DAYANANDA SAGAR COLLEGE OF ENGINEERING

(An Autonomous Institute affiliated to VTU, Belagavi, Approved by AICTE & ISO 9001:2008 Certified)
Accredited by National Assessment & Accreditation Council (NAAC) with 'A' grade, Shavige Malleshwara Hills,Kumaraswamy Layout, Bengaluru-560078.

DEPARTMENT OF ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

DATA VISUALIZATION USING TABLEAU

(SUB-CODE)

LABORATORY MANUAL

20__ - 20__

WHAT IS TABLEAU?

Tableau is an easy to use business intelligence software. It makes data visualization, data analytics, and reporting as easy as dragging and dropping. Anyone can learn to use Tableau without having a prior programming experience. Tableau can combine data from various data sources such as spreadsheets, databases, cloud data, and even big data- all into one program to perform dynamic analysis.

WHY TABLEAU?

Whether it's small or large, profitable or non-profit, every organization needs to analyze its data for optimal decision-making. Analyzing data has never been easier with traditional business intelligence tools.

Here are some of the advantages of using Tableau over the traditional BI tools:

TRADITIONAL METHOD	TABLEAU
Requires specific programming skills	No programming skills required
Focused on only one type of database	Combines different types of database spreadsheets, databases, cloud data, and even big data such as Hadoop
Time consuming	Time saving
Decision makers have to ask the IT people to retrieve any information from the database	Decision makers can directly use the dashboard to retrieve any information from the database
Largely depends on Query languages	Query is done behind the scene
Combining different types of database is difficult	Different types of databases can be combined easily
Not every business intelligence tool offers interactive dashboard	Interactive dashboard is easy to build and it makes data visualization quick and efficient
Comparatively expensive	Comparatively affordable
Mostly designed for large businesses	Perfect BI solution for small, medium, and large businesses, and even for non-profits

STEPS TO DOWNLOAD TABLEAU PUBLIC

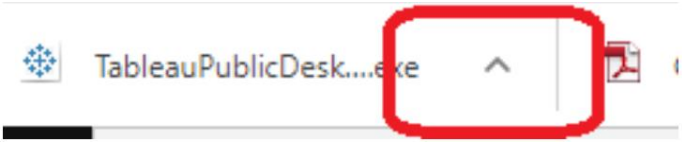
1. Visit the Tableau Public download page

2. Enter your email address and click download

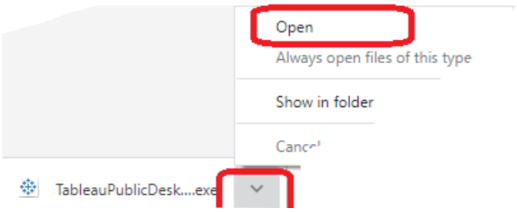
3. Your browser will ask you to save the installation file,

- Navigate to the place where you want to save your install file or to the downloads folder
- Click Save

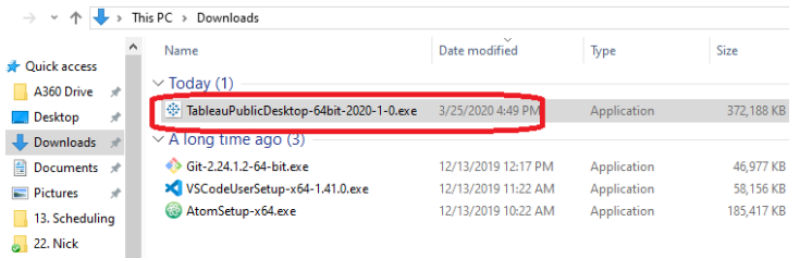
4. Once the file is downloaded click on the arrow next to the file at the bottom of the browser



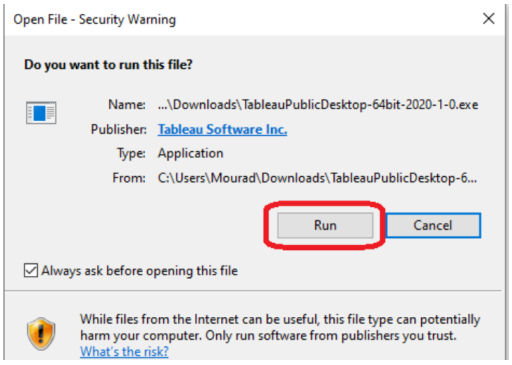
5. Select Open



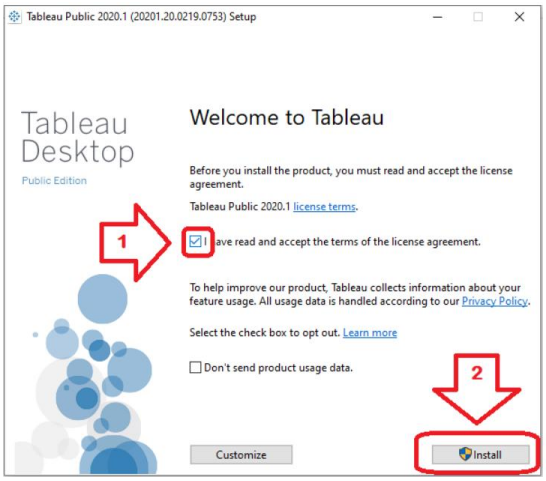
6. If you can't see a file in the browser, navigate to the place where you saved the file using your Windows Explorer (Downloads folder) and open it



7. When Security Message Pops up, Select Run



8. Click “I have read and accepted the terms of the license agreement” and Install



9. If another message pops up “Do you want to allow this app to make changes to your device?” Choose "Yes"

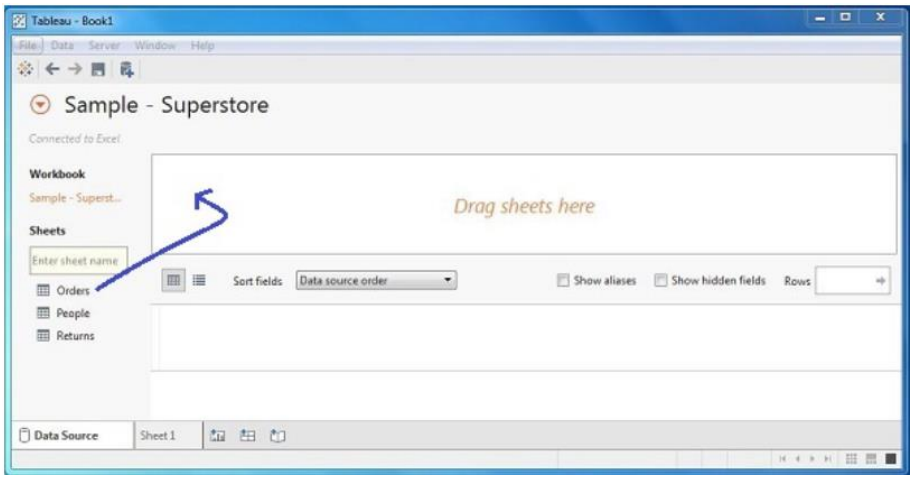
10. Once Tableau installation is finished, you can launch it from your Desktop

TABLEAU: GET STARTED

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

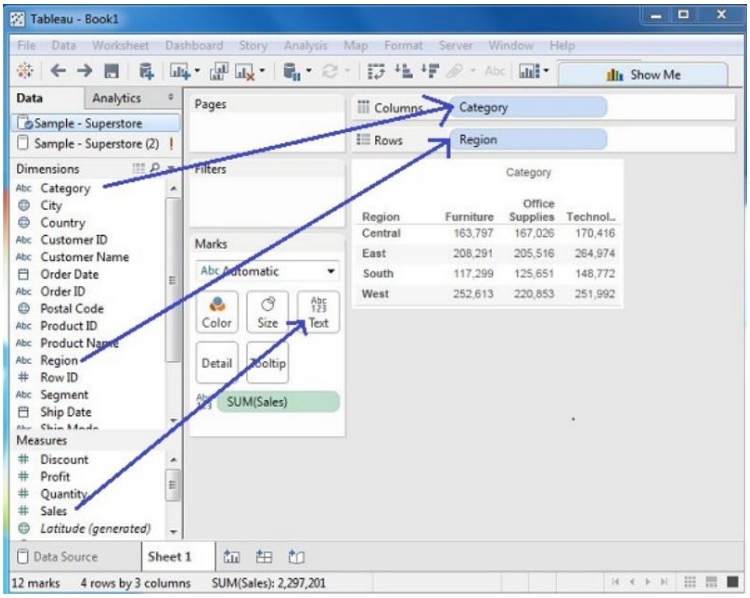
Connect to a Data Source

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**.



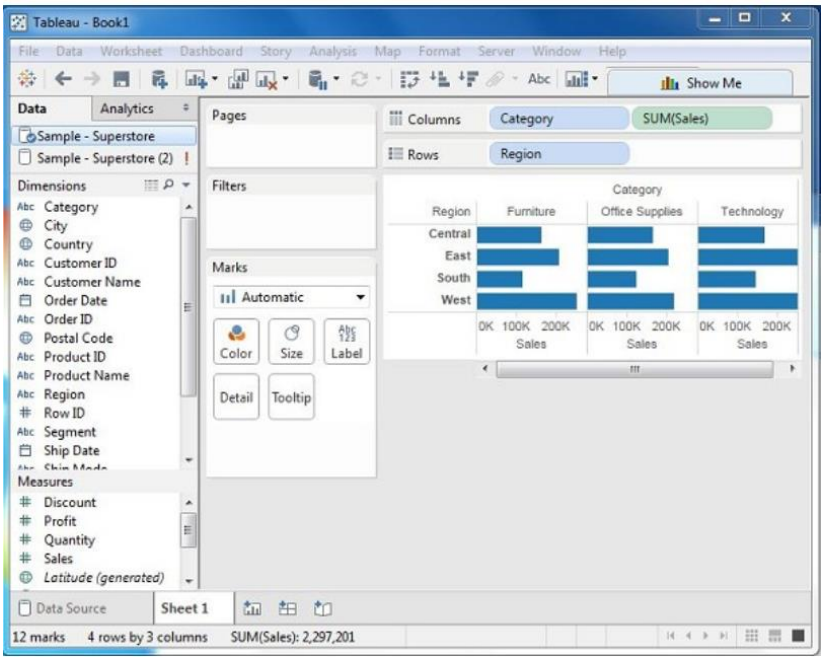
Choose the Dimensions and Measures

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, we 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.Then 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.



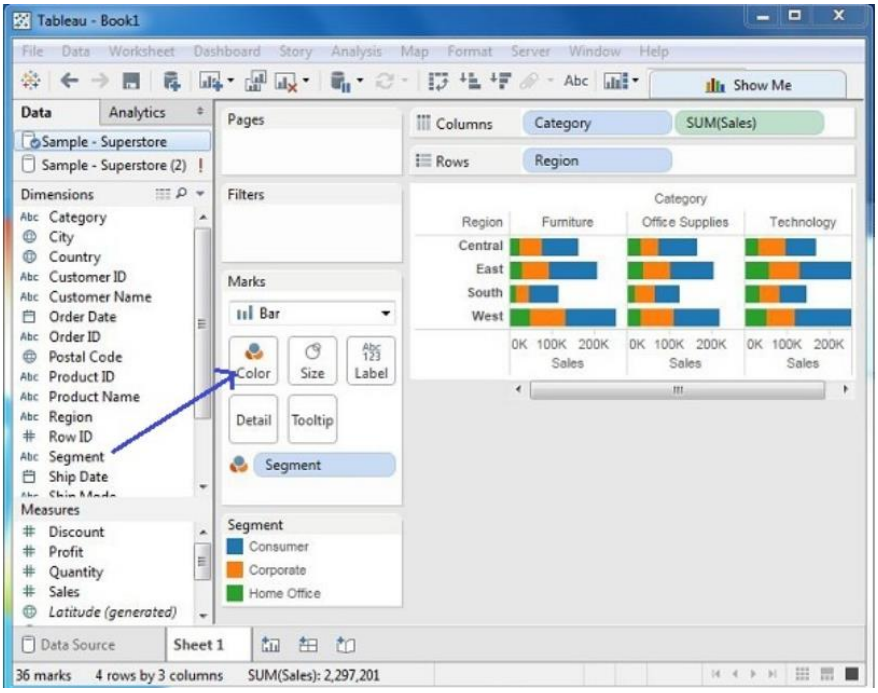
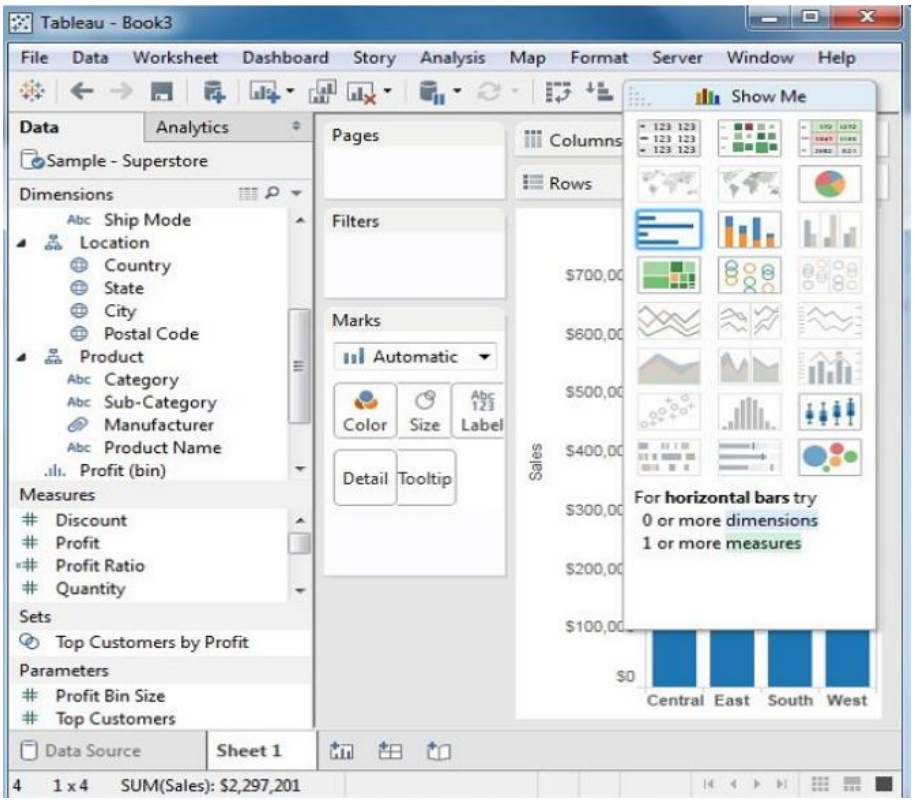


TABLEAU : SHOW ME

As an advanced data visualization tool, Tableau makes the data analysis very easy by providing many analysis techniques without writing any custom code. One such feature is Show Me. It can be used to apply a required view to the existing data in the worksheet. Those views can be a pie chart, scatter plot, or a line chart.

Whenever a worksheet with data is created, it is available in the top right corner as shown in the following figure. Some of the view options will be greyed out depending on the nature of selection in the data pane.

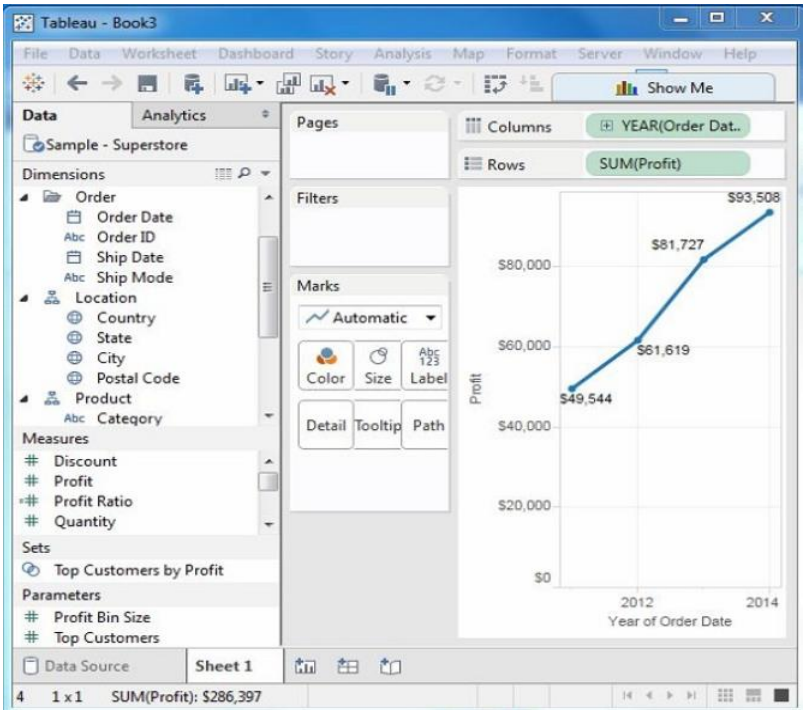


Show Me with Two Fields

The relation between two fields can be visually analyzed easily by using various graphs and charts available in Show Me. In this case, we choose two fields and apply a line chart. Following are the steps −

- **Step 1** − Select the two fields (order date and profit) to be analyzed by holding the control key.
- **Step 2** − Click the Show Me bar and choose line chart.
- **Step 3** − Click the Mark Label button on the scrollbar.

The following diagram shows the line chart created using the above steps.

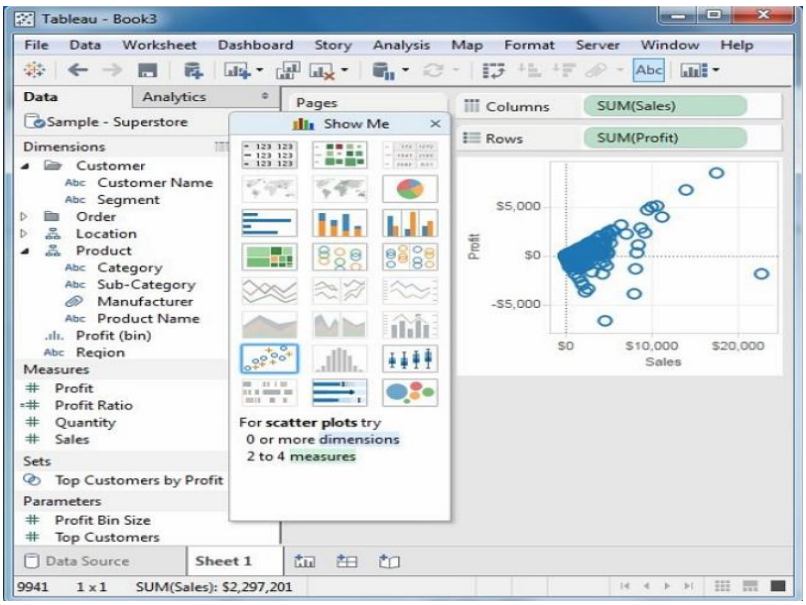


Show Me with Multiple Fields

We can apply a similar technique as above to analyze more than 2 fields. The only difference in this case will be the availability of fewer views in active form. Tableau automatically greys out the views that are not appropriate for the analysis of the fields chosen.

In this case, choose the field's product name, customer name, sales and profit by holding down the control key. As you can observe, most of the views in Show Me are greyed out. From the active views, choose Scatter View.

The following diagram shows the Scatter View chart created.



OVERVIEW OF TABLEAU

Tableau Desktop, Tableau Public, and Tableau Online, all offer Data Visual Creation and choice depends upon the type of work

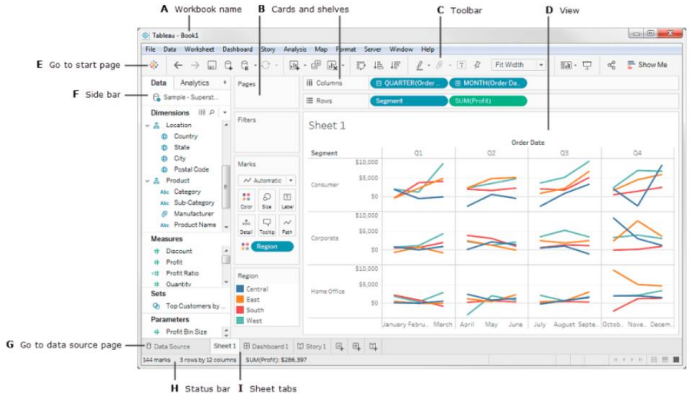
Products	Tableau Desktop	Tableau Public	Tableau Online
Formats Available	Free trial for 14 days	Absolutely Free	Workbooks are created on the Cloud and users can access them from anywhere.
	Free acess to Tableau Desktop for students and academicians for a year		
	Paid Version		
Limitation		User's data and workbooks are made public to all Tableau users.	

LIST OF EXPERIMENTS

Week	Name of the Experiment
1	Getting Started : Tableau Workspace
2	Connecting to Data Source
3	Creating a view
4	Creating a Dashboard
5	Building a Story
6	Tableau integration with R , Python and SQL
7	Saving the workbook
8	Mini Project

PROGRAM 1: GETTING STARTED: TABLEAU WORKSPACE

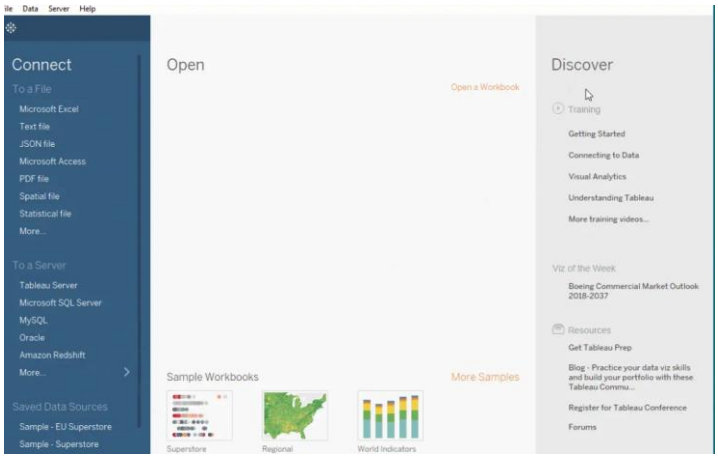
The Tableau workspace is a collection of worksheets, menu bar, toolbar, marks card, shelves and a lot of other elements about which we will learn in sections to come. Sheets can be worksheets, dashboards, or stories. The image below highlights the major components of the workspace.



Connecting to a Data Source

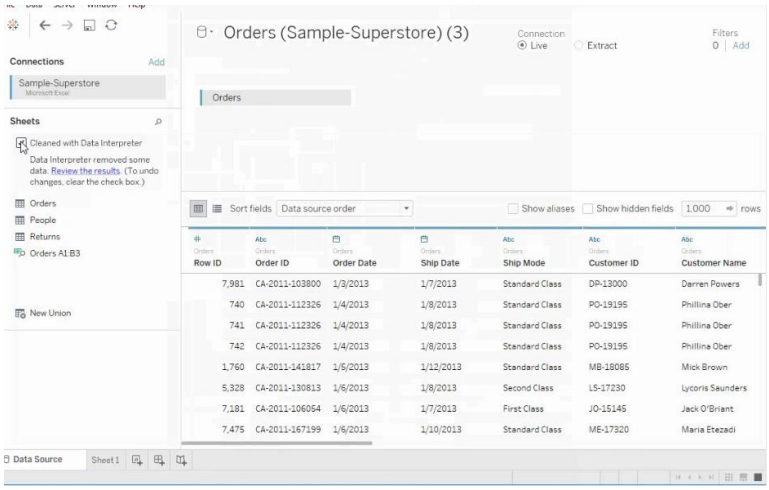
To begin **working with Tableau**, we need to connect Tableau to the data source. Tableau is compatible with a lot of data sources. The data sources supported by Tableau appear on the left side of the opening screen. Some commonly used data sources are excel, text file, relational database or even on a server. One can also connect to a cloud database source such as Google Analytics, Amazon Redshift, etc.

Connecting to the Sample-Superstore data set



Steps

1. Import the Data into tableau workspace from the computer.
2. Under the Sheets Tab, three sheets will become visible namely Orders, People, and Returns. However, we will focus only on Orders data. Double click on Orders Sheet, and it opens up just like a spreadsheet.
3. We observe the first three rows of data looks a bit different and is not in the desired format. Here we make use of **Data Interpreter**, also present under Sheets Tab. By clicking on it, we get a nicely formatted sheet.



Creating a View

We will start by generating a simple chart. In this section, we will get to know our data and will begin to ask questions about the data to gain insights. There are some important terms that we will encounter in this section.

Dimension

Measures

Aggregation

Dimensions are qualitative data, such as a name or date. By default, Tableau automatically classifies data that contains qualitative or categorical information as a dimension, for example, any field with text or date values. These fields generally appear as column headers for rows of data, such as Customer Name or Order Date, and also define the level of granularity that shows in the view.

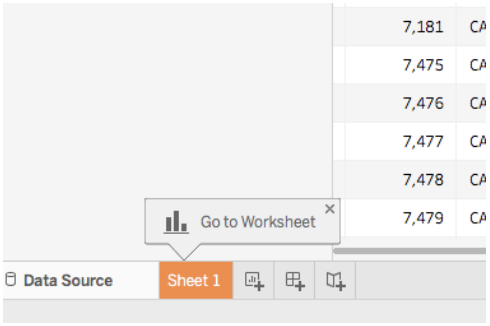
Measures are quantitative numerical data. By default, Tableau treats any field containing this kind of data as a measure, for example, sales transactions or profit. Data that is classified as a measure can be aggregated based on a given dimension, for example, total sales (Measure) by region (Dimension).

Aggregation is the row-level data rolled up to a higher category, such as the sum of sales or total profit.

Tableau automatically sorts the fields in Measures and Dimensions. However, for any anomaly, one can change it manually too.

Steps

1. Go to the worksheet. Click on the tab Sheet 1 at the bottom left of the tableau workspace.



2. Once, you are in the worksheet, from Dimensions under the Data pane, drag the Order Date to the Column shelf.

   On dragging the Order Date to the columns shelf, a column for each year of Orders is created in the dataset. An 'Abc' indicator is visible under each column which implies that text or numerical or text data can be dragged here. On the other hand, if we pulled Sales here, a cross-tab would be created which would show the total Sales for each year.

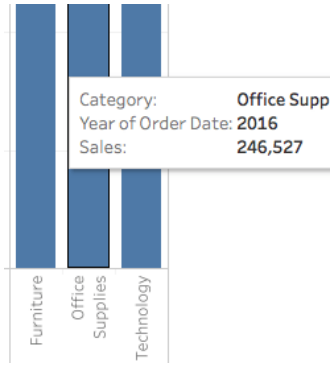
3. Similarly, from the Measures tab, drag the Sales field onto the Rows shelf.

Refining the View

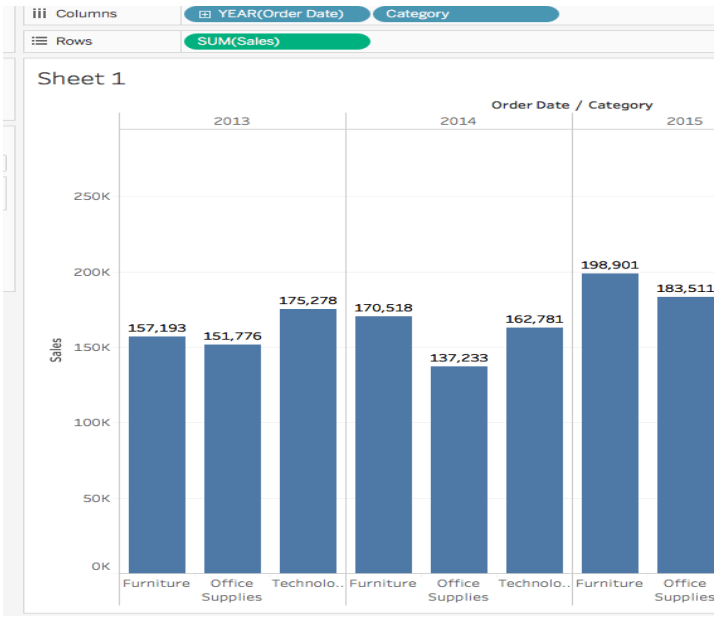
Steps

1. Category is present under the Dimensions pane. Drag it to the columns shelf and place it next to YEAR(Order Date). The Category should be placed to the right of Year. In doing so, the view immediately changes to a bar chart type from a line. The chart shows the overall Sales for every Product by year.

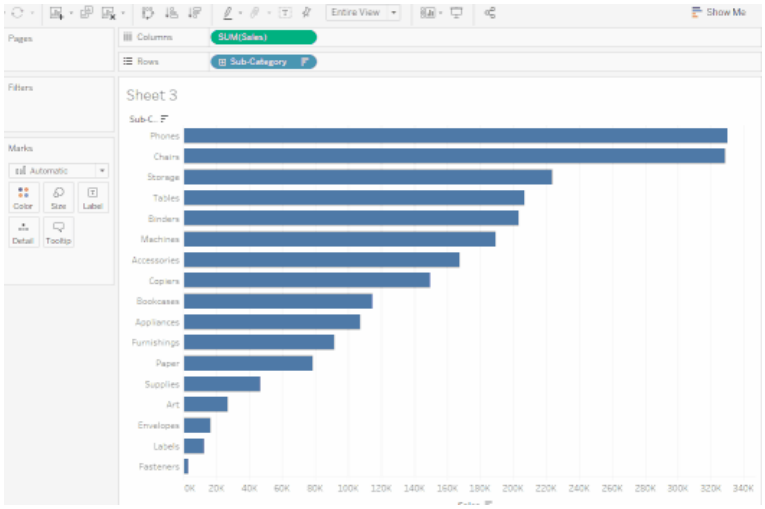
   To view information about each data point (that is, mark) in the view, hover over one of the bars to reveal a tooltip. The tooltip displays total sales for that category. Here is the tooltip for the Office Supplies category for 2016:



   To add labels to the view, click Show Mark Labels on the toolbar.



   The bar chart can be displayed horizontally instead of vertically too. Click Swap on the toolbar for the same.



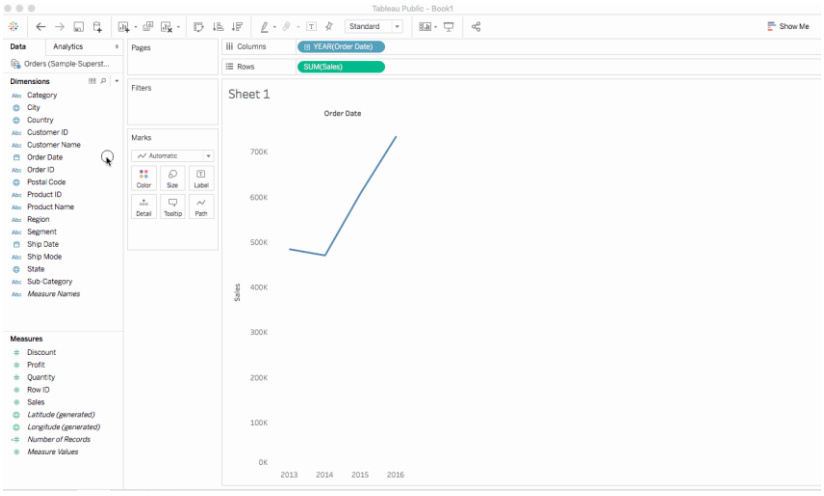
Emphasizing the Results

Adding filters to the view

Filters can be used to include or exclude values in the view. Here we try to add two simple filters to the worksheet to make it easier to look at product sales by sub-category for a specific year.

Steps

In the Data pane, under Dimensions, right-click Order Date and select Show Filter.Repeat for Sub->category field also.



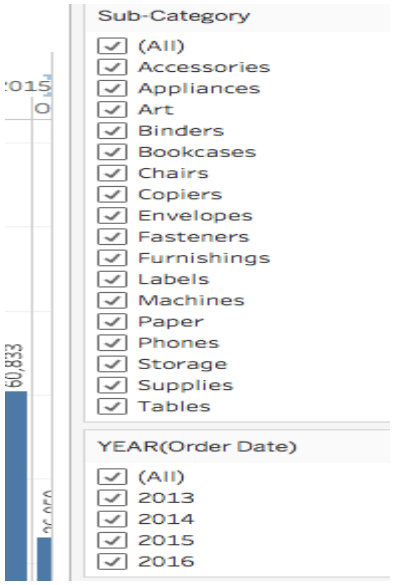
Filters are the type of cards and can be moved around on the worksheet by simple drag and drop

Adding colors to the view

Colors can be helpful in the visual identification of a pattern.

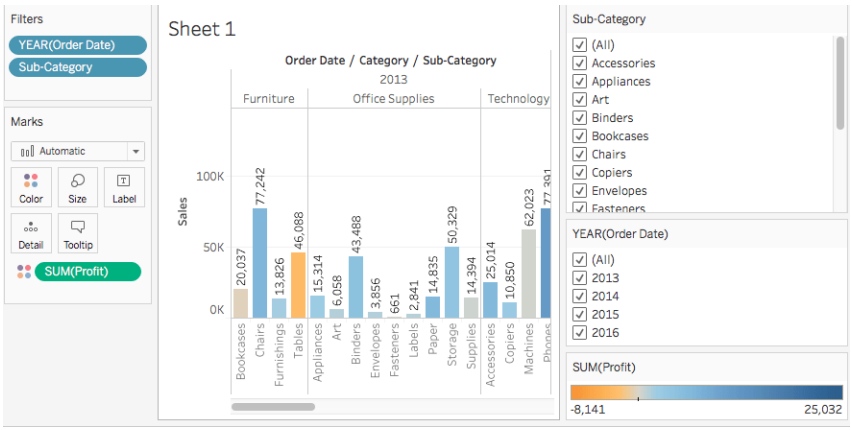
Steps

In the Data pane, under Measures, drag Profit to Color on the Marks card.



It can be seen that Bookcases, Tables and even machine contribute to negative profit, i.e., loss. A powerful insight.

`Hands On`



Map View

Creating a Map View

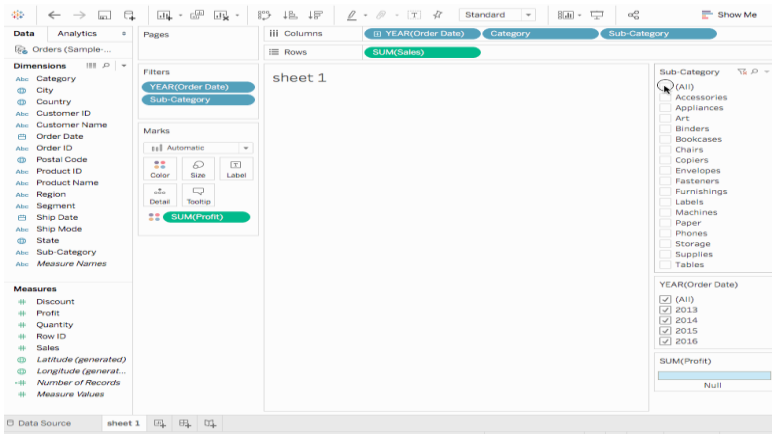
Map views are beneficial when we are looking at geographic data (the Region field). In the current example, Tableau automatically recognizes that the Country, State, City, and Postal Code fields contain geographical information.

`Steps`

1. Create a new worksheet.

2. Add `State` and `Country` under Data pane to `Detail` on the Marks card. We obtain the map view.
3. Drag `Region` to the `Filters` shelf, and then filter down to `South` only. The map view now zooms in to the South region only, and a mark represents each state.
4. Drag the `Sales` measure to the `Color` tab on the Marks card. We obtain a filled map with the colors showing the range of sales in each state.
5. We can change the color scheme by clicking `Color` on the Marks card and selecting `Edit Colors`. We can experiment with the available palettes.
6. We observe that **Florida** is performing the best regarding Sales. If we Hover over Florida, it shows a total of 89,474 USD in sales, as compared to South Carolina, for example, which has only 8,482 USD in sales. Let us gauge the performance by `Profit` now since Profit is a better indicator than Sales alone.
7. Drag `Profit` to `Color` on the Marks card. We now see that Tennessee, North Carolina, and Florida have negative profit, even though it appeared they were doing good in Sales. Rename the sheet as Profit Map

`Hands On`



Getting into the Details

Maps empower us to visualize the data broadly. In the last step, we discovered that we discovered that Tennessee, North Carolina, and Florida have a negative profit. In this section let us draw a Bar chart to explore the reason for the negative profit.

`Steps`

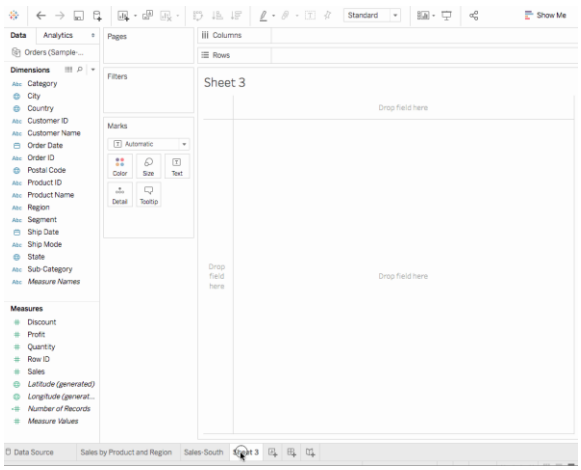
1. Duplicate the Profit Map worksheet and name it Negative Profit Bar Chart.
2. Click `Show Me` on the **Negative Profit Bar Chart** worksheet. `Show Me` presents the number of ways in which a graph can be plotted between items

mentioned in the worksheet. From `Show Me` select the horizontal bar option and the view updates to horizontal from vertical bars instantly.

3. We can select more than one bar at a time by simply clicking and dragging the cursor over them. We want to focus only on the three states, i.e., Tennessee, North Carolina, and Florida. Hence, we will only select the bars pertaining to them.

   Creating Hierarchies
   Hierarchies come in handy when we want to group similar fields so that we can quickly drill down between levels in the viz.

   1. In the Data pane, drag a field and drop it directly on top of another field or right-click the field and select
   2. Drag any additional fields into the hierarchy. Fields can also be re-ordered in the hierarchy by simply dragging them to a new position. In the current viz. we will create the following hierarchies: Location, Order, and Product.



4. On the Rows Shelf, click the plus-shaped icon on the `State` Field to drill-down to the `City` level.

Dashboard

A dashboard is a collection of several views, enabling one to compare a variety of data simultaneously.

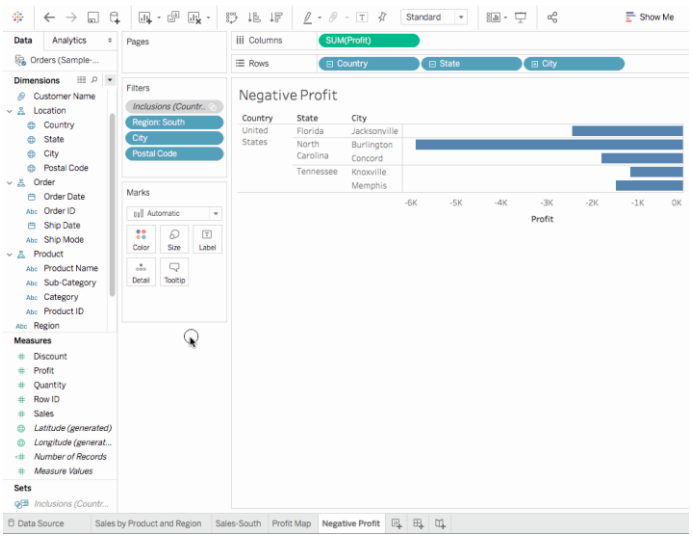
Creating a Dashboard

`Steps`

1. Click the `New dashboard` button.

2. Drag `Sales in the South` to the empty dashboard
3. Drag `Profit Map` to the dashboard, and drop it on top of the Sales in the South view. Both views can be seen at once. To be able to present data in a manner so that others can understand it we can arrange the dashboard to our liking.
4. On the `Sales South` worksheet in the dashboard view, click under the `Region` and clear off the `Show Header`. Repeat the same process for all the other headers. This helps to emphasize only what is needed and hides away the not so important information.
5. On the `Profit Map`, Hide the Title as well and perform the same steps for the `Sales South` map.
6. We can see that the `Sub-Category` filter card and `Year of Order Date` have been repeated on the right-side. Let us get rid of the extra by simply crossing them out. Finally, click on the `Year of Order Date`. A drop-down arrow appears and select the option of `Single Value (Slider)`. Now let the magic unfold. Experiment by choosing different years on the slider and the Sales also vary accordingly.
7. Drag the `SUM(Profit)` filter to the bottom of the dashboard below Sales in South for a better view.

Hands On



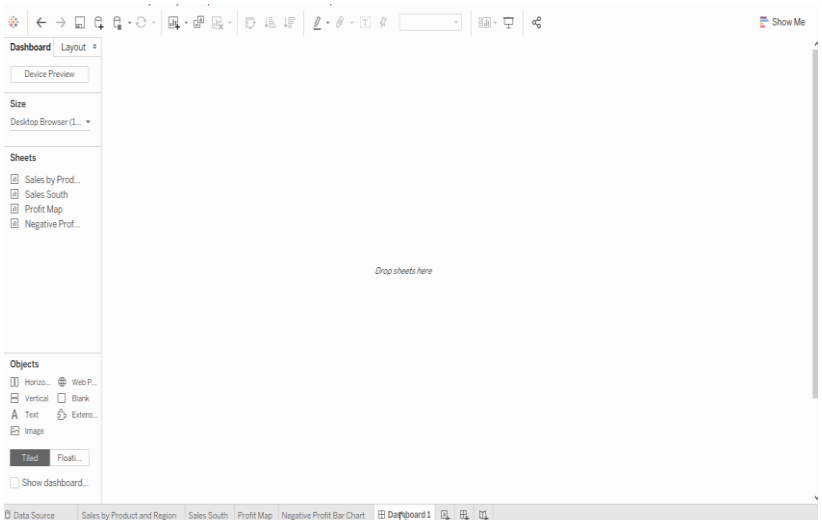
Adding Interactiveness

In order to make the dashboard more interactive like viewing which sub-categories are profitable in which states, a few changes need to be done.

Steps

1. Let's start with the `Profit Map`. On clicking the map, a `Use as filter` icon appears in the upper right. Click on it. If we select any map, Sales corresponding to that state will be highlighted in the `Sales-South` map.
2. For the `Year of Order Date`, click on the drop-down option and go to `Apply to Worksheets > Selected Worksheets`. A dialog box opens up. Select the `All` option followed by `OK`. What does this option do? It applies filters to all the worksheets having the same data source.
3. Explore and experiment. In the visualization below, we can filter the `Sales South` map to view products that are being sold in North Carolina only. We can then easily explore the profits yearly.
4. Rename the Dashboard to `Regional Sales and Profit`.

Hands On



Thus, selling machines in the North Carolina did not bring any profits to the company.

Story

A dashboard is a cool feature, but tableau also offers us to showcase our results in presentation mode in the form of stories about which we will discuss in this section.

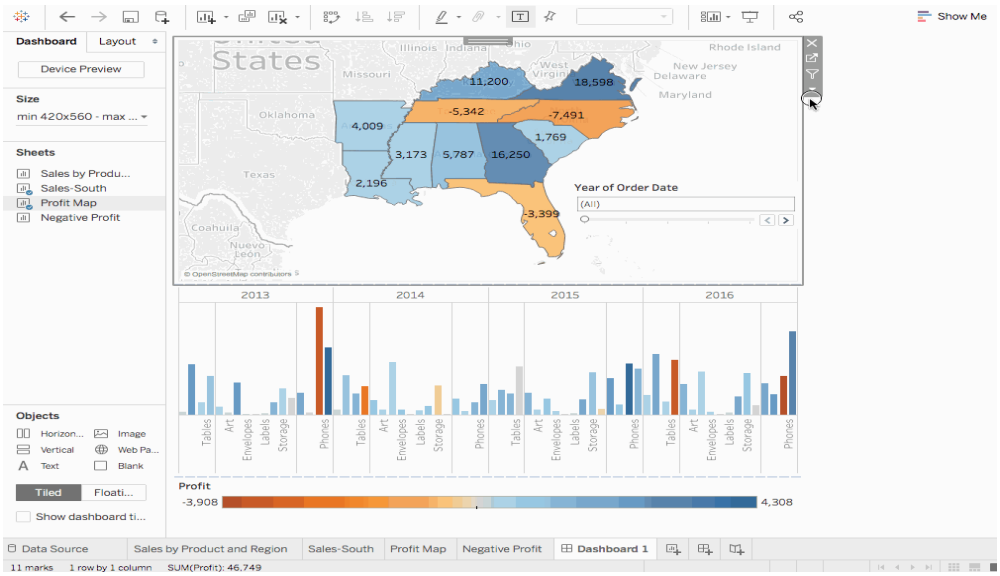
Building a Story

Steps

1. Click the `New story` button.
2. From the Story pane on the left, drag the `Sales in the South` worksheet (created earlier) onto the view.

3. Edit the text in the gray box above the worksheet. This is the caption. Name it as `Sales and profit by year`.
4. Stories are quite specific. Here we will tell a story about selling machines in North Carolina. In the Story pane, click on `Duplicate` to duplicate the first caption, or you may even create a new one.
5. In the `Sub-Category`, filter `select` only `Machines`. This helps to gauge sales and profit of machines by year.
6. Rename the caption to `Machine sales and profit by year`.

Hands On



Steps

1. In the Story pane, select `Blank`. Drag the already created dashboard `Regional Sales and Profit` onto the canvas.
2. Caption it as `Low performing items in the South`.
3. Select `Duplicate` to create another story point with the Regional Profit dashboard. Select North Carolina on the bar chart since we are interested in showing more about it.
4. Select All the years.
5. Add a caption for clarity, like, `Profit in NC : 2013-2016`.
6. Select any year like 2014. Add a caption, for example, `Profit in NC : 2014` and then click on the Duplicate tab. Repeat the same step for all the remaining years.
7. Click on the presentation mode and let the `story` unfold.

Tableau's integration with R, Python & SQL

Tableau and R

R is a popular statistical language used to perform sophisticated predictive analytics, such as linear and nonlinear modeling, statistical tests, time-series analysis, classification, clustering, etc.**(Tableau 8.1 and R)** Using Tableau in conjunction with R has the following advantages:

- Leverages the statistical power of tableau by giving its users access to sophisticated R libraries for gaining better and deeper insights from the data.
- Tableaus's enhanced data exploration options and ability to connect to multiple sources comes in handy for R users.
- Further, it also enables Tableau users to benefit from the usefulness of R language without having to actually know the language.

Setting up Tableau Desktop with R

- **Download and Install `Rserve`.**

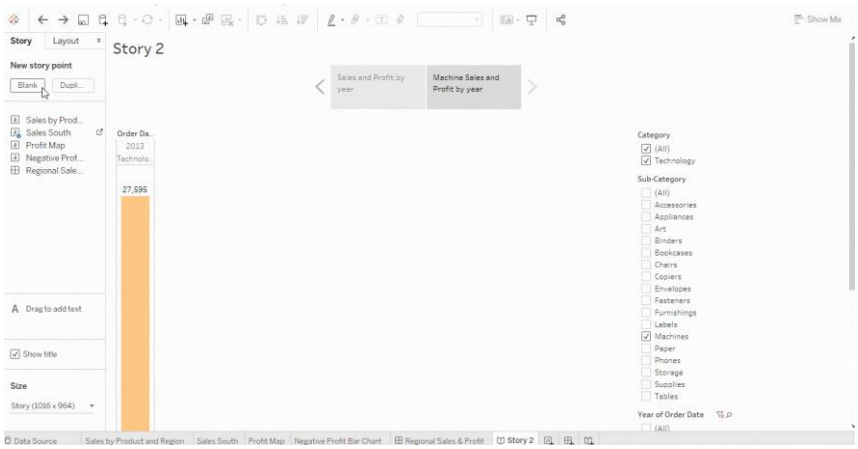
You will need to download and install `Rserve` package for Tableau to connect and utilize the R script functions. In the R console, enter the following commands:

```
install.packages("Rserve")

library(Rserve)

Rserve() / Rserve(args = '—no-save')
```

Connect Tableau to the R Server

After `Rserve` is successfully installed, open Tableau Desktop and follow the below mentioned steps.

1. Go to the `Help > Settings and Preferences and select Manage External Service Connection`.



2. Enter the server name as “Localhost” (or “127.0.0.1”) and a port of “6311”.
3. Click on the “Test Connection” button. You should see a successful message prompt. Click OK to close.

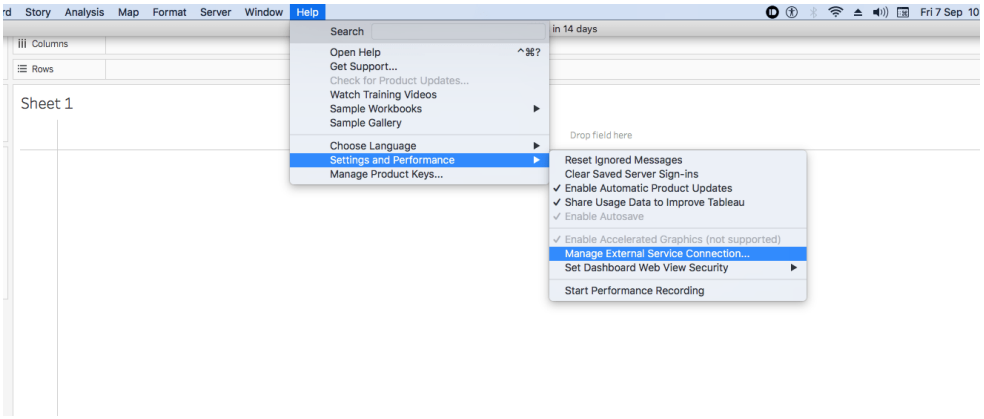


Tableau and Python

Python is a widely used general-purpose programming language. Python provides a large number of libraries to perform statistical analysis, predictive modeling or machine learning. Connecting Tableau with Python is one of the best approaches for predictive analytics. Tabpy is a package developed to do the same. To enable Tableau to harness the power of Python, it can be connected to the TabPy server to execute Python code on the fly and display results in the form of visualizations.

How does Tableau integrate with Python?

When we use TabPy with Tableau, we can define calculated fields in Python, thereby leveraging the power of a large number of machine-learning libraries right from our visualizations.

Setting up Tableau Desktop with Python

Download and Install Tabpy.

Running a Python code within a Tableau workbook requires a Python server to execute it. The TabPy framework is what gets the job done. Download TabPy from Github at the following **link**. Alternatively, you can follow the steps below:

```
conda install -c anaconda tabpy-server
```

Then cd to the directory containing the downloaded tabpy server and run.

```
python setp.py
```

Tableau and SQL Server

There is a hidden value in our Microsoft SQL Server data which lies buried under the standard reports and complex business intelligence tools. **Tableau delivers insight**

everywhere by equipping anyone to do a sophisticated visual analysis of SQL Server data. We can connect Tableau to SQL Server live for tuned, platform-specific queries, or directly bring data into Tableau's analytical engine to take the burden off the database.

Tableau provides an optimized, live connector to SQL Server so that we can create charts, reports, and dashboards while working directly with our data. As we dig into our analysis, Tableau recognizes any schema used in SQL Server, so we don't have to manipulate our data.

Let us walk through an example depicting how to connect SQL server database to Tableau Desktop and then use it to create visualizations.

Steps:

1. Login to the SQL Server
2. Open Tableau Desktop and under Servers, connect to MS SQL.
3. Paste the server name in the dialog box that opens and click ok. This connects Tableau to the SQL Server. Select the database of choice. In this example, we choose the salesDB. We can then select from a list of TABLES too, e.g., Sales Log. The table gets imported into the Tableau environment. Now we can choose to extract the entire data or the portion of it to a new worksheet. We can even specify the number of rows to extract.
4. In the new worksheet we have the extracted data from MS SQL, From here we can work with it like any other Tableau Worksheet.

Problem – 1

TFL Bus Safety

1. For the given Dataset (TFL Bus Safety):

i) Create a bar chart on boroughs field to visualize the trend in the count.

ii) Create a line chart for date of incidence for each month in a quarter, comment on possibilities and suitability of different charts for this timeline.

iii) In above question, apply formatting to display the first letter of the month on X-axis.

iv) Create tree maps of all the data fields except date & year and comment on significance of tree map.

v) Create an interactive dashboard for the above data.

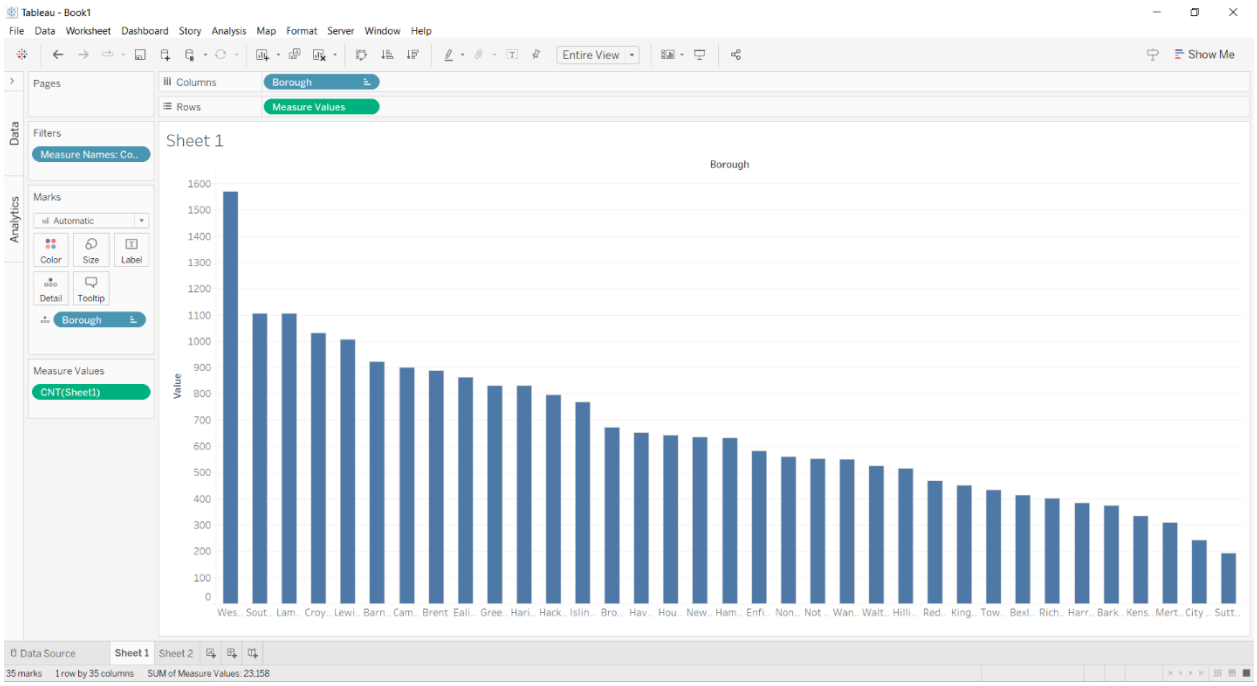
Solutions:

i) a. Drag the 'Borough' field to the columns shelf.

b. Drag 'Measure fields' to the rows.

c. Select 'Bar Chart' under 'Show me' section.

d. Sort the Bar Chart in descending order.



ii) a. Drag the ‘Date of incident’ field to the column field.

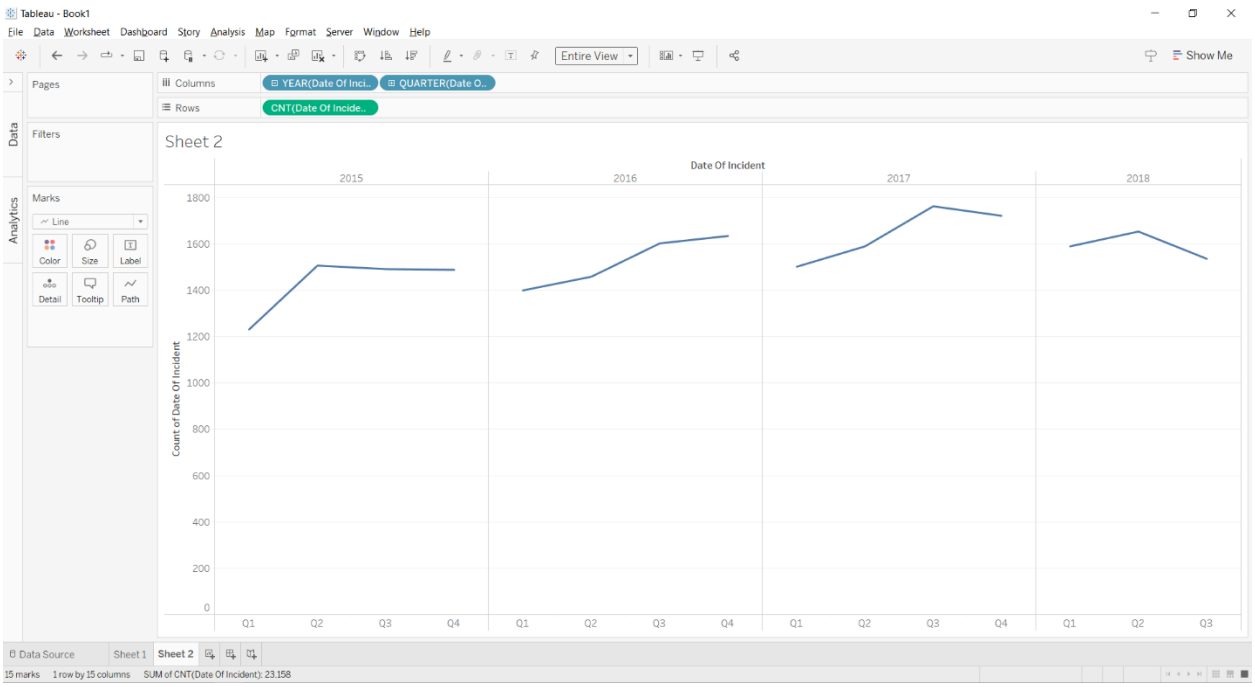
b. Right click on ‘Date of incident’ and select continuous.

c. Drag ‘Number of records’ field to the rows.

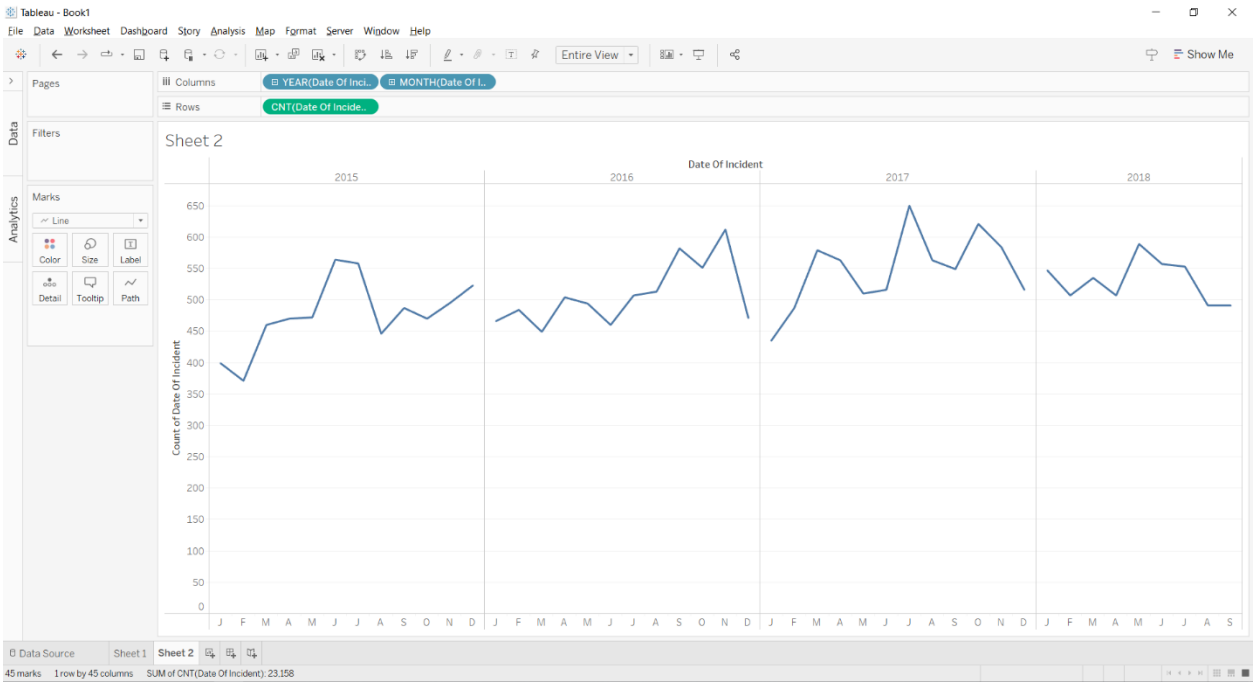
d. Right click on ‘Date of incident’ field in column & select ‘Exact date’.

e. Right click on ‘Date of incident’ field in column & select ‘Quarter’.

f. Drag ‘Quarter’ field to the columns shelf before the ‘Date of incident’ field.



iii) a. Create a line chart by following the previous sub-question.



b. Now, right click on any one label in X-axis, then go to format.

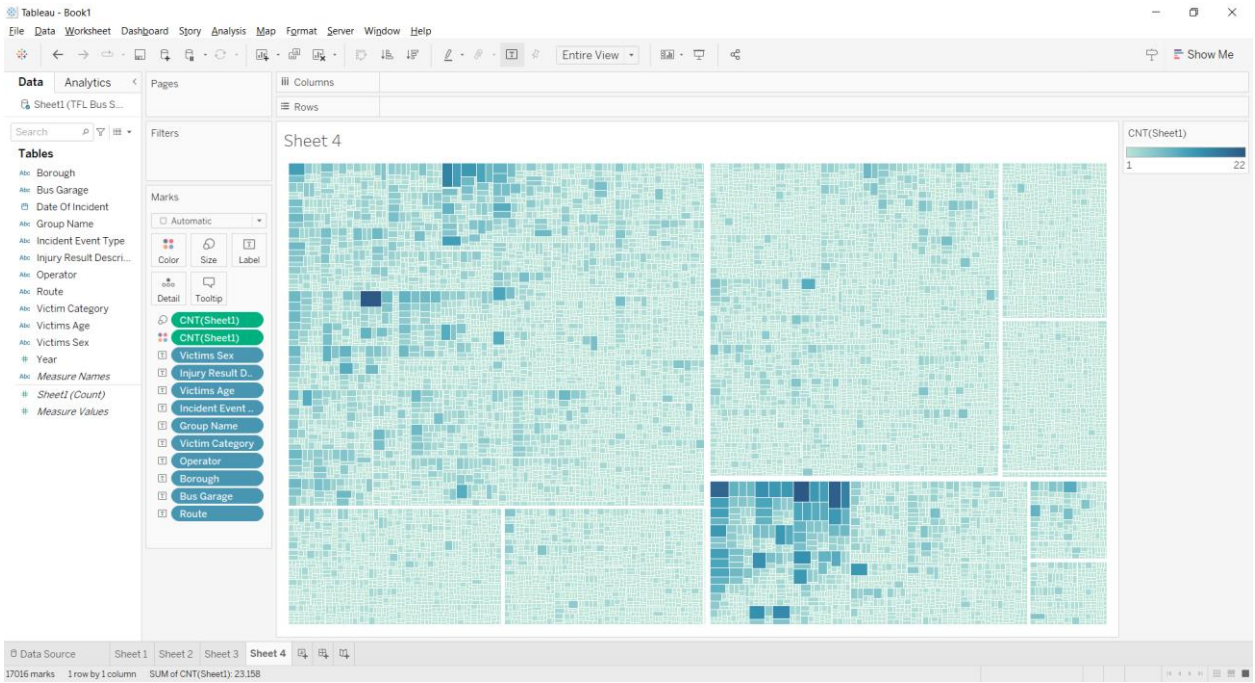
c. Under header, go to default, then select dates.

d. Select first letter under dates.

iv) a. Drag Date to rows & Year to columns.

b. Click on ‘Show Me’ and select the treemaps chart.

c. Customize the treemaps by adding labels, adjusting colors, and arranging the fields as desired.



v) a. Click on ‘New Dashboard’ button in the bottom left corner of the Tableau window.

b. Drag the sheets and drop in the dashboard, select floating windows under Objects, in Dashboard.

c. Rearrange all the sheets, once all the sheets are added.

Tableau can be used to analyze the "TFL Bus Safety" dataset, which includes information about bus incidents. We can create a bar chart to visualize incident counts by borough and a line chart to observe trends in incidents over time. Formatting the x-axis to display the first letter of each month enhances readability. Additionally, treemaps can help explore data fields, excluding date and year, to understand proportions and relationships. An interactive dashboard combines these visualizations for a comprehensive view of the data. Using Tableau, we can gain insights into bus safety trends and develop data analysis skills.

Problem – 2

Sales Revenue Dataset

2. 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.

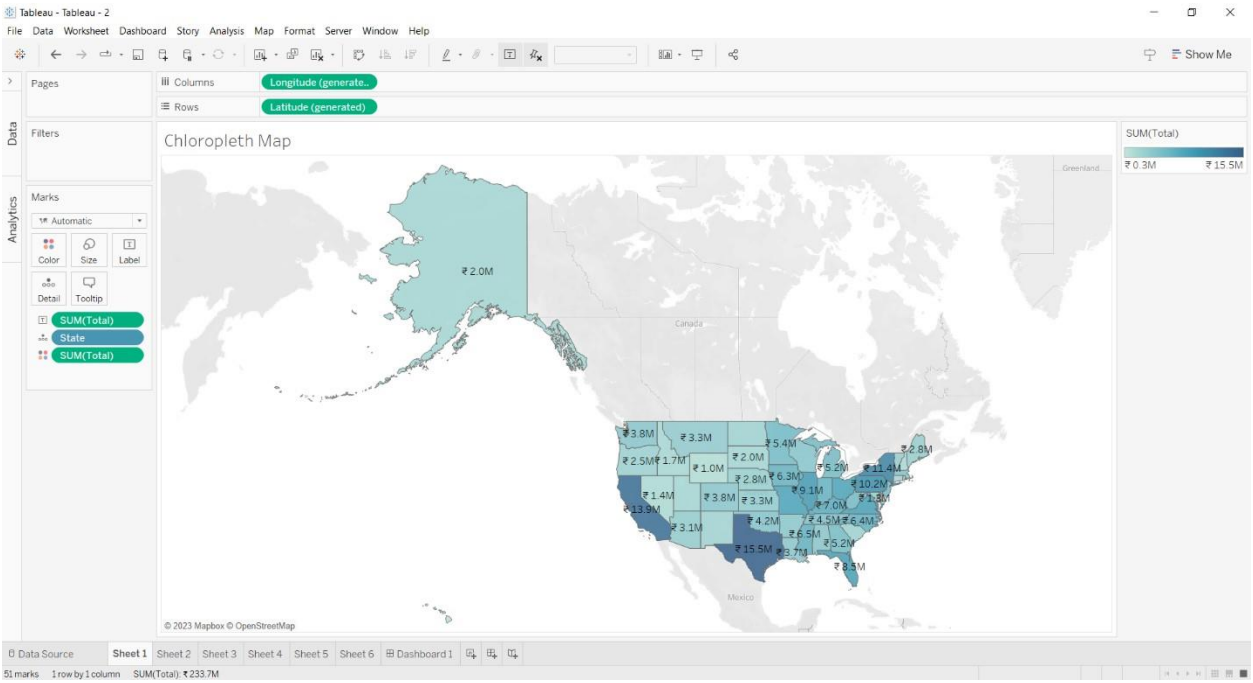
vi) Create a calculated field to show the average revenue per state & display profitable & non-profitable state.

vii) Build a dashboard.

Solutions:

i) a. Drag state to columns, select map, then in map tab, edit map location, change to US, then drag total to label.

b. Drag total to label, then format to millions.

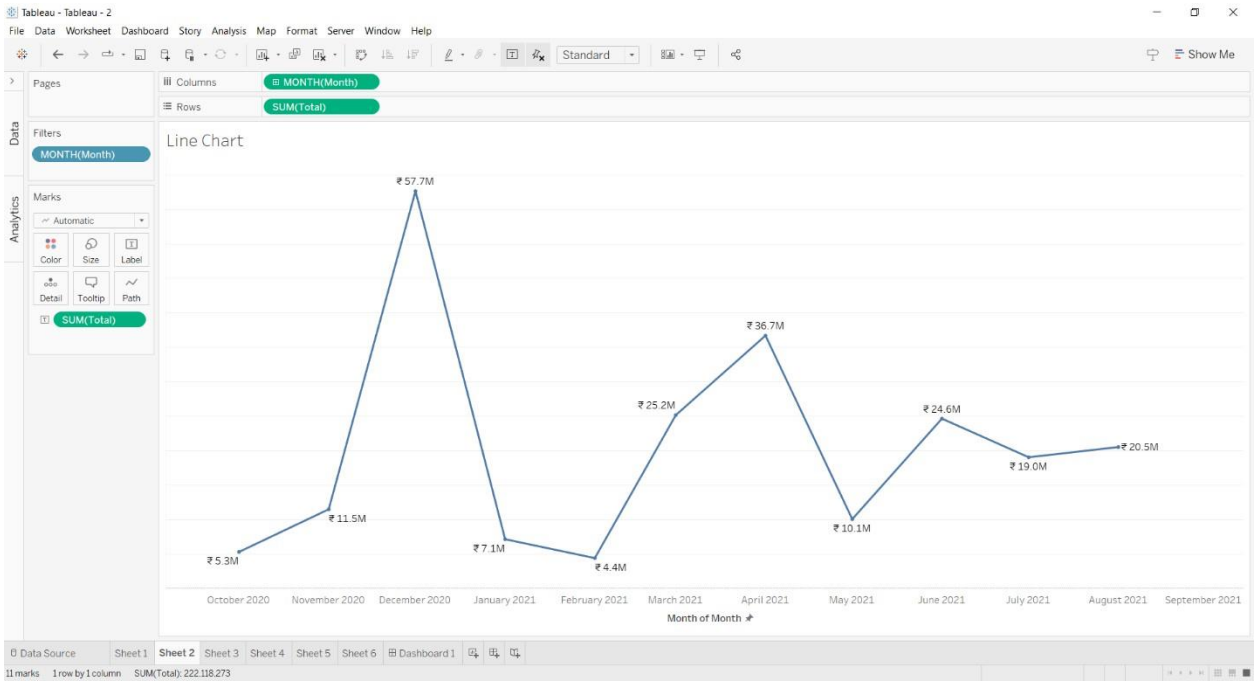


ii) a. Create a line chart to show the revenue based on the month of the year.

b. Drag total to rows.

c. Convert month from string to date, by right click, and drag to column.

d. Under marks, select sum(total), right click, format, then under pane, go to default, under numbers, go to currency(custom), decimal place to 1, display units in Millions.



iii) a. Create a bin of size 10 for the age measure to create a new dimension to show the revenue

b. Drag age to columns, total to rows.

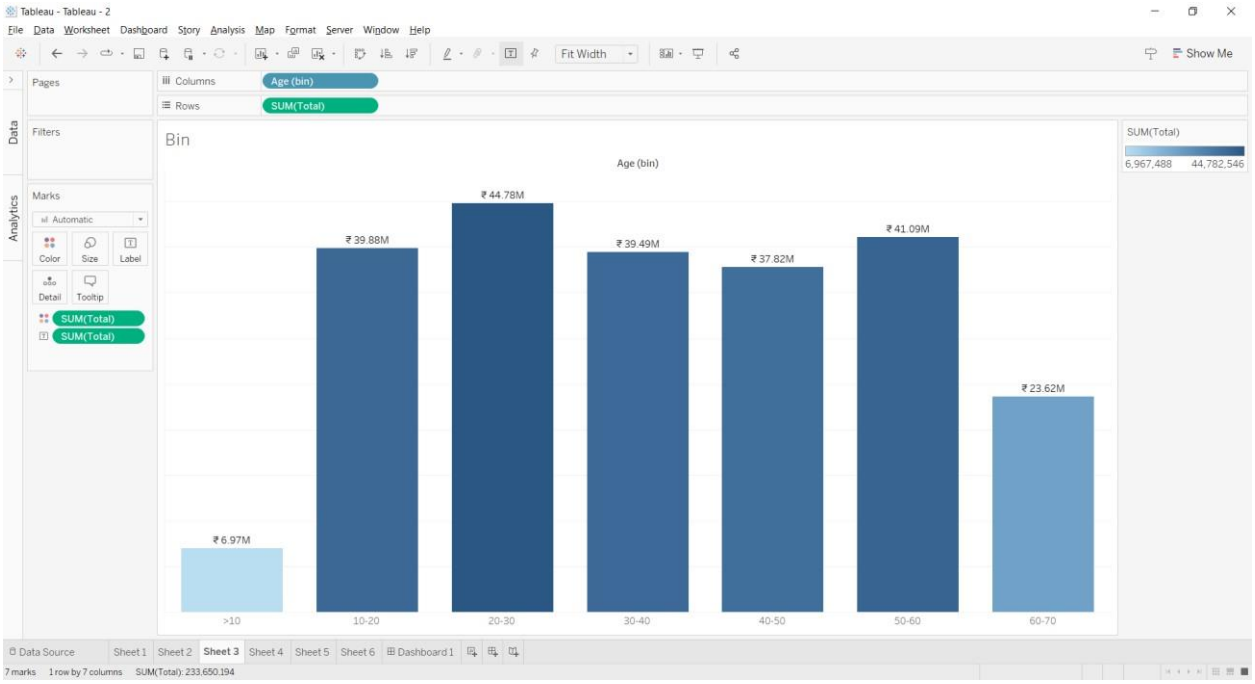
c. Right click age in tables, create, bin, then size of bin to 10.

d. Drag age bin to columns and remove age.

e. Then under each bar, below the axis, right click, go to edit alias, then change each alias from >10, 10-20, 20-30, 30-40, 40-50, 50-60, 60-70 respectively.

f. Right click graph, format, then remove grid lines.

g. Drag total to label, format, then currency, then unit is millions.



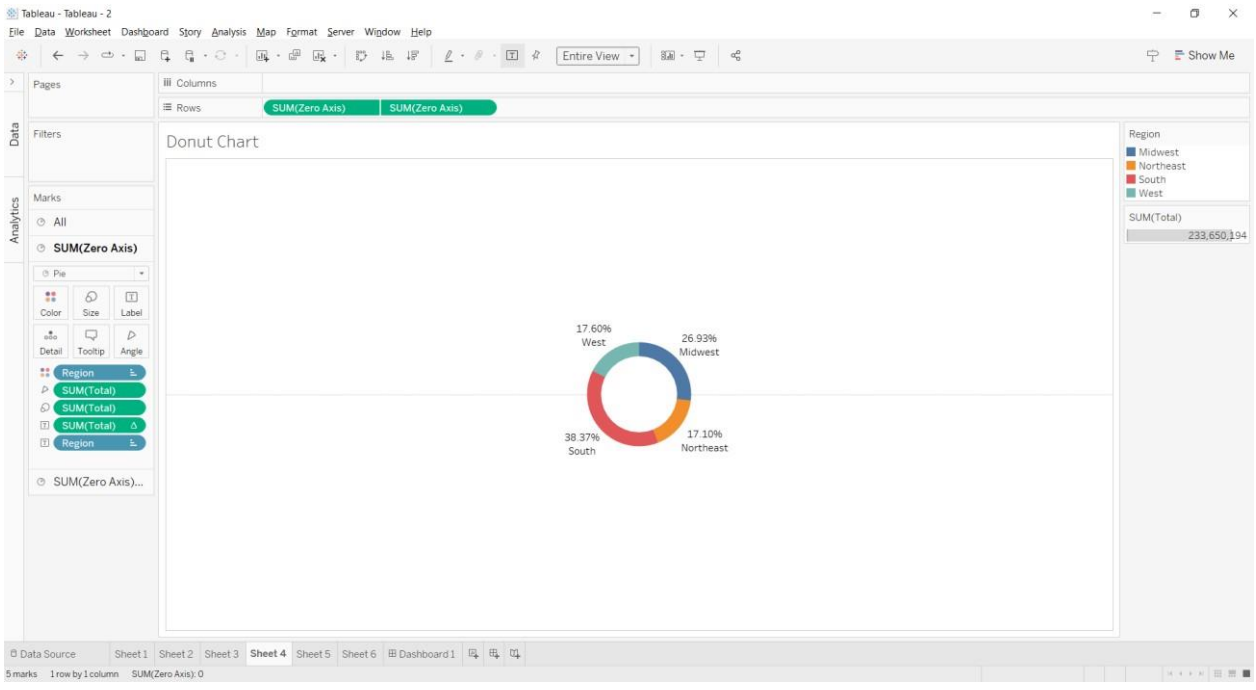
iv) a. Create a donut chart view to show the percentage of revenue per region by creating zero access in the calculated field.

b. Drag Region to column, total to row.

c. Make a pie chart, selecting under ‘Show Me’.

d. Drag region & total to label.

e. In label, sum(total) right click, quick table calculation, percent of total.



f. Create calculated field, rename to Zero Axis, write code as 0, then ok.

g. Drag Zero Axis twice to rows.

h. Then under marks, two fields of Zero Axis will be there, go to second, remove all fields, then in graph, increase size in 1st, decrease in 2nd, then right click on Zero Axis in pie chart, then dual axis, Change color to White.

i. To remove the lines of zero axis, format-> edit -> none in zero lines.

v) a. Create a butterfly chart by reversing the bar chart to compare female & male revenue based on product category.

b. Drag Gender, Total to rows, Category to columns.

c. Make a pie chart, selecting under 'Show Me'.

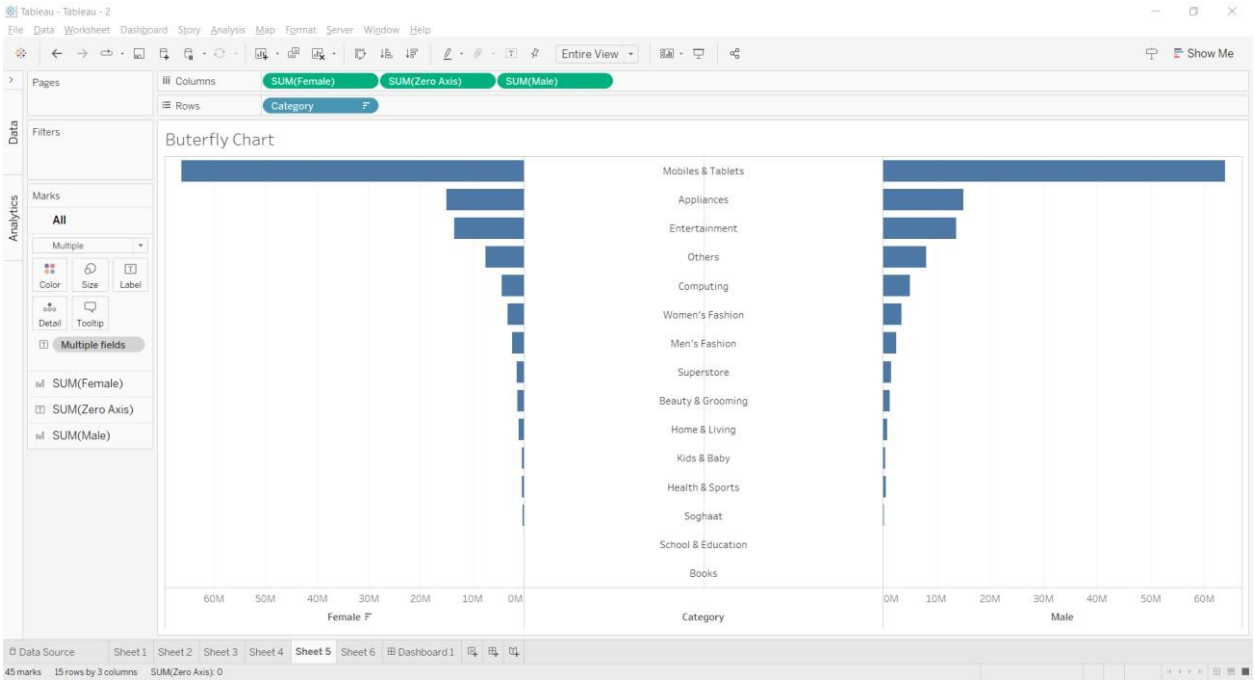
d. Create 2 calculated fields.

e. For female revenue, Code is – if [Gender] = 'F' then [Total]end - Create same for male revenue.

f. Drag female and male to columns, remove total and gender.

g. Drag zero axis between female and male revenue in columns.

h. Rename the Zero Axis as Category by editing, and remove 0 in (tick tab) edit, select none.



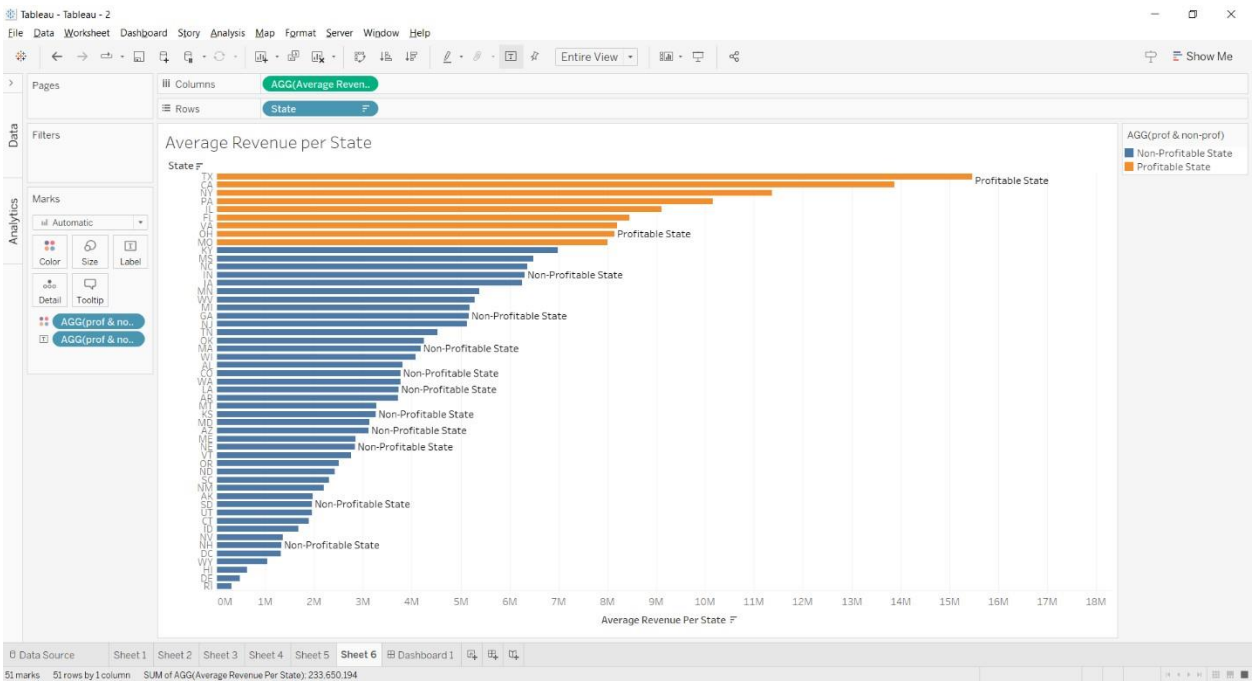
vi) a. Create a calculated field to show the average revenue per state & display profitable & non-profitable state.

b. Create calculated field to calculate average revenue per state, code – avg({include[State] : sum([Total])})

c. Create a calculated field for profitable & non profitable states, code - if([Average Revenue Per State]) >= 8000000 then "Profitable State" else "Non-Profitable State" end .

d. Drag Avg prof and average revenue per state to columns, State to rows.

e. Color the difference for profitable & non-profitable.



vii) a. Create a dashboard, increase the width, and click on “Floating”. - Drag all sheets and arrange properly.

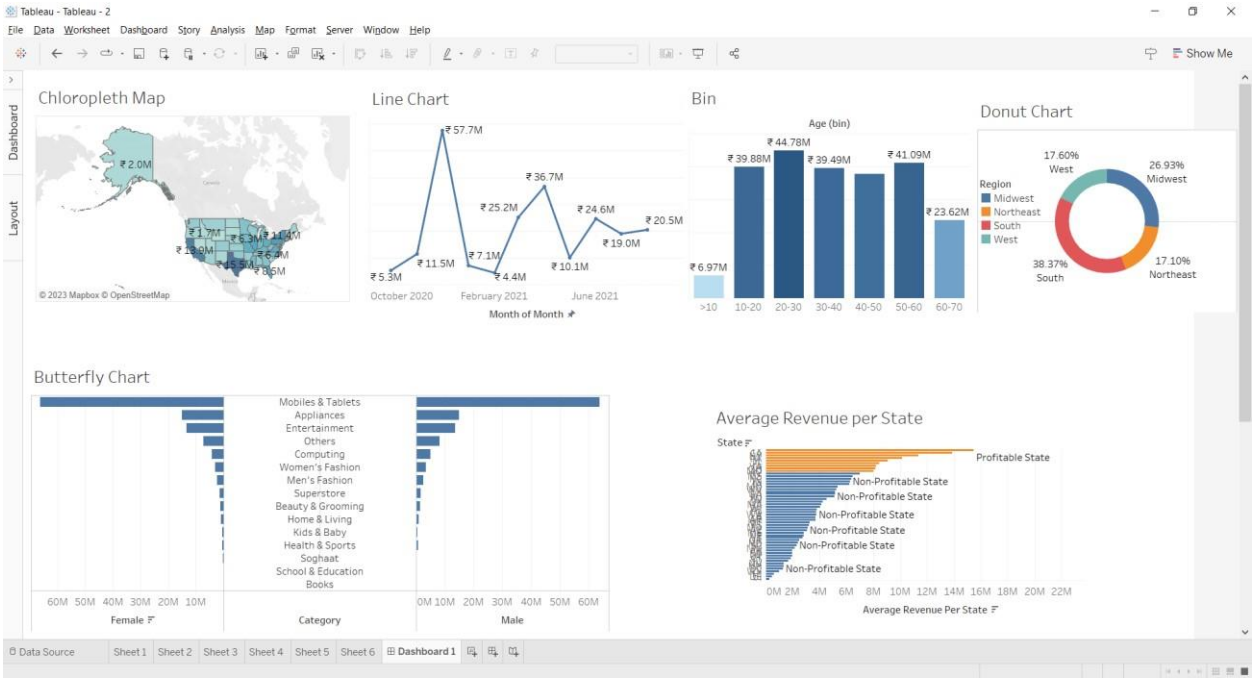


Tableau is a powerful data visualization tool that can be used to analyze large datasets. It can help you to visualize the data in a way that is easy to understand and interpret. Tableau can be used to create a variety of visualizations, including choropleth maps, line charts, bar charts, and donut charts. These visualizations can help you to identify trends and patterns in the data, which can then be used to make informed decisions about your business.

Tableau can be used to answer many questions about your sales data. For example, it can help you to identify which state has the highest revenue, how revenue changes over time, how revenue differs by age group, and what the percentage of revenue is by region. It is a powerful tool that can help you to gain insights into your sales data and make informed decisions about your business.

Problem – 3

GDP Dataset

3. 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.

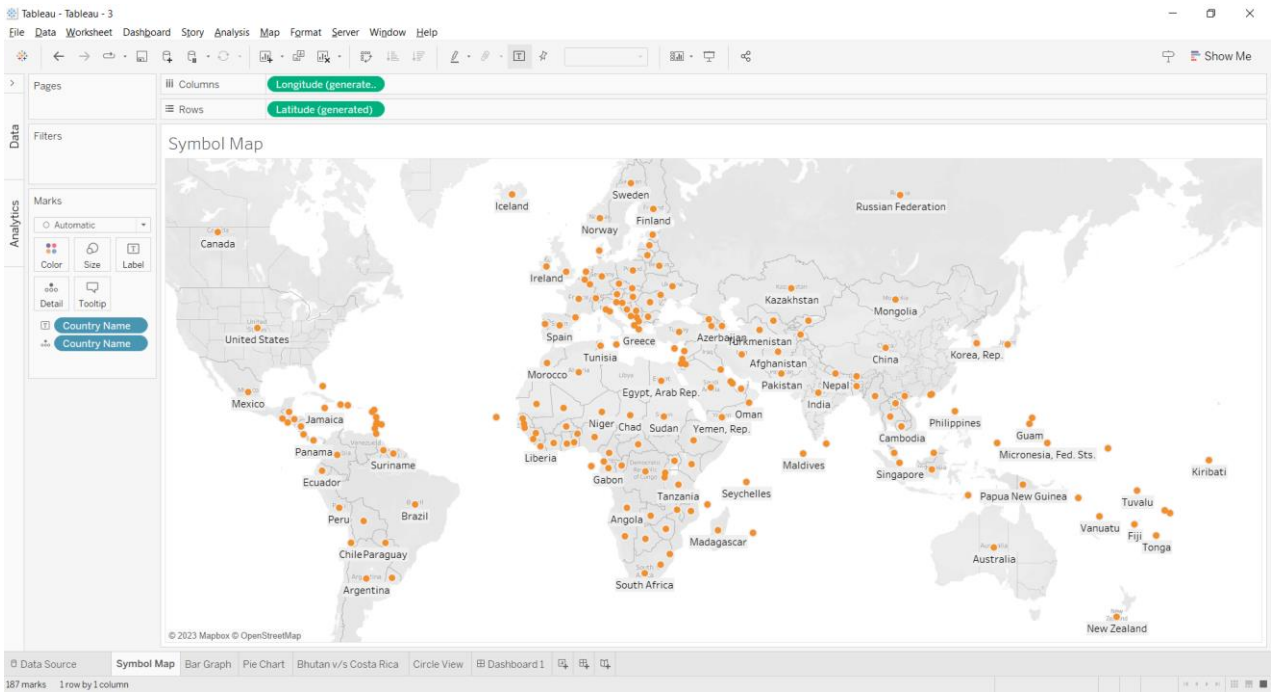
Solutions:

i) a. Drag Country name to columns, latitude of longitude to the rows.

b. Drag country name to label.

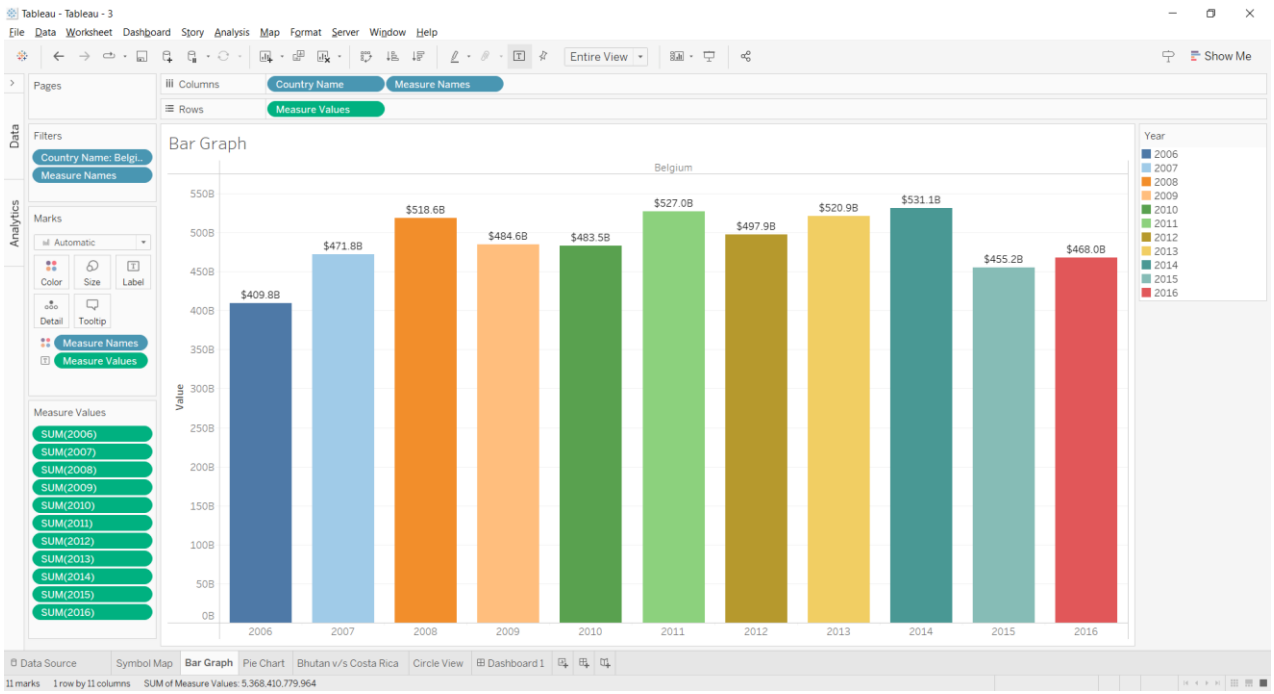
c. Under ‘Show me’, select symbol maps.

d. Drag Country name to color, from drop down, select attribute.

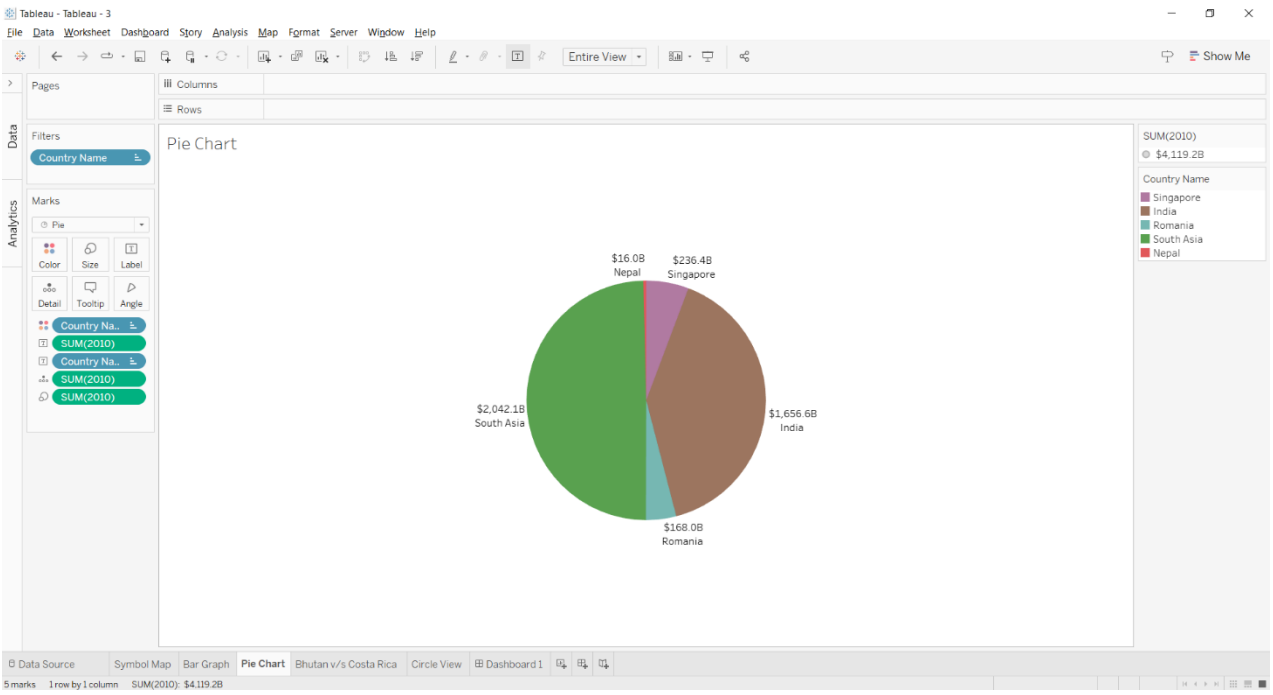


ii) a) Drag Country name to columns, fields from 2006 to 2016 to rows

b) From drop down in Country name present in Columns, select edit filter choose Belgium only and click OK.



iii) a) Drag country name to columns from drop in it, select edit filter and choose India, Nepal, Romania, South Asia, and Singapore

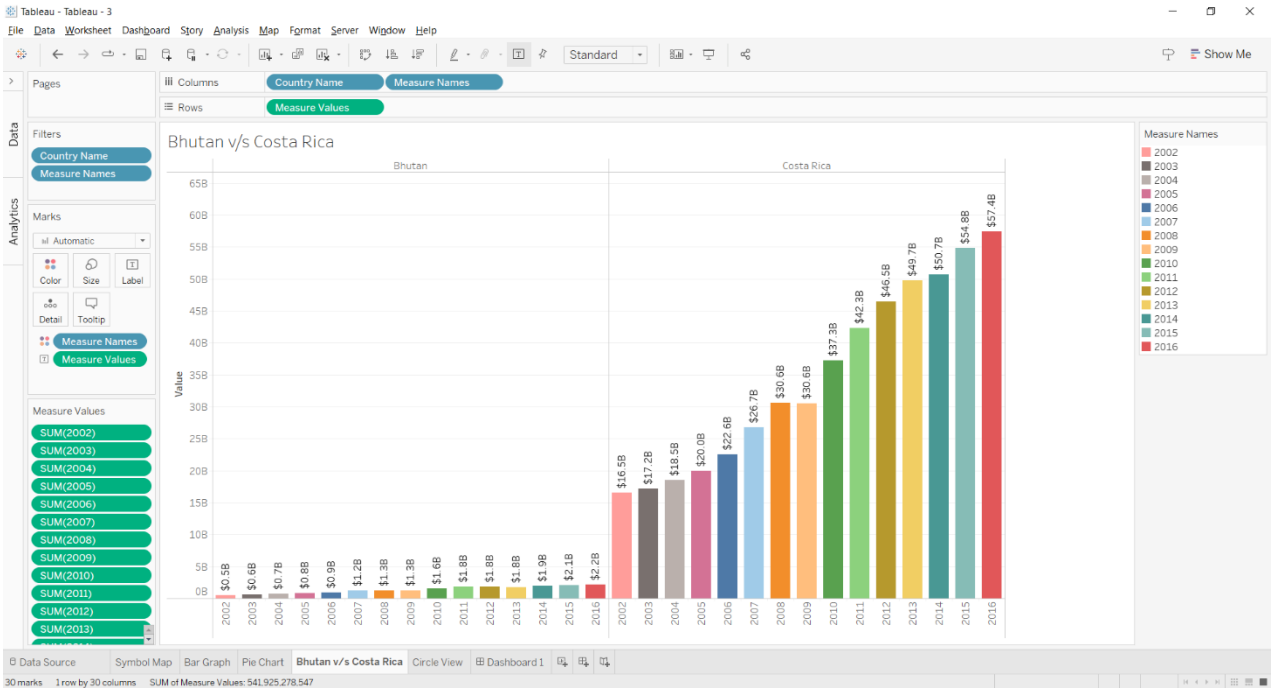


b) Drag Year (2010) to rows.

c) From "Show Me" select Pie chart.

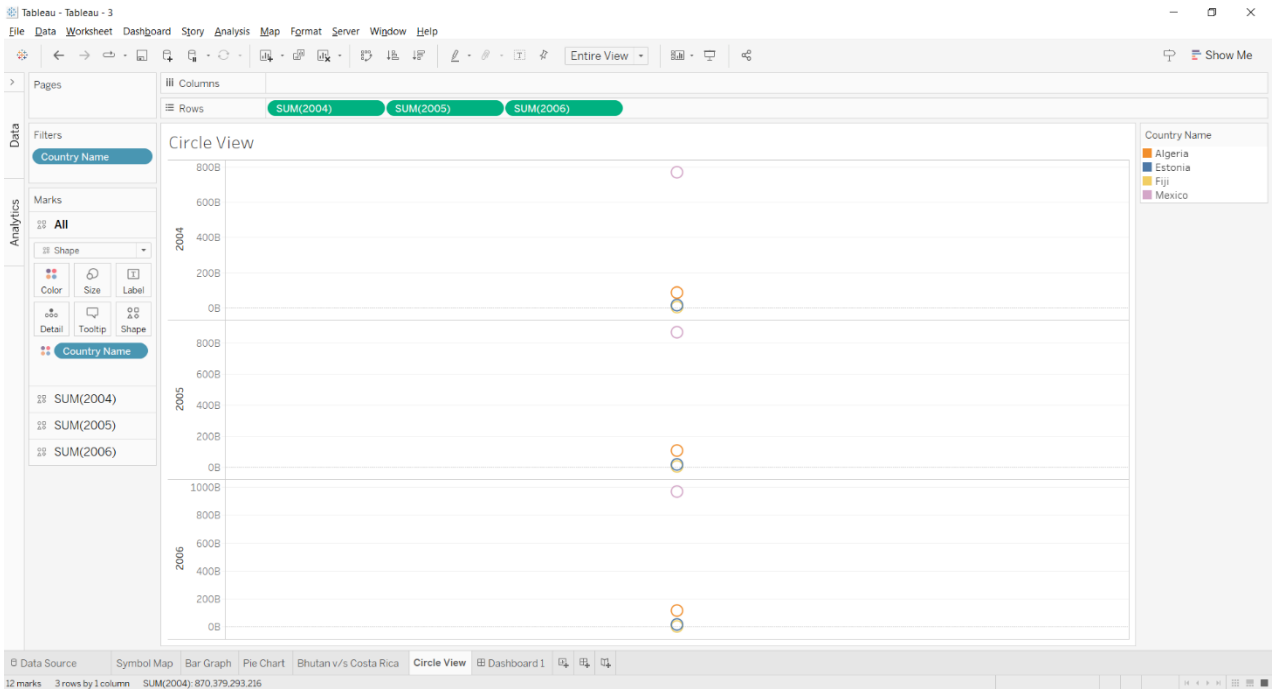
d) Drag Country name, 2010 to Label.

iv) a) Drag Country name to columns from drop down in it, select edit filter and



choose Bhutan & Costa Rica.

b) Drag years from 2008 to 2016 by holding shift key and selecting all at once.



v) a) Drag Country name to columns from drop down, select edit filter and select Mexico, Algeria, Fiji, Estonia and click OK.

b) Drag (Year) 2004, 2005, 2006, to rows

c) From "Show Me" select circle views.

vi) a) Select Dashboard.

b) Change height and width of the dashboard

c) Change objects from tiled to floating.

d) Drag all the sheets and rearrange them.

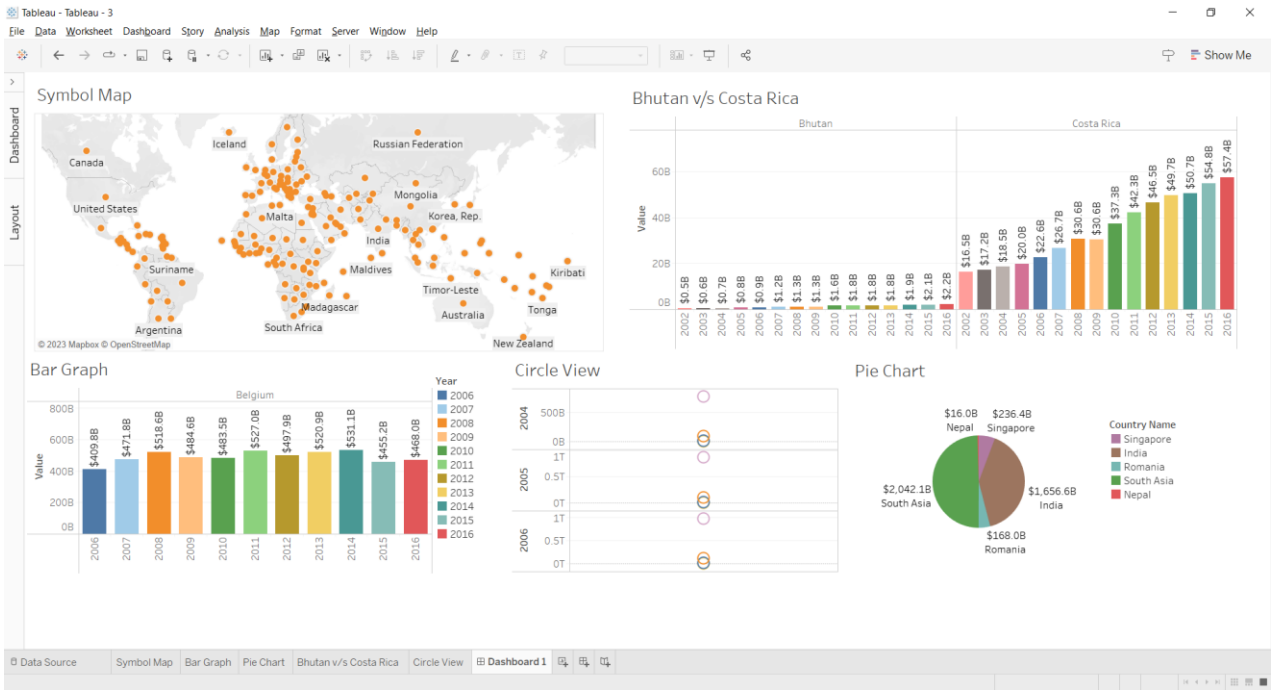


Tableau is essential for sorting and analyzing GDP data due to the data exploration and visualization capabilities. It enables users to easily connect to various data sources visualize GDP tools and sort based on GDP values or other matrices with interactive dashboards and filtering options, stakeholders can explore insights of their own, facilitating better understanding of GDP patterns and disparities. Tableau's geospatial features also help plot GDP data on maps for regional analysis. Additionally, its Scalability & performance ensure efficient handling of large GDP datasets and time series analysis.

Problem – 4

HR Dataset

4. 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.

Solutions:

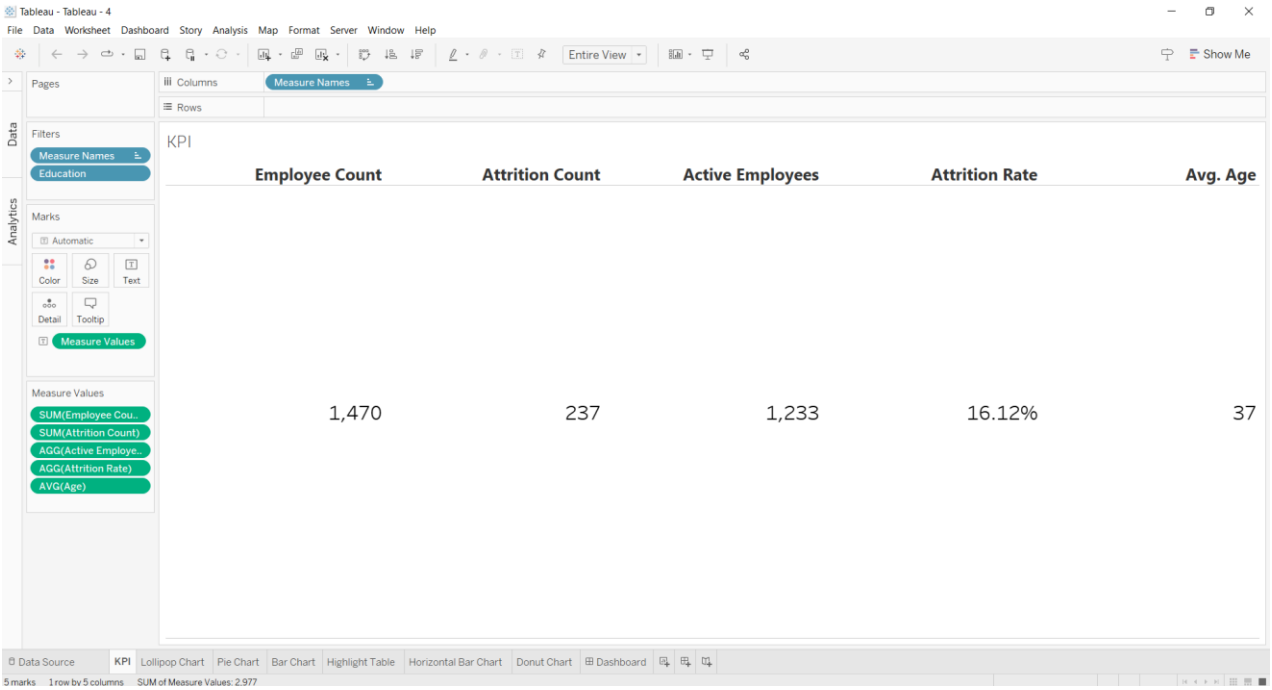
i) a. Employee Count: Drag Employee Count to the Label shelf.

b. Attrition Count: Create a calculated field **Attrition Count** with the formula: IF [Attrition] = 'Yes' THEN 1 ELSE 0 END. Double-click on this calculated field to place it on both the Rows and Columns shelves, then reverse the order.

c. Attrition Rate: Create a calculated field **Attrition Rate** with the formula: SUM([Attrition Count]) / SUM([Employee Count]). Format it as a percentage with 2 decimal places.

d. Active Employees: SUM([Employee Count]) - SUM([Attrition Count]).

e. Average Age: Right-click on the Age field, change measure from SUM to AVG. Format all numbers as decimal places with 0.



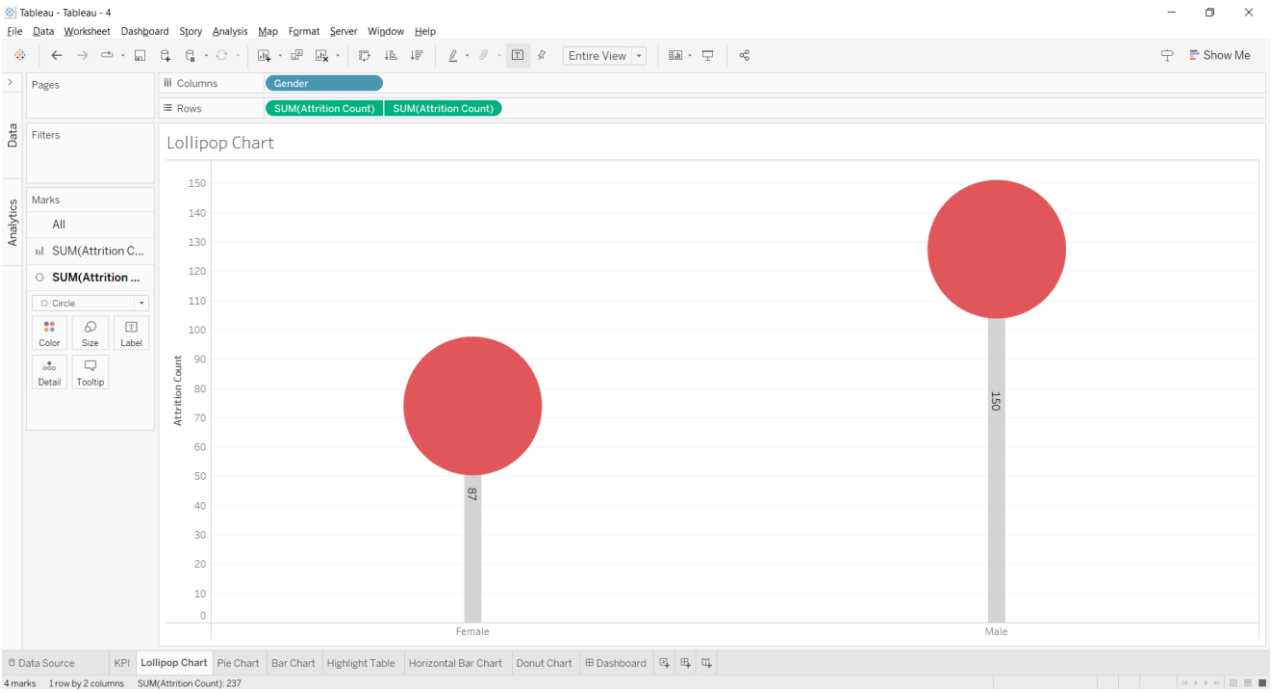
ii) a. Drag Gender to Rows.

b. Drag Attrition Count to Columns twice. Change the second instance to a Circle mark.

c. Create dual axes for both charts (Columns shelf dropdown > Dual Axis).

d. Reduce the size of the circle chart to create the lollipop effect.

e. Add Label to show the count and align it to the center in the format tab.

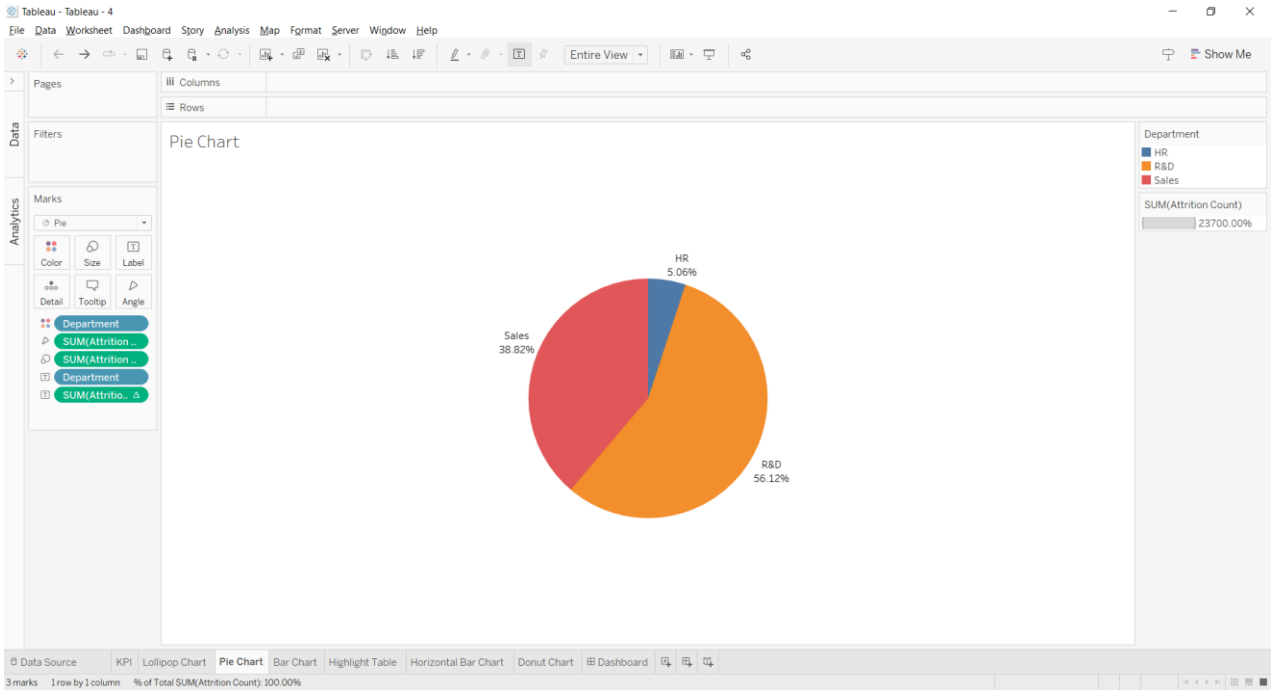


iii) a. Drag Department to Colors shelf.

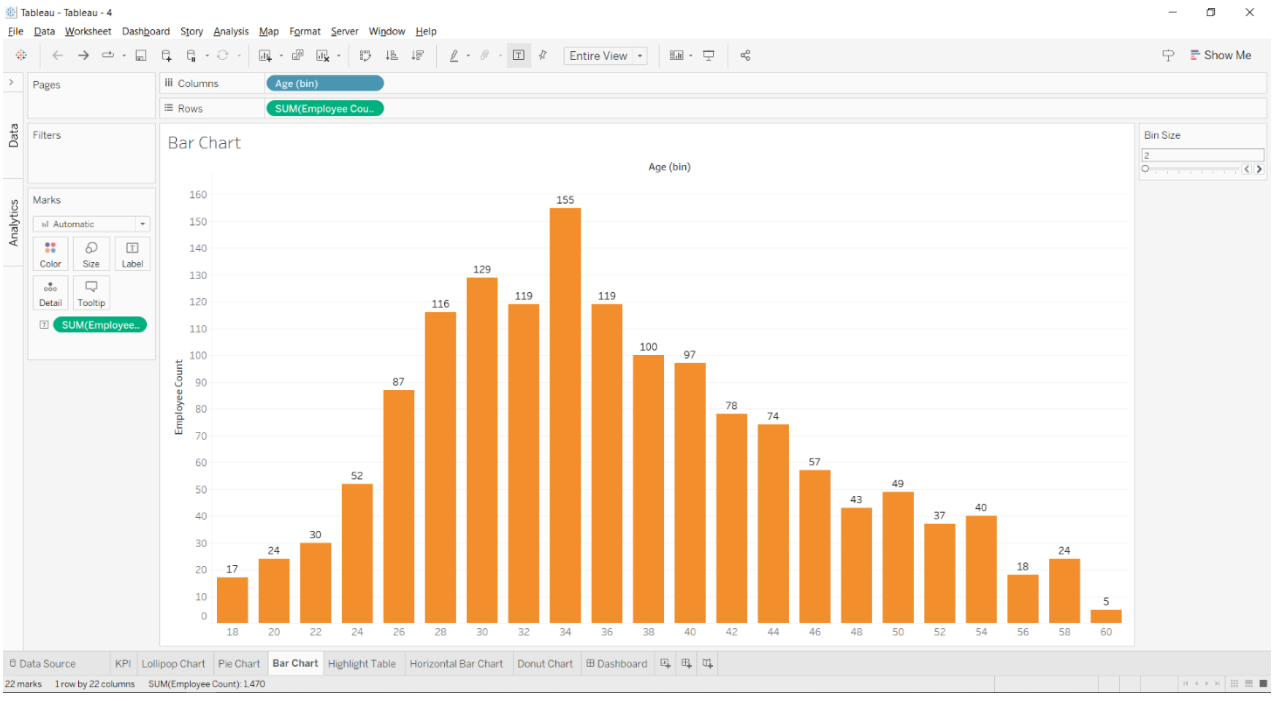
b. Change the mark type to Pie.

c. Label the slices with the Attrition Count and format them as percentages.

d. Add a total label and edit the label accordingly.



iv) a. Create a parameter for Age Bin (dropdown > Create Parameter) with bin size 3, min 2, max 10, step 1.



b. Drag Age Bin to Columns and drag the Employee Count to Rows.

c. Right-click on the Age Bin in Columns, choose Show Parameter Control.

v) a. Drag Job Role to Rows.

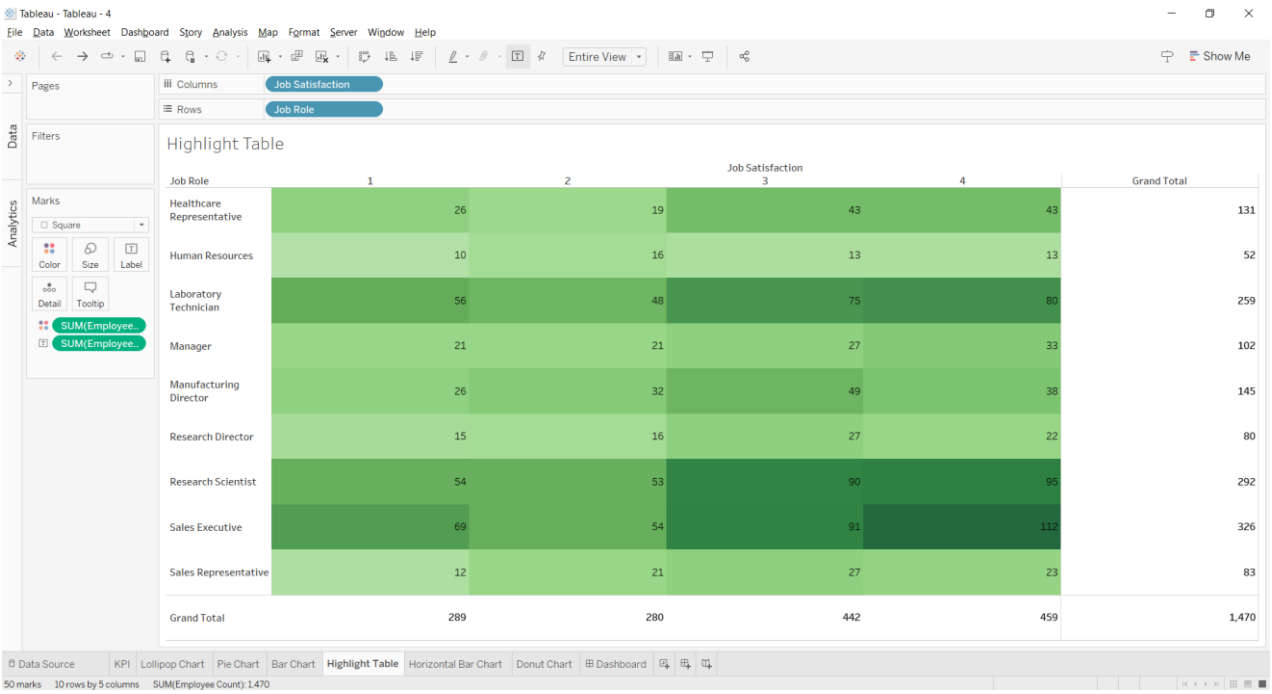
b. Drag Sum of Job Satisfaction to the Columns shelf and then move it to Dimensions.

c. Drag Job Satisfaction to Columns.

d. Drag Employee Count to Text.

e. Analyze > Totals > Show Column and Row Grand Totals.

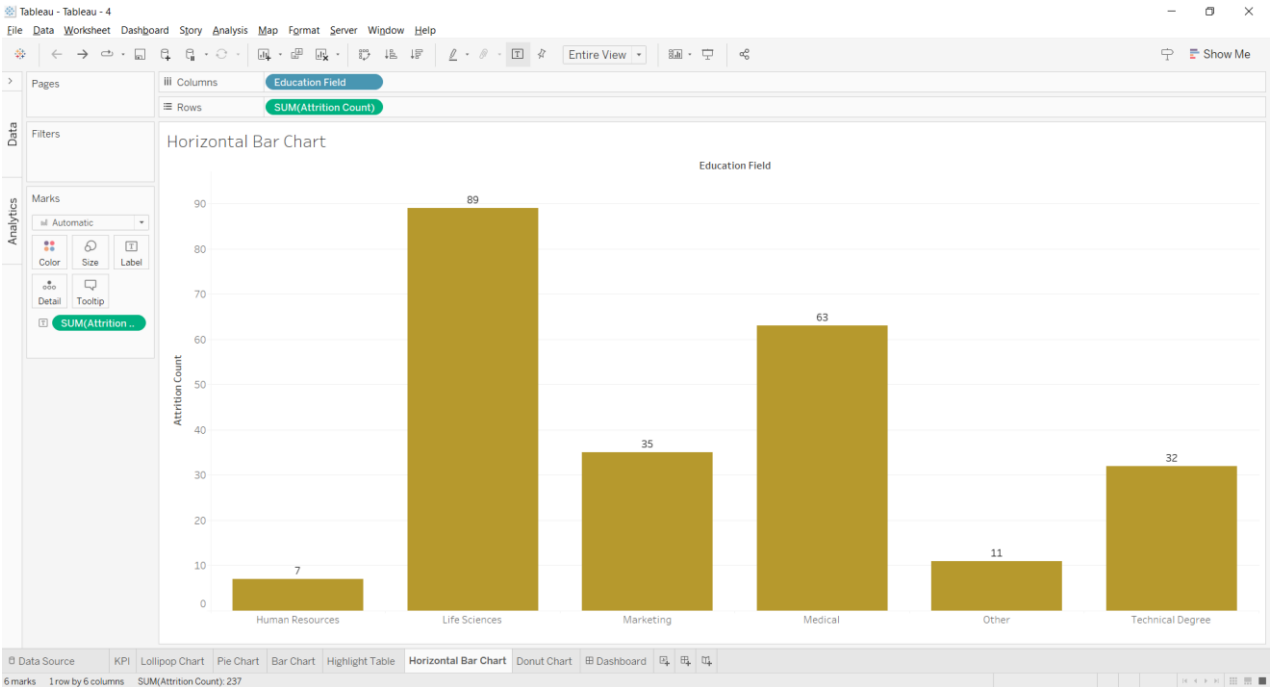
f. Change the chart type to Highlight Table.



vi) a. Drag Education Field to Rows.

b. Drag Sum of Attrition Count to Columns.

c. Label Attrition Count.



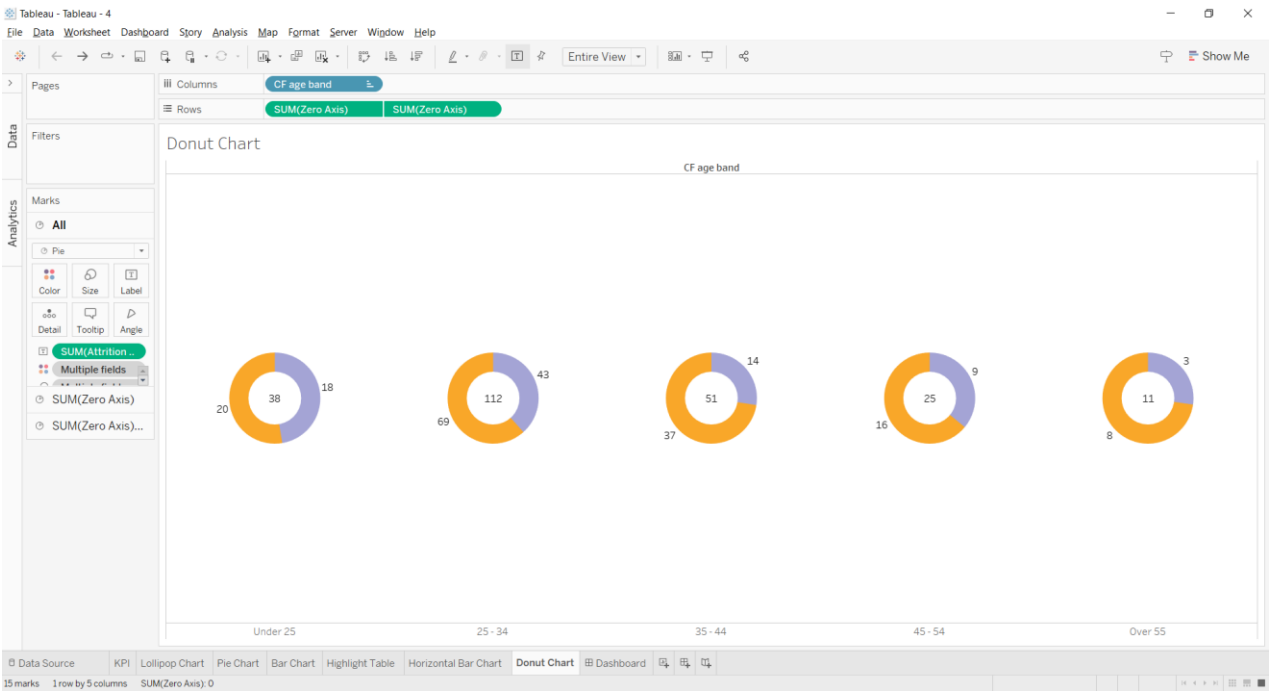
vii) a. Drag Age Band to Columns.

b. Create a pie chart by changing the mark type to Pie and selecting entire view.

c. Drag Gender to Color.

d. Drag Attrition Count to Angle.

e. Create a second instance of Attrition Count and place it on Rows to create a donut chart. Use the Age Band field for sorting by dragging it to the dropdown and selecting Manual.

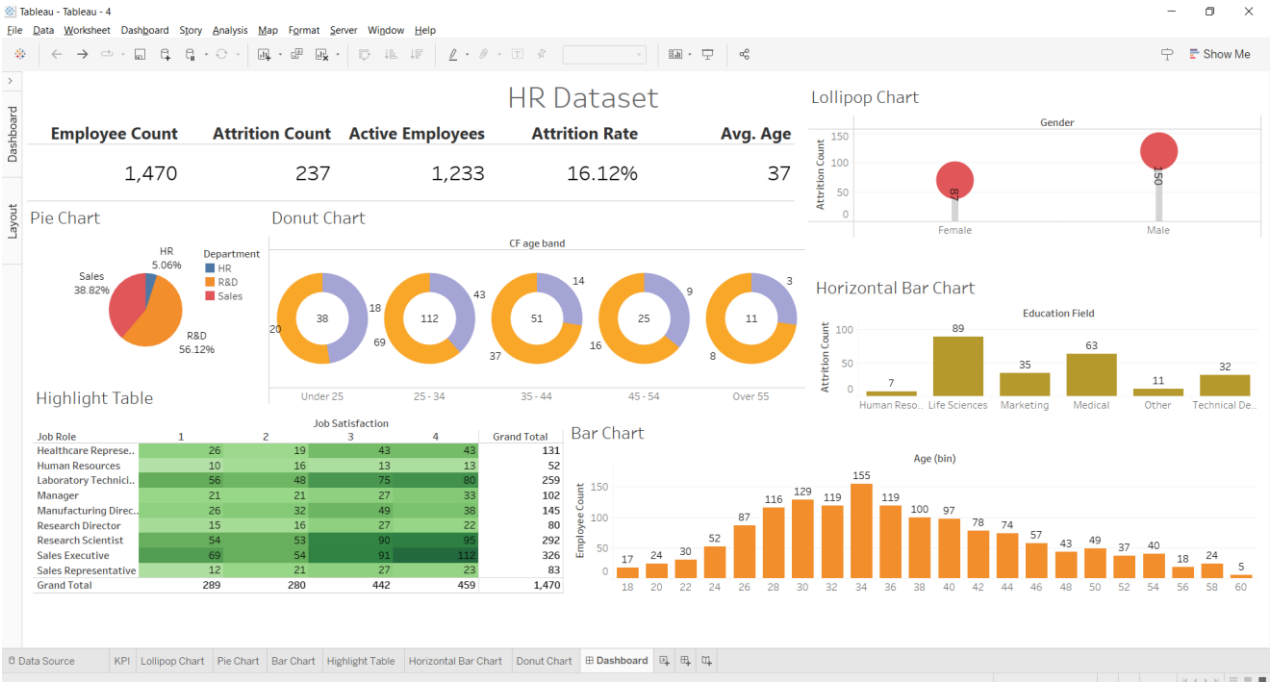


viii) a. Create a new Dashboard.

b. Rename the Dashboard.

c. Change the objects to floating.

d. Drag all the sheets, label & rearrange them.



Analyzing HR Data in Tableau is crucial for organizations to gain valuable insights into their workforce and make informed decisions. By examining variables such as turnover rates, employee performance diversity & compensation, Tableau can help HR teams identify patterns, improve employee satisfaction, and optimize workforce planning. Visualizing demographic data can aid in promoting inclusivity, while salary analysis ensures fair compensation practice. Evaluating factors like engagement career growth, and geographical impact helps foster a productive and satisfied work pace. Through data-driven decision making, Tableau empowers HR professionals and organizational leaders to create an efficient, engaged, and successful workforce.

Problem – 5

Amazon Prime Dataset

5. 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.

Solutions:

i) a. Drag Type and Title to column and rows shelf respectively

b. Select Pie Chart from ‘Show Me' section.

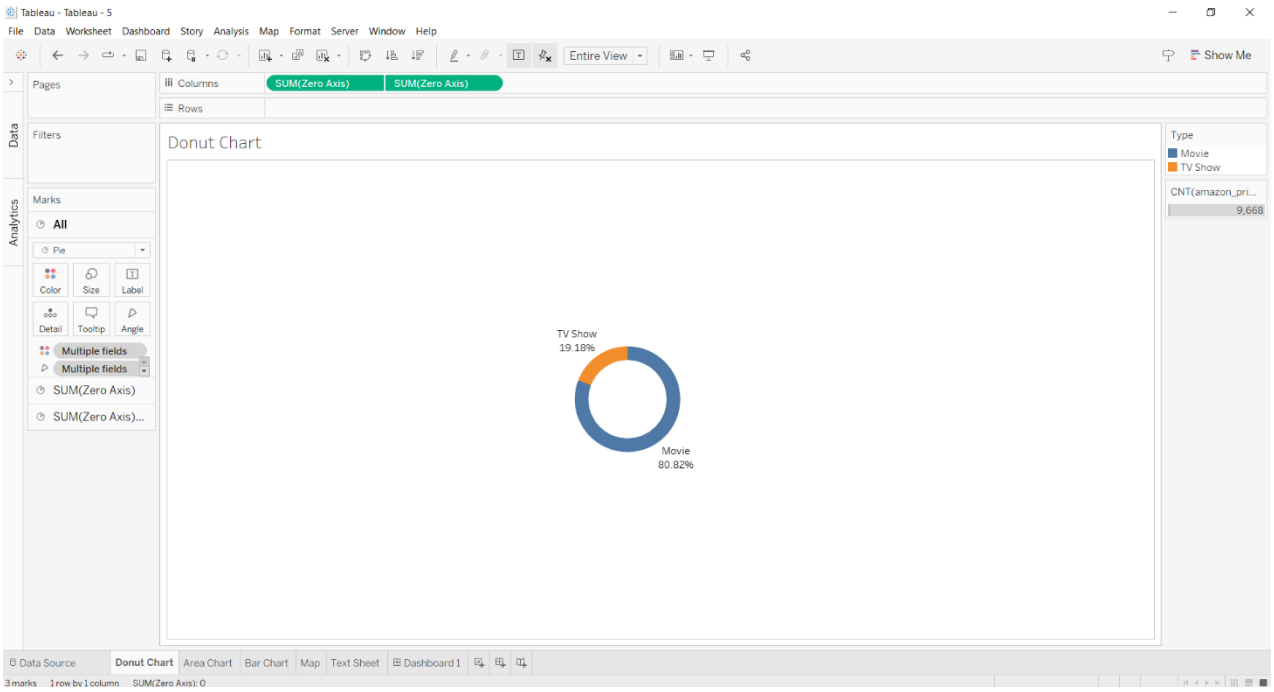
c. Right click on the title - Measure → Count (distinct), type to Color.

d. Drag Type and Title to Label

e. Create calculated field called Zero Axis

f. Drag it twice to rows shelf.

g. Then under Marks, two fields of Zero Axis exist go to second one - remove all fields and decrease its size.



h. Right click on the second Zero Axis and select dual axis.

i. Change the color of the second pie chart to white.

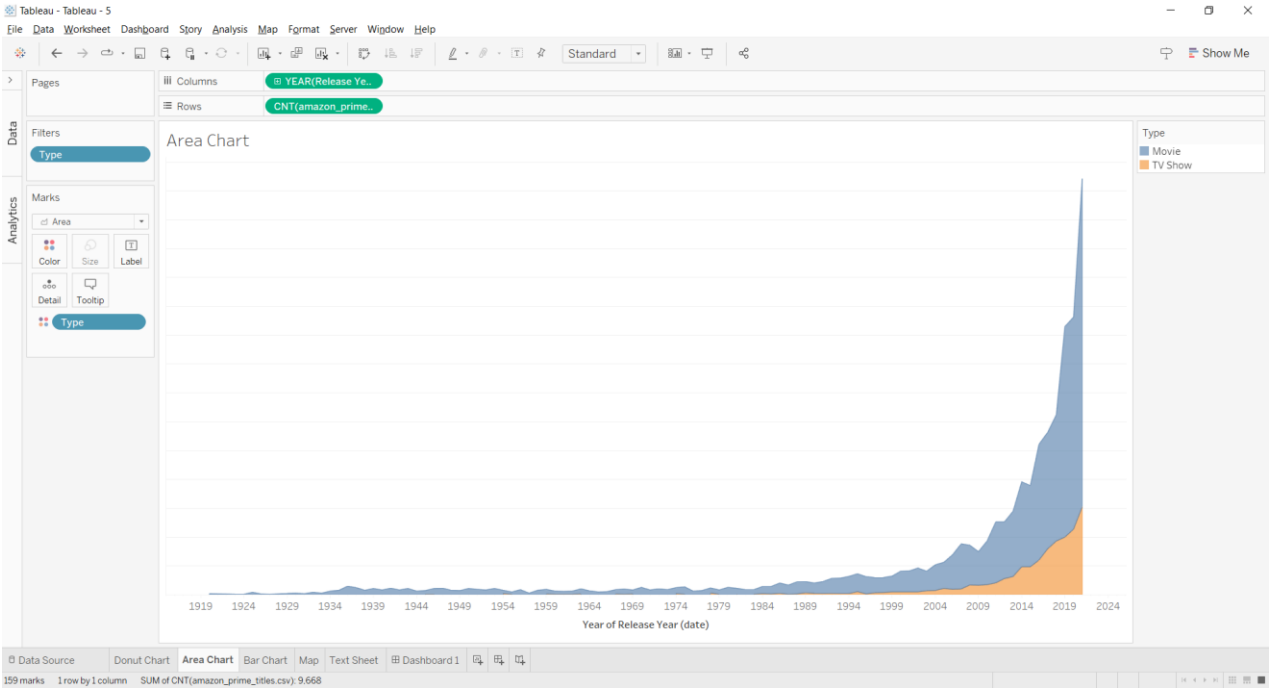
ii) a. Drag and drop 'Release year' and "Show ID" to Columns and rows shelf respectively.

b. Right click on 'Show ID' in rows shelf, click on measure select Count (distinct)

c. In the “Marks" section, select area from the drop-down list in place of automatic.

d. Drag and drop 'Type' to color.

e. Drag and drop 'Type' to Label.



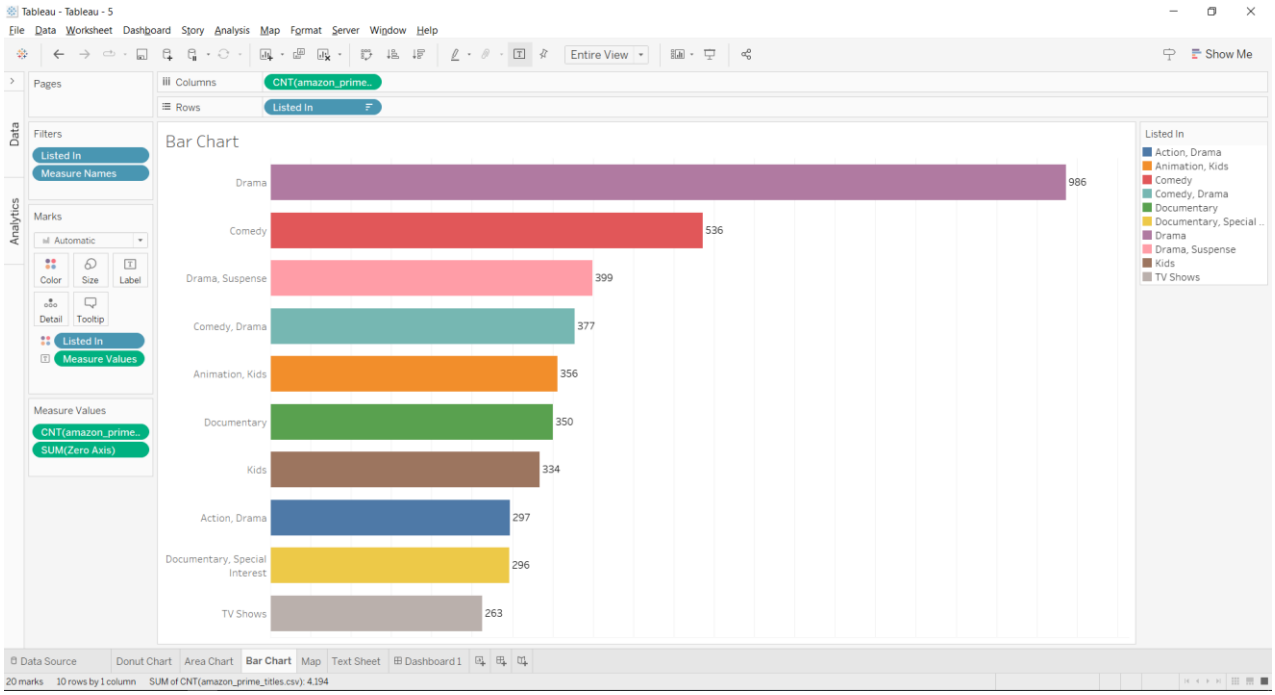
iii) a. Drag Listed In" and "Show ID" to rows and Column shelf respectively

b. Right click on "Show ID” then go to Measure → Count (distinct)

c. Drag 'Listed In' to filter and edit it Accordingly to get top 10 Genres.

d. Drag "Show ID" to label

e. Select Horizontal Bars in Show me drop down to get horizontal bar chart.



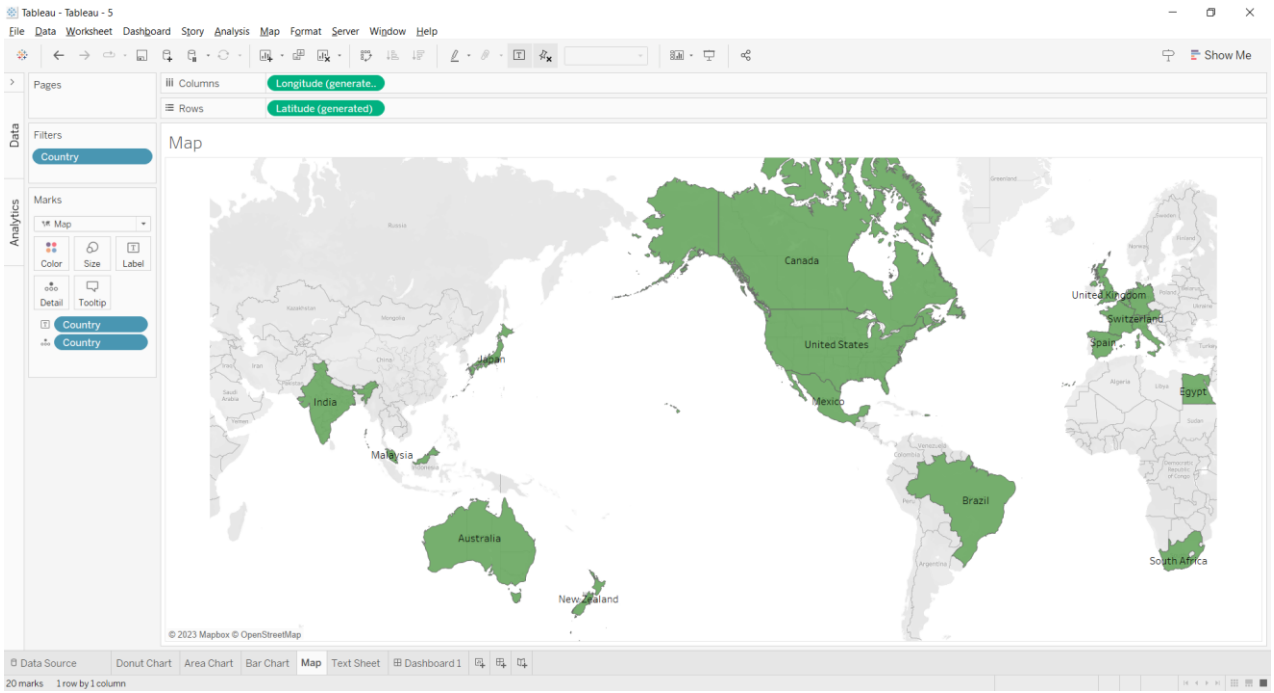
iv) a. Drag Country and Show ID to rows and column shelf respectively.

b. Drag "Country" to filters and remove null values.

c. Drag Type to filters and select TV Shows only.

d. Drag "Country" and "Show ID" to Label.

e. Select Map in place of "automatic from drop down list in Marks.



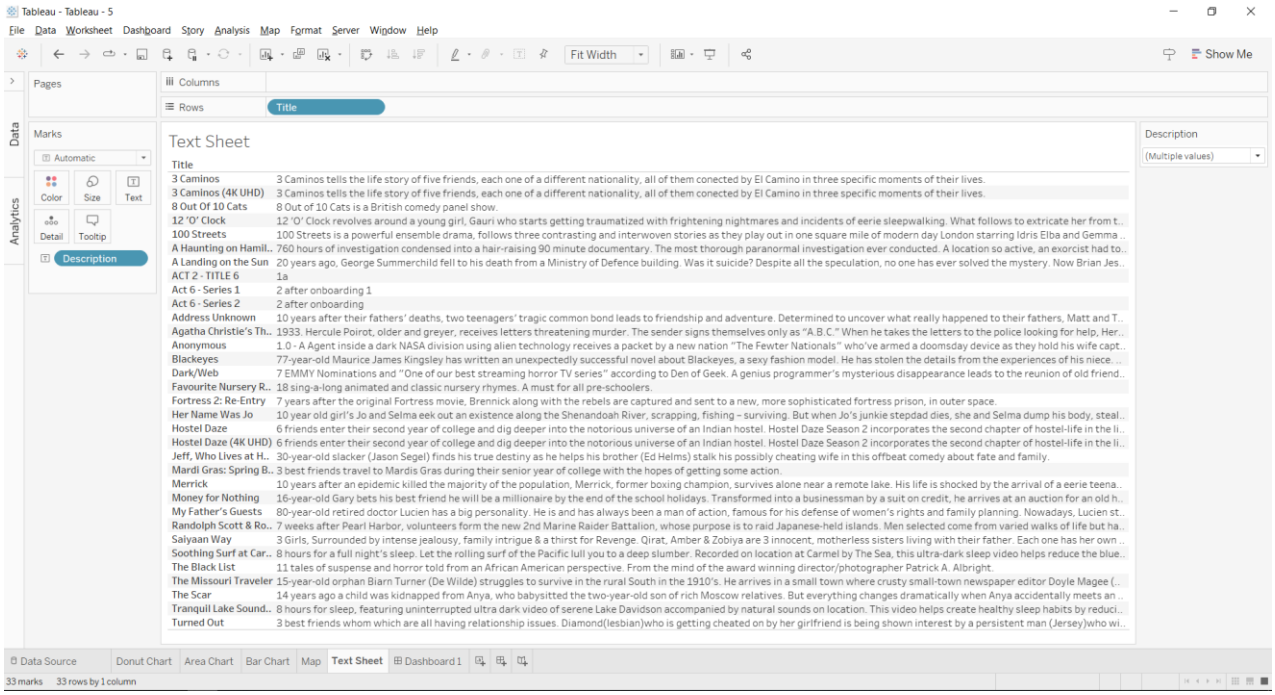
v) a. Drag and drop Title to rows

b. Select "Text" in place of automatic in the Marks section.

c. Add Title to filter. Right click on 'Title' and select Show Filter.

d. Drag "Type" to filter and select Movies only.

e. Drag and drop Description to Text.



vi) a. Click on the Create new Dashboard button located at bottom left corner of Tableau Window.

b. Increase the width of the dashboard.

c. Select floating windows under objects in the dashboard.

d. Drag and drop all sheets and arrange them properly.

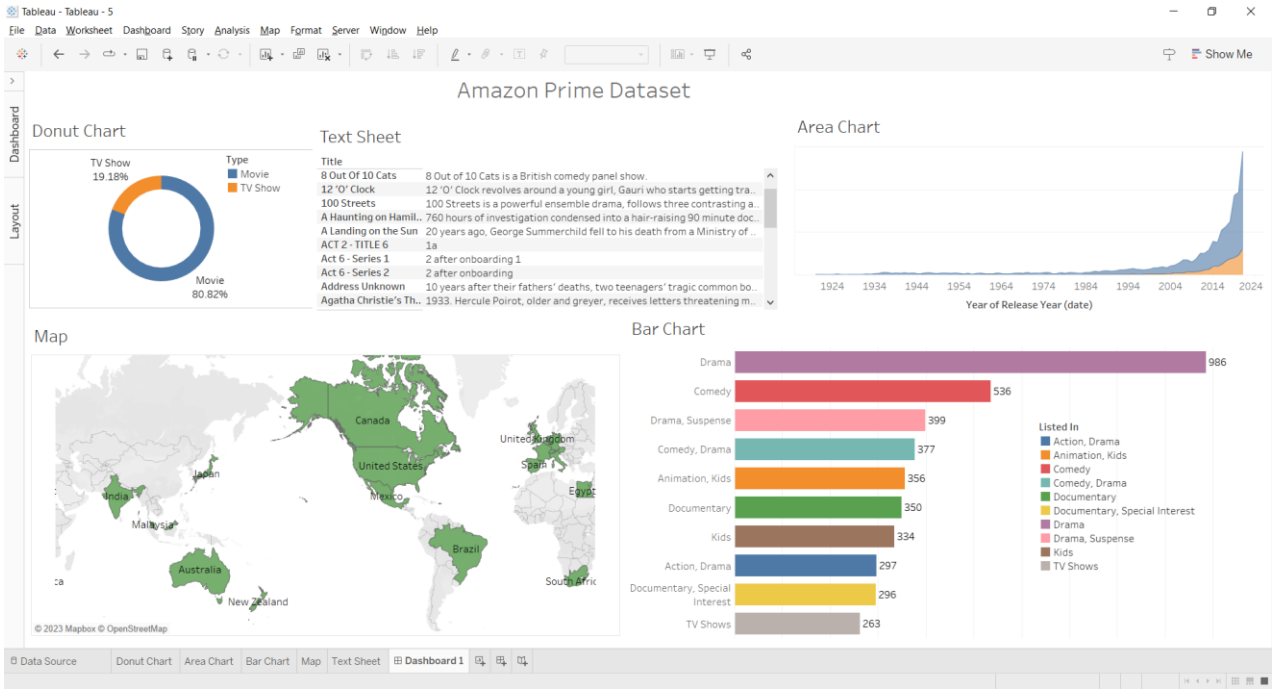


Tableau - Tableau - 5
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Show Me
Dashboard
Layout
Amazon Prime Dataset
Donut Chart
TV Show
19.18%
Type
Movie
TV Show
Movie
80.82%
Text Sheet
Title
8 Out Of 10 Cats
8 Out of 10 Cats is a British comedy panel show.
12 'O' Clock
12 'O' Clock revolves around a young girl, Gauri who starts getting tra..
100 Streets
100 Streets is a powerful ensemble drama, follows three contrasting a..
A Haunting on Hamil..
760 hours of investigation condensed into a hair-raising 90 minute doc..
A Landing on the Sun
20 years ago, George Summerchild fell to his death from a Ministry of ..
ACT 2 - TITLE 6
1a
Act 6 - Series 1
2 after onboarding 1
Act 6 - Series 2
2 after onboarding
Address Unknown
10 years after their fathers' deaths, two teenagers' tragic common bo..
Agatha Christie's Th..
1933. Hercule Poirot, older and greyer, receives letters threatening m..
Area Chart
1924
1934
1944
1954
1964
1974
1984
1994
2004
2014
2024
Year of Release Year (date)
Map
Canada
United States
Mexico
Brazil
Australia
New Zealand
Japan
India
Malaysia
United Kingdom
Spain
Egypt
South Africa
© 2023 Mapbox © OpenStreetMap
Bar Chart
Drama
986
Comedy
536
Drama, Suspense
399
Comedy, Drama
377
Animation, Kids
356
Documentary
350
Kids
334
Action, Drama
297
Documentary, Special Interest
296
TV Shows
263
Listed In
Action, Drama
Animation, Kids
Comedy
Comedy, Drama
Documentary
Documentary, Special Interest
Drama
Drama, Suspense
Kids
TV Shows
Data Source
Donut Chart
Area Chart
Bar Chart
Map
Text Sheet
Dashboard 1