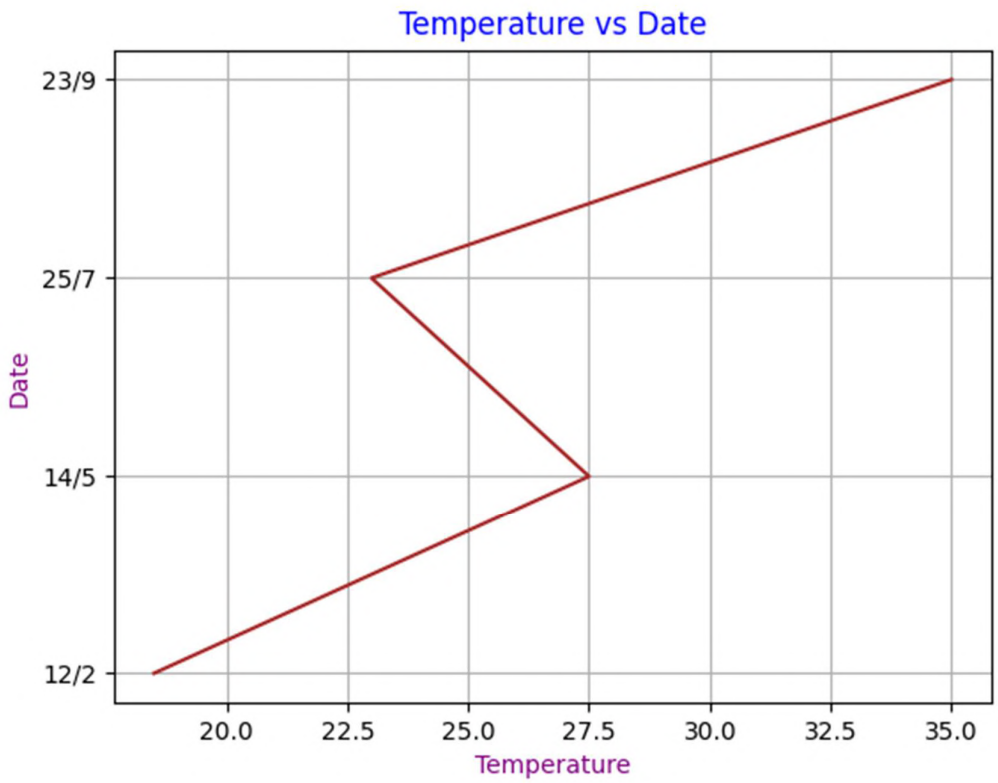
DATA VISUALIZATION LAB

22ISL384

Program – 1

Write a python program to plot Temperature against Date add title, Grids to the plot and also add label on X and Y axis.

```
import matplotlib.pyplot as plt
date=["12/2", "14/5", "25/7", "23/9"]
temp=[18.5, 27.5, 23, 35]
plt.plot(temp, date, color='brown')
plt.xlabel('Temperature', color='purple')
plt.ylabel('Date', color='purple')
plt.title('Temperature vs Date', color='blue')
plt.grid(True)
plt.show()
```



Program – 2

Smile NGO has participated in a three-week cultural mela. Using Pandas, they have stored the sales (in Rs) made day wise for every week in a CSV file named "MelaSales.csv", as shown in Table

Week 1	Week 2	Week 3
5000	4000	4000
5900	3000	5800
6500	5000	3500
3500	5500	2500
4000	3000	3000
5300	4300	5300
7900	5900	6000

Depict the sales for the three weeks using a Line chart. It should have the following:

- Chart title as "Mela Sales Report".
- X axis label as Days.
- Y axis label as "Sales in Rs".
- Line colours are red for week 1, blue for week 2 and brown for week 3.

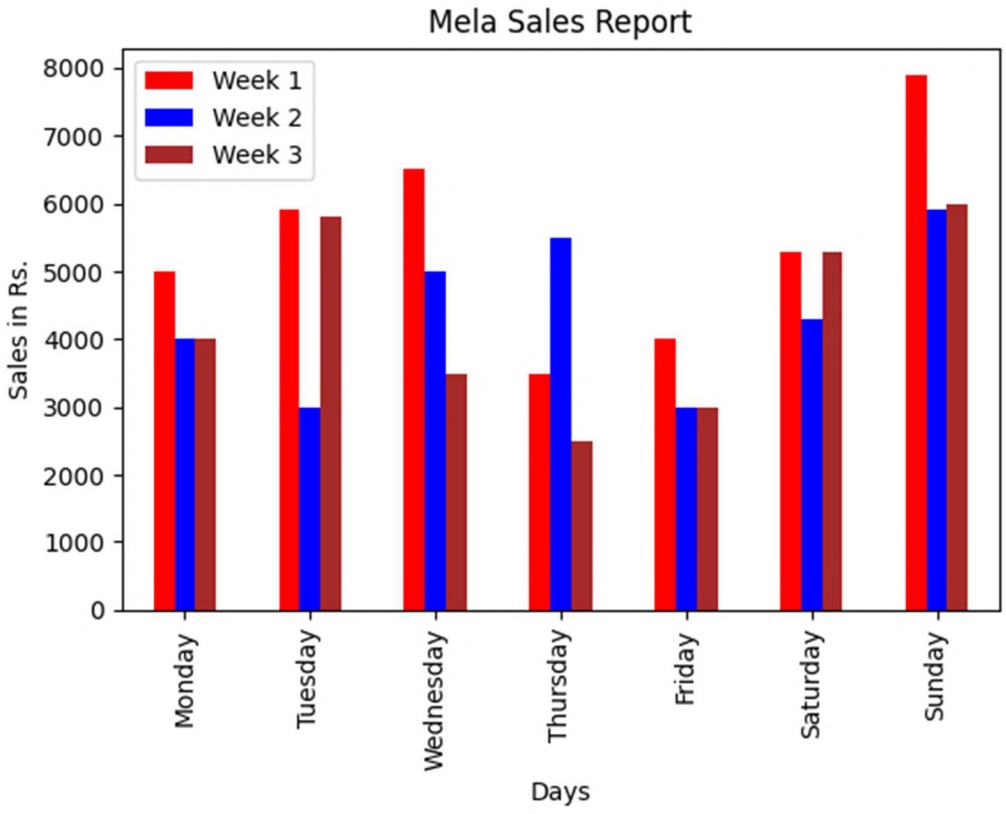
Write a Python script to display Bar graph for the "MelaSales.csv", as shown below with column Day.

Week 1	Week 2	Week 3	Day
5000	4000	4000	Monday
5900	3000	5800	Tuesday
6500	5000	3500	Wednesday
3500	5500	2500	Thursday
4000	3000	3000	Friday
5300	4300	5300	Saturday
7900	5900	6000	Sunday

```
import pandas as pd
import matplotlib.pyplot as plt
df=pd.read_csv('Melasales.csv')
df.plot(kind='line',color=['red','blue','brown'])
plt.title('Mela Sales Report')
plt.xlabel('Days')
plt.ylabel('Sales in Rs.')
plt.show()
```

```
import pandas as pd
import matplotlib.pyplot as plt
df=pd.read_csv('Melasales.csv')
df.plot(kind='bar',x='Day',color=['red','blue','brown'])
plt.title('Mela Sales Report')
plt.xlabel('Days')
plt.ylabel('Sales in Rs.')
plt.show()
```


Mela Sales Report
Week 1
Week 2
Week 3
Sales in Rs.
8000
7000
6000
5000
4000
3000
0
1
2
3
4
5
6
Days

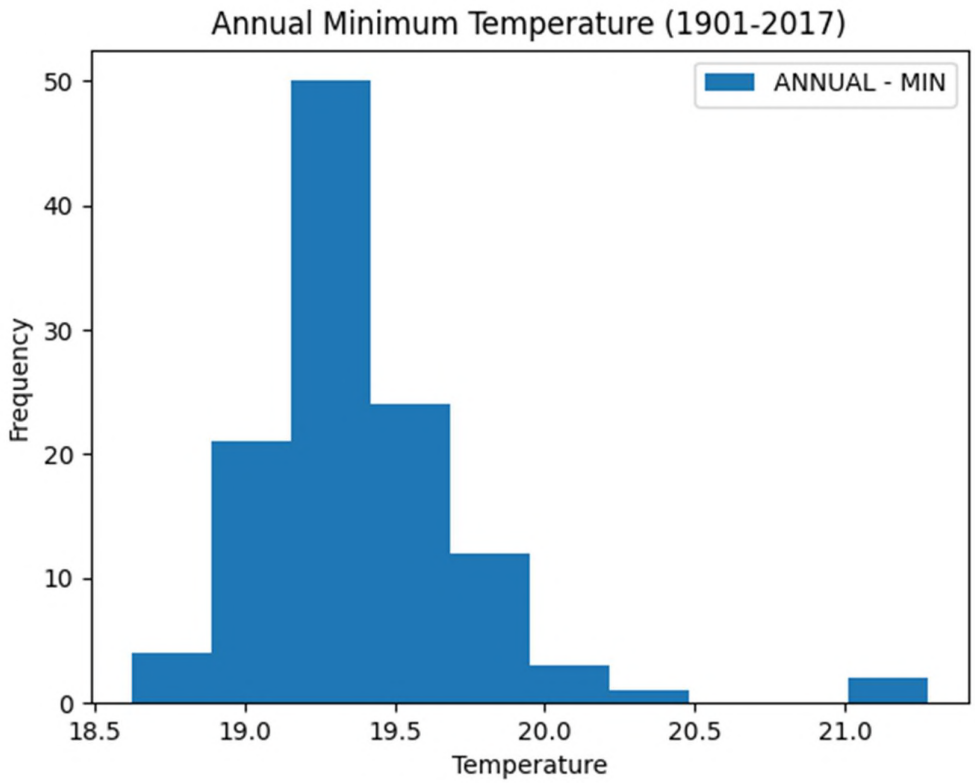

Mela Sales Report
Week 1
Week 2
Week 3
Sales in Rs.
8000
7000
6000
5000
4000
3000
2000
1000
0
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday
Days

Program – 3

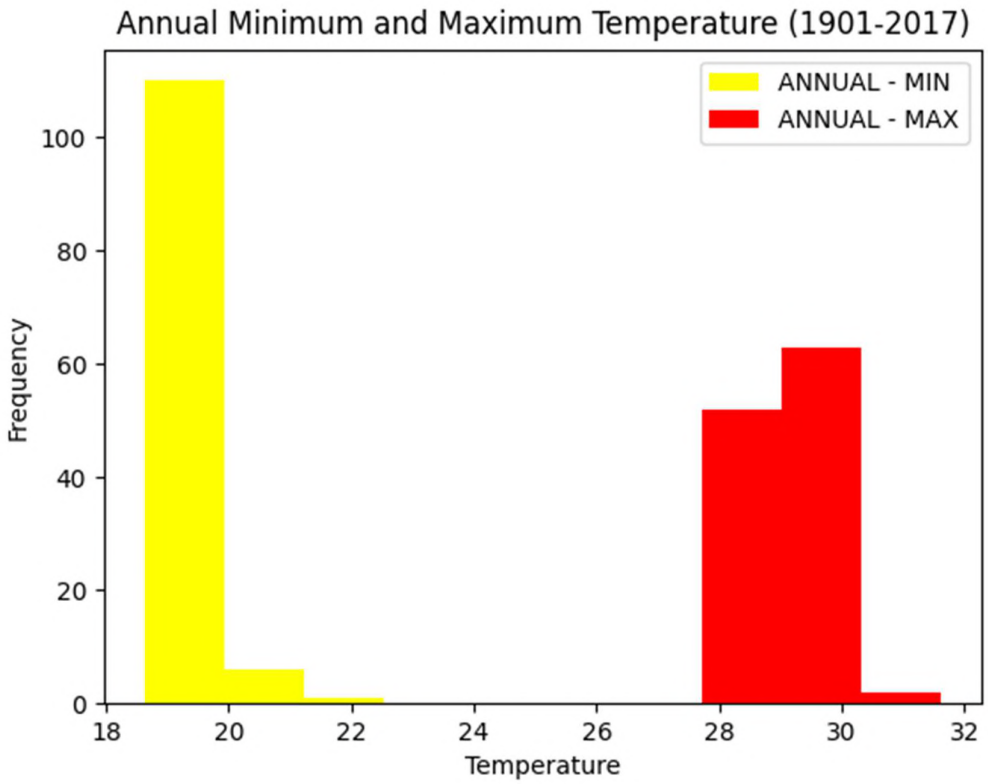
Plot the minimum and maximum temperature on a histogram for ***Seasonal and Annual Min/Max Temp Series - India from 1901 to 2017*** and observe the number of times (frequency)
a particular temperature has occurred from 1901 to 2017. Display two Histogram plots:

1. Only for 'ANNUAL - MIN'
2. For both 'ANNUAL - MIN' and 'ANNUAL - MAX'

```
import pandas as pd
import matplotlib.pyplot as plt
data=pd.read_csv('Max_Min_Seasonal_Temp_IMD-
1901_to_2017.csv',usecols=['ANNUAL - MIN','ANNUAL - MAX'])
df=pd.DataFrame(data)
df.plot(kind='hist',y='ANNUAL - MIN',title='Annual Minimum Temperature
(1901-2017)')
plt.xlabel('Temperature')
plt.ylabel('Frequency')
plt.show()
```



```
import pandas as pd
import matplotlib.pyplot as plt
data=pd.read_csv('Max_Min_Seasonal_Temp_IMD-
1901_to_2017.csv',usecols=['ANNUAL - MIN','ANNUAL - MAX'])
df=pd.DataFrame(data)
df.plot(kind='hist',title='Annual Minimum and Maximum Temperature (1901-
2017)',color=['yellow','red'])
plt.xlabel('Temperature')
plt.ylabel('Frequency')
plt.show()
```



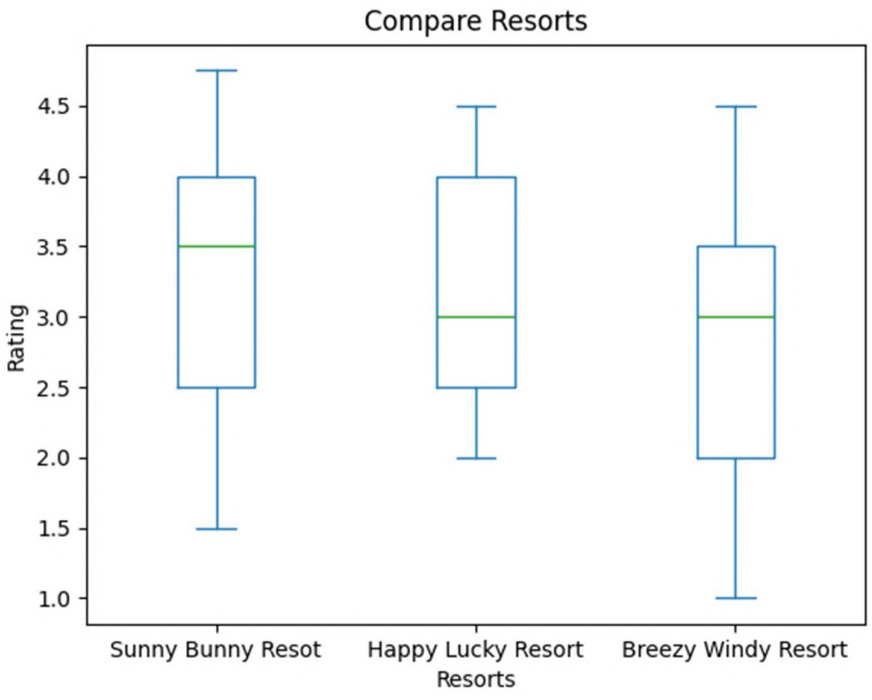
Program – 4

To keep improving their services, XYZ group of resorts have asked all the three resorts to get feedback form filled by their customers at the time of Check-Out. After getting ratings on a scale of (1–5) on factors such as Food, Service, Ambience, Activities, Distance from tourist spots. They calculate the average rating and store it in a CSV file.

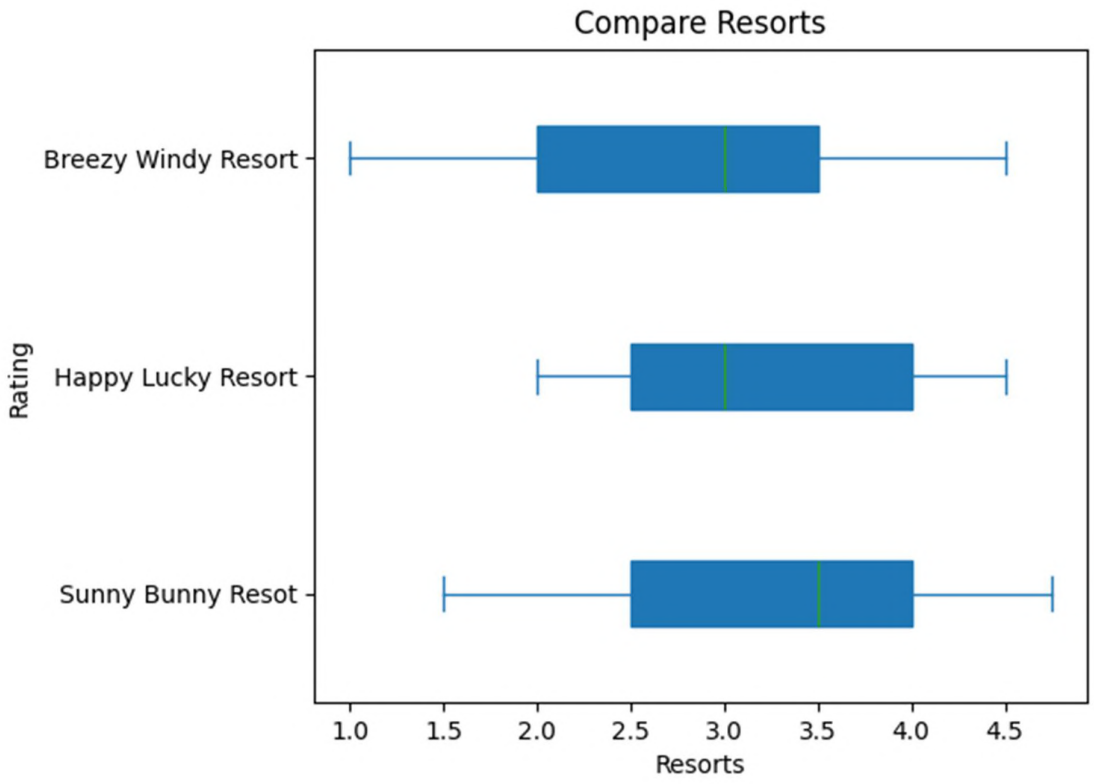
Year	Sunny Bunny Resort	Happy Lucky Resort	Breezy WIndy Resort
2014	4.75	3	4.5
2015	2.5	4	2
2016	3.5	2.5	3
2017	4	2	3.5
2018	1.5	4.5	1

This year, to award the best resort they have decided to analyse the ratings of the past 5 years for each of the resorts. Plot the data using Boxplot. Display the whisker in horizontal and vertical direction.

```
import pandas as pd
import matplotlib.pyplot as plt
data=pd.read_csv('L:\Python\DataVisualization\Resort.csv')
df=pd.DataFrame(data)
df.plot(kind='box',x='Year',title="Compare Resorts")
plt.xlabel('Resorts')
plt.ylabel('Rating')
plt.show()
```



```
import pandas as pd
import matplotlib.pyplot as plt
data=pd.read_csv('L:\Python\DataVisualization\Resort.csv')
df=pd.DataFrame(data)
df.plot(kind='box',x='Year',title="Compare
Resorts",vert=False,patch_artist=True)
plt.xlabel('Resorts')
plt.ylabel('Rating')
plt.show()
```



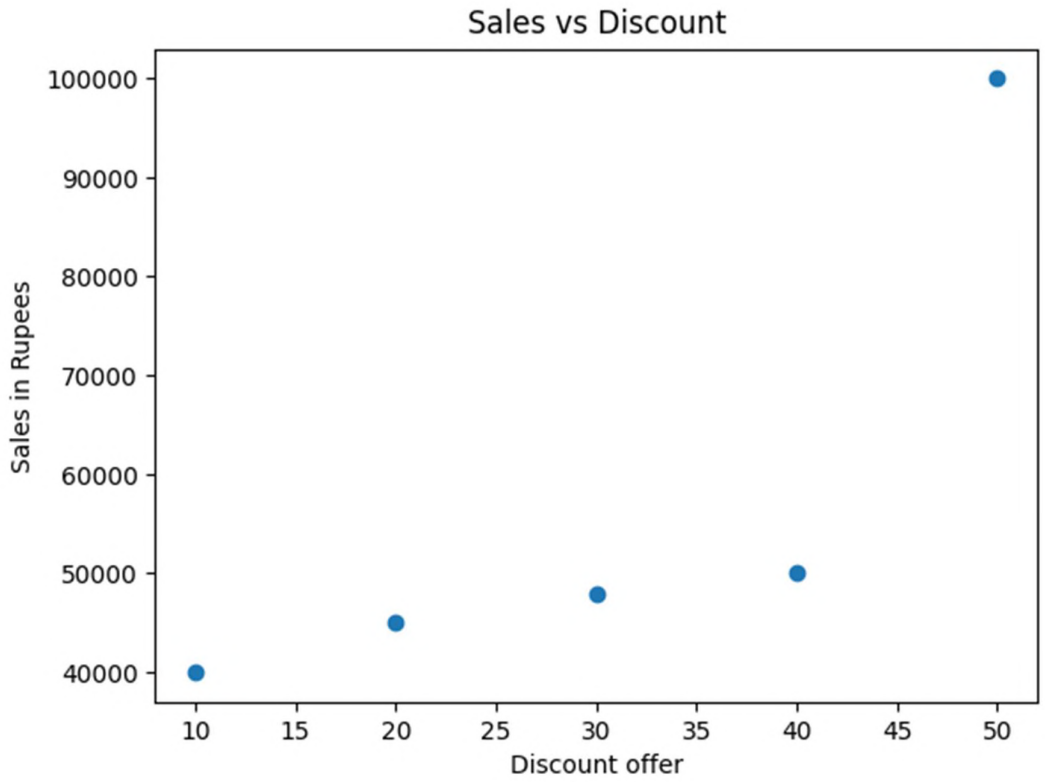
Program – 5

A person sells designer bags and wallets. During the sales season, he gave discounts ranging from 10% to 50% over a period of 5 weeks. He recorded his sales for each type of discount in an array.

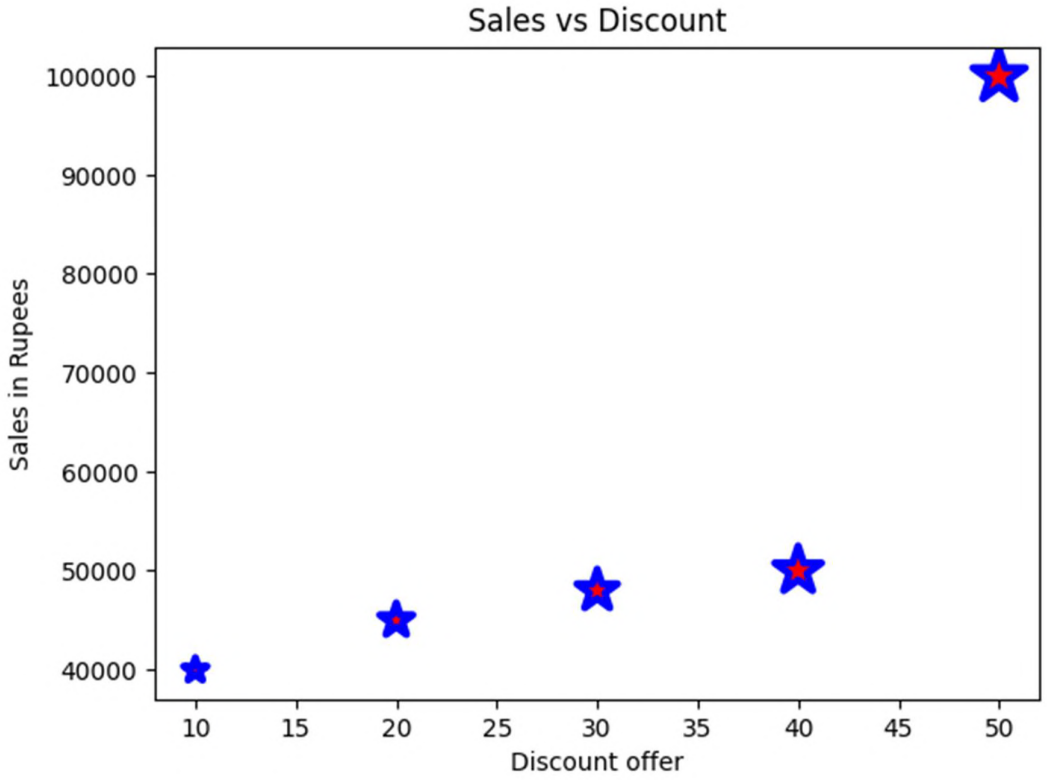
1. Draw a scatter plot to show a relationship between the discount offered and sales made.
2. Opt for displaying the size of the bubble as 10 times the discount

discount= [10,20,30,40,50]
saleInRs=[40000,45000,48000,50000,100000]

```
import numpy as np
import matplotlib.pyplot as plt
discount=np.array([10,20,30,40,50])
sale=np.array([40000,45000,48000,50000,100000])
plt.scatter(x=discount,y=sale)
plt.title('Sales vs Discount')
plt.xlabel('Discount offer')
plt.ylabel('Sales in Rupees')
plt.show()
```



```
import numpy as np
import matplotlib.pyplot as plt
discount=np.array([10,20,30,40,50])
sale=np.array([40000,45000,48000,50000,100000])
size=discount*10
plt.scatter(x=discount,y=sale,s=size,color='Red',linewidth=3,marker='*',
edgecolor='blue')
plt.title('Sales vs Discount')
plt.xlabel('Discount offer')
plt.ylabel('Sales in Rupees')
plt.show()
```



Program – 6

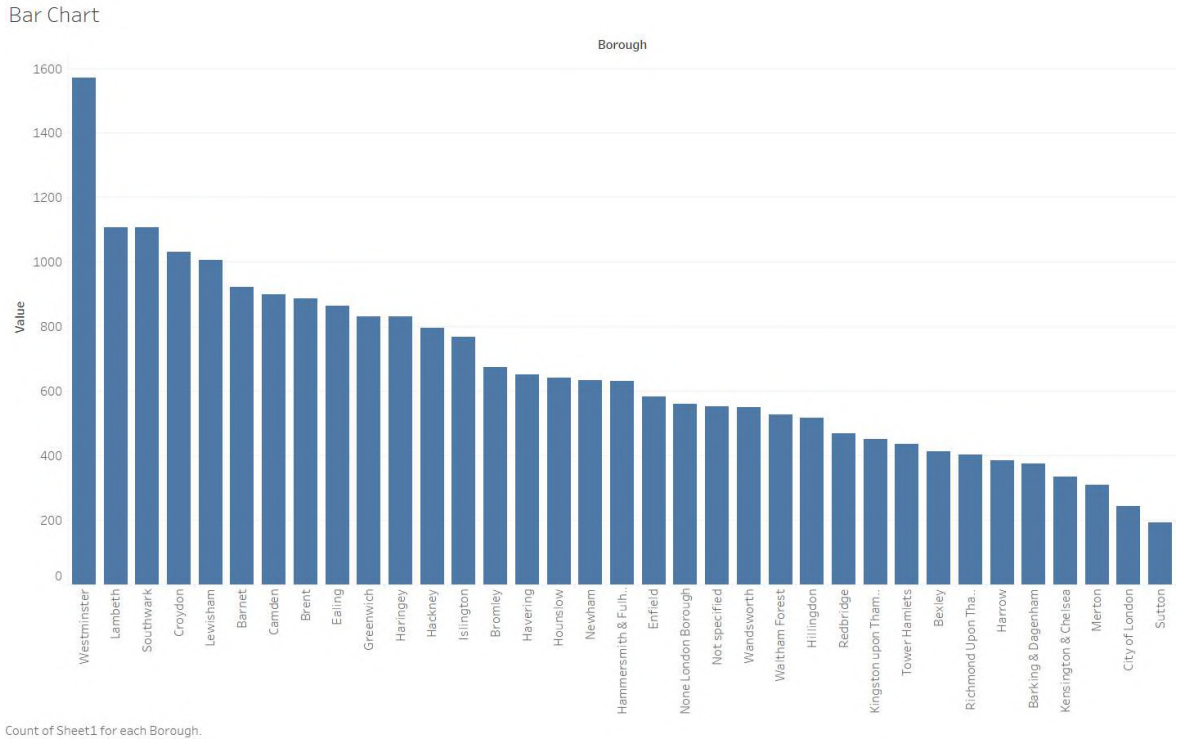
For the given Dataset (TFL Bus Safety):

1. Create a bar chart on boroughs field to visualize the trend in the count.
2. Create a line chart for date of incidence for each quarter in a year, comment on possibilities and suitability of different charts for this timeline.
3. Create a line chart for date of incidence for each month in a year apply formatting to display the first letter of the month on X-axis.
4. Create tree maps of all the data fields except date & year and comment on significance of tree map.
5. Build an interactive dashboard for the above data.

Steps:

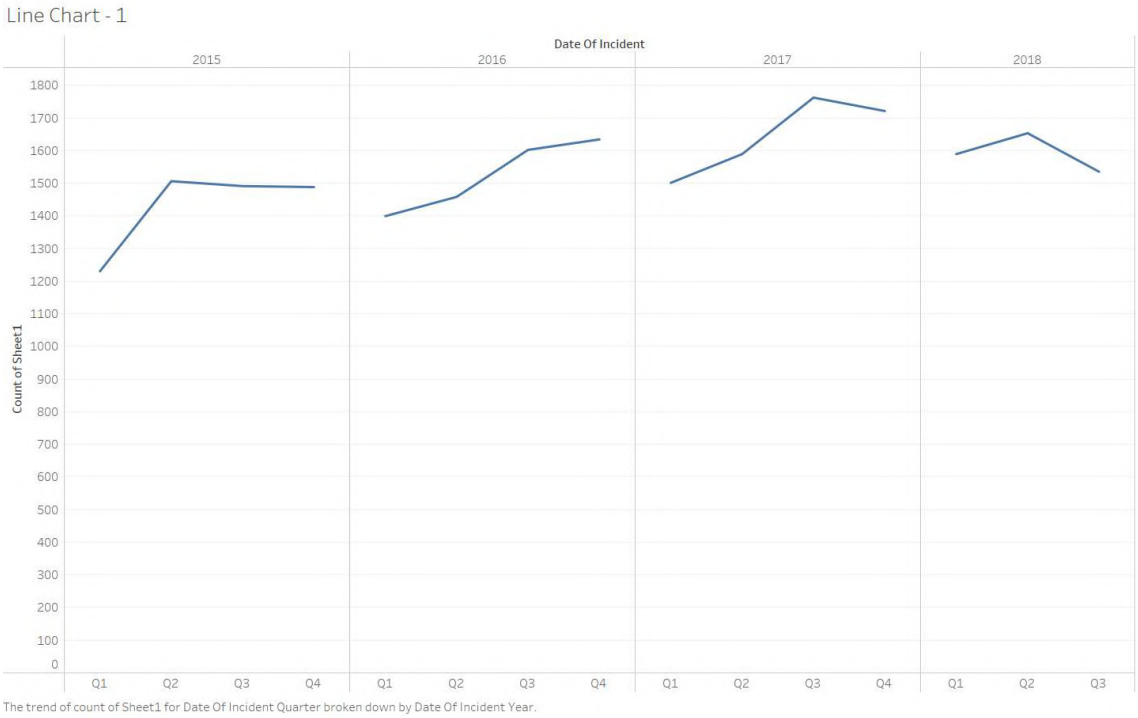
1)

- Drag the 'Borough' field to the columns shelf.
- Drag 'Measure Values' or 'Count' to the rows.
- Select 'Bar Chart' under 'Show me' section.
- Sort the Bar Chart in descending order with Sort By as 'Field'.
- Rename the worksheet as 'Bar Chart'.



2)

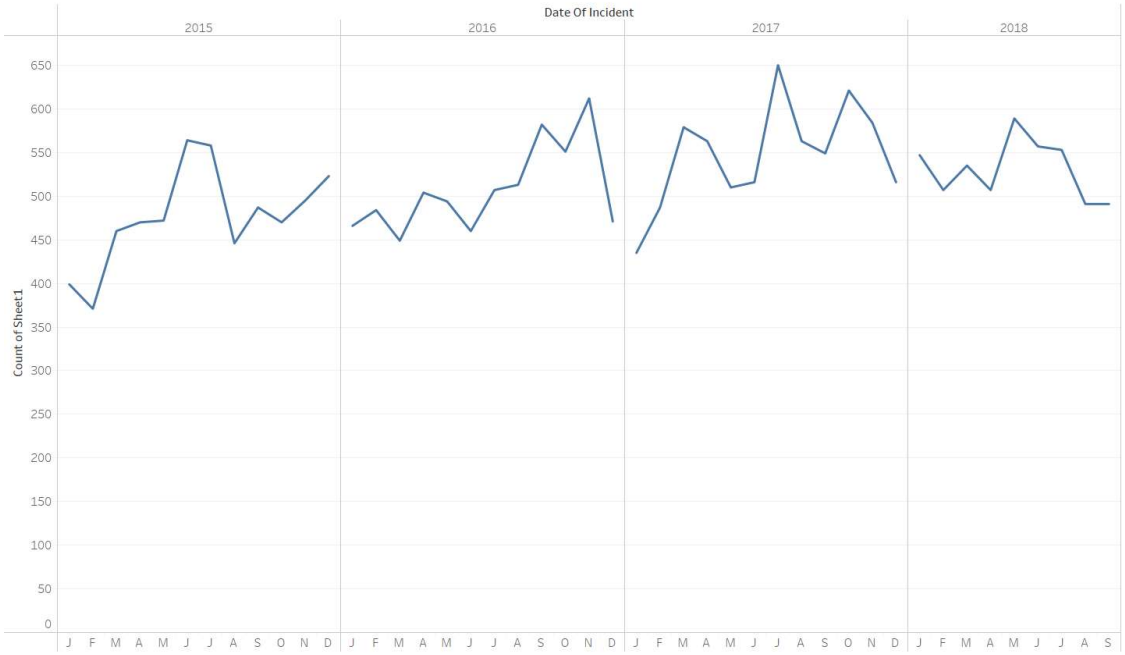
- Drag and drop the ‘Date of incident’ field to the column.
- Drag the ‘Date of incident’ field to the column field it automatically selects quarter.
- Drag and drop ‘Sheet1(Count)’ field to the rows.
- Rename the worksheet as ‘Line Chart – 1’.



3)

- Create a line chart by following the previous sub-question.
- Right click on ‘Date of incident’ field in column & select ‘Month’.
- Now, right click on any one label in X-axis, then go to format.
- Under header, go to default, then select dates.
- Select first letter under dates.
- Rename the worksheet as ‘Line Chart – 2’.

Line Chart - 2

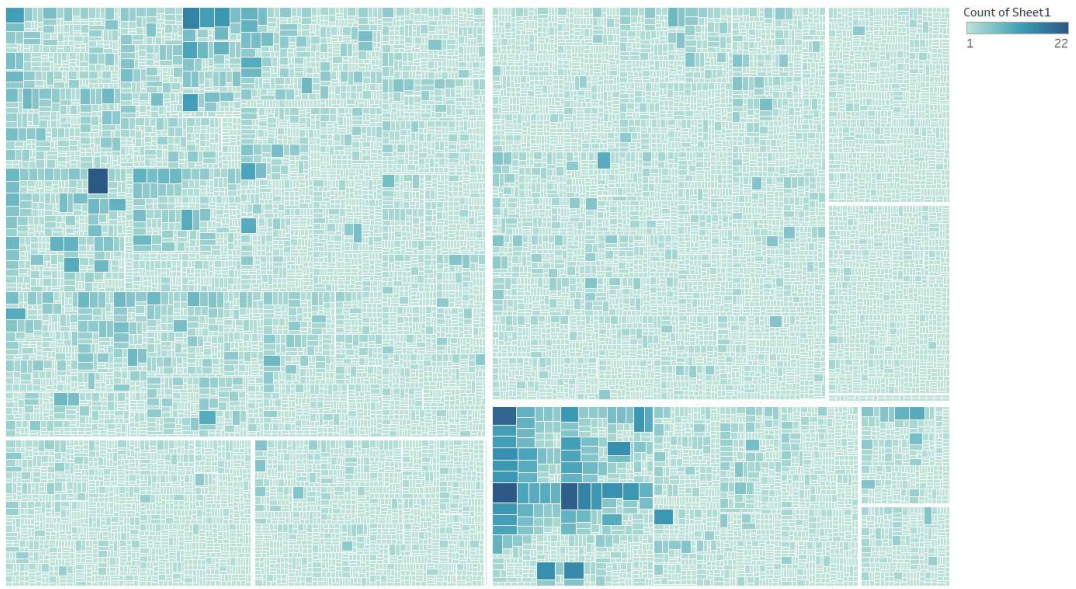


The trend of count of Sheet1 for Date Of Incident Month broken down by Date Of Incident Year.

4)

- Drag & drop all fields one by one except 'Date' & 'Year' to marks section.
- Click on 'Show Me' and select the tree maps chart.
- Rename the Worksheet as 'Tree Map'.

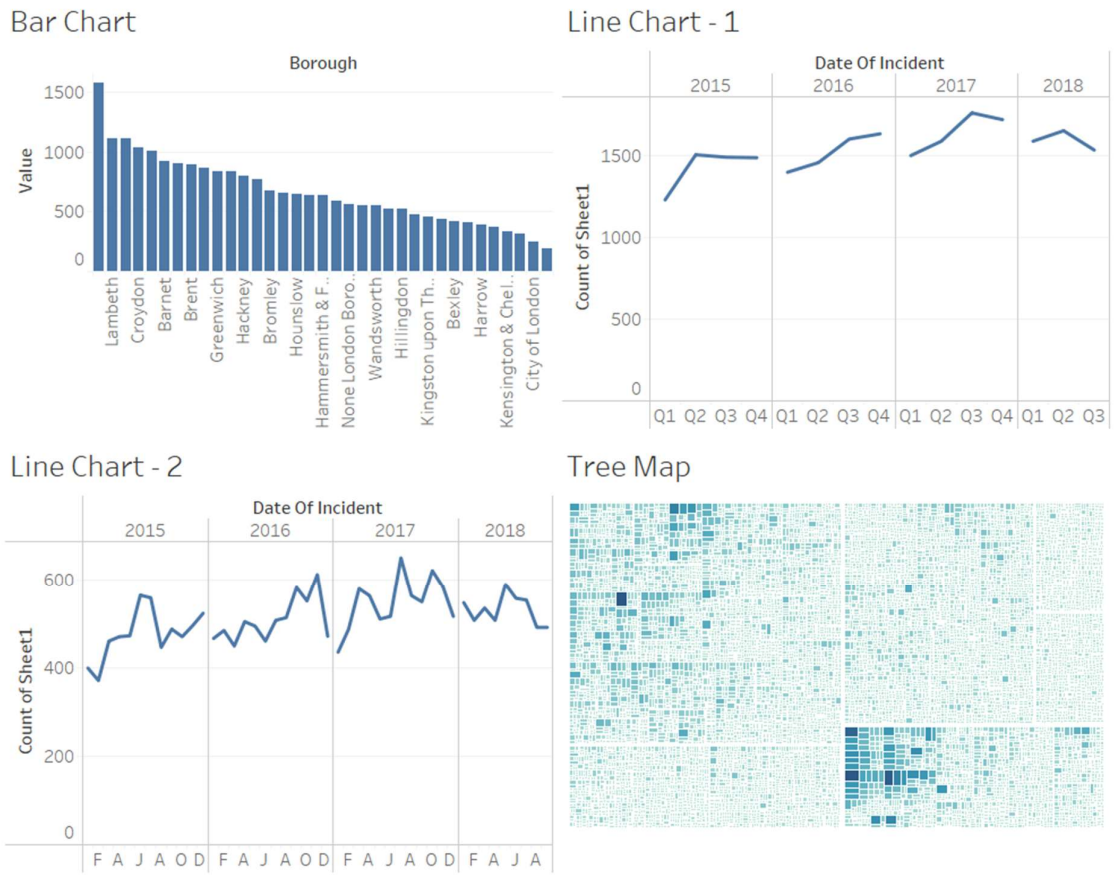
Tree Map



Victims Sex, Injury Result Description, Victims Age, Incident Event Type, Group Name, Victim Category, Operator, Borough, Bus Garage and Route. Color shows count of Sheet1. Size shows count of Sheet1. The marks are labeled by Victims Sex, Injury Result Description, Victims Age, Incident Event Type, Group Name, Victim Category, Operator, Borough, Bus Garage and Route.

5)

- Click on ‘New Dashboard’ button in the bottom left corner of the Tableau window.
- Select floating windows under Objects, in Dashboard.
- Drag all the worksheets and drop them in the dashboard.
- Rearrange all the worksheets, once all the worksheets are added.



Program – 7

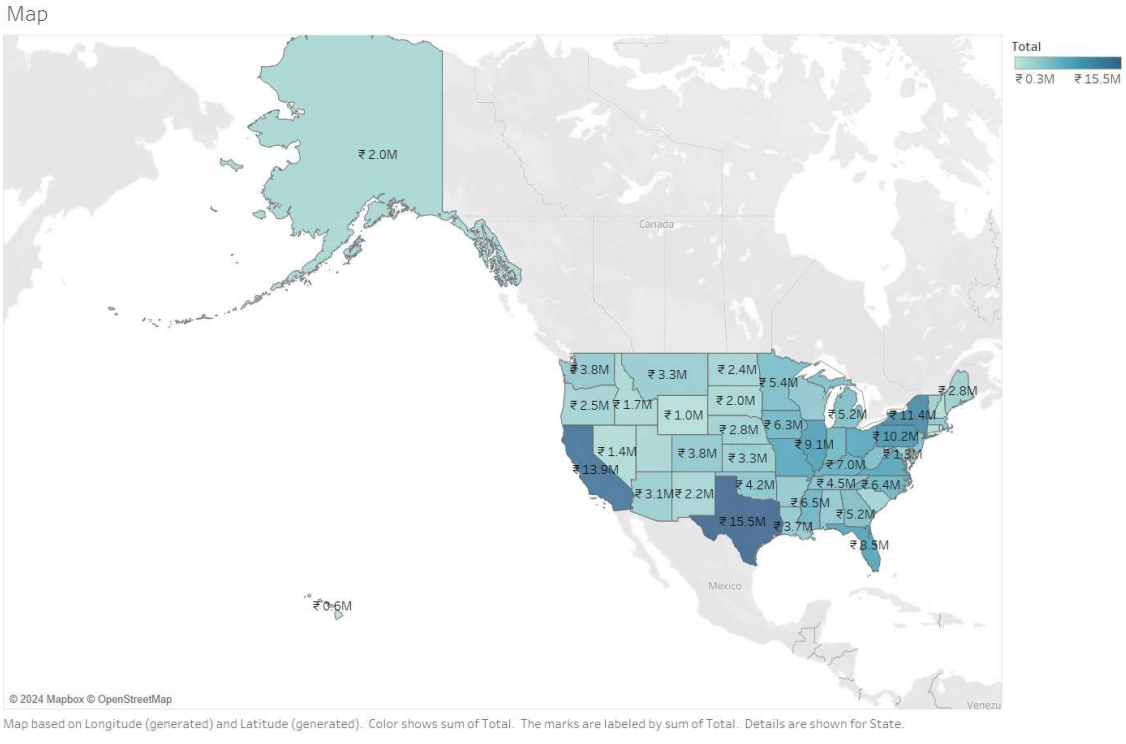
Analysis of revenue in Sales dataset:

1. Create a choropleth map (fill the map) to spot the special trends to show the state which has the highest revenue.
2. Create a line chart to show the revenue based on the month of the year (Continuous data based on date).
3. Create a donut chart view to show the percentage of revenue per region by creating zero axis in the calculated field.
4. Create a butterfly chart by reversing the bar chart to compare female & male revenue based on product category (Calculated fields based on total).
5. Build an interactive dashboard for the above data.

Steps:

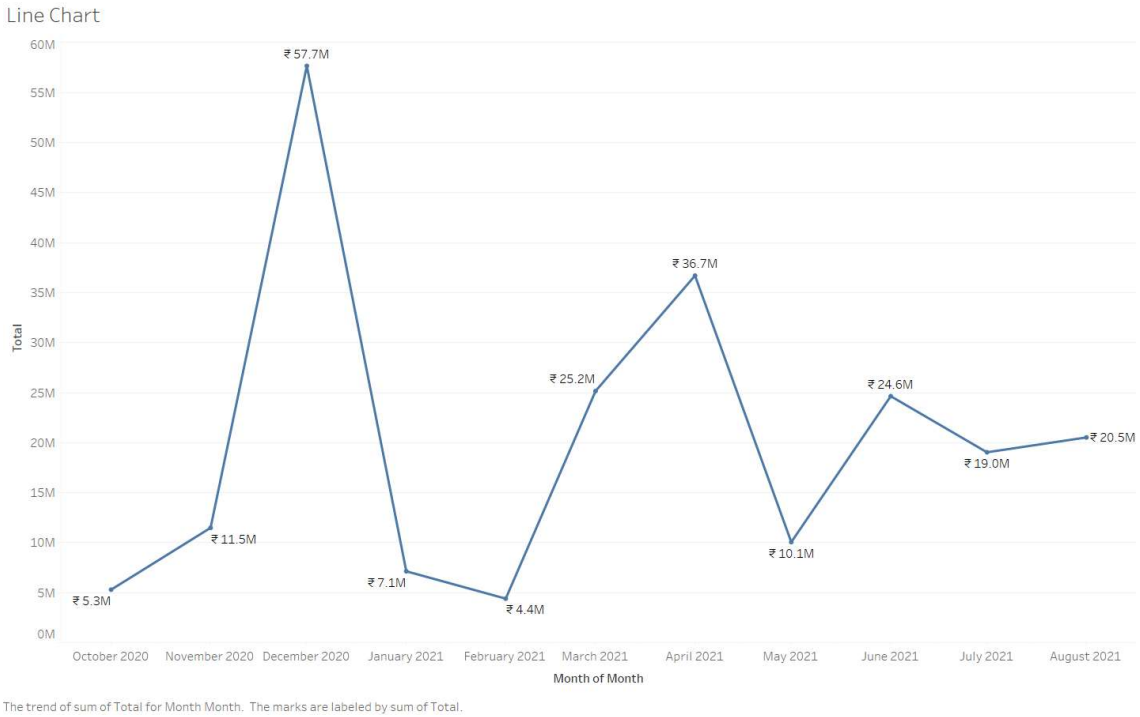
1)

- Double click on 'State' field under data.
- Under Map tab click, edit locations and select 'United States'.
- Drag and drop 'Total' to label and color under marks section.
- 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.
- Rename the worksheet as 'Map'.



2)

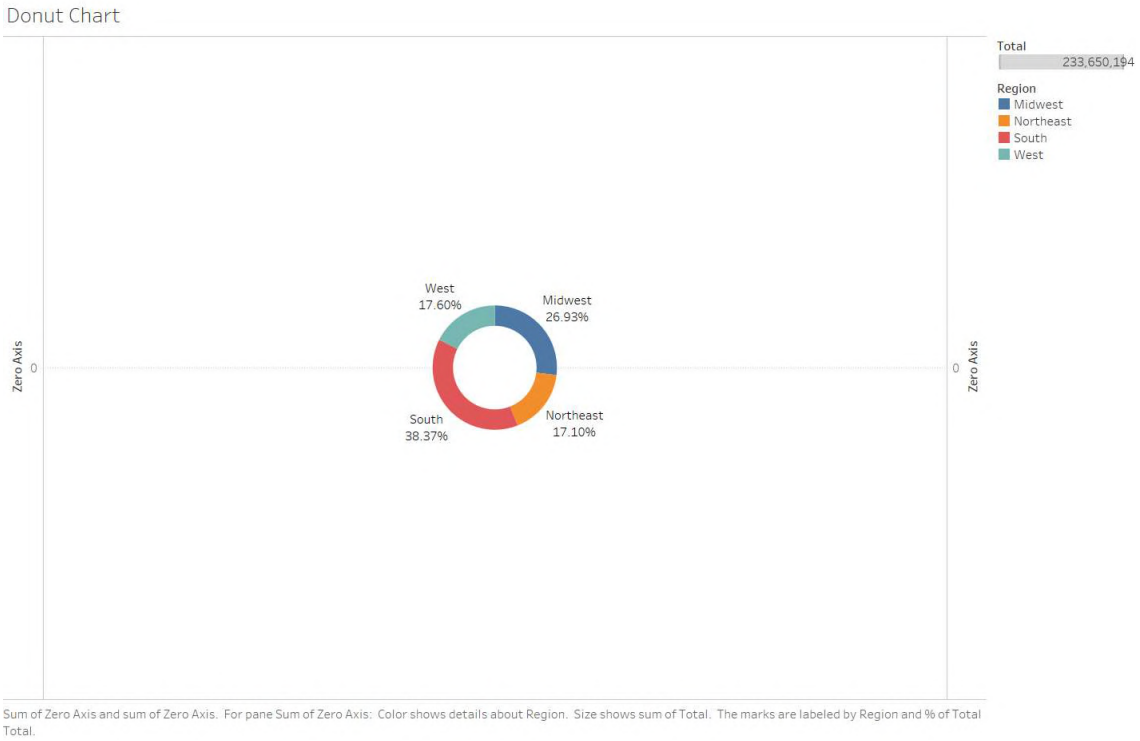
- Drag and drop ‘Total’ to rows.
- Convert month from string to date, by right click, then change data type and drag to column.
- Drag and drop ‘Total’ to label under marks section.
- Under marks, select sum(total), right click, format, then under pane, go to default, under numbers, go to currency(custom), decimal place to 1, display units in Millions.
- Right Click on Month in columns and select Month i.e. continuous.
- Rename the worksheet as ‘Line Chart’.



3)

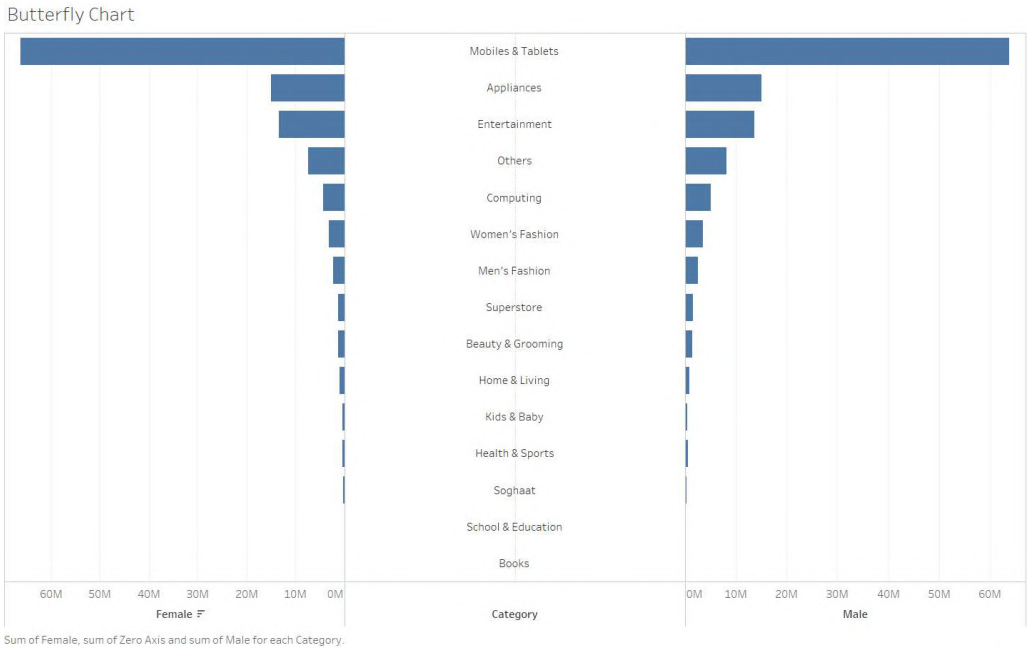
- Drag and drop ‘Region’ to column and ‘Total’ to rows.
- Make a pie chart, selecting under ‘Show Me’.
- Drag and drop ‘Region’ & ‘Total’ to label under marks section.
- In label, sum(total) right click, quick table calculation, percent of total.
- Create calculated field, rename to ‘Zero Axis’, write code as 0, then ok.
- Drag and drop ‘Zero Axis’ twice to rows.

- Then under marks, two fields of Zero Axis will be present, go to second, remove all fields, change the color to white then increase the size in 1st, decrease in 2nd Zero Axis.
- Right click on second Zero Axis in rows, then click dual axis and select Entire View in Fit tab.
- Rename the worksheet as 'Donut Chart'.



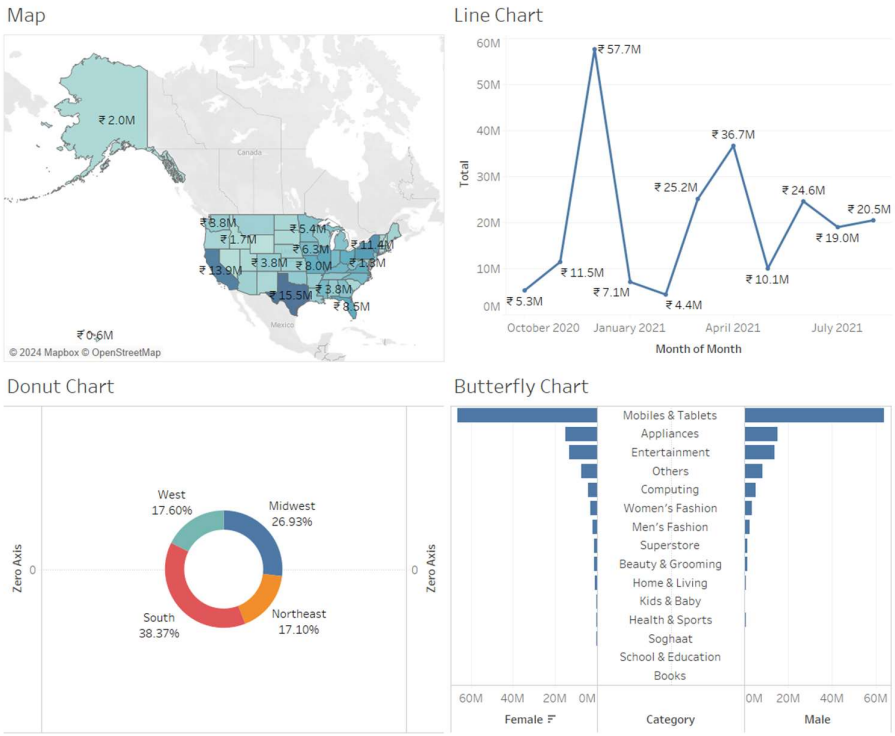
4)

- Drag and drop 'Category' to rows shelf.
- Create 2 calculated fields 'Female' and 'Male'.
- For female revenue, Code is – if [Gender] = 'F' then [Total]end - Create same for male revenue by changing 'F' to 'M'.
- Drag and drop 'Female' and 'Male' to columns shelf.
- Drag and drop 'Zero axis' between female and male revenue in columns.
- Select Zero Axis under Marks Section, Select Text in place of automatic and Drag Category to text.
- Edit the Female Axis and select reversed.
- Sort the Category and select entire view in fit tab, then disable show header in Y axis.
- Rename the 'Zero Axis' as 'Category' by editing, and remove 0 (tick tab).
- Rename the worksheet as 'Butterfly Chart'.



5)

- Click on 'New Dashboard' button in the bottom left corner of the Tableau window.
- Select floating windows under Objects, in Dashboard.
- Drag all the worksheets and drop them in the dashboard.
- Rearrange all the worksheets, once all the worksheets are added.



Program – 8

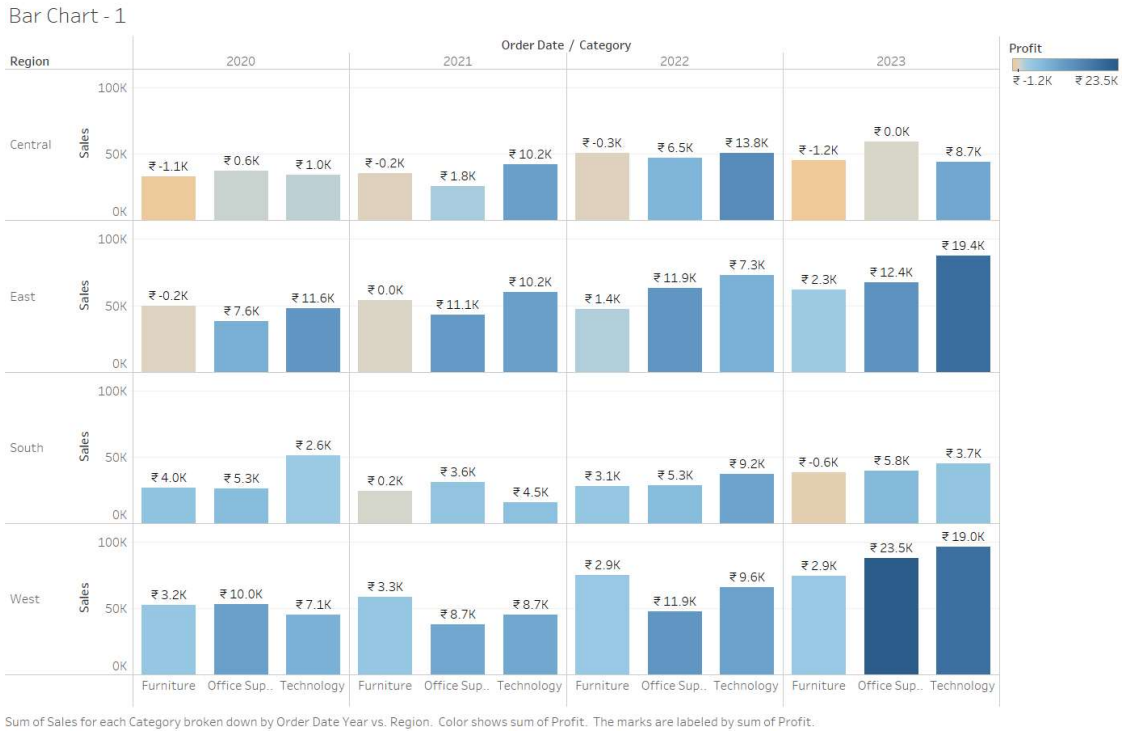
Analysis of sample superstore

1. Create a bar chart to visualize the sales of product of each category by year with respect to region and also show the profit (Text).
2. Create a bar chart to visualize the sales of product all categories of years 2022 and 2023.
3. Create a horizontal bar chart to visualize the profit by category.
4. Create a map to visualize the profit by state.
5. Build an interactive dashboard for the above data.

Steps:

1)

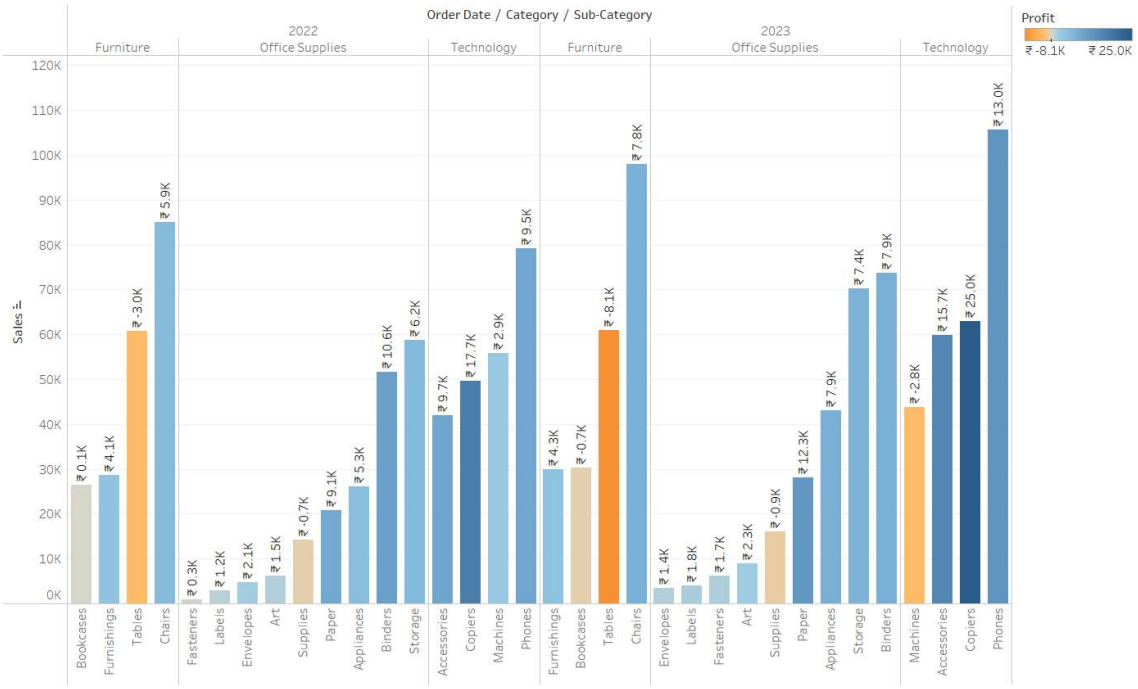
- Drag and drop ‘Order Date’ and ‘Category’ to columns.
- Drag and drop ‘Sales’ to rows.
- Drag and drop ‘Profit’ to color and label in marks section.
- Under marks, select sum(profit) of label, right click, format, then under pane, go to default, under numbers, go to currency(custom), decimal place to 1, display units in Thousands(K).
- Drag & drop ‘Region’ to rows before sales and select entire view in fit tab.
- Rename the worksheet as ‘Bar Chart – 1’.



2)

- Drag ‘Order Date’ and ‘Category’ and ‘Sub-Category’ to columns.
- Drag and drop ‘Sales’ to rows.
- Right click on Year (Order Date) in columns, select filter and select only 2022 and 2023.
- Drag and drop ‘Profit’ to color and label in marks section.
- Under marks, select sum(profit) of label, right click, format, then under pane, go to default, under numbers, go to currency(custom), decimal place to 1, display units in Thousands (K).
- Rename the worksheet as ‘Bar Chart – 2’.

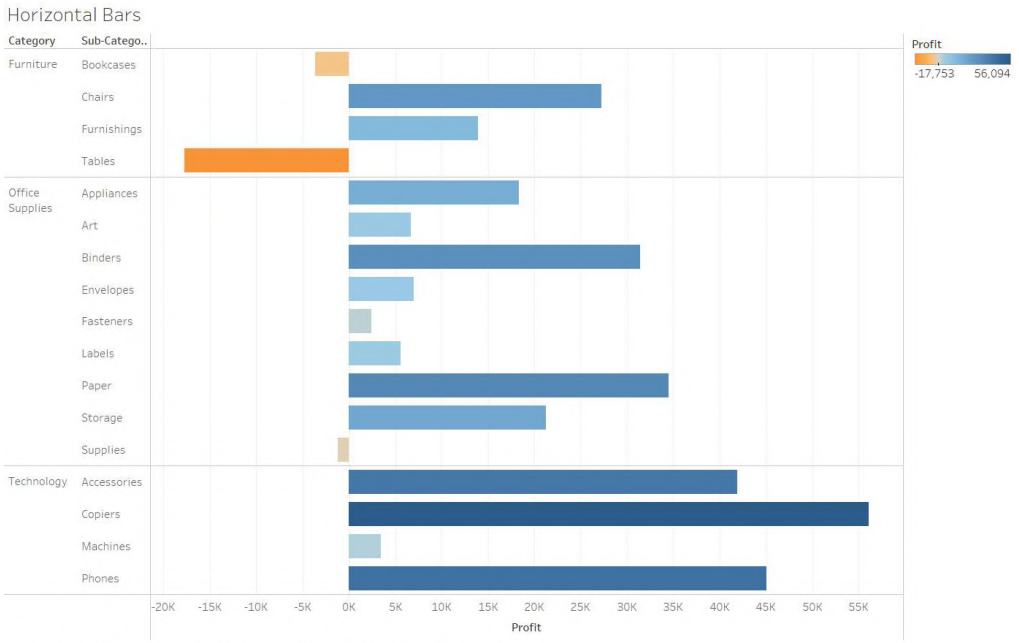
Bar Chart - 2



Sum of Sales for each Sub-Category broken down by Order Date Year and Category. Color shows sum of Profit. The marks are labeled by sum of Profit. The view is filtered on Order Date Year, which keeps 2022 and 2023.

3)

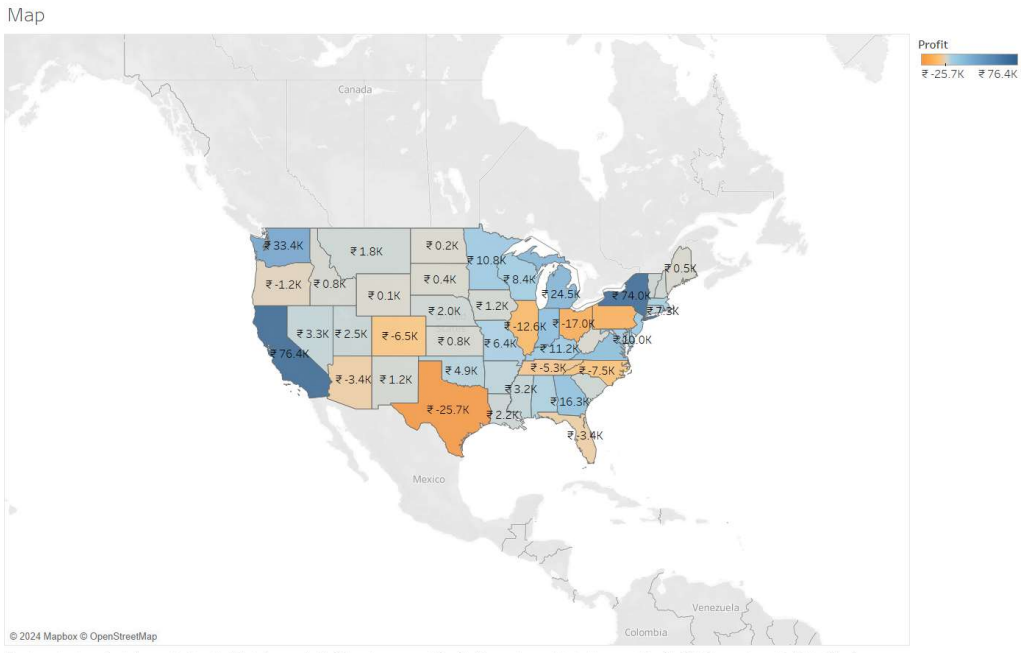
- Drag and drop ‘Category’ field to rows and add ‘Sub-Category’ also.
- Drag and drop ‘Profit’ to text under marks section.
- Select Horizontal bars under show me.
- Drag and drop ‘Profit’ to colour under marks section.
- Rename the worksheet as ‘Horizontal Bars’.



Sum of Profit for each Sub-Category broken down by Category. Color shows sum of Profit.

4)

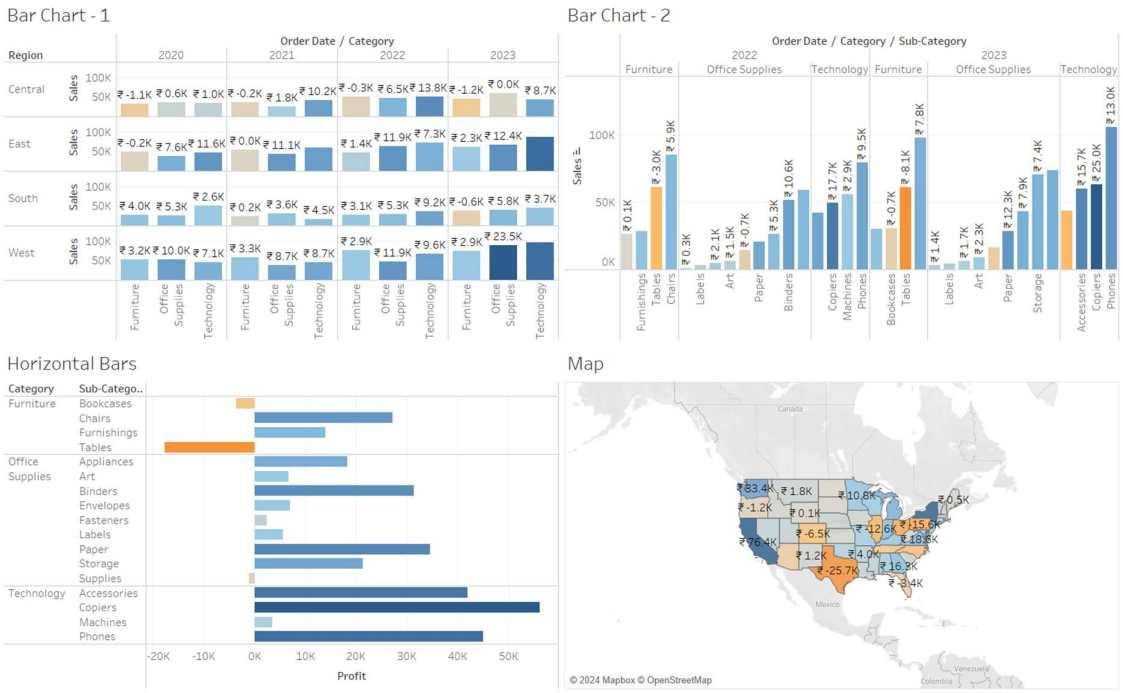
- Drag and drop the 'State/Province' field to Details under marks section.
- Under Map tab click, edit locations and select 'United States'.
- Drag and drop 'Profit' to size and label under marks section.
- Select 'Sum (Profit)', right click, format, then under pane, go to default, under numbers, go to currency(custom), decimal place to 1, display units in Thousands (K).
- Rename the worksheet as 'Map'.



Map based on Longitude (generated) and Latitude (generated). Color shows sum of Profit. The marks are labeled by sum of Profit. Details are shown for State/Province.

5)

- Click on ‘New Dashboard’ button in the bottom left corner of the Tableau window.
- Select floating windows under Objects, in Dashboard.
- Drag all the worksheets and drop them in the dashboard.
- Rearrange all the worksheets, once all the worksheets are added.



Program – 9

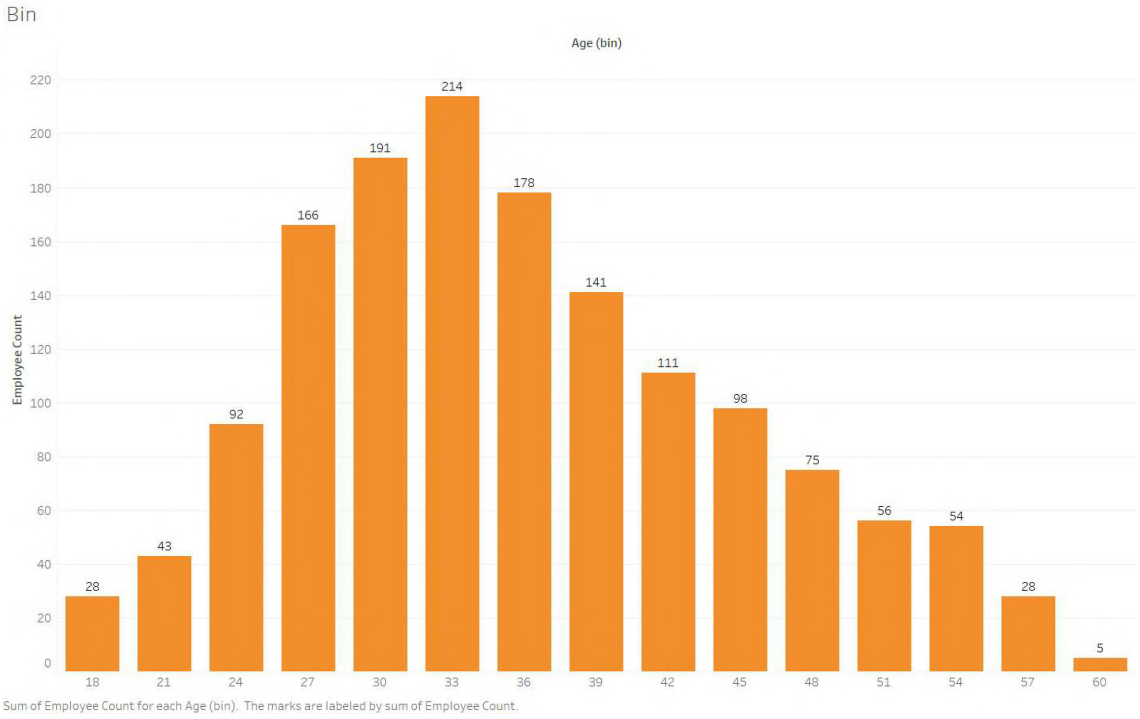
Analysis of HR Dataset:

1. Create a bin of size 3 for the age group to create a new dimension to show the number of employees.
2. Create a highlight table to show the Job Satisfaction (Discrete) Rating for each job role based on employee count.
3. 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,
4. Create multiple donut chart to show the Attrition Count by Gender for different Age group.
5. Build an interactive dashboard for the above data.

Steps:

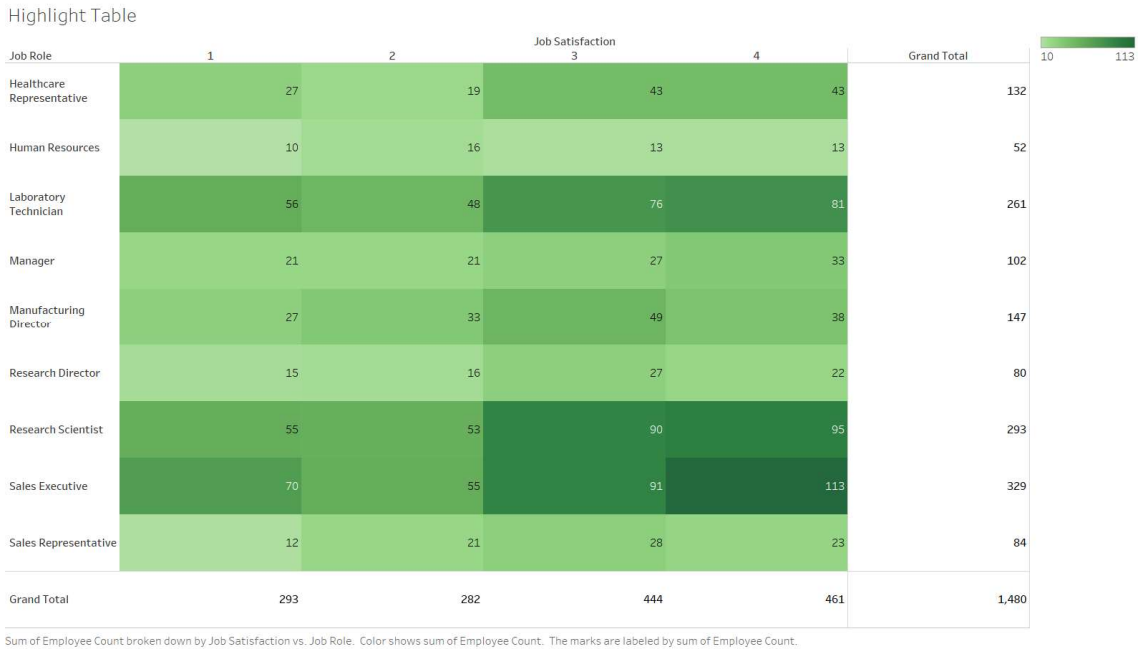
1)

- Right click on the 'Age' tab and select Create and select "Bins" in it.
- Rename it to 'Age(bin)' and change the bin size to 3 and save it.
- Drag 'Age(bin)' to Columns and drag the 'Employee Count' to Rows.
- Drag the 'Employee Count' on the label option under the Marks dropdown.
- Rename the worksheet to 'Bin'.



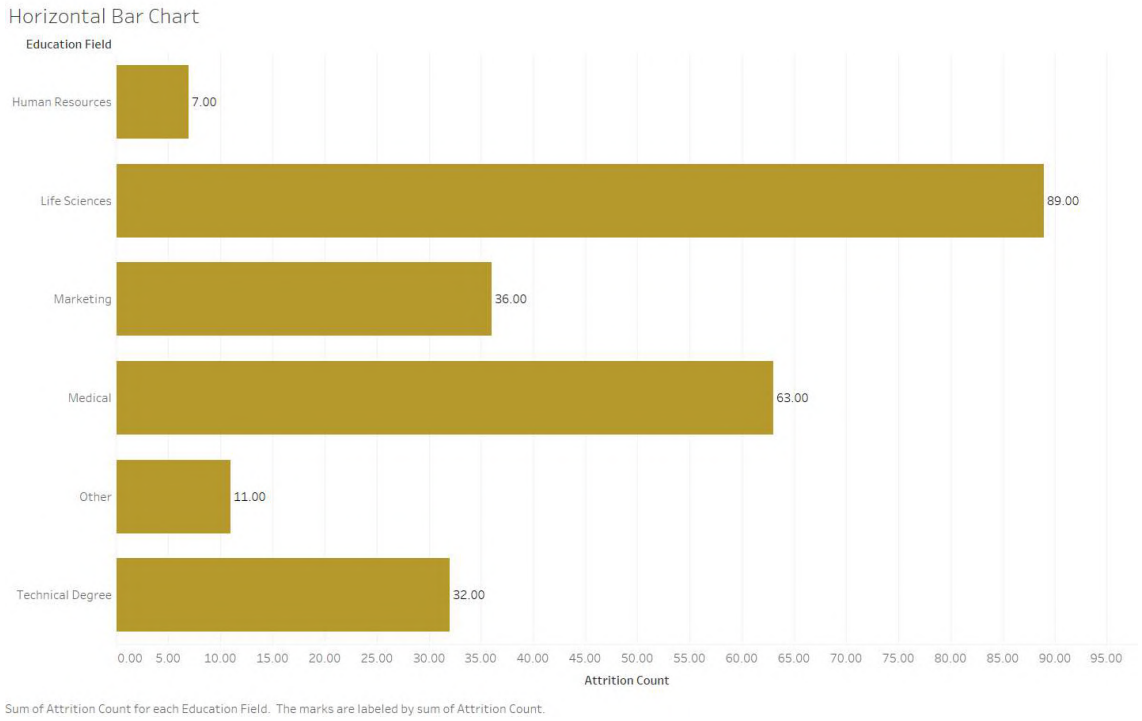
2)

- Drag and drop ‘Job Role’ to Rows.
- Drag and drop ‘Job Satisfaction’ to the Columns make it discrete and select dimension.
- Under the Show Me change the chart type to Highlight Table.
- Drag and drop ‘Employee Count’ to Color and Label.
- Rename the worksheet to ‘Highlight Table’.



3)

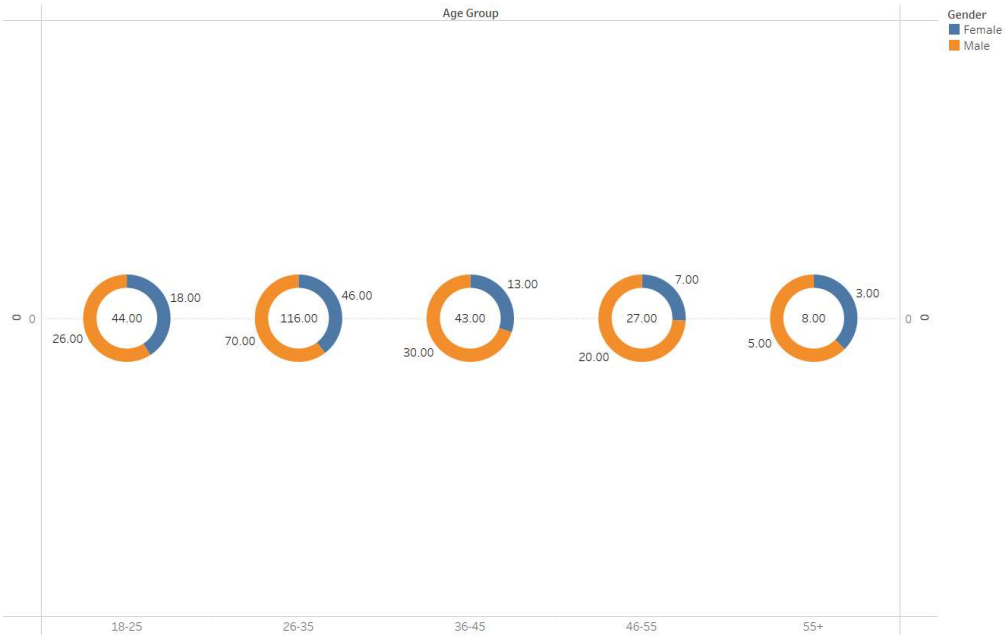
- Drag and drop ‘Education Field’ to Rows.
- Create a calculated field ‘Attrition Count’ with the formula -
  IF [Attrition] = ‘Yes’ THEN 1 ELSE 0 END.
- Drag and drop ‘Attrition Count’ to Columns.
- Drag and drop Attrition count to label option under marks dropdown.
- Rename the worksheet to ‘Horizontal Bar Chart’.



4)

- Drag and drop ‘Age Group’ to Columns.
- Create a pie chart by changing the mark type from automatic to Pie and select entire view in fit tab.
- Double click on the empty space in row tab and type 0 and press enter.
- Hold ctrl and drag and drop the SUM (0) beside it in the row tab.
- Drag and drop ‘Gender’ to color under marks section.
- Drag and drop ‘Attrition Count’ to Angle and Label.
- Remove all the tabs under the SUM (0)(2) dropdown and change the color to white.
- Right click on second SUM (0) in row and select ‘Dual Axis’.
- Vary the size of both the circles to make it look like donut.
- Rename the worksheet to ‘Donut chart’.

Donut Chart

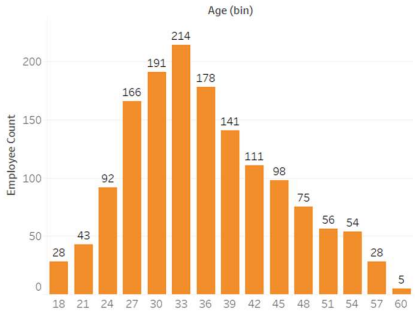


Sum of 0 and sum of 0 for each Age Group. The marks are labeled by sum of Attrition Count. For pane Sum of 0: Color shows details about Gender.

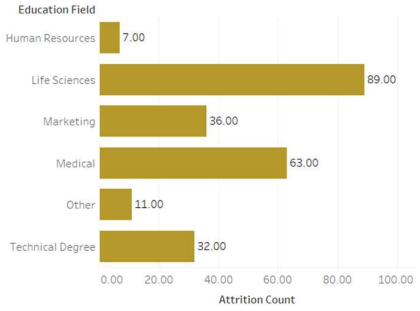
5)

- Click on 'New Dashboard' button in the bottom left corner of the Tableau window.
- Select floating windows under Objects, in Dashboard.
- Drag all the worksheets and drop them in the dashboard.
- Rearrange all the worksheets, once all the worksheets are added.

Bin



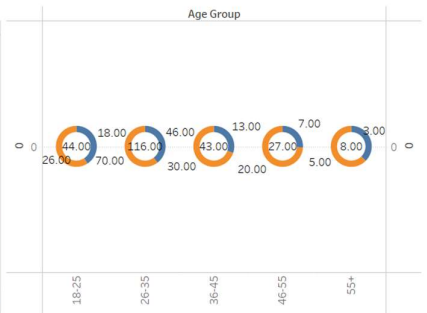
Horizontal Bar Chart



Highlight Table

Job Role	Job Satisfaction 1	2	3	4	Grand Total
Healthcare Represe..	27	19	43	43	132
Human Resources	10	16	13	13	52
Laboratory Technici..	56	48	76	81	261
Manager	21	21	27	33	102
Manufacturing Direc..	27	33	49	38	147
Research Director	15	16	27	22	80
Research Scientist	55	53	90	95	293
Sales Executive	70	55	91	113	329
Sales Representative	12	21	28	23	84
Grand Total	293	282	444	461	1,480

Donut Chart



Problem – 10

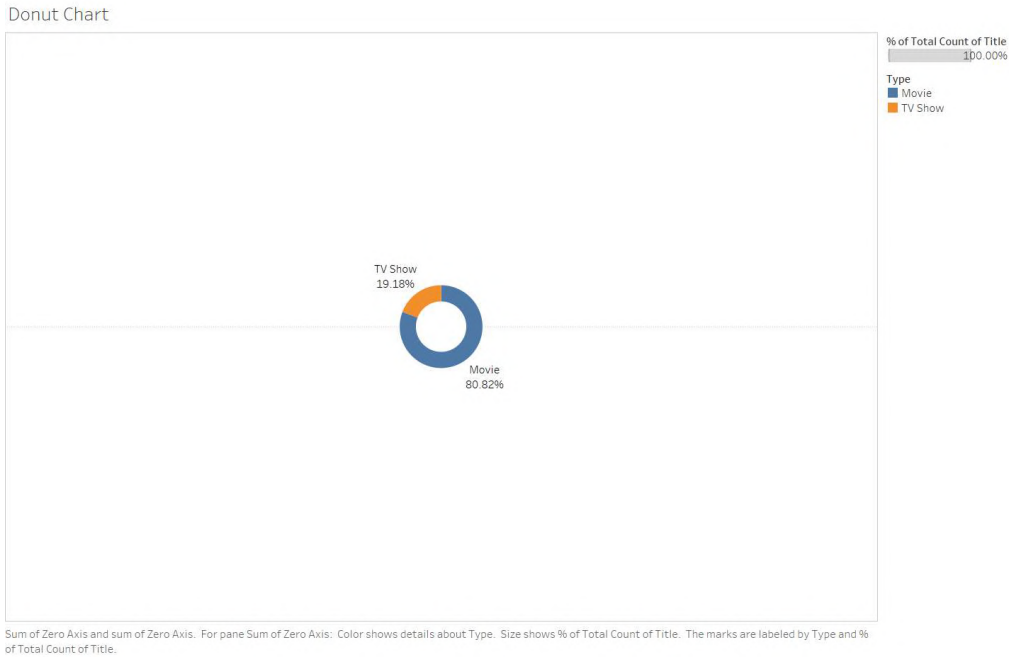
Analysis of Amazon Prime Dataset:

1. Create a Donut chart to show the percentage of movie and tv shows.
2. Create an area chart to shows by release year and type.
3. Create a horizontal bar chart to show Top 10 genre.
4. Create a map to display total shows by country.
5. Build an interactive Dashboard.

Steps:

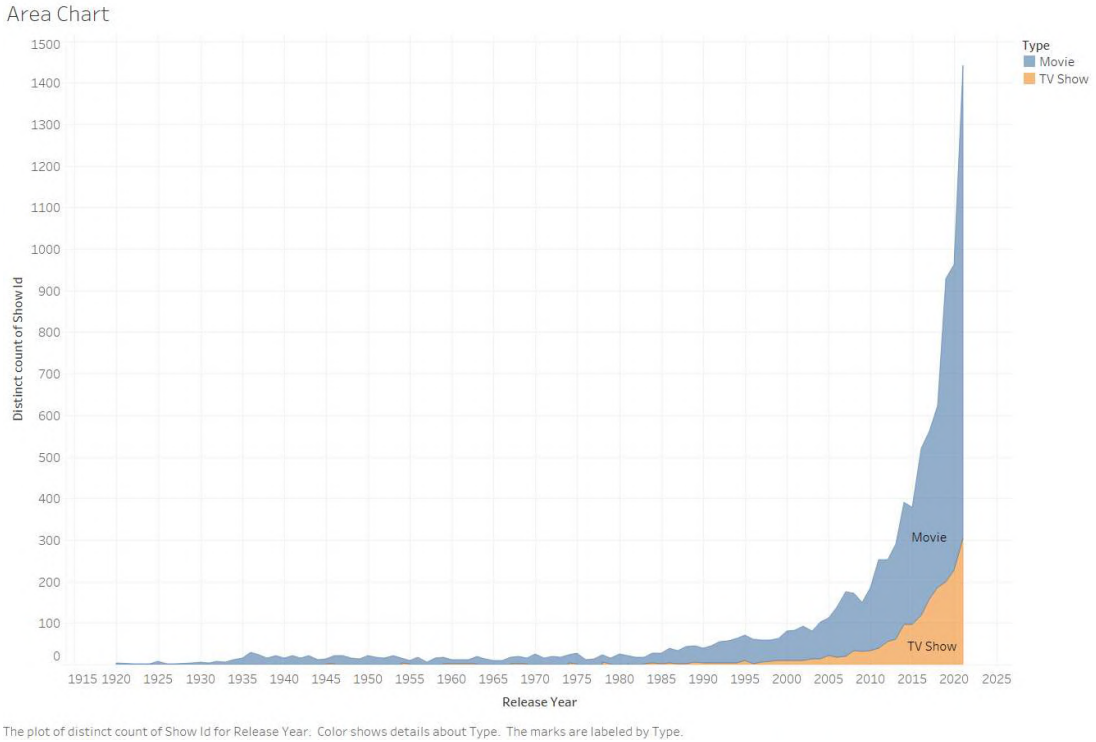
1)

- Drag and drop ‘Type’ to Color.
- In the “Marks" section, select Pie from the drop-down list in place of automatic.
- Drag and drop ‘Title’ to Size.
- Right click on 'Title' in Marks shelf, click on measure select Count and Percent of total under Quick table calculation.
- Drag ‘Type’ and ‘Title’ from marks section using control and drop to label.
- Create calculated field called ‘Zero Axis’.
- Drag it twice to rows shelf.
- Then under Marks, two fields of Zero Axis exist go to second one - remove all fields, change colour to white and decrease its size and also increase the size of first one.
- Right click on the second ‘Zero Axis’ and select ‘Dual axis’.
- Rename the worksheet as ‘Donut Chart’.



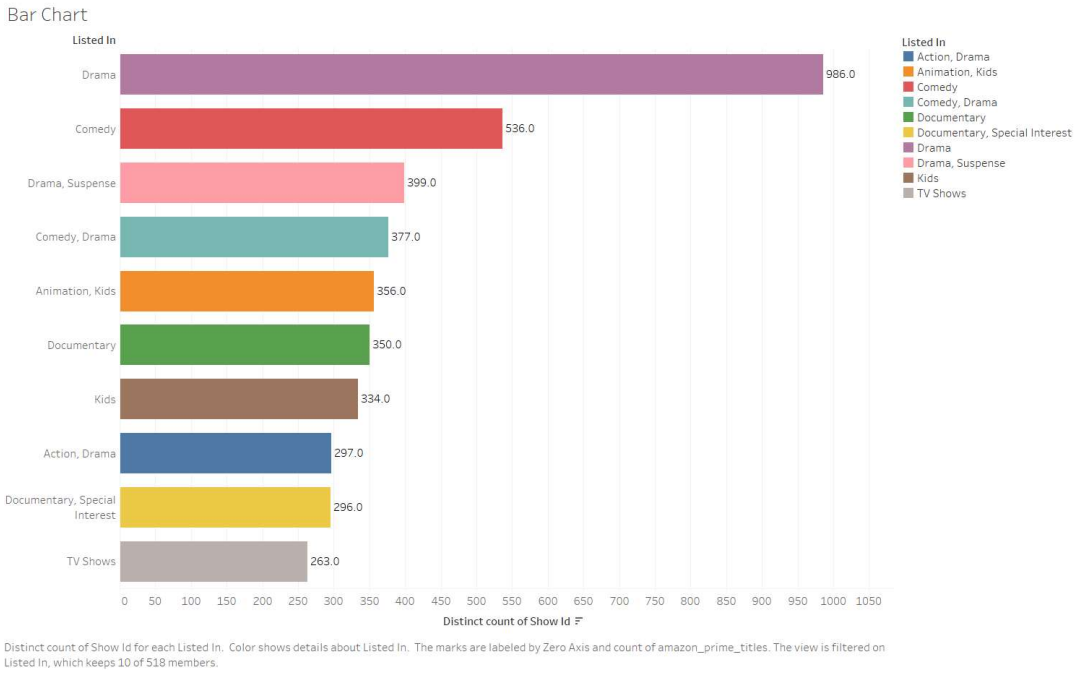
2)

- Drag and drop ‘Release year’ and ‘Show ID’ to Columns and rows shelf respectively.
- Right click on ‘Show ID’ in rows shelf, click on measure select Count (distinct)
- In the “Marks” section, select area from the drop-down list in place of automatic.
- Drag and drop ‘Type’ to colour.
- Drag and drop ‘Type’ to Label.
- Rename the worksheet as ‘Area Chart’.



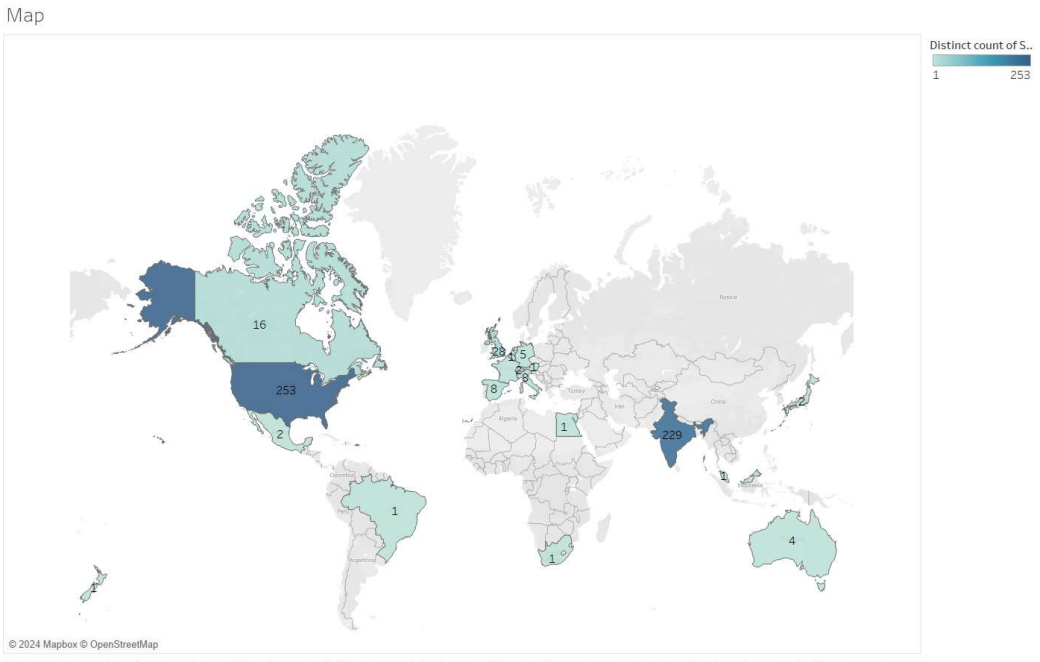
3)

- Drag “Listed In" and "Show ID" to rows and Column shelf respectively.
- Right click on "Show ID” then go to Measure → Count (distinct)
- Drag 'Listed In' to filter and edit it Accordingly to get top 10 Genres and sort them.
- Drag and drop "Listed In" to colour.
- Drag and drop “Measure Values” to label.
- Rename the worksheet as ‘Bar Chart’.



4)

- Double Click on country.
- Drag "Country" to filters and remove null values.
- Drag Show Id to Text, Right click on "Show ID" then go to Measure → Count (distinct).
- Hold ctrl, drag and drop the Show Id to Label.
- Rename the worksheet as 'Map'.



5)

- Click on ‘New Dashboard’ button in the bottom left corner of the Tableau window.
- Select floating windows under Objects, in Dashboard.
- Drag all the worksheets and drop them in the dashboard.
- Rearrange all the worksheets, once all the worksheets are added.