| | **Marwadi University** <br> **Faculty of Engineering and Technology** <br> **Department of Information and Communication Technology** | | |
|---|---|---|---|
| **Subject: Data Visualization and Dashboard (01CT0410)** | **Aim: Draw the scatter plot using plotly, matplotlib and seaborn libraries in python on random dataset.** | | |
| **Experiment No: 07** | **Date:12-02-2024** | **Enrollment No: 92200133030** | |

**Aim: Draw the scatter plot using plotly, matplotlib and seaborn libraries in python on random dataset.**

**IDE:** Excel, Tableau, Spyder IDE

**Theory:**

**Scatter plot**

Scatter and bubble charts are a great way to display a lot of related data in one chart. In scatter charts, the x-axis displays one numeric field and the y-axis displays another, making it easy to see the relationship between the two values for all the items in the chart.

In a scatter graph, both horizontal and vertical axes are value axes that plot numeric data. Typically, the independent variable is on the x-axis, and the dependent variable on the y-axis. The chart displays values at the intersection of an x and y axis, combined into single data points.

The main purpose of a scatter plot is to show how strong the relationship, or correlation, between the two variables is. The tighter the data points fall along a straight line, the higher the correlation.
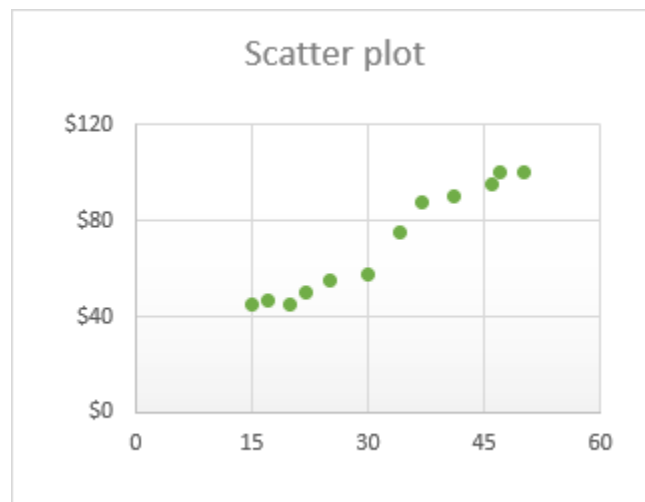


**How to arrange data for a scatter chart**

For ease of use, the independent variable should be in the left column as this column is going to be plotted on the x axis. The dependent variable (the one affected by the independent variable) should be in the right column, and it will be plotted on the y axis.

In our example, we are going to visualize the relationship between the advertising budget for a certain month (independent variable) and the number of items sold (dependent variable), so we arrange the data accordingly:

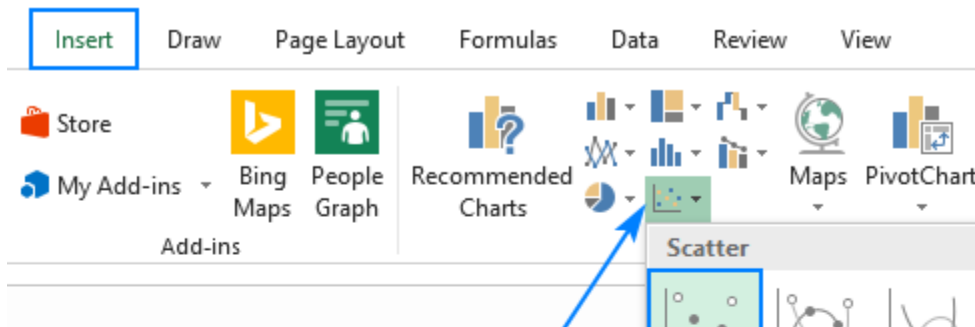| | Marwadi University |
| :---: | :--- |
| **Marwadi University** | **Faculty of Engineering and Technology** |
| | **Department of Information and Communication Technology** |
| **Subject: Data Visualization and Dashboard (01CT0410)** | **Aim: Draw the scatter plot using plotly, matplotlib and seaborn libraries in python on random dataset.** |
| **Experiment No: 07** | **Date:12-02-2024** | **Enrollment No: 92200133030** |

| | B | C | D |
| :--- | :--- | ---: | ---: |
| | **Month** | **Advertising** | **Items sold** |
| | Jan | $45 | 15 |
| | Feb | $55 | 25 |
| | Mar | $47 | 17 |
| | Apr | $75 | 34 |
| | May | $90 | 41 |
| | Jun | $100 | 47 |
| | Jul | $100 | 50 |
| | Aug | $95 | 46 |
| | Sep | $88 | 37 |
| | Oct | $50 | 22 |
| | Nov | $45 | 20 |
| | Dec | $58 | 30 |

**How to create a scatter plot in Excel**

With the source data correctly organized, making a scatter plot in Excel takes these two quick steps:

1. Select two columns with numeric data, including the column headers. In our case, it is the range C1:D13. Do not select any other columns to avoid confusing Excel.

2. Go to the *Inset* tab > *Chats* group, click the **Scatter** chart icon, and select the desired template. To insert a classic scatter graph, click the first thumbnail:

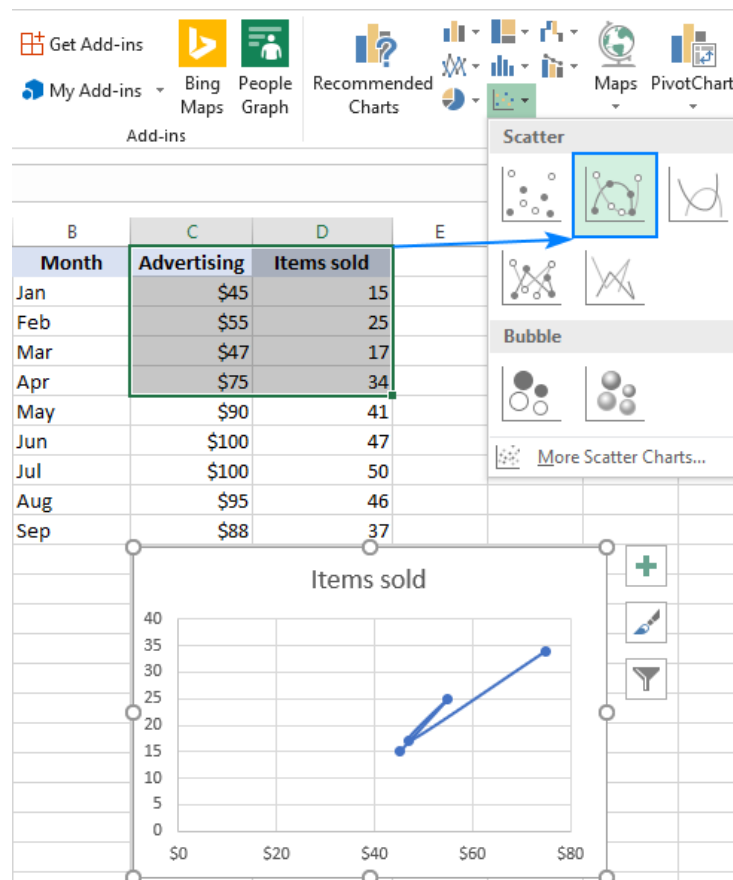|  | **Marwadi University** |
|---|---|
| | **Faculty of Engineering and Technology** |
| | **Department of Information and Communication Technology** |
| **Subject: Data Visualization and Dashboard (01CT0410)** | **Aim: Draw the scatter plot using plotly, matplotlib and seaborn libraries in python on random dataset.** |
| **Experiment No: 07** | **Date:12-02-2024** | **Enrollment No: 92200133030** |



The scatter diagram will be immediately inserted in your worksheet:
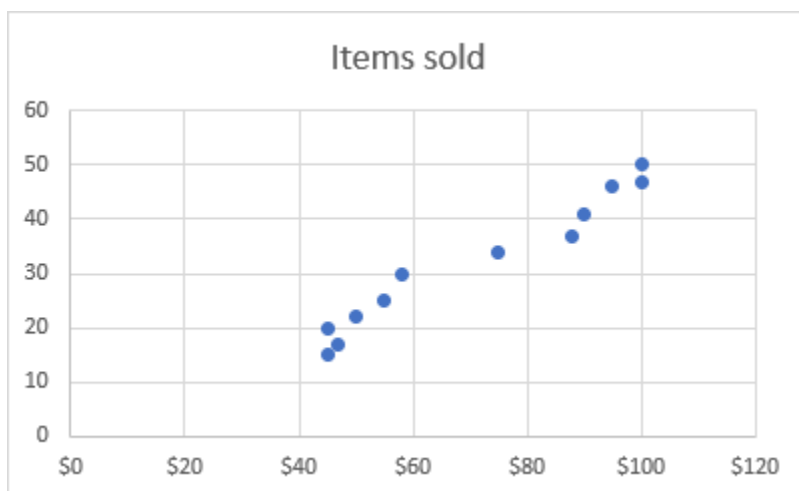


Basically, you may consider the work done. Or, you can customize some elements of your graph to make it look more beautiful and to convey the correlation between the two variables clearer.

| | Marwadi University |
|---|---|
| ![Marwadi University logo] | **Marwadi University**<br>**Faculty of Engineering and Technology**<br>**Department of Information and Communication Technology** |
| **Subject: Data Visualization and Dashboard (01CT0410)** | **Aim: Draw the scatter plot using plotly, matplotlib and seaborn libraries in python on random dataset.** |
| **Experiment No: 07** | **Date:12-02-2024** | **Enrollment No: 92200133030** |



**Scatter chart types**

Besides the classic scatter plot shown in the above example, a few more templates are available:
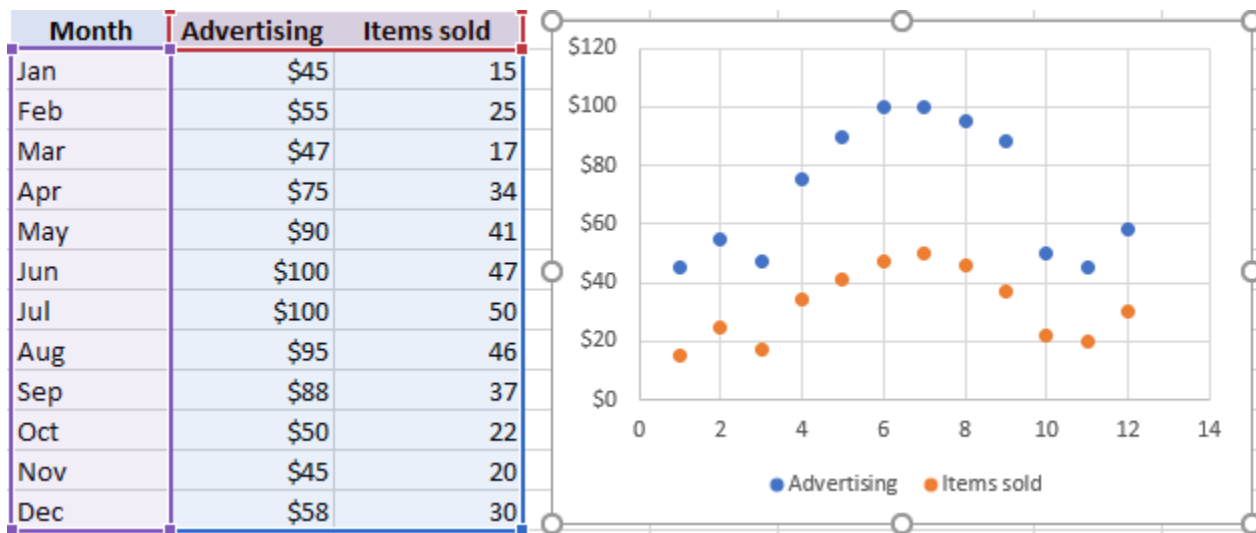
- Scatter with smooth lines and markers
- Scatter with smooth lines
- Scatter with straight lines and markers
- Scatter with straight lines

**Scatter with lines** is best to be used when you have few data points. For example, here's how you can represent the data for the first four months by using the scatter graph with smooth lines and markers:
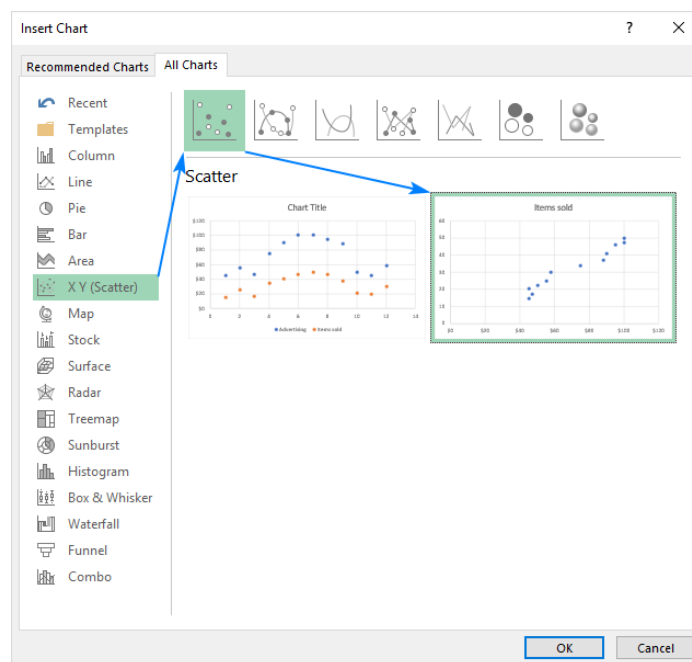
The Excel XY plot templates can also draw **each variable separately**, presenting the same relationships in a different way. For this, you should select 3 columns with data - the leftmost column with text values (labels), and the two columns with numbers.

| | Marwadi University Faculty of Engineering and Technology Department of Information and Communication Technology | | |
|---|---|---|---|
| Subject: Data Visualization and Dashboard (01CT0410) | Aim: Draw the scatter plot using plotly, matplotlib and seaborn libraries in python on random dataset. | | |
| Experiment No: 07 | Date:12-02-2024 | Enrollment No: 92200133030 | |

In our example, the blue dots represent advertising cost, and orange dots represent the items sold:

| Month | Advertising | Items sold |
|---|---|---|
| Jan | $45 | 15 |
| Feb | $55 | 25 |
| Mar | $47 | 17 |
| Apr | $75 | 34 |
| May | $90 | 41 |
| Jun | $100 | 47 |
| Jul | $100 | 50 |
| Aug | $95 | 46 |
| Sep | $88 | 37 |
| Oct | $50 | 22 |
| Nov | $45 | 20 |
| Dec | $58 | 30 |

To view all available scatter types in one place, select your data, click the Scatter (X, Y) icon on the ribbon, and then click More Scatter Charts… This will open the Inset Chart dialog box with the XY (Scatter) type selected, and you switch between the different templates at the top to see which one provides the best graphic representation of your data:

| ![Marwadi University logo] | **Marwadi University** |
|---|---|
| | **Faculty of Engineering and Technology** |
| | **Department of Information and Communication Technology** |
| **Subject: Data Visualization and Dashboard (01CT0410)** | **Aim: Draw the scatter plot using plotly, matplotlib and seaborn libraries in python on random dataset.** |
| **Experiment No: 07** | **Date:12-02-2024** | **Enrollment No: 92200133030** |

**Pre Lab Exercise:**

a. What is scatter plot used for?

_____

_____

_____

_____

_____

_____

b. What type of data is a scatter plot?

_____

_____
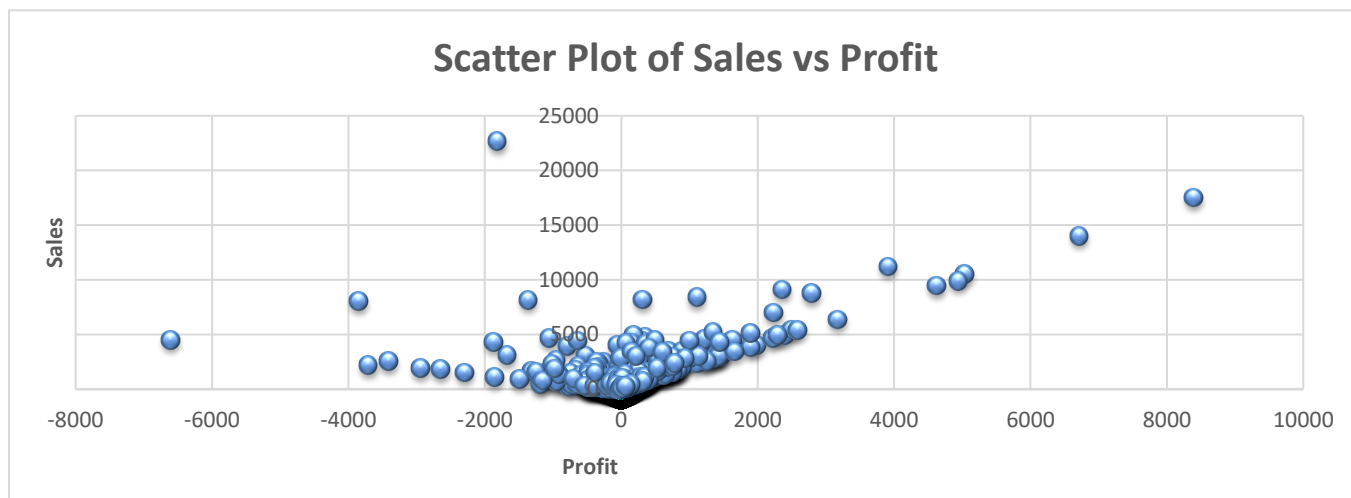
_____

_____

_____

_____

c. What is the XY scatter plot?

_____

_____

_____

**Pre-Lab Tasks:-**
Perform the following tasks using Excel:

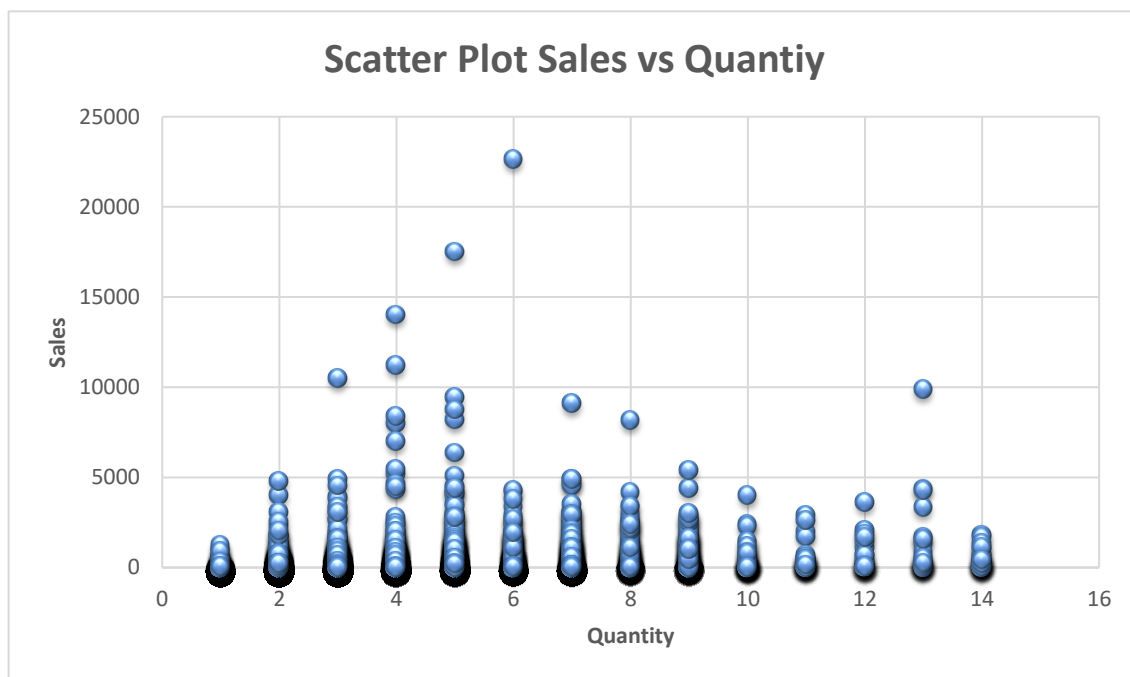Task 1: Create a scatterplot to compare sales to profit in the Superstore dataset:
**Results:-**

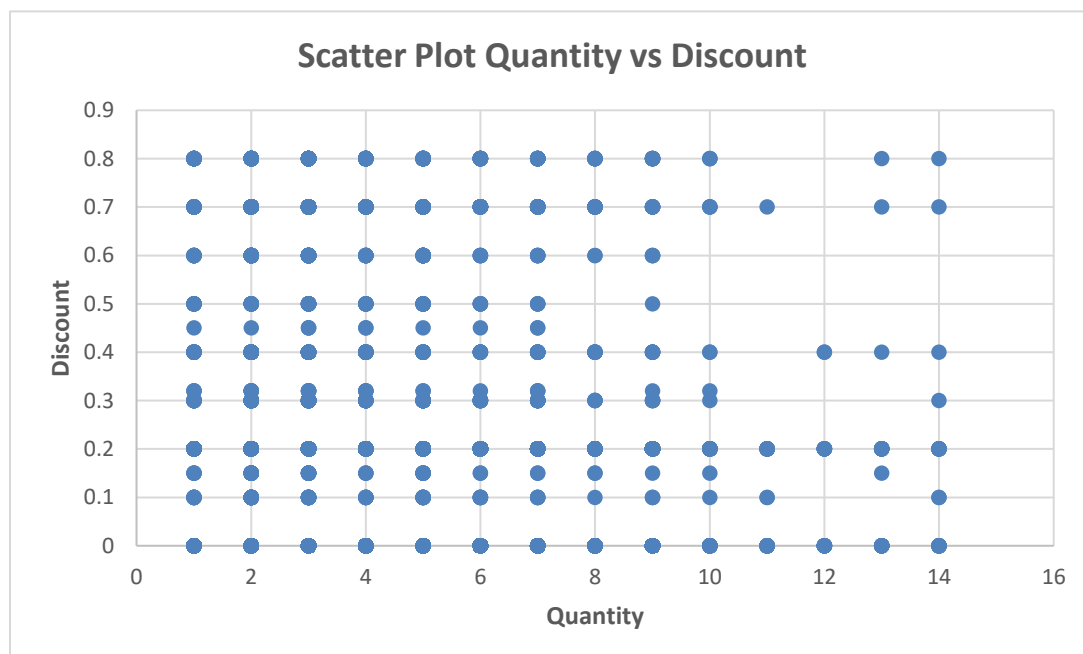| ![Marwadi University Logo] | **Marwadi University** |
| | **Faculty of Engineering and Technology** |
| | **Department of Information and Communication Technology** |
| **Subject: Data Visualization and Dashboard (01CT0410)** | **Aim: Draw the scatter plot using plotly, matplotlib and seaborn libraries in python on random dataset.** |
| **Experiment No: 07** | **Date:12-02-2024** | **Enrollment No: 92200133030** |

Task 2: Create a scatterplot to compare sales to quantity in the Superstore dataset:
**Results:-**



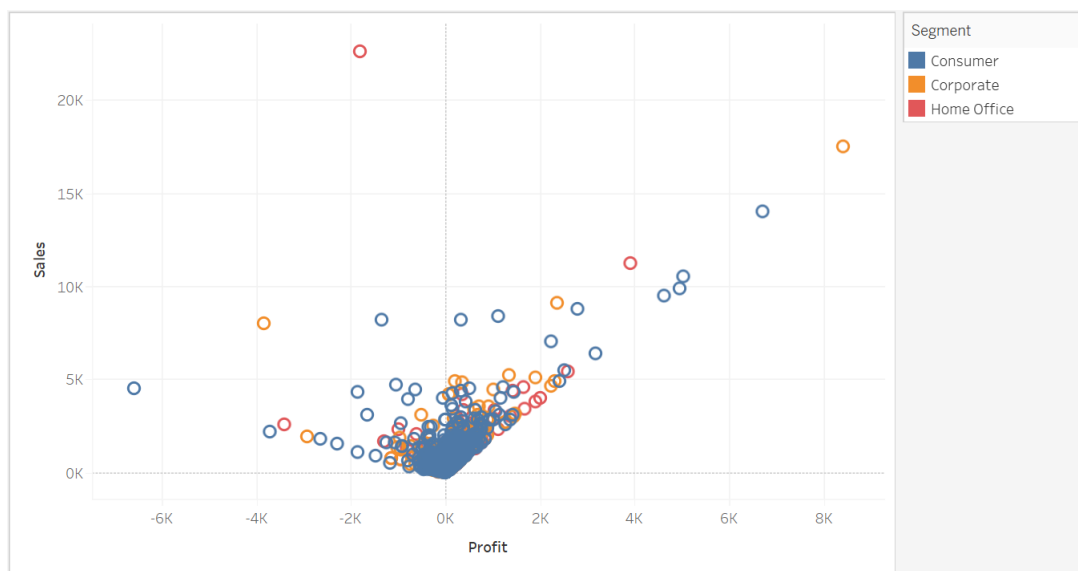Task 3: Create a scatterplot to compare Quantity to discount in the Superstore dataset:
**Results:-**

|  | **Marwadi University** |
| --- | --- |
| | **Faculty of Engineering and Technology** |
| | **Department of Information and Communication Technology** |
| **Subject: Data Visualization and Dashboard (01CT0410)** | **Aim: Draw the scatter plot using plotly, matplotlib and seaborn libraries in python on random dataset.** |
| **Experiment No: 07** | **Date:12-02-2024** | **Enrollment No: 92200133030** |

**In-Lab Tasks:-**

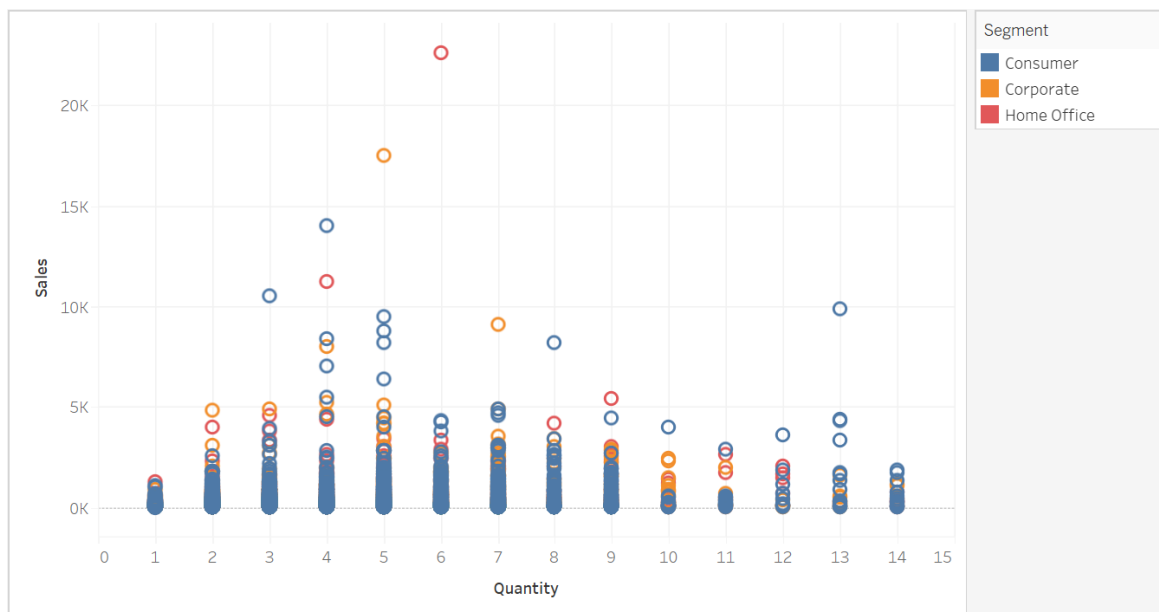Perform the following tasks using Tableau:

Task 1: Create a scatterplot to compare sales to profit in the Superstore dataset:

**Results:-**



Task 2: Create a scatterplot to compare sales to quantity in the Superstore dataset:

**Results:-**

| ![Marwadi University Logo] | **Marwadi University** |
|---|---|
| | **Faculty of Engineering and Technology** |
| | **Department of Information and Communication Technology** |
| **Subject: Data Visualization and Dashboard (01CT0410)** | **Aim: Draw the scatter plot using plotly, matplotlib and seaborn libraries in python on random dataset.** |
| **Experiment No: 07** | **Date:12-02-2024** | **Enrollment No: 92200133030** |

Task 3:  Create a scatterplot to compare Quantity to discount in the Superstore dataset:
**Results:-**
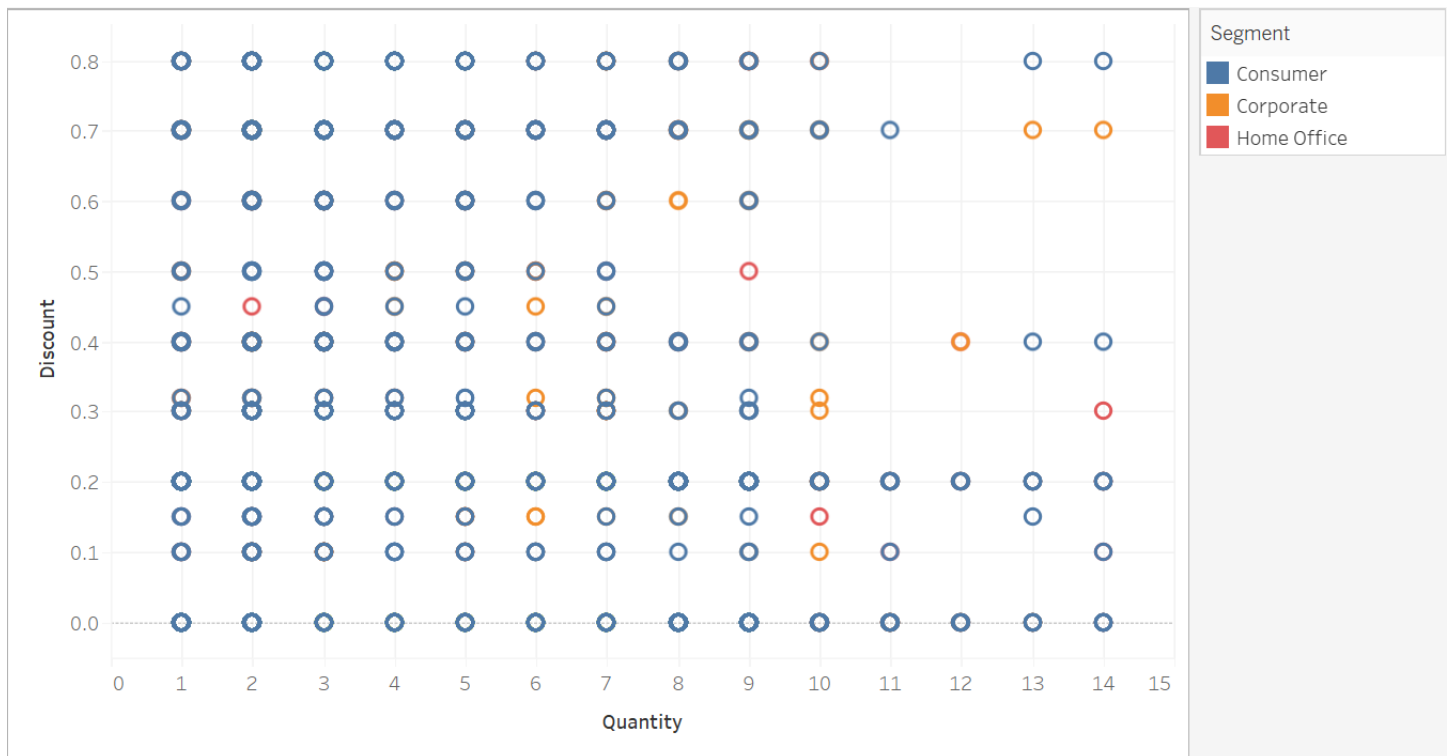


**Post-Lab Tasks:-**
Perform the following tasks using Python:
**Pre-Requisites :-**
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

Dataset = pd.read_excel( "D:/Aryan Data/Usefull Data/Semester - 4/Data Visulization and Dashboards/Lab Manual/Exp-7 Creating Scatter plot/Sample - Superstore.xlsx",'Orders')

| | **Marwadi University** |
| :---: | :--- |
| ![Marwadi University Logo] | **Faculty of Engineering and Technology** |
| | **Department of Information and Communication Technology** |
| **Subject: Data Visualization and Dashboard (01CT0410)** | **Aim: Draw the scatter plot using plotly, matplotlib and seaborn libraries in python on random dataset.** |
| **Experiment No: 07** | **Date:12-02-2024** | **Enrollment No: 92200133030** |

Task 1: Create a scatterplot to compare sales to profit in the Superstore dataset:

**Code:-**

```
plt.figure(figsize=(16,8))
sns.scatterplot(data=Dataset, x=Dataset['Profit'] , y=Dataset['Sales'], hue=Dataset['Segment'])
plt.title("Scatter Plot of Superstore Sales vs Profit")
plt.xlabel("Sales")
plt.ylabel("Profit")
plt.legend()
plt.show()
```
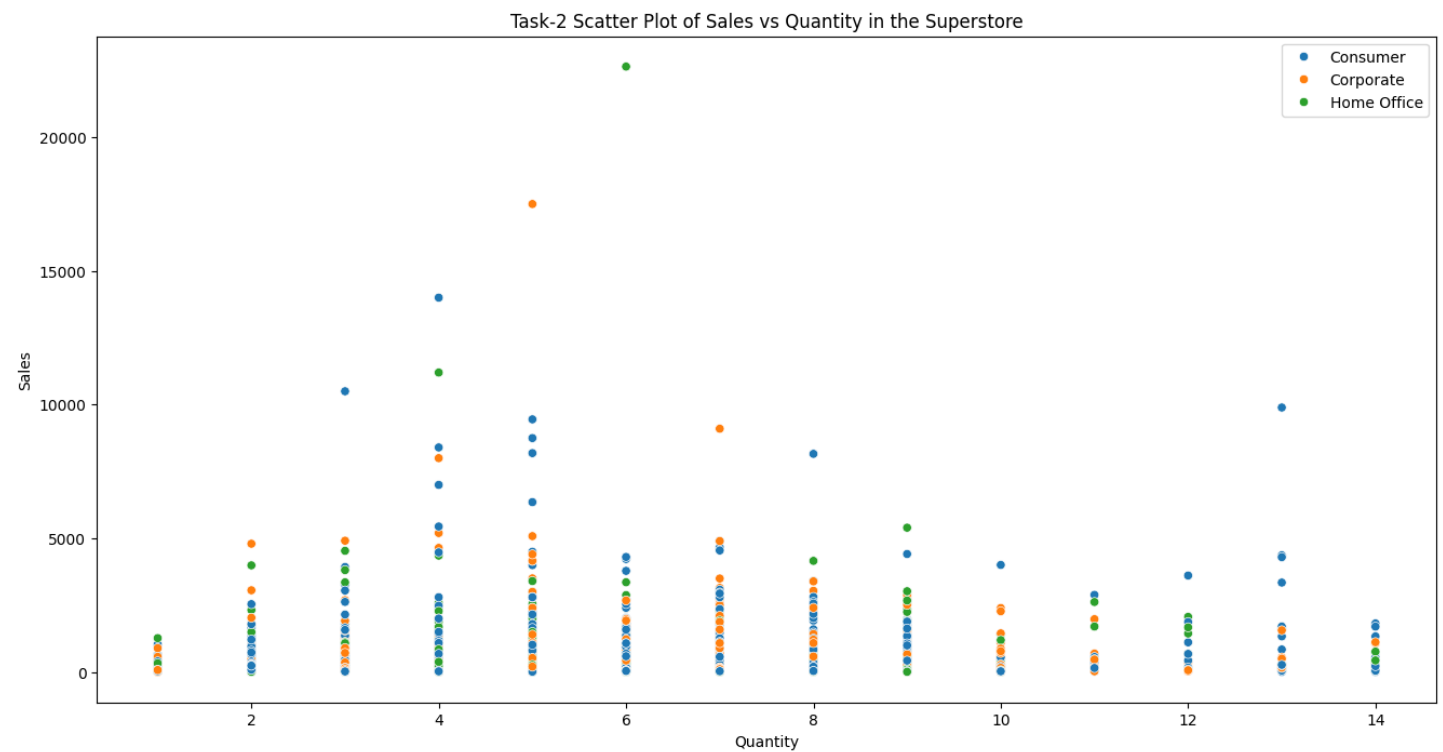
**Results:-**
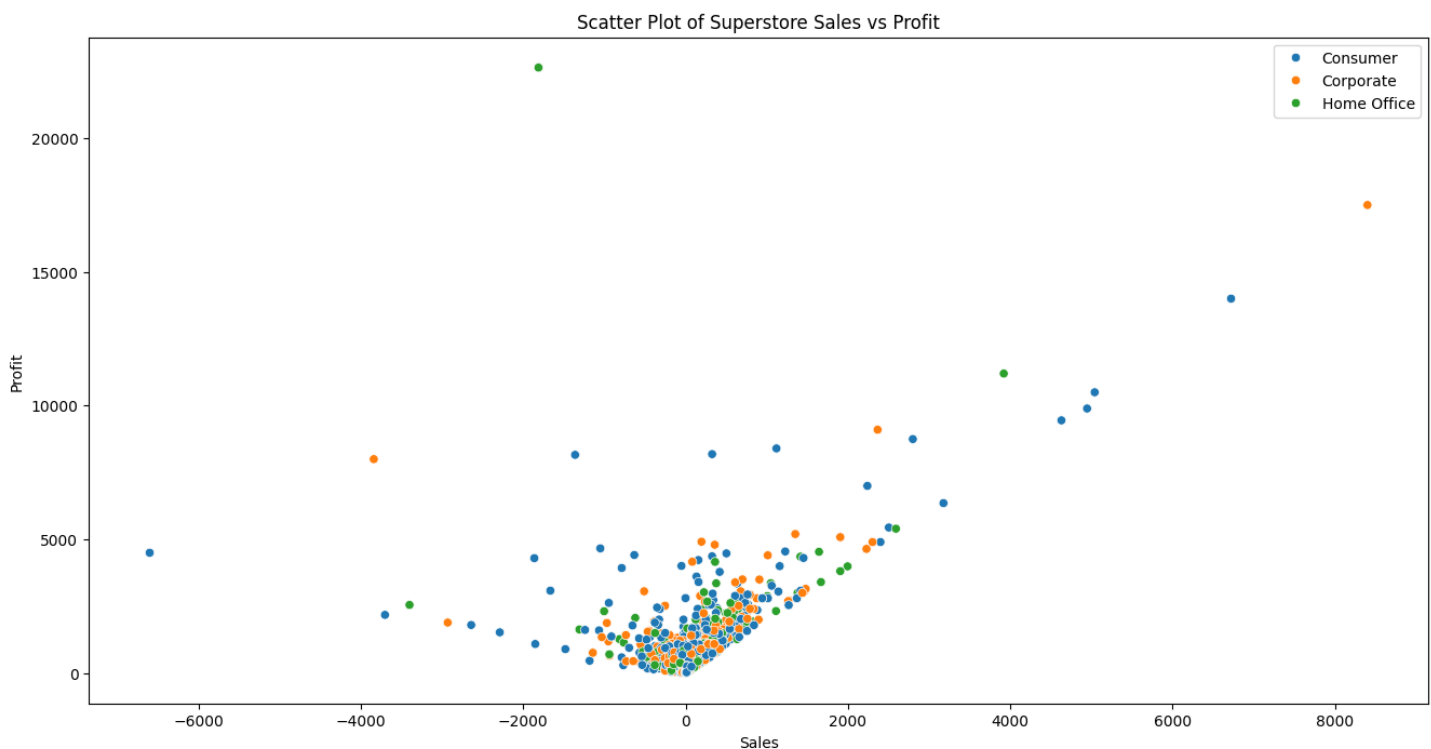


Task 2: Create a scatterplot to compare sales to quantity in the Superstore dataset:

**Code:-**

```
plt.figure(figsize=(16, 8))
sns.scatterplot(x=Dataset["Quantity"], y=Dataset["Sales"], data=Dataset, hue=Dataset["Segment"])
plt.title("Task-2 Scatter Plot of Sales vs Quantity in the Superstore")
plt.legend()
plt.show()
```

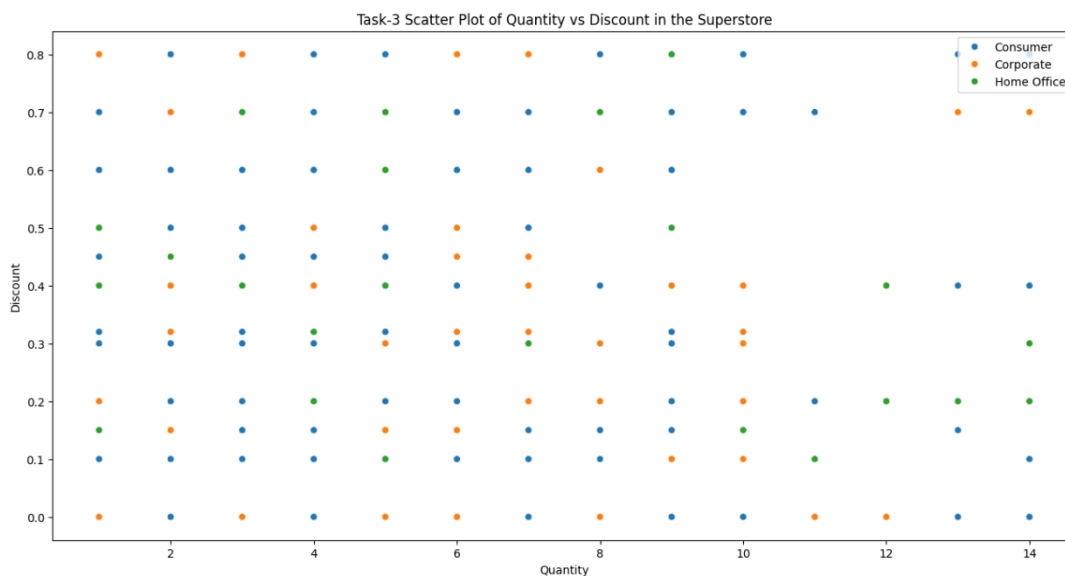|  | **Marwadi University** **Faculty of Engineering and Technology** **Department of Information and Communication Technology** |
|---|---|
| **Subject: Data Visualization and Dashboard (01CT0410)** | **Aim: Draw the scatter plot using plotly, matplotlib and seaborn libraries in python on random dataset.** |
| **Experiment No: 07** | **Date:12-02-2024** | **Enrollment No: 92200133030** |

**Results:-**



Task 3: Create a scatterplot to compare Quantity to discount in the Superstore dataset:

**Code:-**

```
plt.figure(figsize=(16, 8))
sns.scatterplot(x=Dataset["Quantity"], y=Dataset["Discount"], data=Dataset, hue=Dataset["Segment"])
plt.title("Task-3 Scatter Plot of Quantity vs Discount in the Superstore")
plt.legend()
plt.show()
```

| | **Marwadi University**<br>**Faculty of Engineering and Technology**<br>**Department of Information and Communication Technology** | |
| --- | --- | --- |
| **Subject: Data Visualization and Dashboard (01CT0410)** | **Aim: Draw the scatter plot using plotly, matplotlib and seaborn libraries in python on random dataset.** | |
| **Experiment No: 07** | **Date:12-02-2024** | **Enrollment No: 92200133030** |

**Results:-**



Reference Link: http://www.science.smith.edu/~jcrouser/SDS136/labs/lab3/

**Observation and Result Analysis:**

Write the final observation and process corresponding to each task

1. _____
   _____
   _____

2. _____
   _____
   _____

3. _____
   _____
   _____

**Post Lab Exercise:**

**Python Implementation**

**Reference Link**https://www.kaggle.com/code/ashwinshetgaonkar/super-store-analysis-data-visual-seaborn