Data Visualization Using Matplotlib (DAY-4)

Matplotlib is a powerfull plotting library for python that enables the creation of static, animated, and interactive visualizations.

* ! pip install matplotlib
* import matplotlib.pyplot as plt
* x = [1,2,3,4,5]
  y = [1,4,9,16,25]
  plt.plot(x,y)

then the plot will be created

if we want to add the labels

```
plt.xlabel('x axis')
plt.ylabel('y axis')
plt.title("Basic Line plot")
plt.show()
```

* Create a Customized line plot

```
x = [1,2,3,4,5]

y = [1,4,9,16,25]

plt.plot(x, y, color = 'red', linestyle = '--',
         marker = 'o', linewidth = 3, markersize = 9)

plt.grid(True)
```

* Multiple plots :

```
x = [1,2,3,4,5]

y1 = [1,4,9,16,25]

y2 = [1,2,3,4,5]

plt.figure(figsize = (9,5))

plt.subplot(2,2,1)

plt.plot(x,y1, color = 'green')

plt.title("plot 1")

plt.subplot(2,2,2)

plt.plot(y1, x, color = 'red')

plt.title("plot 2")

plt.subplot(2,2,3)

plt.plot(x,y2, color = 'blue')

plt.title("plot 3")
```

* Create a bar Graph

```
Categories = ['A', 'B', 'C', 'D', 'E']
Values = [5, 7, 3, 8, 6]
plt.bar (Categories, Values, color = 'purple')
plt.xlabel ('Categories')
plt.ylabel ('Values')
plt.title ('Bar plot')
plt.show ()
```

Histograms :-

↳ They are used to represent the distribution of a dataset,

→ They divide the data into bins & count the no. of data points in each bin

* `data = [1,2,2,3,3,3,4,4,4,4,5,5,5,5,5]`

`plt.hist(data, color = "Orange", edgecolor = "black")`

* Create a scatter plot

```
x = [1, 2, 3, 4, 5]

y = [2, 3, 4, 5, 6]

plt.scatter(x, y, color = "blue", marker = 'x')
```

* Create a pie chart

```
labels = ['A', 'B', 'C', 'D']

Sizes = [30, 20, 40, 10]

Colors = ['gold', 'yellowgreen', 'lightcoral', 'lightskyblue']

explode = (0.2, 0, 0, 0)

plt.pie(Sizes, explode = explode, label = labels, colors = colors,
        autopct = '%1.1f%%', shadow = True)
```

* Sales Data Visualization

```
import pandas as pd

Sales_data_df = pd.read_csv('filename.csv')

Sales_data_df.head(5)

Sales_data_df.info()
```

total - sales by - product = sales - data. df . groupby ['Product Category']
['Total Revenue'] . sum()

Print (total - sales - by - product)

total_sales - by - product . plot (kind = 'bar' , color = 'red')

* plot sales trend over time.

Sales_trend . = sales - data.df . groupby ('Data') ['Total Revenue']
. sum () . reset_index ()

Plt.plot (Sales_trend ['Date'], Sales_trend ['Total Revenue'])