

2nd Mar Assignment

March 11, 2023

1 Assignment 26

Q1: What is Matplotlib? Why is it used? Name five plots that can be plotted using the Pyplot module of Matplotlib.

Ans. Matplotlib is a popular data visualization library in Python that allows developers to create various types of graphs, charts, and plots. It is widely used for data analysis, scientific research, and other visualization tasks.

Some of the main reasons why Matplotlib is used are:

1. It is a powerful tool for creating high-quality visualizations and plots.
2. It is open-source and freely available, which makes it accessible to a wide range of users.
3. It is highly customizable, allowing users to create plots with different styles, colors, and formatting options.
4. It integrates well with other Python libraries such as NumPy and Pandas, making it a versatile tool for data analysis.

Five plots that can be plotted using the Pyplot module of Matplotlib are:

1. Line plot: A line plot is a type of chart that displays data as a series of points connected by straight lines. It is commonly used to show trends over time.
2. Bar plot: A bar plot is a chart that displays data as a series of bars, with the height or length of each bar representing the data value. It is commonly used to compare data across different categories.
3. Scatter plot: A scatter plot is a chart that displays data as a collection of points, where each point represents the value of two different variables. It is commonly used to show the relationship between two variables.
4. Histogram: A histogram is a chart that displays data as a series of bars, with the width of each bar representing a range of data values. It is commonly used to show the distribution of a dataset.
5. Pie chart: A pie chart is a chart that displays data as a series of wedges, with the size of each wedge representing the proportion of the data value.

Q2: What is a scatter plot? Use the following code to generate data for x and y. Using this generated data plot a scatter plot.

```
import numpy as np
```

```
np.random.seed(3)
```

```
x = 3 + np.random.normal(0, 2, 50)
```

```
y = 3 + np.random.normal(0, 2, len(x))
```

Note: Also add title, xlabel, and ylabel to the plot.

Ans. Scatter plot: A scatter plot is a chart that displays data as a collection of points, where each point represents the value of two different variables. It is commonly used to show the relationship between two variables

```
[1]: import numpy as np
      np.random.seed(3)
      x = 3 + np.random.normal(0, 2, 50)
      y = 3 + np.random.normal(0, 2, len(x))
```

```
[2]: x
```

```
[2]: array([ 6.57725695,  3.8730197 ,  3.19299494, -0.72698541,  2.44522359,
            2.29048204,  2.83451704,  1.74599865,  2.91236366,  2.04556394,
            0.37227049,  4.76924476,  4.76263608,  6.41914613,  3.10006728,
            2.19064517,  1.9092801 , -0.09295463,  4.96473487,  0.79786474,
            0.62990695,  2.5887002 ,  5.97229671,  3.47343253,  0.95242972,
            1.5740136 ,  4.25048993,  2.67897327,  1.4623273 ,  2.53993856,
            4.49011253,  6.95222157,  0.51175334,  1.74716618,  1.39246781,
           -1.83816635,  1.15241596,  0.95224848,  5.24795592,  2.73617153,
           -0.24657089,  4.2933509 ,  2.28745848, -0.48628207,  1.80670072,
            1.82281124,  1.2522354 ,  3.05942763, -1.49651554,  2.46447627])
```

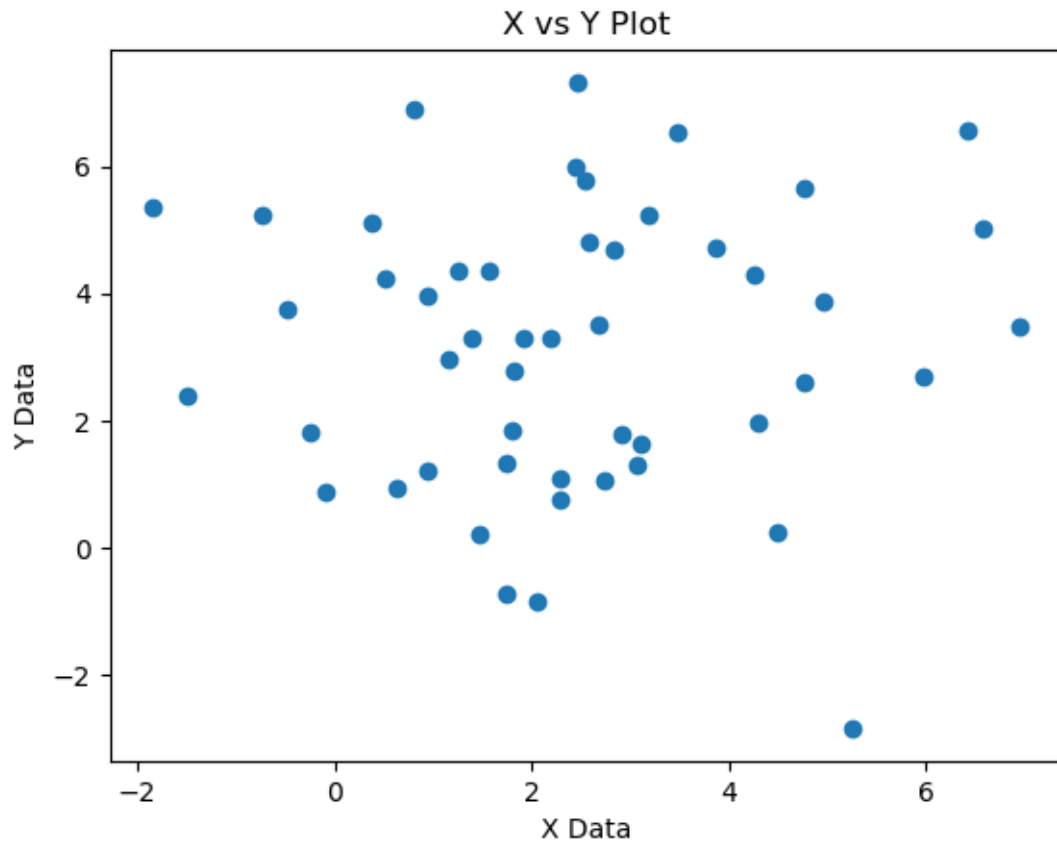
```
[3]: y
```

```
[3]: array([ 5.02636688,  4.70559568,  5.216375 ,  5.23878131,  5.97508626,
            0.76339863,  4.69166681, -0.72177906,  1.79422979, -0.82894409,
            5.09629502,  5.66747564,  2.60517064,  6.54929006,  1.65054498,
            3.30123373,  3.30589141,  0.87160945,  3.87589322,  6.87795692,
            0.95013825,  4.79867689,  2.6909863 ,  6.53925461,  3.9675767 ,
            4.3524328 ,  4.28632656,  3.49817341,  0.20847299,  5.78332582,
            0.25866197,  3.47712638,  4.22815418,  1.32417545,  3.29012643,
            5.33576457,  2.95179106,  1.22268516, -2.8314755 ,  1.05631899,
            1.81784252,  1.96716526,  1.08000764,  3.75459047,  1.85058316,
            2.78109133,  4.3581432 ,  1.28912566,  2.39958785,  7.31629868])
```

```
[4]: import matplotlib.pyplot as plt
```

```
[8]: plt.scatter(x,y)
plt.xlabel('X Data')
plt.ylabel('Y Data')
plt.title('X vs Y Plot')
```

```
[8]: Text(0.5, 1.0, 'X vs Y Plot')
```



Q3: Why is the subplot() function used? Draw four line plots using the subplot() function.

Use the following data:

```
import numpy as np
```

For line 1: `x = np.array([0, 1, 2, 3, 4, 5])` and `y = np.array([0, 100, 200, 300, 400, 500])`

For line 2: `x = np.array([0, 1, 2, 3, 4, 5])` and `y = np.array([50, 20, 40, 20, 60, 70])`

For line 3: `x = np.array([0, 1, 2, 3, 4, 5])` and `y = np.array([10, 20, 30, 40, 50, 60])`

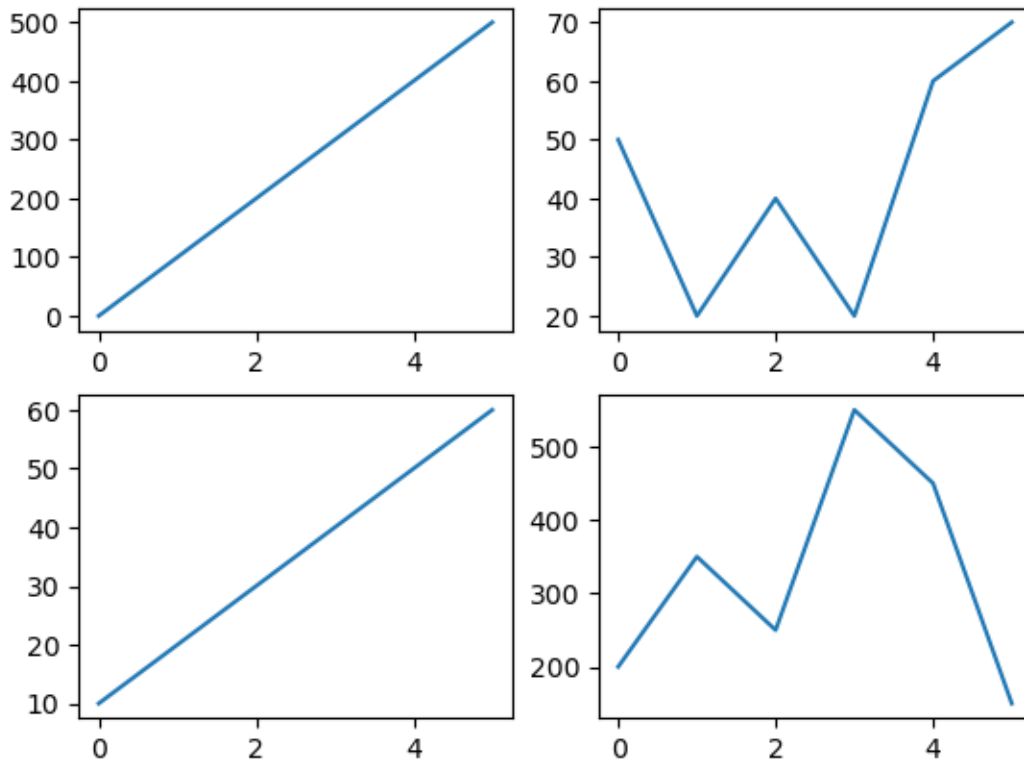
For line 4: `x = np.array([0, 1, 2, 3, 4, 5])` and `y = np.array([200, 350, 250, 550, 450, 150])`

Ans.Subplot is a python inbuilt function that is used to plot multiple graph on the same canvas. It mainly takes two arguments as rows and columns

```
[10]: import numpy as np
x1 = np.array([0, 1, 2, 3, 4, 5])
y1 = np.array([0, 100, 200, 300, 400, 500])
x2 = np.array([0, 1, 2, 3, 4, 5])
y2 = np.array([50, 20, 40, 20, 60, 70])
x3 = np.array([0, 1, 2, 3, 4, 5])
y3 = np.array([10, 20, 30, 40, 50, 60])
x4 = np.array([0, 1, 2, 3, 4, 5])
y4 = np.array([200, 350, 250, 550, 450, 150])
```

```
[14]: import matplotlib.pyplot as plt
fig, axes = plt.subplots(nrows=2,ncols=2)
axes[0,0].plot(x1,y1)
axes[0,1].plot(x2,y2)
axes[1,0].plot(x3,y3)
axes[1,1].plot(x4,y4)
```

```
[14]: [<matplotlib.lines.Line2D at 0x7f2919b1d360>]
```



Q4: What is a bar plot? Why is it used? Using the following data plot a bar plot and a horizontal bar plot.

```
import numpy as np
```

```
company = np.array(["Apple", "Microsoft", "Google", "AMD"])
```

```
profit = np.array([3000, 8000, 1000, 10000])
```

Ans.Bar plot: A bar plot is a chart that displays data as a series of bars, with the height or length of each bar representing the data value. It is commonly used to compare data across different categories.

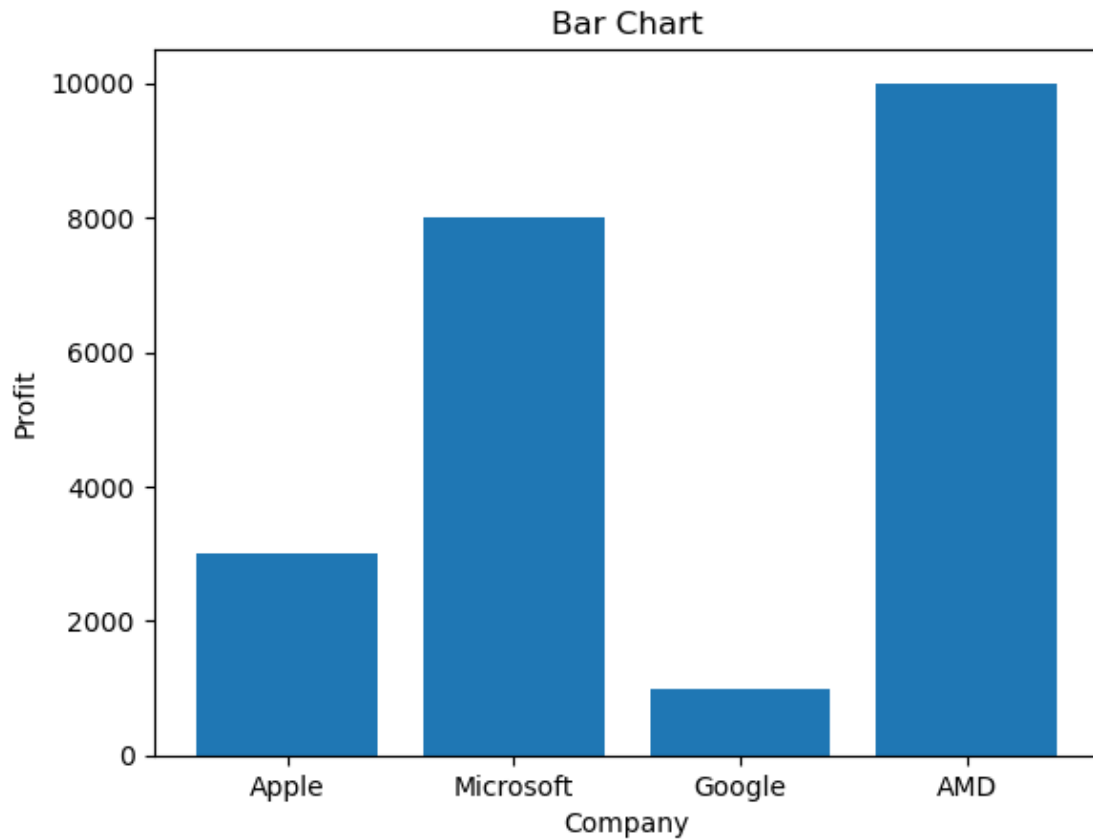
Uses - A bar chart is used when you want to show a distribution of data points or perform a comparison of metric values across different subgroups of your data.

```
[15]: import numpy as np
      company = np.array(["Apple", "Microsoft", "Google", "AMD"])
      profit = np.array([3000, 8000, 1000, 10000])
```

```
[17]: import matplotlib.pyplot as plt
      plt.bar(company,profit)
```

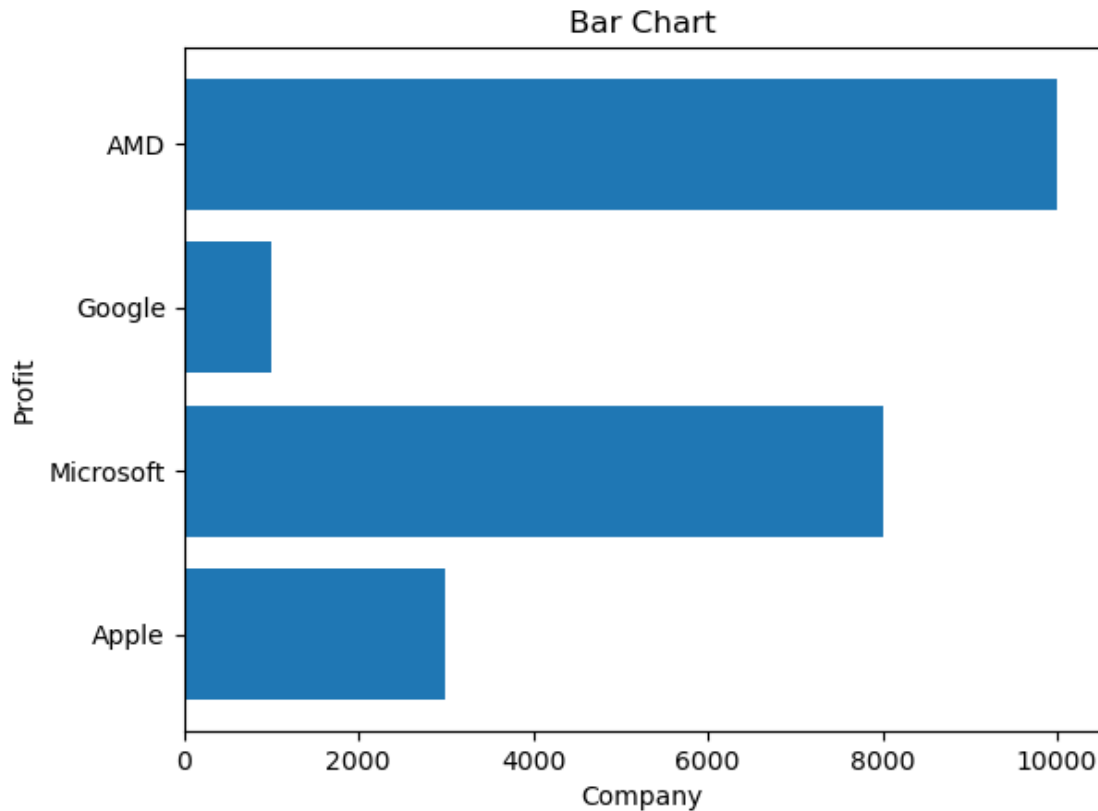
```
plt.xlabel('Company')
plt.ylabel('Profit')
plt.title('Bar Chart')
```

```
[17]: Text(0.5, 1.0, 'Bar Chart')
```



```
[18]: import matplotlib.pyplot as plt
plt.barh(company,profit)
plt.xlabel('Company')
plt.ylabel('Profit')
plt.title('Bar Chart')
```

```
[18]: Text(0.5, 1.0, 'Bar Chart')
```



Q5: What is a box plot? Why is it used? Using the following data plot a box plot.

```
box1 = np.random.normal(100, 10, 200)
```

```
box2 = np.random.normal(90, 20, 200)
```

Ans. A box plot, also known as a box-and-whisker plot, is a graphical representation of numerical data through their quartiles. It consists of a rectangular box that represents the middle 50% of the data, with a vertical line inside the box representing the median.

Box plots are used to summarize and compare distributions of data, particularly when comparing multiple groups. They provide a visual representation of the median, spread, and skewness of the data.

```
[19]: box1 = np.random.normal(100, 10, 200)
      box2 = np.random.normal(90, 20, 200)
```

```
[21]: plt.boxplot([box1, box2])
```

```
[21]: {'whiskers': [<matplotlib.lines.Line2D at 0x7f29186fe9b0>,
                  <matplotlib.lines.Line2D at 0x7f29186fec50>,
```

```

<matplotlib.lines.Line2D at 0x7f29186ff7f0>,
<matplotlib.lines.Line2D at 0x7f29186ffa90>],
'caps': [<matplotlib.lines.Line2D at 0x7f29186feef0>,
<matplotlib.lines.Line2D at 0x7f29186ff190>,
<matplotlib.lines.Line2D at 0x7f29186ffd30>,
<matplotlib.lines.Line2D at 0x7f29186fffd0>],
'boxes': [<matplotlib.lines.Line2D at 0x7f29186fe710>,
<matplotlib.lines.Line2D at 0x7f29190d2f50>],
'medians': [<matplotlib.lines.Line2D at 0x7f29186ff430>,
<matplotlib.lines.Line2D at 0x7f29187302b0>],
'fliers': [<matplotlib.lines.Line2D at 0x7f29186ff6d0>,
<matplotlib.lines.Line2D at 0x7f2918730550>],
'means': []}

```

