KIT-KALAIGNARKARUNANIDHI INSTITUTE OF TECHNOLOGY

(AN AUTONOMOUS INSTITUTION)

(Accredited by NAAC&NBA with ‘A’ Grade)

Approved by AICTE & Affiliated to Anna University, Chennai)

Kannampalayam Post, Coimbatore-641402

Department of Artificial Intelligence and Data Science

B19ADP702 – DATA VISULIZATION LABORATORY

Name : …………………………………………………

Batch : ………………   Reg.No : ……………….……..

Branch : …………….   Year : ……………………….

KIT-KALAIGNARKARUNANIDHI INSTITUTE OF TECHNOLOGY
(ANAUTONOMOUS INSTITUTION)
(Accredited by NAAC & NBA with 'A' Grade)
Approved by AICTE & Affiliated to Anna University, Chennai)
Kannampalayam Post, Coimbatore-641402

Department of Artificial Intelligence and Data Science

BONAFIDE CERTIFICATE

Name : ….…………………………………………………………………………………………

Roll No. : ……………………………………… Reg. No.: …………………………………

Branch : B.Tech – Artificial Intelligence and Data Science

Certified that this is Bonafide record work done by Mr./Ms. ……………………………………

of IV – Year Artificial Intelligence and Data Science during the academic year 2024-2025.

FACULTY IN-CHARGE **HOD**

Submitted for the University Practical Examination held on ……………………….

INTERNAL EXAMINER **EXTERNAL EXAMINER**

Instructions for Laboratory Classes

1. Enter the lab with record work book & necessary things.
2. Enter the lab without bags and footwear.
3. Footwear should be kept in the outside shoe rack neatly.
4. Maintain silence during the Lab hours.
5. Read and follow the work instructions inside the laboratory.
6. Handle the computer systems with care.
7. Shutdown the Computer properly and arrange chairs in order before leaving the lab.
8. The program should be written on the left side pages of the record workbook.
9. The record work book should be completed in all aspects and submitted in the next class itself.
10. Experiment number with date should be written at the top left-hand corner of the record work book page.
11. Strictly follow the uniform dress code for Laboratory classes.
12. Maintain punctuality for lab classes.
13. Avoid eatables inside and maintain the cleanliness of the lab.

VISION

To produce competent professionals to the dynamic needs of the emerging field of Artificial Intelligence and Data Science

MISSION

- To empower students with the knowledge and skills necessary to create intelligent systems and innovative solutions that address societal issues.
- Providing technical knowledge on par with Industry to the students through qualified faculty members having knowledge in recent trends and technologies.
- To produce competent engineers who are both professional and life-skills oriented.
- Providing opportunities for students to improve their research skills in order to address a variety of societal concerns through innovative projects.

PROGRAMMEOUTCOMES (POs)

Students graduating from Artificial Intelligence and Data Science should be able to:

PO1 Engineering knowledge : Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex Artificial Intelligence and Data Science problems.

PO2 Problem analysis: Identify, formulate, research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and Artificial Intelligence and Data Sciences.

PO3 Design/development to solutions: Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations in the field of Artificial Intelligence and Data Science.

PO4 Conduct investigations of complex problems: Using research-based knowledge and Artificial Intelligence & Data Science oriented research methodologies including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.

PO5 Modern tool usage: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex Artificial Intelligence and Data Science Engineering activities with an understanding of the limitations.

PO6 The engineer and society: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.

PO7 Environment and sustainability: Understand the impact of the professional Artificial Intelligence and Data Science Engineering solutions in societal and environmental contexts, and demonstrate the knowledge, and need for the sustainable development.

PO8 Ethics: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.

PO9 Individual and team work: Function effectively as an individual, and as a member or leader in diverse teams and in multidisciplinary settings.

PO10 Communication: Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

PO11 Projectmanagementandfinance:Demonstrate knowledge and understanding of the Artificial Intelligence and Data Science engineering and management principles and apply these to one's own work, as a member and leader in a team and, to manage projects in multidisciplinary environments.

PO12 Life-long learning: Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

PROGRAM EDUCATIONAL OBJECTIVES (PEOs)

PEO1: Graduates will have a strong foundation in mathematics, programming, machine learning, artificial intelligence, and data science, as well as advanced skills in these areas to solve technical problems.

PEO2: Graduates will have the capability to apply their knowledge and skills to identify and solve the issues in real world Artificial Intelligence and Data Science related applications.

PEO3: Graduates will be able to engage in life-long learning by completing advanced software Technologies, certificates, and/or other professional development.

PROGRAM SPECIFIC OUTCOME(PSOs)

Graduates of Artificial Intelligence and Data Science Programmed should be able to:

PSO1: Apply fundamental concepts of Artificial Intelligence and Data Science according to the environmental needs.

PSO2: Ability to develop skills to address and solve Artificial Intelligence based social and environmental problem using Data Science to deal multidisciplinary projects using modern tools.

COURSE OUTCOMES:

Students will be able to:

Course Outcome	**Knowledge Level**
CO1: Make use Python, R and Tableau for data visualization	K3
CO2: Apply data visuals to convey trends in data over time using tableau.	K3
CO3: Construct effective data visuals to solve complex problems.	K3
CO4: Choose and work with different plotting libraries.	K3
CO5: Build creative and effective visualizations Dashboards.	K3

CO/PO &PSO		**PO1** **(K3)**	**PO2** **(K4)**	**PO3** **(K5)**	**PO4** **(K5)**	**PO5** **(K6)**	**PO6** **(K3)** **(A3)**	**PO7** **(K2)** **(A3)**	**PO8** **(K3)** **(A3)**	**PO9** **(A3)**	**PO10** **(A3)**	**PO11** **(K3)** **(A3)**	**PO12** **(A3)**	**PSO1** **(K3,A3)**	**PSO2** **(K3,A** **3)**
CO1	K3	3	3	2	1	3	-	-	-	-	-	-	2	3	3
CO2	K3	3	3	2	1	3	-	-	-	-	-	-	2	3	3
CO3	K3	3	3	2	1	3	-	-	-	-	-	-	2	3	3
CO4	K3	3	3	2	1	3	-	-	-	-	-	-	2	3	3
CO5	K3	3	3	2	1	3	-	-	-	-	-	-	2	3	3
Weighted Average		3	3	2	1	3	-	-	-	-	-	-	2	3	3

SYLLABUS

LIST OF EXPERIMENTS

1. Introduction to various Data Visualization tools.
2. Basic Visualization in Python.
3. Basic Visualization in R.
4. Introduction to Tableau and Installation.
5. Connecting to Data and preparing data for visualization in Tableau.
6. Data Aggregation and Statistical functions in Tableau.
7. Data Visualizations in Tableau.
8. Basic Dashboards in Tableau.

Total hours: 45

Practical Record Book Index Page

Sl. No.	Date	Name of the Experiment	Page Number	Aim (10 Marks)	Program /procedure (40 Marks)	Output & Inference (15 Marks)	Viva-Voce (10Marks)	Total (75Marks)	Signature of the Faculty Member

Model Exam Marks (25):__________Total (100):_______

Signature of the Faculty Member

Practical Record Book Index Page

Sl. No.	Date	Name of the Experiment	Page Number	Aim (20 Marks)	Program /procedure (30 Marks)	Output & Inference (15Marks)	Viva-Voce (10Marks)	Total (75Marks)	Signature of the Faculty Member

Model Exam Marks (25):________________Total (100):_____

Signature of the Faculty Member

S.NO	EXPERIMENT	PREREQUISITES	LEARNING OBJECTIVES
1	FOR ALL EXPERIMENTS	PROGRAMMING IN PYTHON	1.Understand the importance of data visualization for business intelligence and decision making. 2. Know approaches to understand visual perception. 3.Learn about categories of visualization and application areas. 4.Familiarize with the data visualization tools 5.Gain knowledge of effective data visuals to solve workplace problems.

Ex. No:	1	Introduction to various Data Visualization tools
Date:		

AIM

Study experiment on knowing the various visualization tools.

DESCRIPTION

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.

The following are some common types of data visualizations:

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

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.

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

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

Viva Questions:

1. Give Two Top companies where Tableau is Used for Data Visualization.

2. What are the two uses of Google Charts?

3. Zoho analytics Specific use.

4. List the best 10 tools of Data Visualization Software of 2024 .

5. Is Excel a data visualization tool?

6. What are Matplotlib, Seaborn ?

7. Is Oracle a data visualization tool?

8**.**Why does NASA use Matlab?

9.How is data visualization used in healthcare?

10. How is data visualization used in healthcare?

RESULT

Thus the study experiment on various Data visualization tools is successfully completed.

Ex. No:	**2**	**Basic Visualization in Python**
Date:		

Aim :

To implement different visualization effects with Python on data using matplotlib seaborn.

Description :

Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python. It presents data in 2D graphics. Seaborn is a visualization library that is built on top of Matplotlib. Matplotlib can be installed using the following command:

pip install matplotlib

Once the module installed, it must be imported into the program using the following **command import matplotlib as mpl,** where mpl is the alias name given to matplotlib library.matplotlib. **Pyplot** is a state-based interface to matplotlib.

matplotlib.pyplot is a collection of functions that make matplotlib
work like MATLAB.

Each pyplot function makes some change to creates a figure,
creates a plotting area in a figure, plots some lines in a plotting area,
decorates the plot with labels etc.

pyplot can be imported into the program using following command

import matplotlib.pyplot as plt

Following are some of the basic data visualization plots

1. Line plots
2. Area plots
3. Histograms
4. Bar charts
5. Pie charts
6. Box plots
7. Scatter plots

I) Line Plots:

A line plot is used to represent quantitative values over a continuous interval or time period. It is generally used to depict trends on how the data has changed over time.

Program:

```
import matplotlib.pyplot as plt
x = [1, 2, 3, 4, 5, 6]
y = [1, 5, 3, 5, 7, 8]
plt.plot(x, y)
plt.show()
```

Output :

II) Area Plots:

An Area Plot is also called as Area Chart which is used to display magnitude and proportion of multiple variables.

Program:

```
import matplotlib.pyplot as plt
days = [1,2,3,4,5]
sleeping =[7,8,6,11,7]
eating = [2,3,4,3,2]
working =[7,8,7,2,2]
playing = [8,5,7,8,13]
plt.plot([],[],color='m', label='Sleeping', linewidth=5)
plt.plot([],[],color='c', label='Eating', linewidth=5)
plt.plot([],[],color='r', label='Working', linewidth=5)
plt.plot([],[],color='k', label='Playing', linewidth=5)
plt.stackplot(days, sleeping,eating,working,playing, colors=['m','c','r','k'])
plt.xlabel('x')
plt.ylabel('y')
plt.title('Stack Plot')
plt.legend()
plt.show()
```

Output :

iii) Histograms:

Histograms represents the frequency distribution of a dataset. It is a graph showing the number of observations within each given interval.

Program:

```
import matplotlib.pyplot as plt
population_age=[22,55,62,45,21,22,34,42,42,4,2,102,95,85,55,110,120,70,65,55,111,115,80]
bins = [0,10,20,30,40,50,60,70,80,90,100]
plt.hist(population_age, bins, histtype='bar', rwidth=0.8)
```

```
 plt.xlabel('age groups')
plt.ylabel('Number of people')
 plt.title('Histogram')
 plt.show()
```

Output:

iv) Bar Charts:

A Bar chart or bar graph is a chart or graph that presents categorical data with rectangular bars with heights or lengths proportional to the values that they represent.

A bar plot is a way of representing data where the length of the bars represents the magnitude/size of the feature/variable.

Program:

```
import matplotlib.pyplot as plt
plt.bar([0.25,1.25,2.25,3.25,4.25],[50,40,70,80,20],label="BMW",width=.5)
plt.bar([.75,1.75,2.75,3.75,4.75],[80,20,20,50,60],label="Audi", color='r',width=.5)
plt.legend()
plt.xlabel('Days')
plt.ylabel('Distance (kms)')
plt.title('Information')
plt.show()
```

Output:

V) **Pie Charts:**

A Pie chart is a circular statistical chart, which is divided into sectors to illustrate numerical proportion.

Program :

```
import matplotlib.pyplot as plt
```

```
days = [1,2,3,4,5]
sleeping =[7,8,6,11,7]
eating = [2,3,4,3,2]
working =[7,8,7,2,2]
playing = [8,5,7,8,13]
slices = [7,2,2,13]
activities = ['sleeping','eating','working','playing'] cols = ['c','m','r','b']
plt.pie(slices,labels=activities, colors=cols,start angle=90,shadow=True, explode=(0,0.1,0,0),
 autopct='%1.1f%%')
plt.title('Pie Plot')
 plt.show()
```

Output:

VI) Box Plots:

A Box plot (or box-and-whisker plot) shows the distribution of quantitative data in a way that facilitates comparisons between variables or across levels of a categorical variable.

Box plot shows the quartiles of the dataset while the whiskers extend encompass the rest of the distribution but leave out the points that are the outliers.

Program:

```
 import matplotlib.pyplot as plt
  x=[1,2,3,4,5,6,7]
  y=[1,2,4,5,3,6,9]
  z=[x,y] plt.boxplot(z,labels=[‘A’,’B’],showmeans=True)
  plt.show()
```

Output:

VII) Scatter Plots:

A Scatter chart, also called a scatter plot, is a chart that shows the relationship between two variables.

Program:

```
 import matplotlib.pyplot as plt
 x=[1,1.5,2,2.5,3,3.5,3.6]
 y=[7.5,8,8.5,9,9.5,10,10.5]
 x1=[8,8.5,9,9.5,10,10.5,11]
 y1=[3,3.5,3.7,4,4.5,5,5.2]
plt.scatter(x,y, label='high income low saving',color='r')
plt.scatter(x1,y1,label='low income high  savings',color='b')
plt.xlabel('saving*100')
plt.ylabel('income*1000')
plt.title('Scatter Plot')
plt.legend()
plt.show()
```

Output

Viva Questions :

1. **List the few advantages of Matplotlib .**

2. **Where in Data visualization Charts can be used ?**

3. **Histograms are used for what in visualization.**

4.what is the use of box plot visualization ?

5.In which of these situations are scatter plots useful?

Result:

Thus, the implementation of various visualization using python Matplotlib was executed Successfully

Ex. No:	3	Basic Visualization in R
Date:		

AIM

To implement the various fundamental data visualizations using R language .

DESCRIPTION:

ggplot2 is an open-source data visualization package for the statistical programming language R. ggplot is enriched with customized features to make visualization better. ggplot2 is a system for declaratively creating graphics, based on the Grammar Of Graphics .

The ggplot2 package can be easily installed using the following R function:

install. packages(ggplot2)

then the following command must be used in program to use ggplot package: library(ggplot2)

Consider the following dataset named **surveys.**

All the visualizations mentioned above are applied on this dataset.

```
Surveys<-data.frame(
record_id=c(1,2,3,4,5),
month=c(7,7,7,7,7),
day=c(16,16,16,17,17),
year=c(1977,1977,1977,1977,1977),
plot_id=c(2,3,2,7,3),
species_id=c(NL,NL,DM,DM,DM),
sex=c(M,M,F,M,M),
hindfoot_length=c(32,33,37,     36,35))
weight = c(20, 22, 25, 23, 24) # Added weight variable for demonstration
```

)

The visualizations are carried out for the following:

1. Scatter plots
2. Line plots
3. Box plots
4. Histograms
5. Bar charts

1.Scatter plot

Program:

```
library(ggplot2)

 Surveys <- data.frame(
  record_id = c(1, 2, 3, 4, 5),
  month = c(7, 7, 7, 7, 7),
  day = c(16, 16, 16, 17, 17),
  year = c(1977, 1977, 1977, 1977, 1977),
  plot_id = c(2, 3, 2, 7, 3),
  species_id = c("NL", "NL", "DM", "DM", "DM"),
  sex = c("M", "M", "F", "M", "M"),
  hindfoot_length = c(32, 33, 37, 36, 35),
  weight = c(20, 22, 25, 23, 24) # Added weight variable for demonstration
)

# Scatter plot
ggplot(data = Surveys, mapping = aes(x = hindfoot_length, y = weight)) +
  geom_point(alpha = 0.1, color = "blue")
```

OutPut:

2) **Histogram**

Program

```
    library(ggplot2)
   Surveys <- data.frame(
   record_id = c(1, 2, 3, 4, 5),
    month = c(7, 7, 7, 7, 7),
    day = c(16, 16, 16, 17, 17),
    year = c(1977, 1977, 1977, 1977, 1977),
    plot_id = c(2, 3, 2, 7, 3),
    species_id = c("NL", "NL", "DM", "DM", "DM"),
    sex = c("M", "M", "F", "M", "M"),
    hindfoot_length = c(32, 33, 37, 36, 35),
    weight = c(20, 22, 25, 23, 24) # Added weight variable for demonstration
    )

 # Create a histogram of the weight variable
 ggplot(data = Surveys, aes(x = weight)) +
  geom_histogram(binwidth = 1, fill = "blue", color = "black") +
  labs(x = "Weight", y = "Frequency", title = "Histogram of Weight")
```

Output:

3**) Bar Chart**

Program

```
  library(ggplot2)
  library(dplyr)

 # Create the data frame
 Surveys <- data.frame(
 record_id = c(1, 2, 3, 4, 5),
 month = c(7, 7, 7, 7, 7),
day = c(16, 16, 16, 17, 17),
year = c(1977, 1977, 1977, 1977, 1977),
plot_id = c(2, 3, 2, 7, 3),
species_id = c("NL", "NL", "DM", "DM", "DM"),
sex = c("M", "M", "F", "M", "M"),
hindfoot length = c(32, 33, 37, 36, 35),
weight = c(20, 22, 25, 23, 24) # Added weight variable for demonstration
 )

# Summarize the total weight for each species
Surveys_summary <- Surveys %>%
group_by(species_id) %>%
summarise(total_weight = sum(weight))

# Create the bar chart using ggplot2
ggplot(Surveys_summary, aes(x = species_id, y = total_weight)) +
geom_col(fill = "Red") +
 labs(title = "Total Weight by Species", x = "Species ID", y = "Total Weight")
```

Output:

4.Box Plot:

Program

```
library(ggplot2)
# Load required package
 library(ggplot2)

 # Create the data frame
   Surveys <- data.frame(
   record_id = c(1, 2, 3, 4, 5),
   month = c(7, 7, 7, 7, 7),
   day = c(16, 16, 16, 17, 17),
 year = c(1977, 1977, 1977, 1977, 1977),
  plot_id = c(2, 3, 2, 7, 3),
  species_id = c("NL", "NL", "DM", "DM", "DM"),
  sex = c("M", "M", "F", "M", "M"),
  hindfoot_length = c(32, 33, 37, 36, 35),
  weight = c(20, 22, 25, 23, 24) # Added weight variable for demonstration
   )

 # Create the box plot using ggplot2
ggplot(data = Surveys, mapping = aes(x = species_id, y = weight)) +
geom_boxplot() +
labs(title = "Box Plot of Weight by Species", x = "Species ID", y = "Weight")
```

Output:

5)Line Plot

Program

```
  # Load required packages
    library(ggplot2)
    library(dplyr)

    # Create example Surveys dataframe
    Surveys <- data.frame(
    record_id = c(1, 2, 3, 4, 5, 6, 7, 8, 9, 10),
    month = c(7, 7, 7, 7, 7, 8, 8, 8, 8, 8),
    day = c(16, 16, 16, 17, 17, 16, 16, 16, 17, 17),
    year = c(1977, 1977, 1978, 1978, 1978, 1979, 1979, 1979, 1980, 1980),
    plot_id = c(2, 3, 2, 7, 3, 2, 3, 2, 7, 3),
   species_id = c("NL", "NL", "DM", "DM", "DM", "NL", "NL", "DM", "DM", "DM"),
   sex = c("M", "M", "F", "M", "M", "F", "M", "M", "F", "M"),
   hindfoot_length = c(32, 33, 37, 36, 35, 34, 32, 38, 37, 36),
   weight = c(20, 22, 25, 23, 24, 21, 23, 26, 24, 25)
   )
  # Summarize counts per year and species
  yearly_counts <- Surveys %>%
  group_by(year, species_id) %>%
  summarise(n = n())

   # Print the summarized dataframe
   print(yearly_counts)
 # Create the line plot using ggplot2
ggplot(data = yearly_counts, aes(x = year, y = n, group = species_id, colour = species_id)) +
geom_line() +
labs(title = "Yearly Counts by Species", x = "Year", y = "Count", colour = "Species ID")
```

Output:

Viva Questions :

1. **What is the Use of R ?**

2. **What is the difference between plot and ggplot2 in R?**

3. **What is a Geom in ggplot?**

4. **List the most popular companies use R for visualization .**

5. **What are the advantages of R language ?**

Result:

Thus, the implementation of various visualization using ggplot2 in R language was implemented Successfully**.**

Ex. No:	4	Introduction to Tableau and Installation
Date:		

AIM

To implement the Tableau installation tool for performing the Data visualization.

DESCRIPTION:

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 can be downloaded from the following website:

https://www.tableau.com/products/public/download

after downloading, the following is the screen appears.

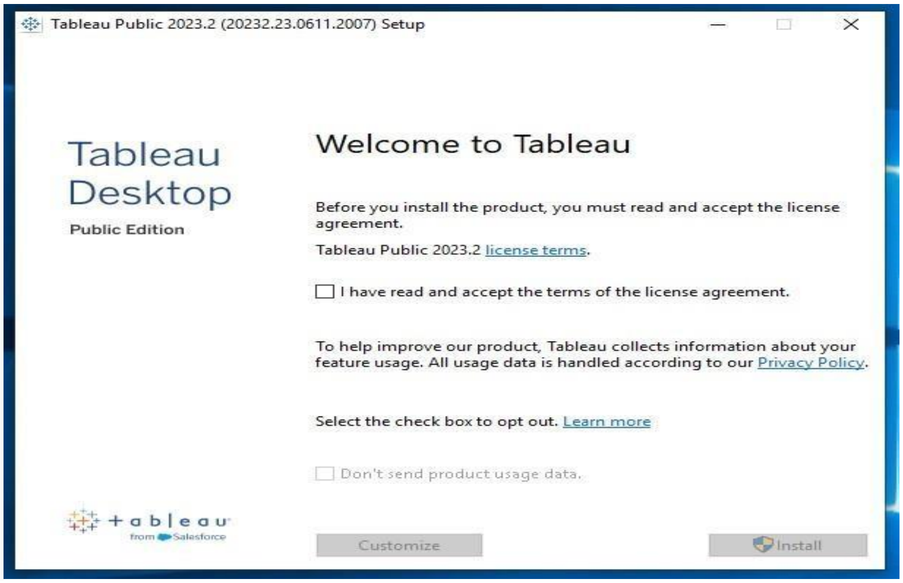


Tableau Public 2023.2 (20232.23.0611.2007) Setup

Tableau Desktop

Public Edition

Welcome to Tableau

Before you install the product, you must read and accept the license agreement.

Tableau Public 2023.2 license terms.

□ I have read and accept the terms of the license agreement.

To help improve our product, Tableau collects information about your feature usage. All usage data is handled according to our Privacy Policy.

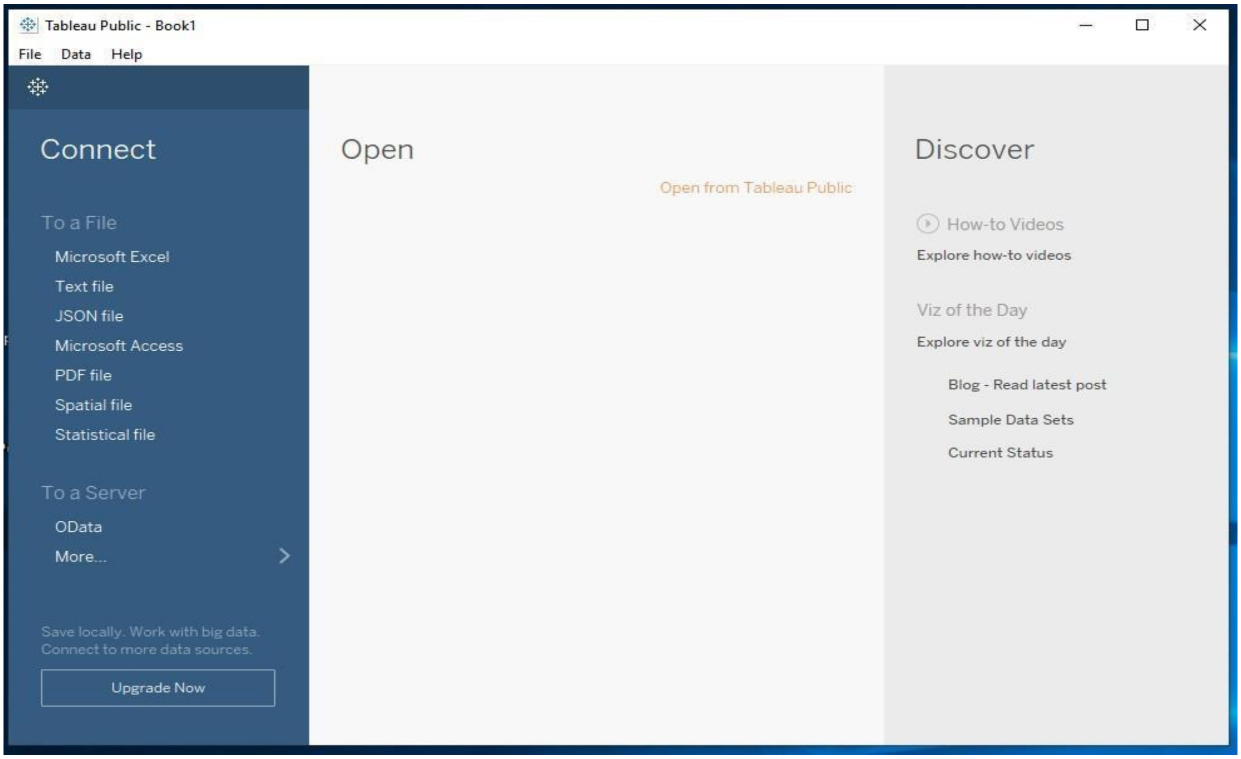
Select the check box to opt out. Learn more

□ Don't send product usage data.

Click the licence agreement checkbox and then click on install button.

After installation,click on Tableau Public icon to run Tableau.

Following is the Tableau Public home screen.



Viva Questions

1. **What is three types Tableau ?**

2. **What are the 7 data types in Tableau?**

3. **Is Tableau a coding language?**

4. **Is Tableau a tool or software?**

5.List few Best Features of Tableau .

Result :

Thus the Tableau for Public was installed successfully .

	5	Connecting to Data and preparing data for visualization in Tableau
Date:		

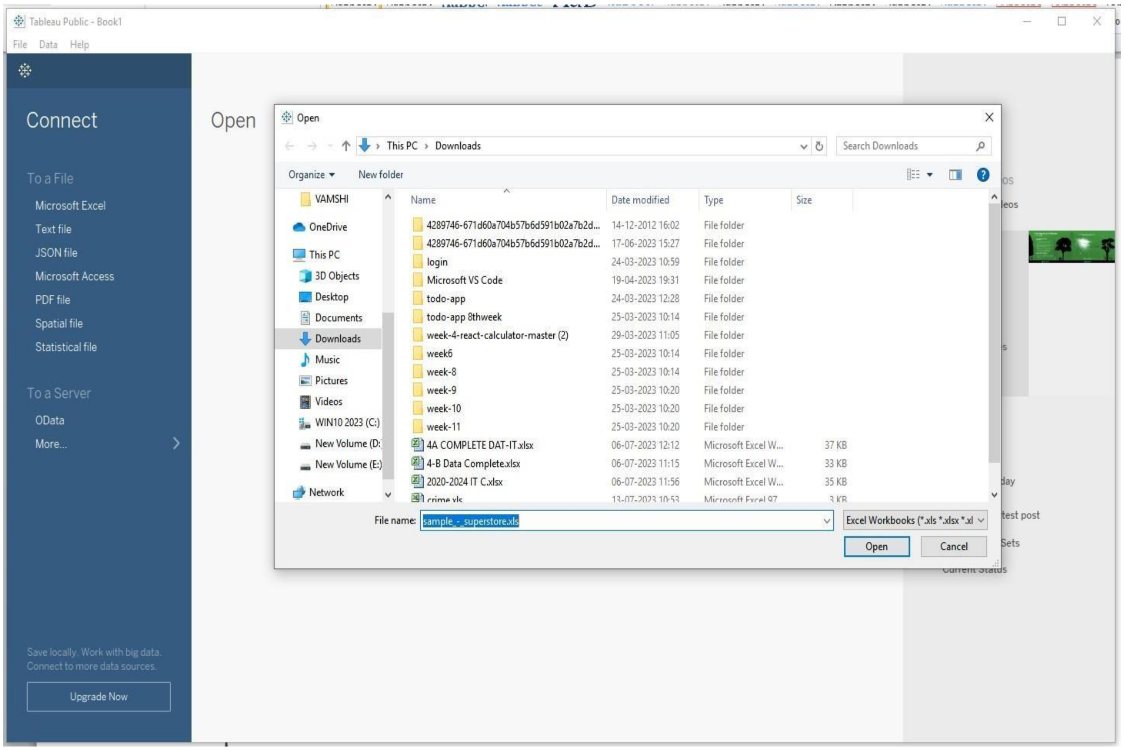
Aim

To connect the data from the external source and prepare it for visualization

DESCRIPTION:

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 data set was loaded into Tableau as follows:

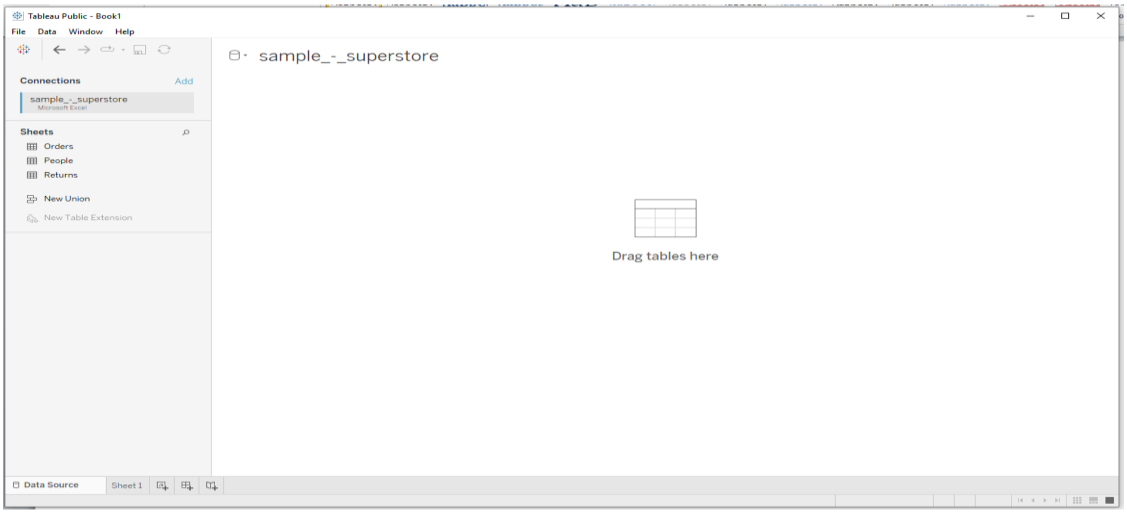


Name	Date modified	Type	Size
4289746-671d60a704b57b6d591b02a7b2d...	14-12-2012 16:02	File folder	
4289746-671d60a704b57b6d591b02a7b2d...	17-06-2023 15:27	File folder	
login	24-03-2023 10:59	File folder	
Microsoft VS Code	19-04-2023 19:31	File folder	
todo-app	24-03-2023 12:28	File folder	
todo-app 8thweek	25-03-2023 10:14	File folder	
week-4-react-calculator-master (2)	29-03-2023 11:05	File folder	
week6	25-03-2023 10:14	File folder	
week-8	25-03-2023 10:14	File folder	
week-9	25-03-2023 10:20	File folder	
week-10	25-03-2023 10:20	File folder	
week-11	25-03-2023 10:20	File folder	
4A COMPLETE DAT-IT.xlsx	06-07-2023 12:12	Microsoft Excel W...	37 KB
4-B Data Complete.xlsx	06-07-2023 11:15	Microsoft Excel W...	33 KB
2020-2024 IT C.xlsx	06-07-2023 11:56	Microsoft Excel W...	35 KB
crime.xls	13-07-2023 10:53	Microsoft Excel 97...	3 KB

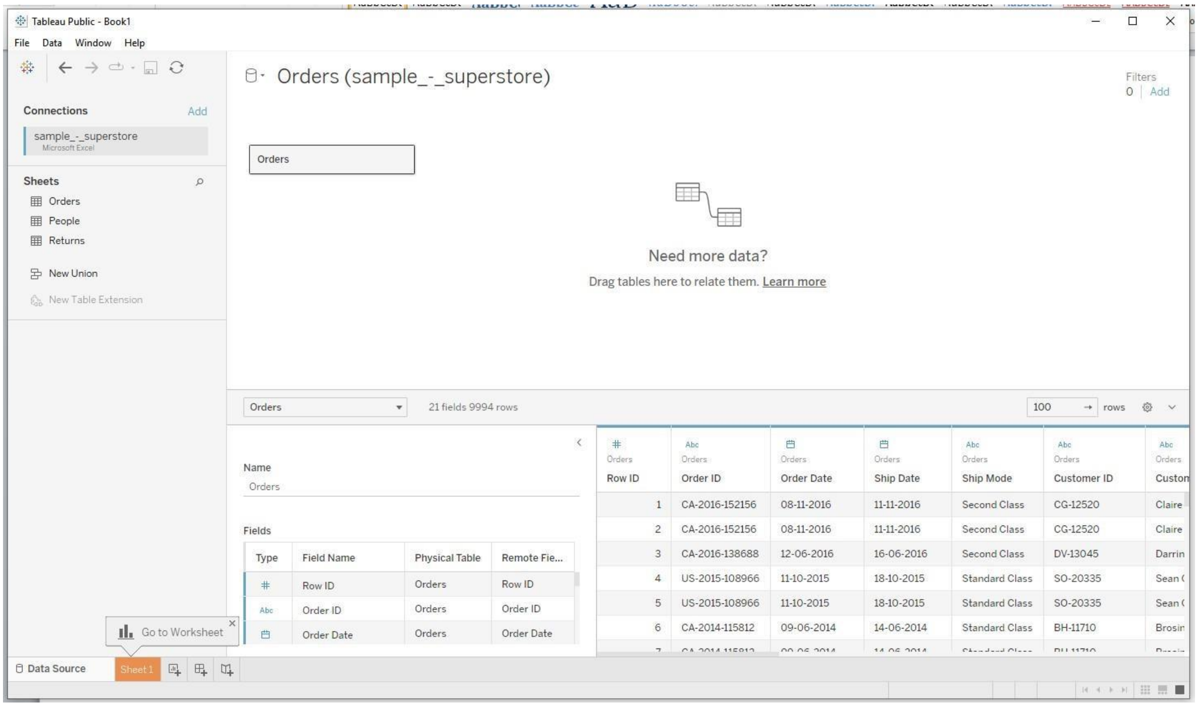
File name: sample_-_superstore.xls

Excel Workbooks (*.xls *.xlsx *.xl

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.



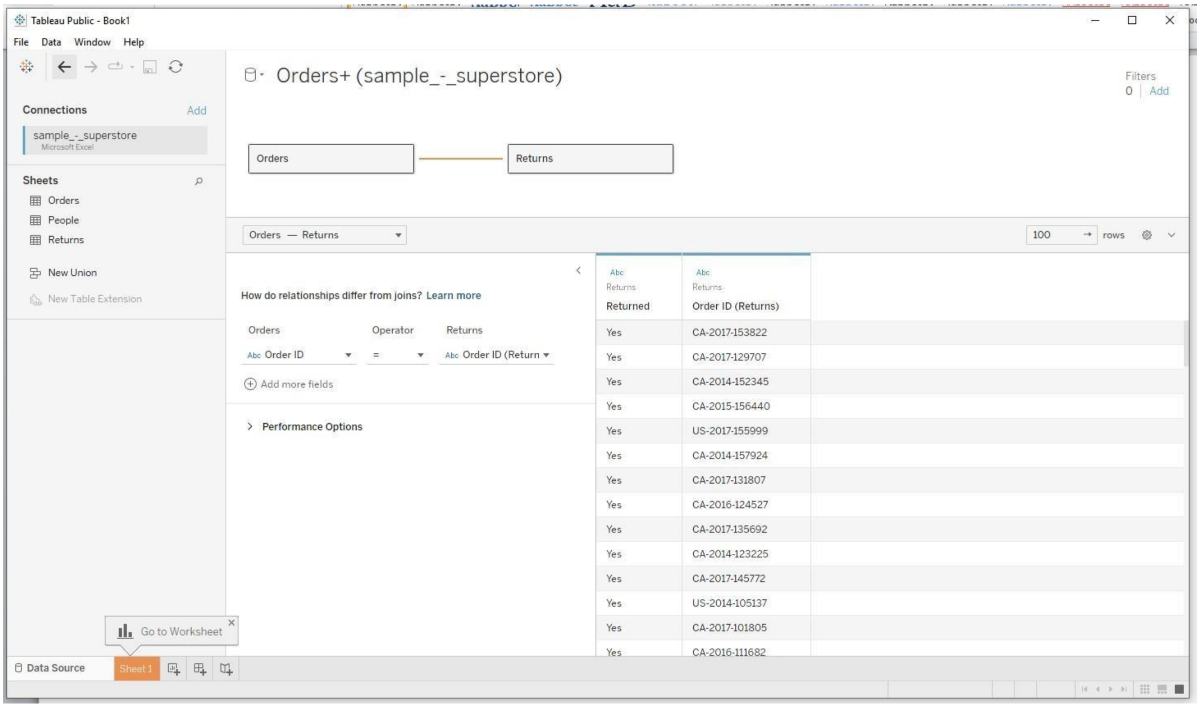
Name
Orders

Fields

Type	Field Name	Physical Table	Remote Fie...
#	Row ID	Orders	Row ID
Abc	Order ID	Orders	Order ID
📅	Order Date	Orders	Order Date

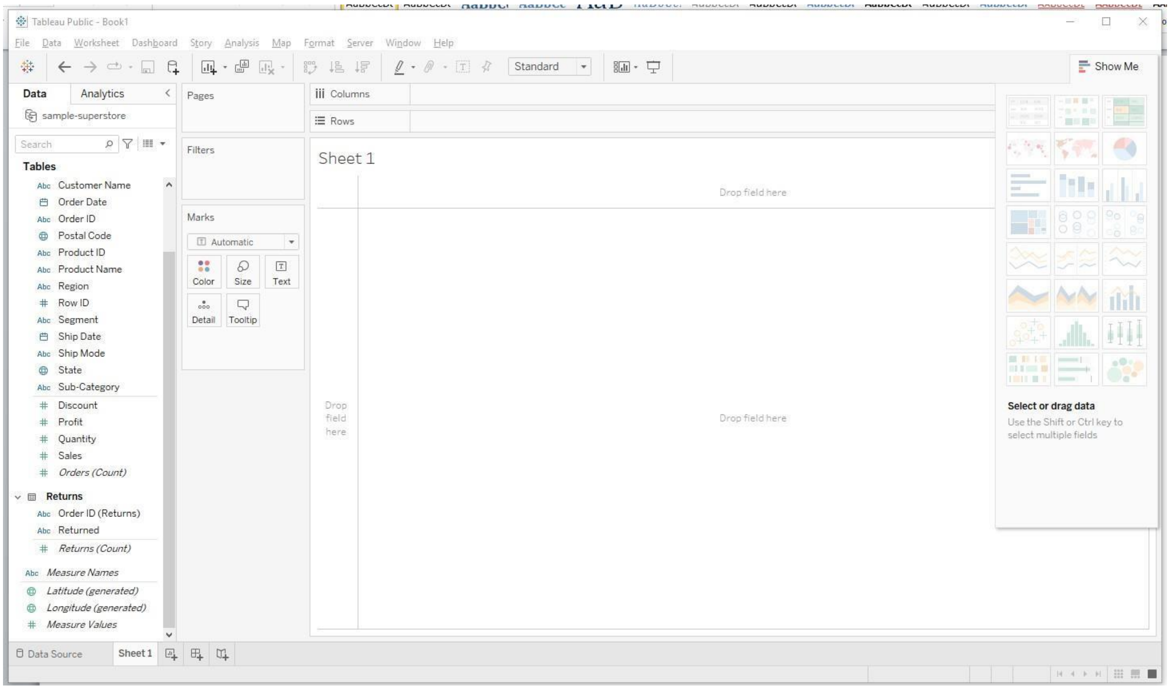
# Orders Row ID	Abc Orders Order ID	📅 Orders Order Date	📅 Orders Ship Date	Abc Orders Ship Mode	Abc Orders Customer ID	Abc Orders Custon
1	CA-2016-152156	08-11-2016	11-11-2016	Second Class	CG-12520	Claire
2	CA-2016-152156	08-11-2016	11-11-2016	Second Class	CG-12520	Claire
3	CA-2016-138688	12-06-2016	16-06-2016	Second Class	DV-13045	Darrin
4	US-2015-108966	11-10-2015	18-10-2015	Standard Class	SO-20335	Sean (
5	US-2015-108966	11-10-2015	18-10-2015	Standard Class	SO-20335	Sean (
6	CA-2014-115812	09-06-2014	14-06-2014	Standard Class	BH-11710	Brosin

Now drag Returns table onto the Canvas to the right of Orders table. This shows the relation between the two tables Orders and Returns.



Returned	Order ID (Returns)
Yes	CA-2017-153822
Yes	CA-2017-129707
Yes	CA-2014-152345
Yes	CA-2015-156440
Yes	US-2017-155999
Yes	CA-2014-157924
Yes	CA-2017-131807
Yes	CA-2016-124527
Yes	CA-2017-135692
Yes	CA-2014-123225
Yes	CA-2017-145772
Yes	US-2014-105137
Yes	CA-2017-101805
Yes	CA-2016-111682

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.

VIVA VOICE

1. **What is the use of connect in tableau**

2. **How to load the data into tableau sheet**

3. **What are dimensions in tableau with example?**

4. **What is a measurement in tableau?**

5. **What is the use of SHOW ME in tableau?**

Result:

I have learnt thus the implementation for loading the prepared data was executed successfully

Ex. No:	**6**	**Data aggregation and statistical functions**
Date:		

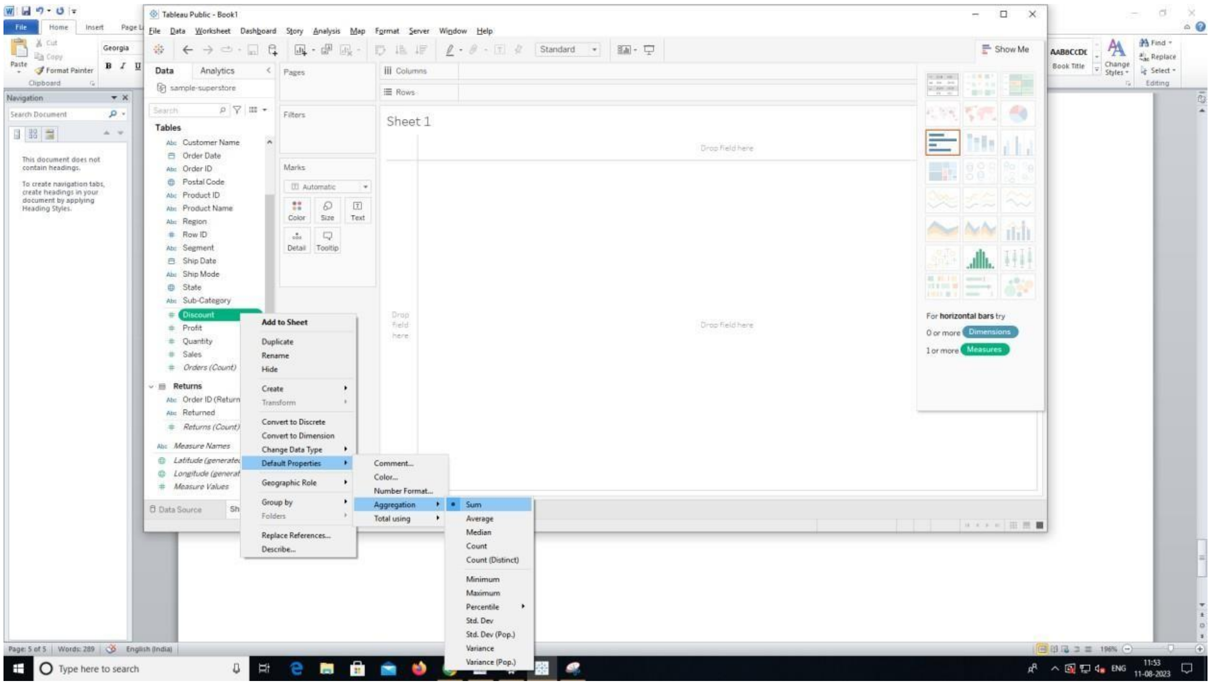
AIM

To implement the data aggregation and statistical functions

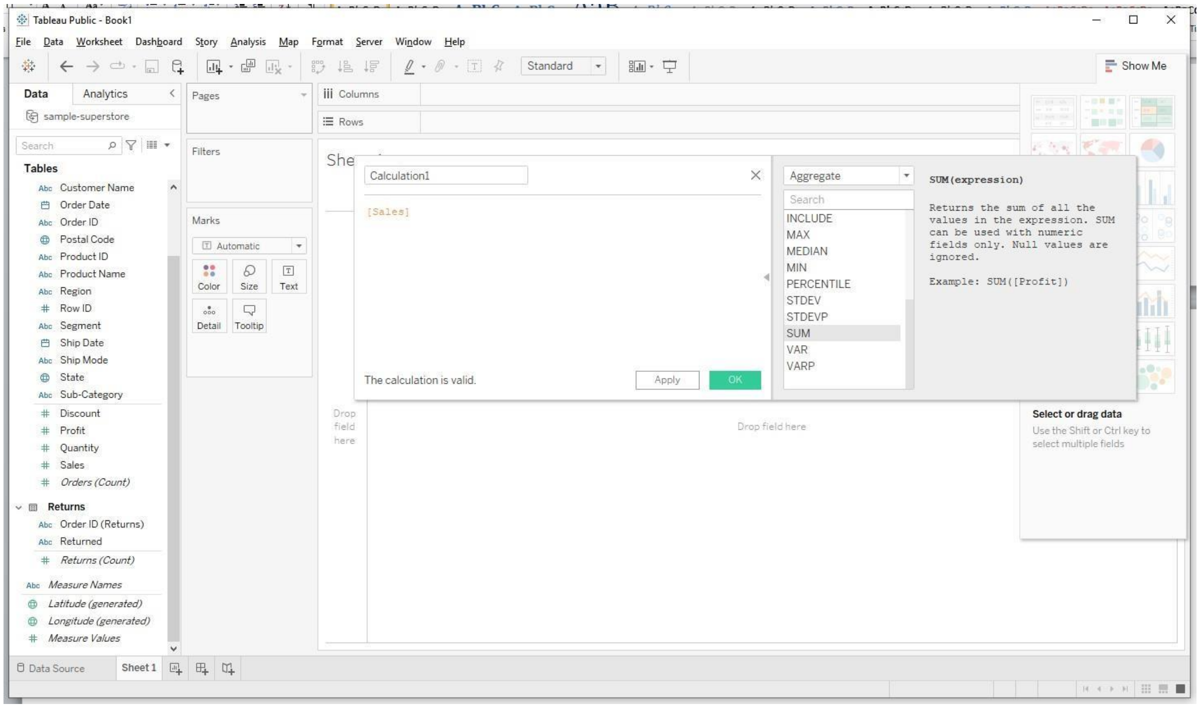
DESCRIPTION:

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.



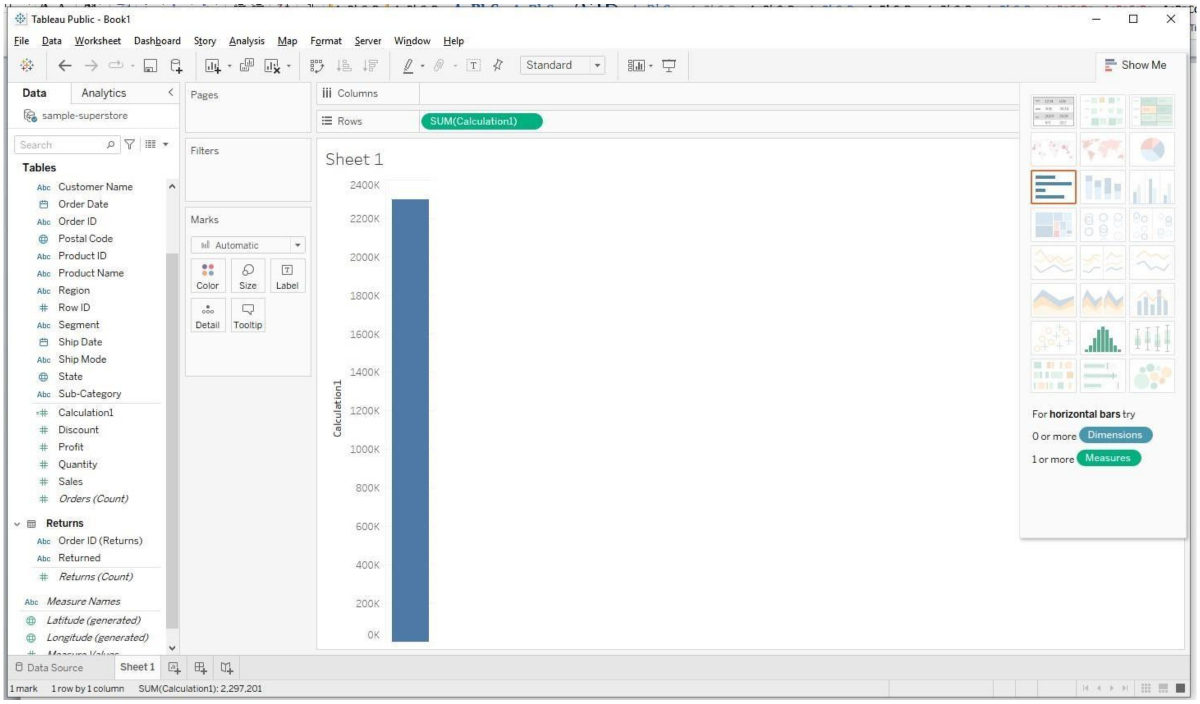
Calculation1

```
[Sales]
```

The calculation is valid.

```
SUM(expression)

Returns the sum of all the
values in the expression. SUM
can be used with numeric
fields only. Null values are
ignored.

Example: SUM([Profit])
```

Then click on apply and results are shown below:



1 mark 1 row by 1 column SUM(Calculation1): 2,297,201

In the same way we can apply any aggregate or statistical function on data with the help of calculated fields.

Output:

VIVA VOICE

1. **What is aggregation in data visualization?**

2. **What is the default aggregation measure in tableau?**

3. **Can tableau do standard devatation?**

4. **How to show percentile in tableau?**

5. **How to calculate formula in tableau?**

Result:

Thus the implementation data aggregation and statistical functions are profound using tableau Is completed and executed successfully

Ex. No:	7	Data Visualization
Date:		

AIM

To implement various types of Visualization using tableau

Description:

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.

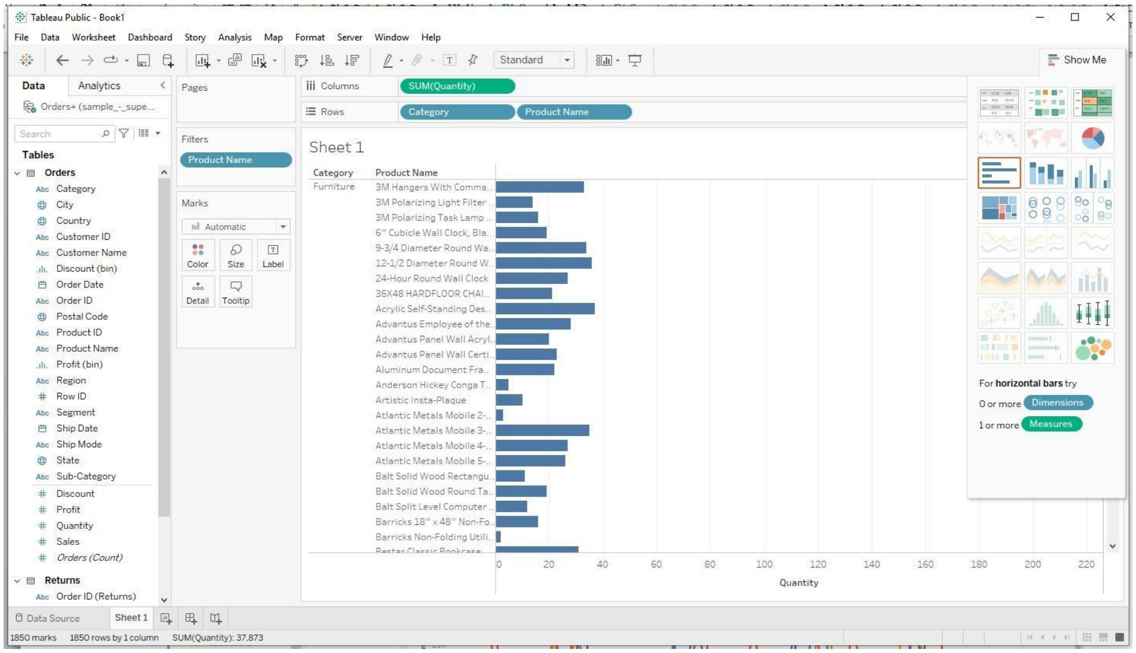


Tableau Public - Book1

Columns: SUM(Quantity)

Rows: Category, Product Name

Filters: Product Name

Marks: Automatic

For **horizontal bars** try

0 or more Dimensions

1 or more Measures

1850 marks 1850 rows by 1 column SUM(Quantity): 37,873

In similar to above, stacked bar graph can be created and the result is shown below

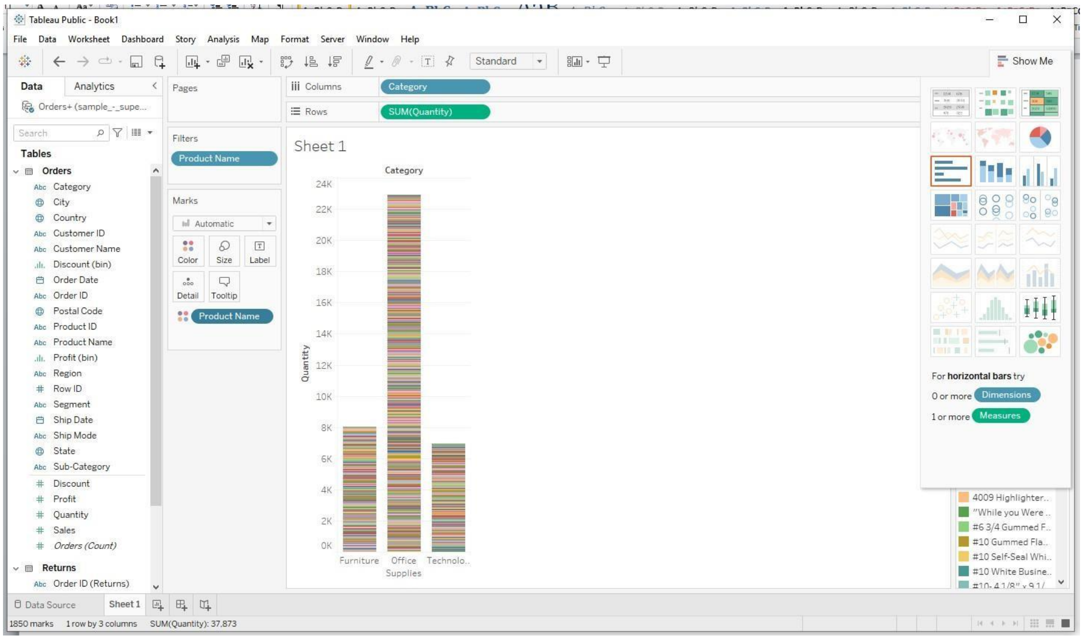


Tableau Public - Book1

Columns: Category

Rows: SUM(Quantity)

Filters: Product Name

Marks: Automatic; Color: Product Name

1850 marks 1 row by 3 columns SUM(Quantity): 37,873

Side-by-side bar chart can be created in following way.

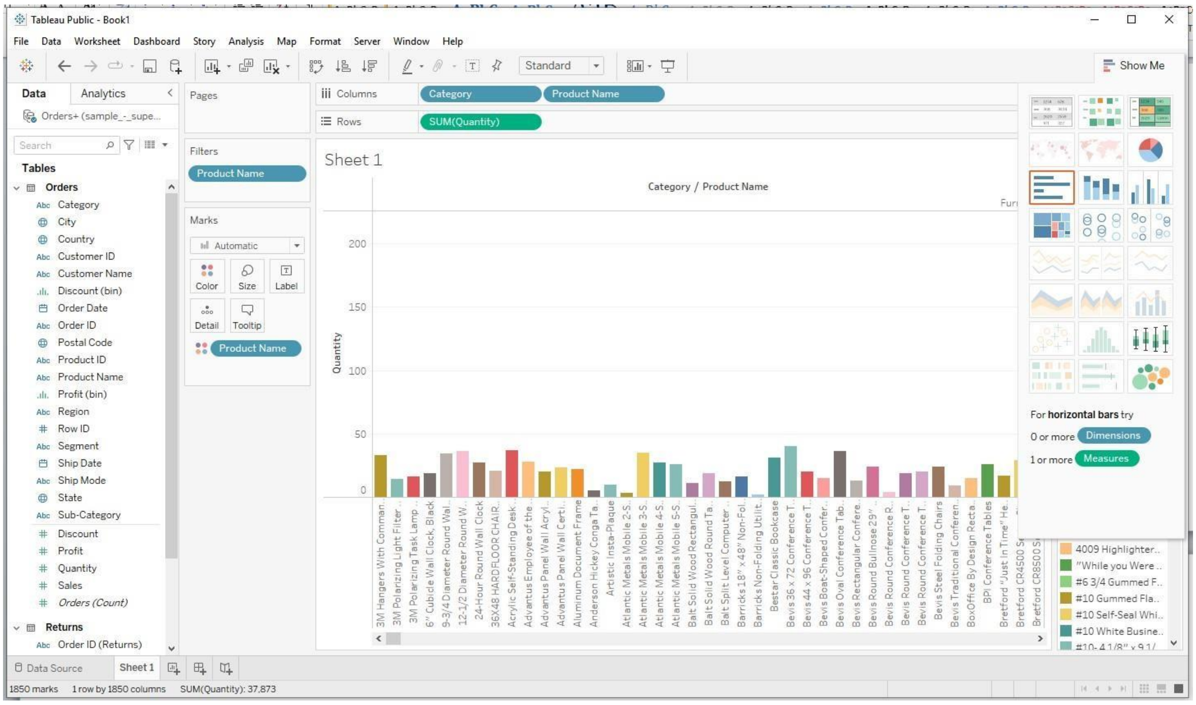


Tableau Public - Book1

Columns: Category, Product Name

Rows: SUM(Quantity)

Filters: Product Name

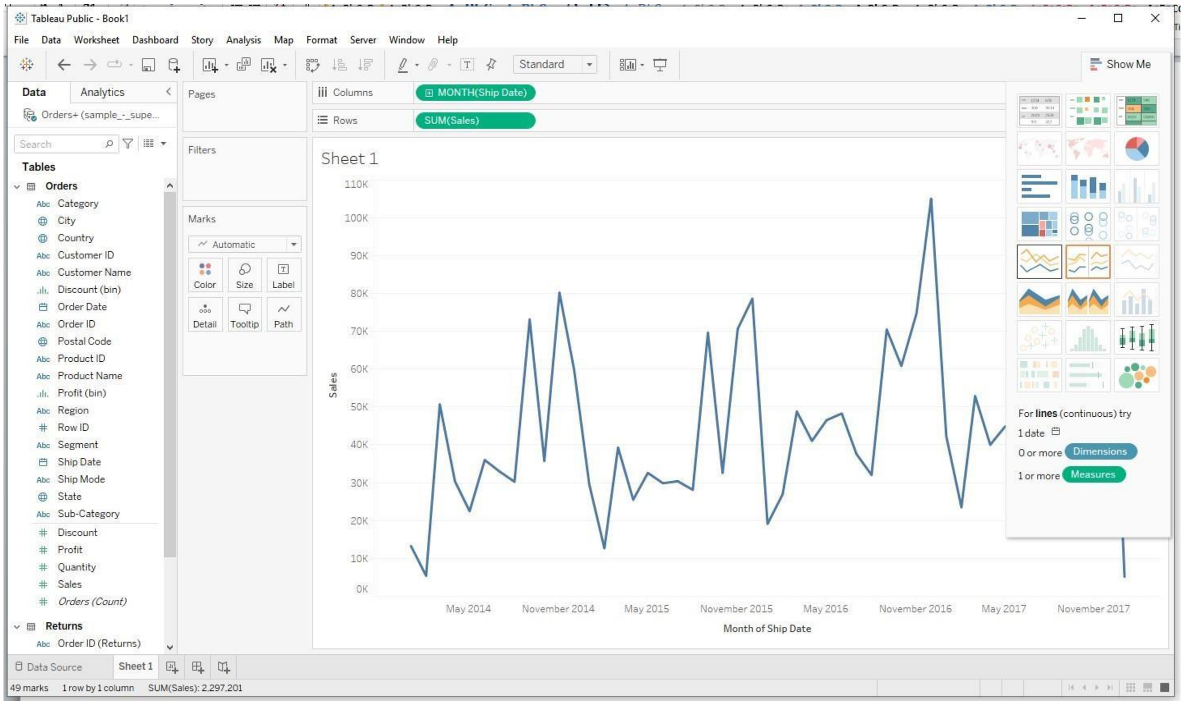
Marks: Automatic; Color: Product Name

1850 marks 1 row by 1850 columns SUM(Quantity): 37,873

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Orders+ (sample_-_supe...
Tables
Orders
Category
City
Country
Customer ID
Customer Name
Discount (bin)
Order Date
Order ID
Postal Code
Product ID
Product Name
Profit (bin)
Region
Row ID
Segment
Ship Date
Ship Mode
State
Sub-Category
Discount
Profit
Quantity
Sales
Orders (Count)
Returns
Order ID (Returns)
Pages
Color
Size
Label
Detail
Tooltip
Sheet 1
Category / Product Name
Furn
Quantity
200
150
100
50
0
3M Hangers With Comman..
3M Polarizing Light Filter
3M Polarizing Task Lamp ..
6" Cubicle Wall Clock, Black
9-3/4 Diameter Round Wal..
12-1/2 Diameter Round W..
24-Hour Round Wall Clock
36X48 HARDFLOOR CHAIR..
Acrylic Self-Standing Desk..
Advantus Employee of the ..
Advantus Panel Wall Acryl..
Advantus Panel Wall Certi..
Aluminum Document Frame
Anderson Hickey Conga Ta..
Artistic Insta-Plaque
Atlantic Metals Mobile 2-S..
Atlantic Metals Mobile 3-S..
Atlantic Metals Mobile 4-S..
Atlantic Metals Mobile 5-S..
Bait Solid Wood Rectangul..
Bait Solid Wood Round Ta..
Bait Split Level Computer..
Barricks 18" x 48" Non-Fol..
Barricks Non-Folding Utilit..
Bestar Classic Bookcase
Bevis 36 x 72 Conference T..
Bevis 44 x 96 Conference T..
Bevis Boat-Shaped Confer..
Bevis Oval Conference Tab..
Bevis Rectangular Confere..
Bevis Round Bullnose 29"..
Bevis Round Conference R..
Bevis Round Conference T..
Bevis Round Conference T..
Bevis Steel Folding Chairs
Bevis Traditional Conferen..
BoxOffice By Design Recta..
BPI Conference Tables
Bretford "Just In Time" He..
Bretford CR4500 Se..
Bretford CR8500 Se..
For horizontal bars try
0 or more Dimensions
1 or more Measures
4009 Highlighter..
"While you Were ..
#6 3/4 Gummed F..
#10 Gummed Fla..
#10 Self-Seal Whi..
#10 White Busine..
Data Source
Sheet 1

Line graph: Line graph can be continuous or discrete.

Continuous line graph is shown below:



Columns: MONTH(Ship Date)

Rows: SUM(Sales)

For **lines** (continuous) try

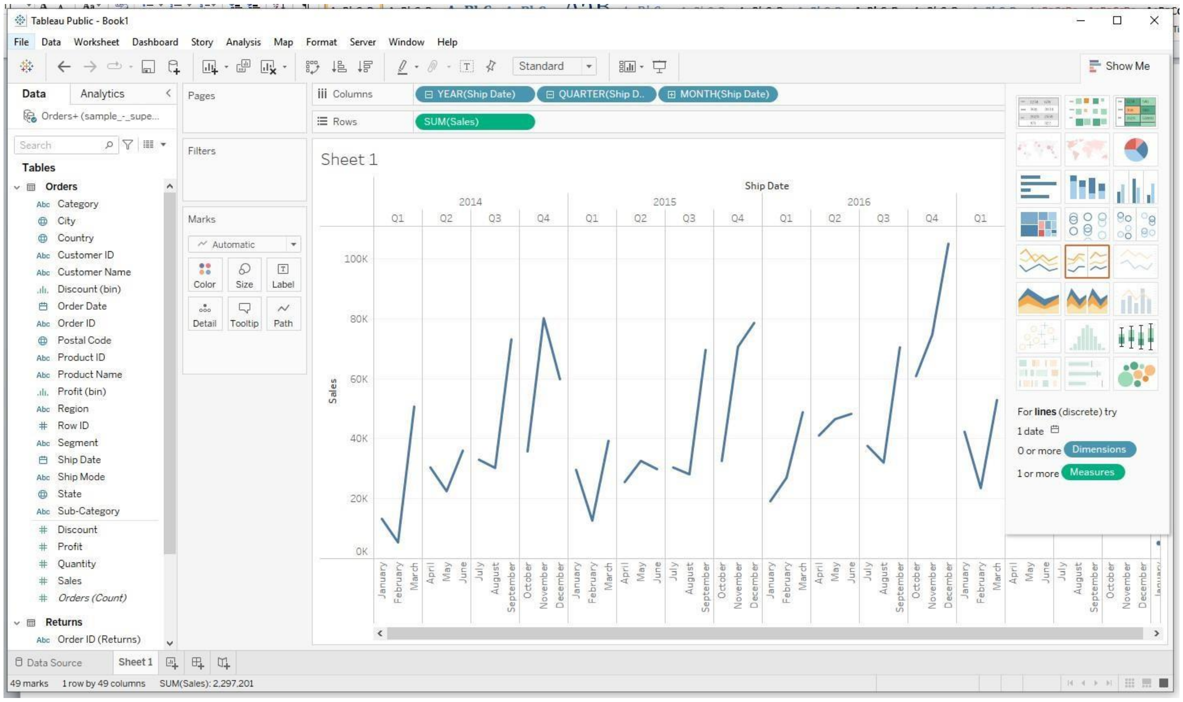
1 date

0 or more Dimensions

1 or more Measures

49 marks 1 row by 1 column SUM(Sales): 2,297,201

Discrete line graph is shown below:



Columns: YEAR(Ship Date) QUARTER(Ship D.. MONTH(Ship Date)

Rows: SUM(Sales)

For **lines** (discrete) try

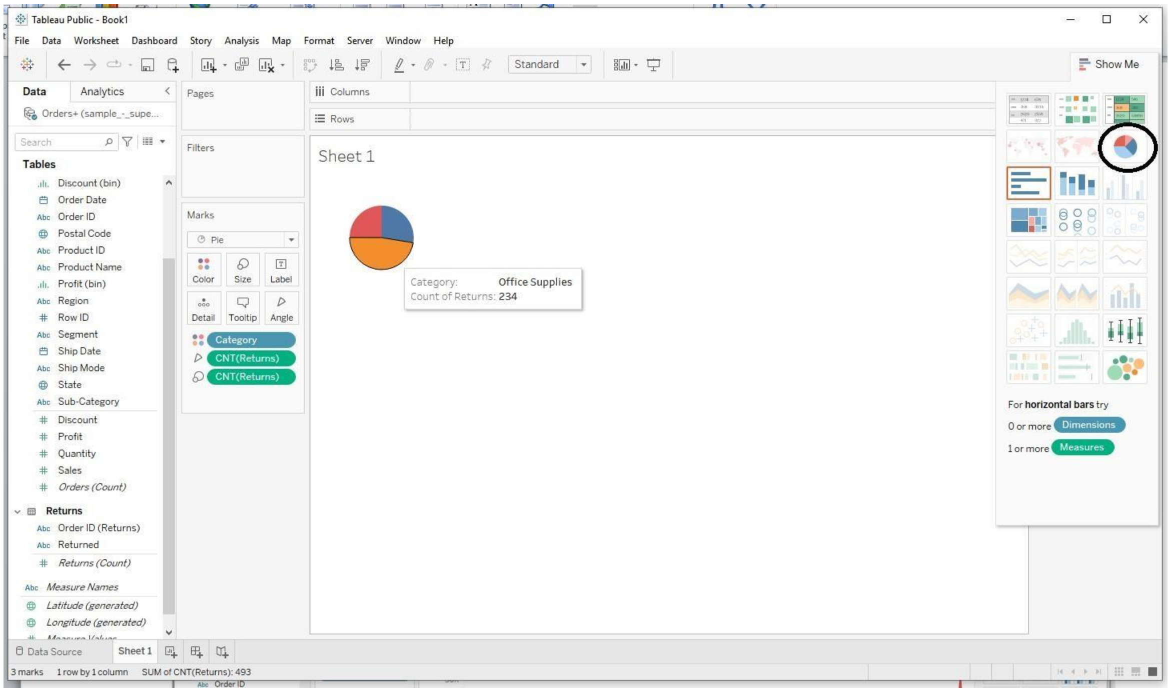
1 date

0 or more Dimensions

1 or more Measures

49 marks 1 row by 49 columns SUM(Sales): 2,297,201

Pie chart:



Marks: Pie

Category

CNT(Returns)

CNT(Returns)

Category: Office Supplies
Count of Returns: 234

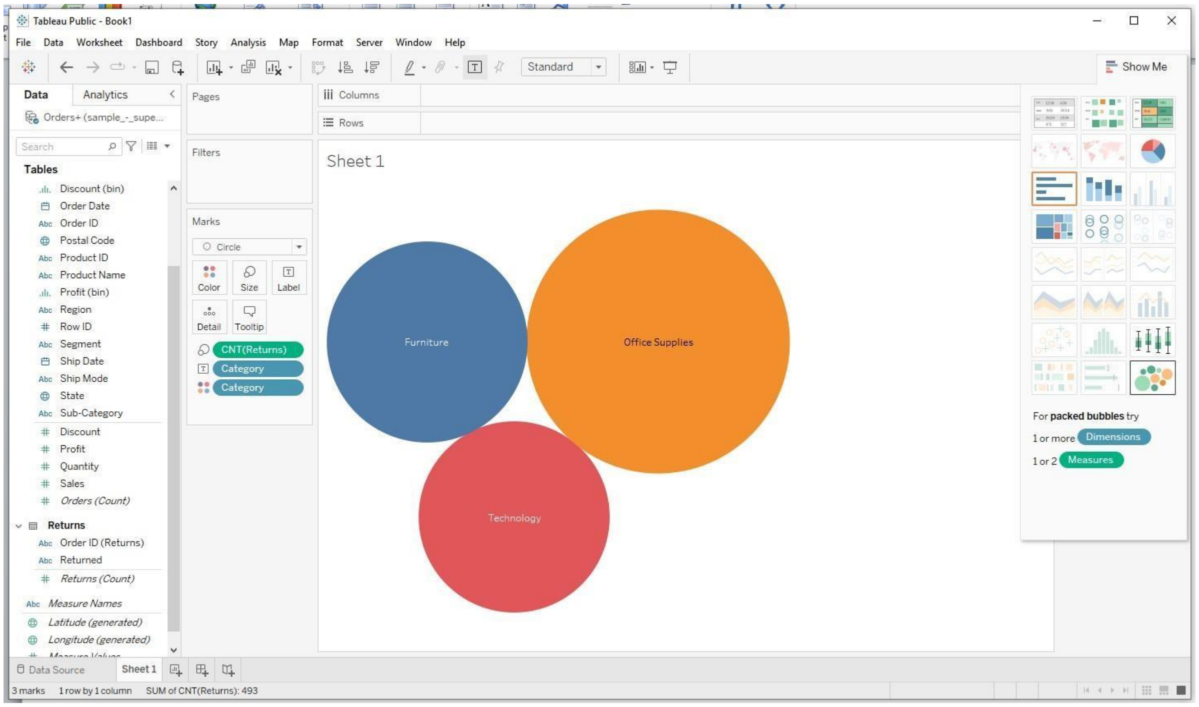
For **horizontal bars** try

0 or more Dimensions

1 or more Measures

3 marks 1 row by 1 column SUM of CNT(Returns): 493

Bubble chart:



Marks: Circle

CNT(Returns)

Category

Category

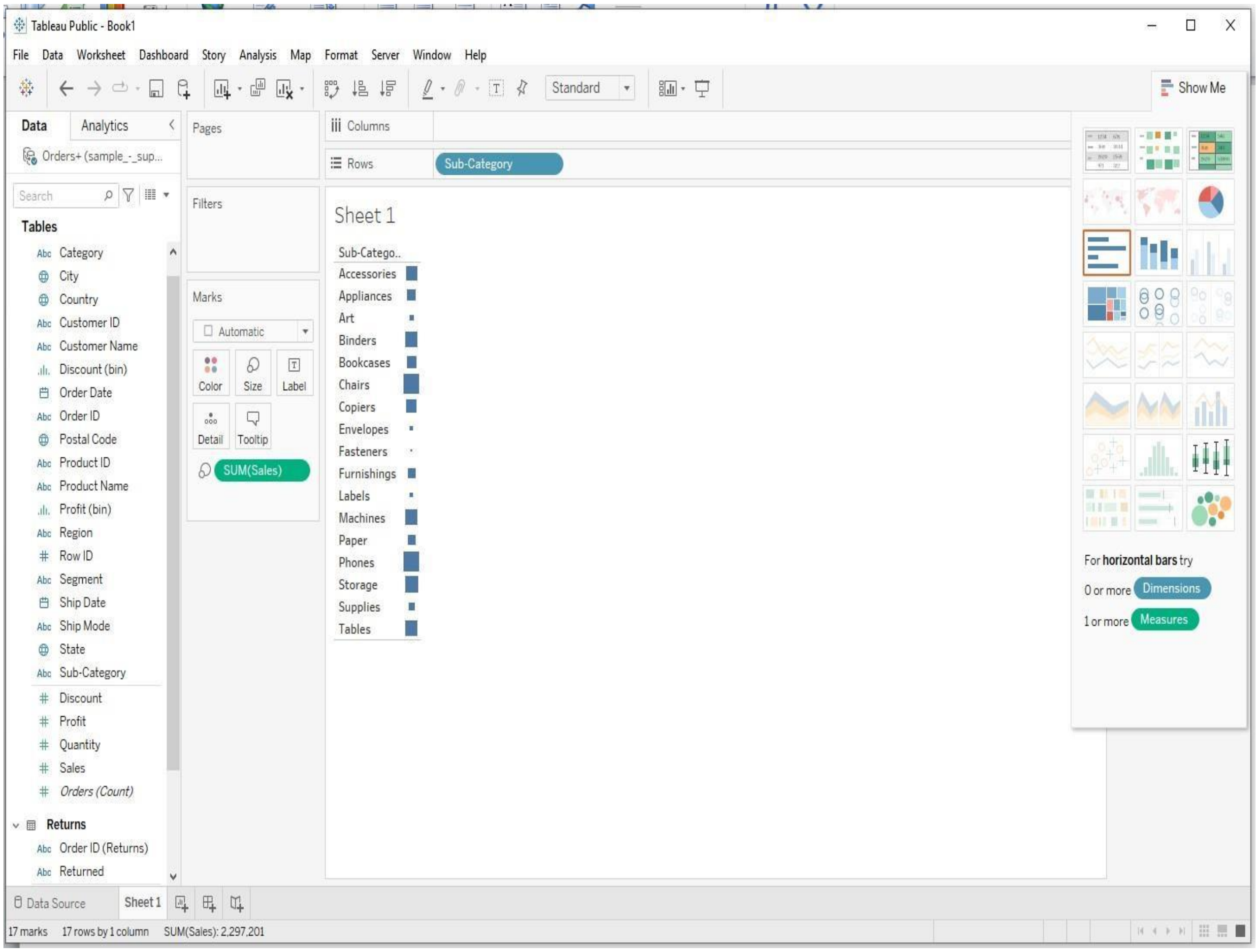
For **packed bubbles** try

1 or more Dimensions

1 or 2 Measures

3 marks 1 row by 1 column SUM of CNT(Returns): 493

Heat map:



Sub-Catego..
Accessories
Appliances
Art
Binders
Bookcases
Chairs
Copiers
Envelopes
Fasteners
Furnishings
Labels
Machines
Paper
Phones
Storage
Supplies
Tables

17 marks 17 rows by 1 column SUM(Sales): 2,297,201

VIVA VOICE

1. **Is tableau a tool or a software?**

2.List few advantages of tableau.

3.What is tableau called now?

4.Give some tableau dashboard examples

5.how many visualizations are there in tableau

RESULT

Thus the implementation of visualization using different types in tableau is completed and executed successfully

Ex. No:	8	Dashboards
Date:		

AIM

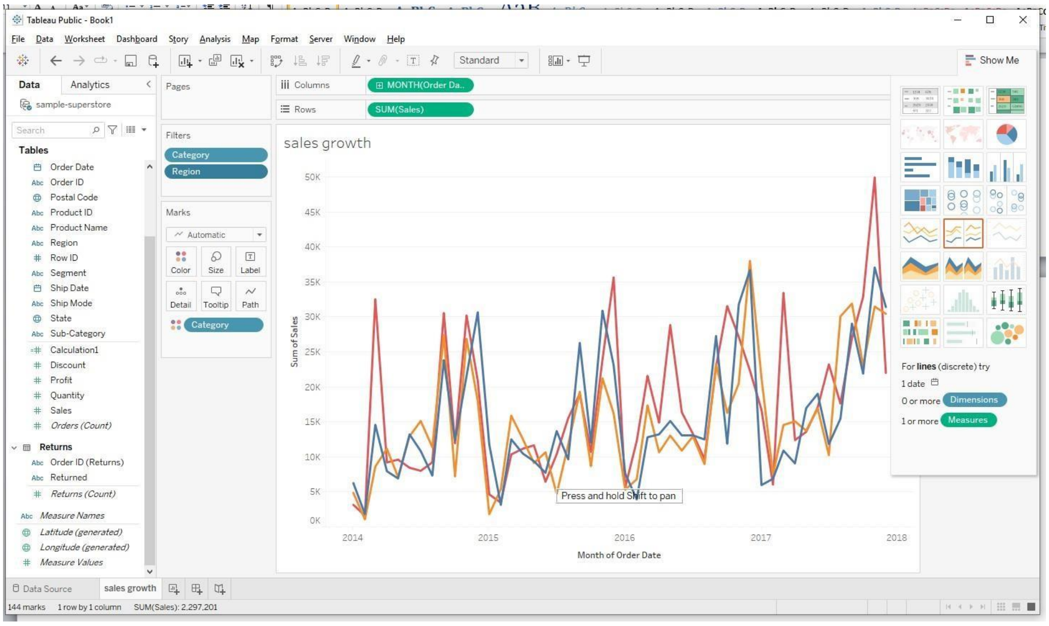
To implements Dashboards using tableau

Description:

A 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 and understand 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.



Columns: MONTH(Order Da..

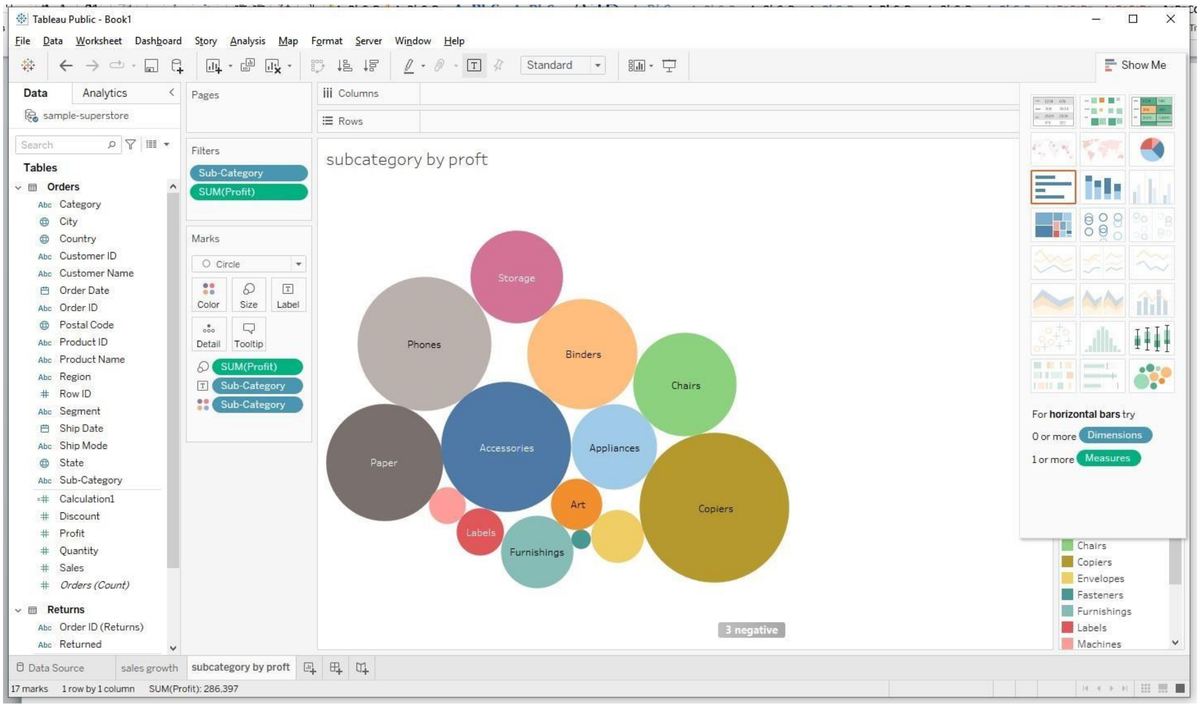
Rows: SUM(Sales)

Filters: Category, Region

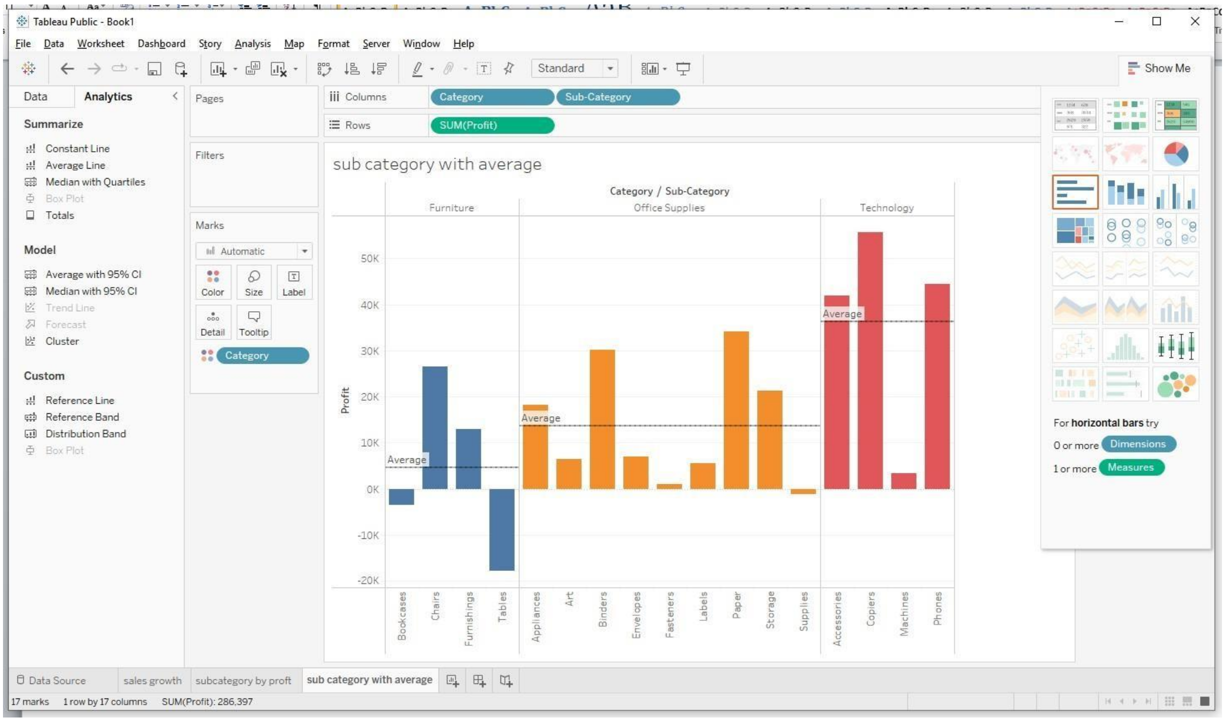
Marks: Automatic; Color: Category

144 marks 1 row by 1 column SUM(Sales): 2,297,201

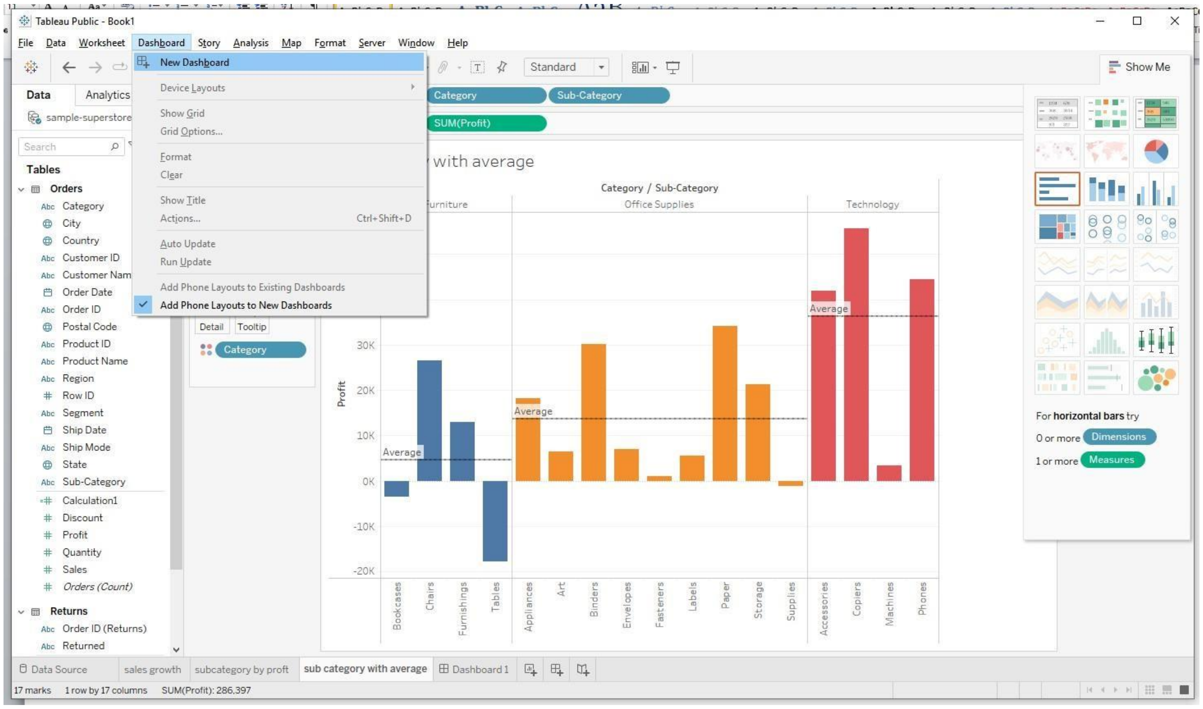
Now go to 2nd sheet for creating the 2nd view. The second view is shown below. A bubble chart was drawn between profit and subcategory. Then rename the sheet.



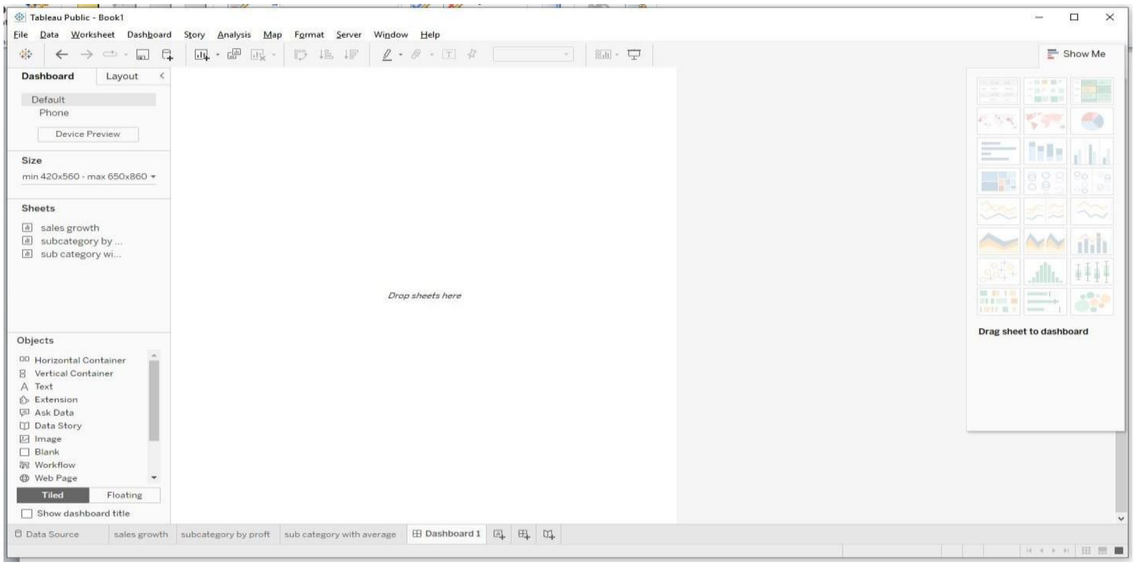
Next 3rd view is created as follows for profit for each subcategory in the category with averages.



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. Dahsboard can be customized in terms of its appearance by the user if requied. Dashboard once created can be saved on users system and can be retrieved whenever required.

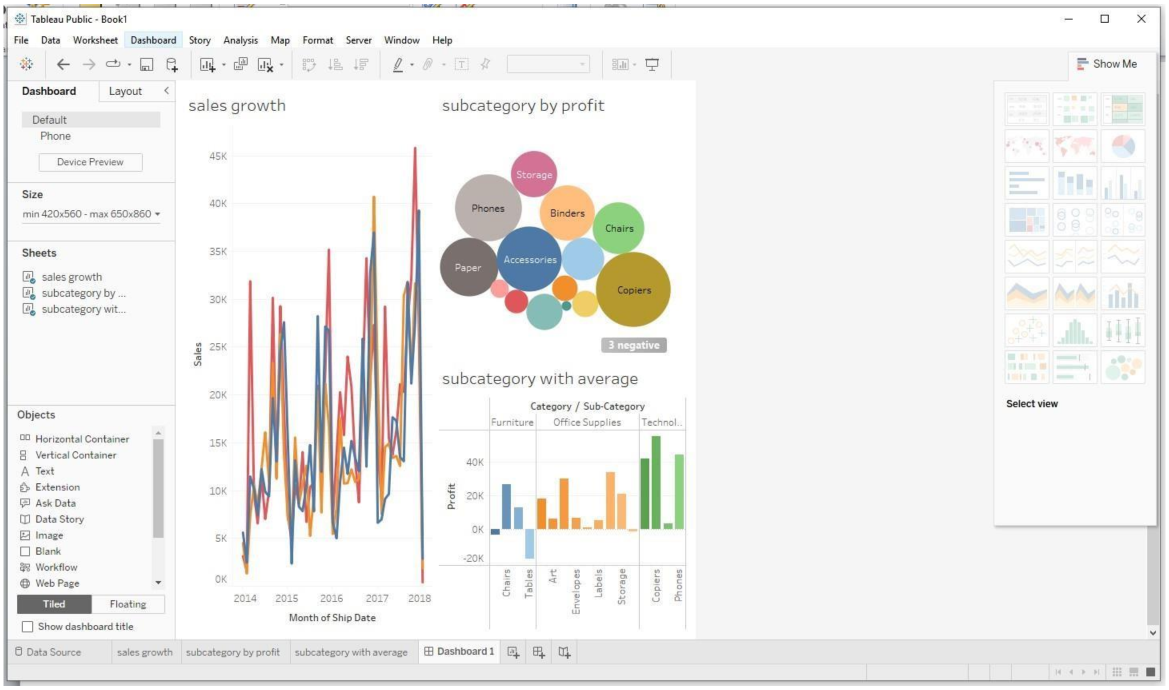


Tableau Public - Book1
File
Data
Worksheet
Dashboard
Story
Analysis
Map
Format
Server
Window
Help
Show Me
Dashboard
Layout
Default
Phone
Device Preview
Size
min 420x560 - max 650x860
Sheets
sales growth
subcategory by ...
subcategory wit...
Objects
Horizontal Container
Vertical Container
Text
Extension
Ask Data
Data Story
Image
Blank
Workflow
Web Page
Tiled
Floating
□ Show dashboard title
sales growth
45K
40K
35K
30K
25K
20K
15K
10K
5K
0K
Sales
2014
2015
2016
2017
2018
Month of Ship Date
subcategory by profit
Storage
Phones
Binders
Chairs
Paper
Accessories
Copiers
3 negative
subcategory with average
Category / Sub-Category
Furniture
Office Supplies
Technol..
40K
20K
0K
-20K
Profit
Chairs
Tables
Art
Envelopes
Labels
Storage
Copiers
Phones
Select view
Data Source
sales growth
subcategory by profit
subcategory with average
Dashboard 1

OUTPUT

VIVA VOCE:

1. **What are dashboards in tableau?**

2. **What is the use of dashboard in tableau?**

3. **List the difference between dashboard and report in tableau?**

4. **What are the components of tableau?**

5. **List any 5 benefits of dashboard?**

RESULT

Thus, the implementation of creation of dashboard is completed and executed successfully

CONTENT BEYOND SYLLABUS

Ex. No:	**9**	**TIME SERIES ANALYSIS**
Date:		

AIM

To Analysis Time series visualization

What Is a Time Series?

A time series is a set of data points that are collected over a period of time, usually at regular intervals. The most common type of time series data is financial data, such as stock prices or exchange rates. However, time series can also be used to track other types of information, such as meteorological data or sales figures.

Time series can be either univariate or multivariate.

- Univariate Time Series Data: Based on one variable, such as stock prices or the number of cases of a disease
- Multivariate Time Series Data: Based on multiple variables, such as weather data (which could include variables such as temperature, humidity, and rainfall)

Time series are often graphed to visualize the data, and they can be analyzed using statistical methods. Time series analysis can be used for forecasting future values, and it is a powerful tool for understanding complex data.

Types of Time Series Data

There are two main types of time series data:

Continuous data: This type of data is collected at regular intervals and can be represented by a line on a graph. For example, data from a thermometer would be considered continuous data.

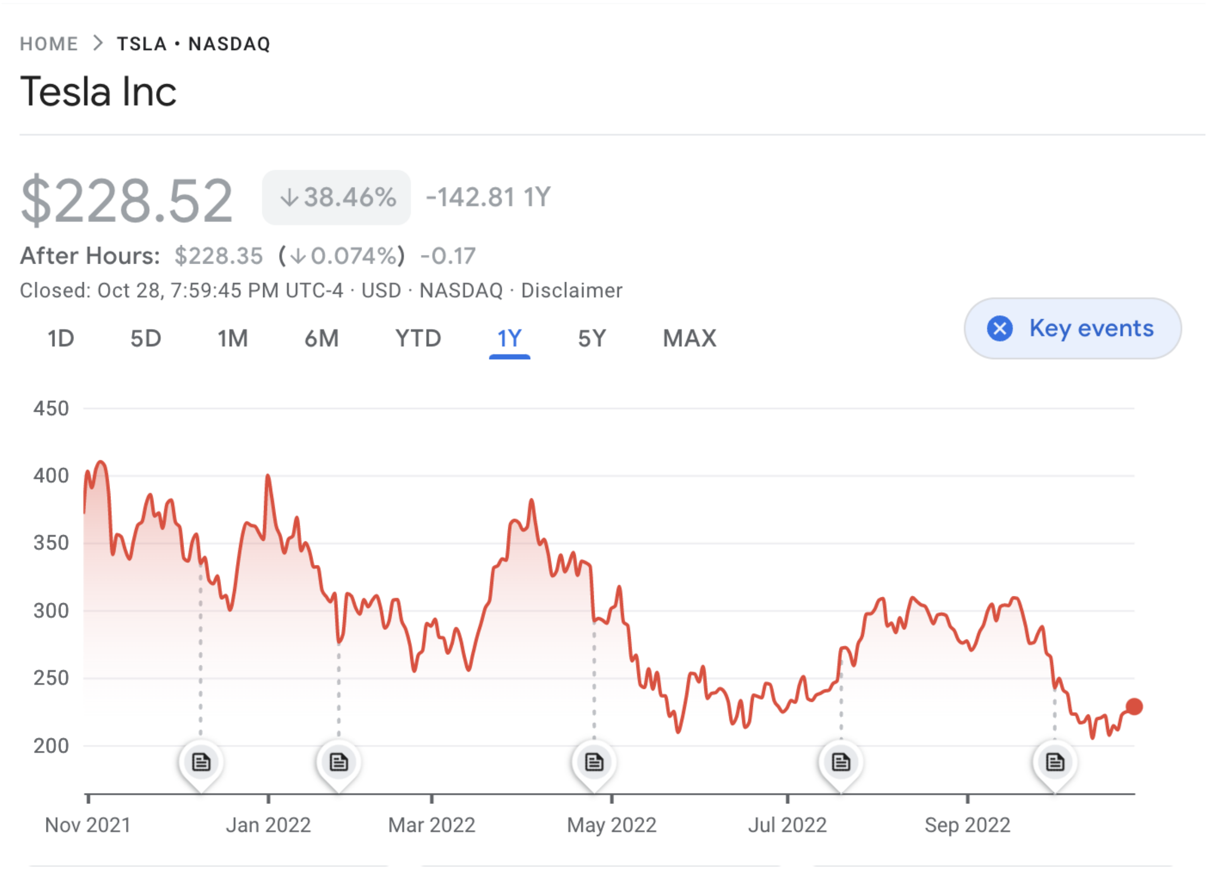
Discrete data: This type of data is collected at specific points in time and can be represented by a dot on a graph. For example, data from a survey would be considered discrete data.

Continuous data is more common than discrete data since most real-world phenomena are continuous.

For example, the population of a country is a continuous variable that changes over time—it doesn’t jump from one value to another. In contrast, the results of an election are discrete since there are only a finite number of outcomes (e.g., candidate A wins, candidate B wins, etc.).

The Best Way to Visualize Time Series Data

Time series line graphs are the best way to visualize data that changes over time. This is because line graphs show how a variable changes from one point in time to another, making it easy to see trends and patterns.



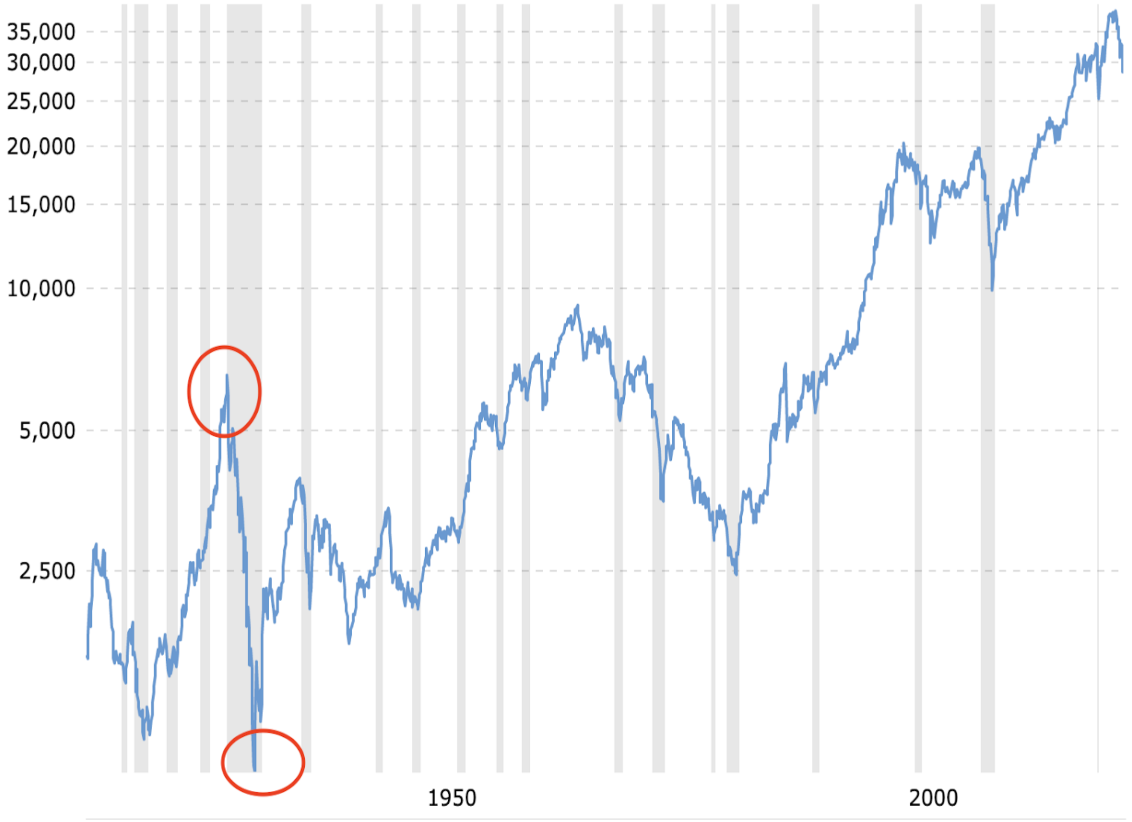
HOME > TSLA • NASDAQ

Tesla Inc

$228.52 ↓38.46% -142.81 1Y

After Hours: $228.35 (↓0.074%) -0.17

Closed: Oct 28, 7:59:45 PM UTC-4 · USD · NASDAQ · Disclaimer



Line graphs are also useful for identifying specific points in time when there was a sudden change in the data (known as an anomaly).

When to Use Other Temporal Visualizations

Other types of graphs can be used to visualize time series data, but they are less common.

For example, you could use a scatter plot to visualize how two variables are related. This would happen in cases where you have multivariate time series data.

For example, consider the following scatter plot of stock prices and interest rates:

The scatter plot shows that there is a positive relationship between stock prices and interest rates—as one variable increases, the other variable also tends to increase.

This relationship would be much harder to see if the data were presented in a line graph.

Another example would be a bar graph, which could be used to compare different time periods. For example, you could use a bar graph to compare the stock prices of two different companies:

The bar graph makes it easy to see that Company A's stock price is higher than Company B's stock price.

However, the bar graph is not as useful for seeing trends over time since it doesn't show how the data changes from one point in time to another. Best Platforms to Visualize Data

There are a few different platforms that businesses and data scientists use to visualize data:

Tableau

R

Microsoft Power BI

Excel

Python

Seven Time Series Data Visualization Examples

Line Graphs

Bar Graphs

Gantt Charts

Heat Maps

Time Series Data Visualization using Python

We will use Python libraries for visualizing the data.

Importing the Libraries

We will import all the libraries that we will be using throughout this article in one place so that do not have to import every time we use it this will save both our time and effort.

- **Numpy** – A Python library that is used for numerical mathematical computation and handling multidimensional ndarray, it also has a very large collection of mathematical functions to operate on this array.
- **Pandas** – A Python library built on top of NumPy for effective matrix multiplication and dataframe manipulation, it is also used for data cleaning, data merging, data reshaping, and data aggregation.
- **Matplotlib** – It is used for plotting 2D and 3D visualization plots, it also supports a variety of output formats including graphs for data.

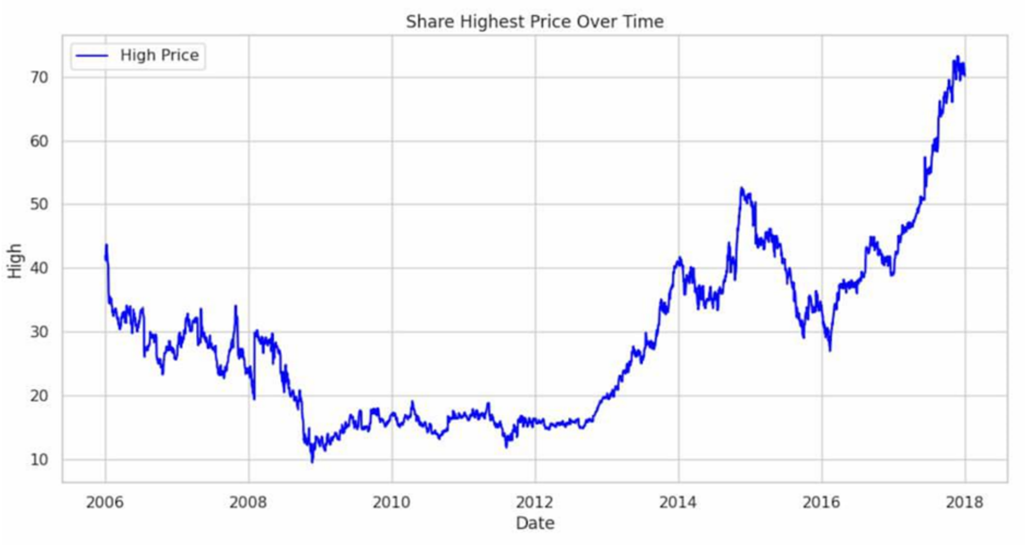
```
import pandas as pd

import numpy as np

import seaborn as sns

import matplotlib.pyplot as plt

from statsmodels.graphics.tsaplots import plot_acf

from statsmodels.tsa.stattools import adfuller
```

Loading The Dataset

To load the dataset into a dataframe we will use the pandas read_csv() function. We will use head() function to print the first five rows of the dataset. Here we will use the 'parse_dates' parameter in the read_csv function to convert the 'Date' column to the DatetimeIndex format. By default, Dates are stored in string format which is not the right format for time series data analysis.

```
# reading the dataset using read_csv

df = pd.read_csv("stock_data.csv",

                 parse_dates=True,

                 index_col="Date")


# displaying the first five rows of dataset

df.head()
```

OutPut:



Result :

Thus the Time series analysis for visualization using python has been executed successfully.