization.
- **Marks:** Controls the appearance of the data. You can set marks to be circles, bars, or other shapes and control size, color, and label.
- **Filters:** Used to limit the data displayed in the view.
- **Pages Shelf:** Used for creating animations or segmenting your view by categories.

2.Basic Functionalities:

a. Basic Visualization (Bar Chart of Global Sales by Genre):

- In your worksheet, drag Genre to the Columns shelf.

- Drag Global_Sales to the Rows shelf.

- You should see a bar chart. If the data isn't aggregating correctly, check if the aggregation is set to SUM by right-clicking Global_Sales -> Measure -> Sum.
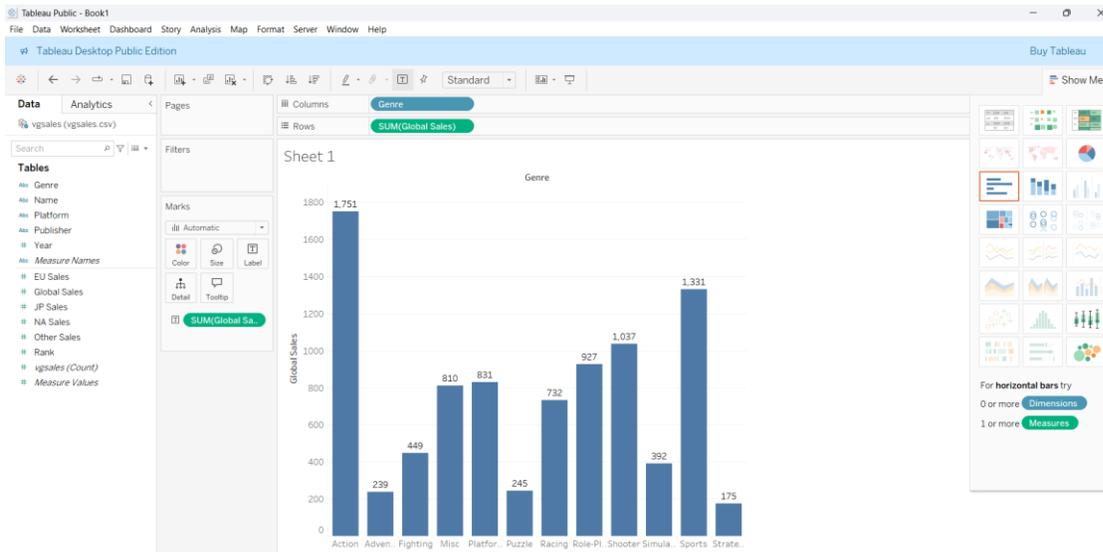
a. Sorting:

- Click on the Global_Sales axis and sort descending to show the genres with the most sales first.
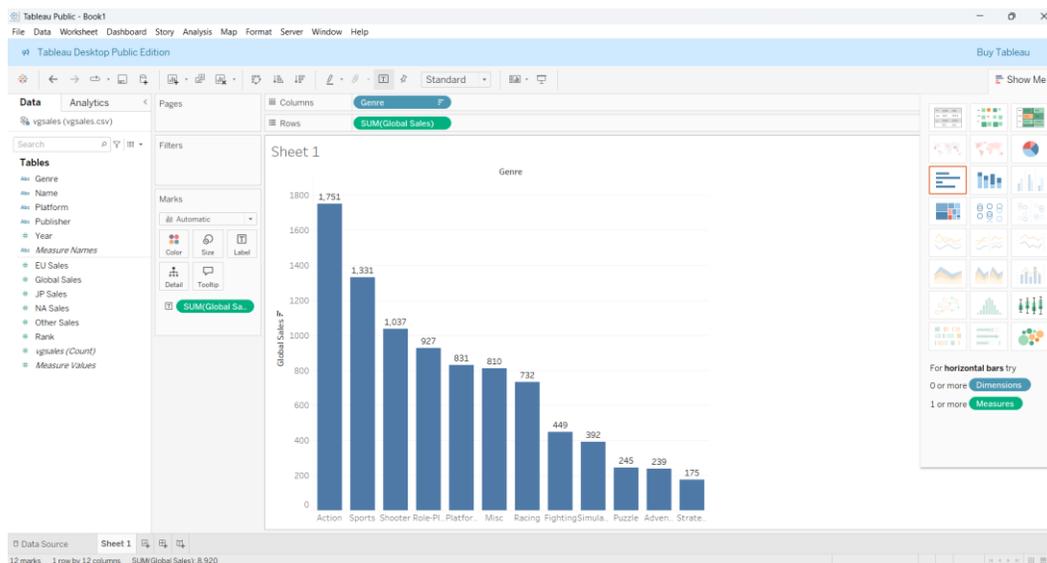
Filtering:

- Drag Year to the Filters shelf.
- Choose the range of years you want to display (e.g., 2000-2016).
- Add Year to the Pages shelf to create a dynamic view of how sales changed over time