

# Day02 - Session 3 - Data Visualization using Matplotlib

## Topics

1. Matplotlib

## Note:

1. First Clean the Environment (Go to "Kernel" Menu --> "Restart & Clean Output")
2. To execute the code --> Click on a cell and press ctrl + enter key

## ▼ Data Visualization

1. Graphical representation of information and Data
2. Graph, Chart, and Map make it easier to understand the patterns in Data

## Matplotlib

Matplotlib is an amazing visualization library in Python for 2D plots of arrays.

## To Save image on your Machine

```
plt.savefig("path\name")
```

## ▼ Import matplotlib

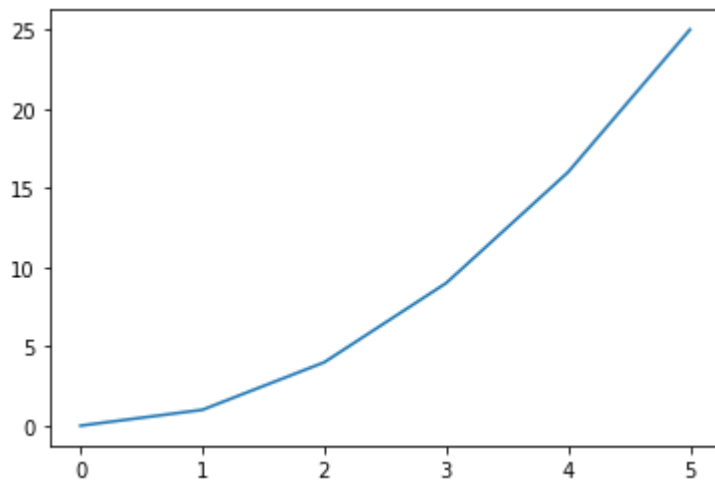
```
import matplotlib.pyplot as plt
```

## ▼ 1 Line Plot

### ▼ 1.1 Plot a Line

```
x_values = [0,1,2,3,4,5]
squares = [0,1,4,9,16,25]
plt.plot(x_values,squares)

plt.savefig("testimage.jpg")
```



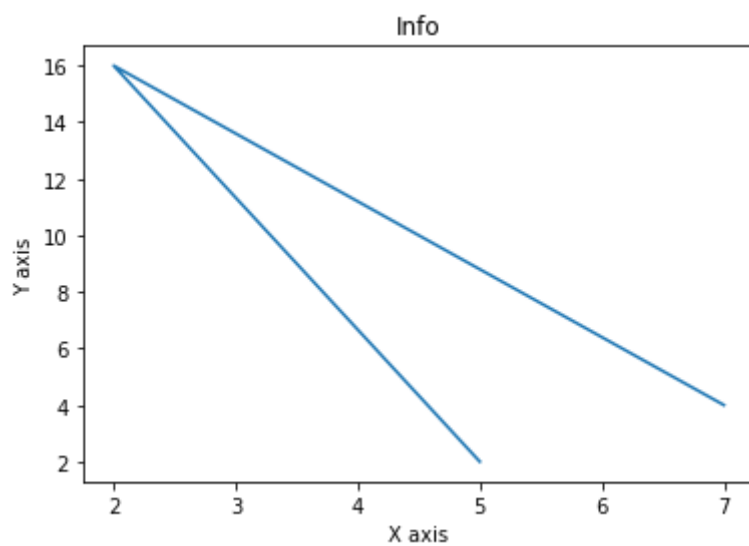
## ▼ 1.2 Plot a Line with Details

```
x = [5,2,7]
y = [2,16,4]
```

```
plt.plot(x,y)
```

```
plt.title('Info')
plt.ylabel('Y axis')
plt.xlabel('X axis')
```

```
plt.show()
```



## ▼ 1.3 Plot Multiple Lines

```
x = [1,2,3,4,5]
y = [50,40,70,80,20]
v2 = [80,20,20,50,60]
```

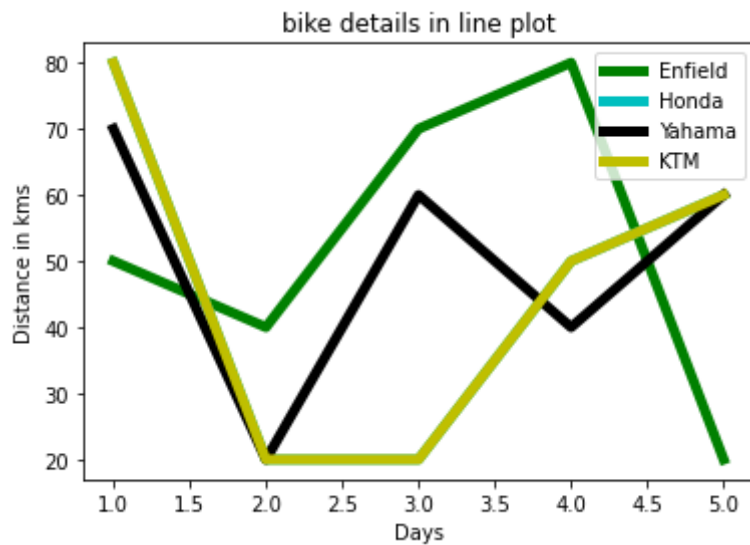
```

x = [1,2,3,4,5]
y3 = [70,20,60,40,60]
y4 = [80,20,20,50,60]

plt.plot(x,y,'g',label='Enfield', linewidth=5)
plt.plot(x,y2,'c',label='Honda',linewidth=5)
plt.plot(x,y3,'k',label='Yahama',linewidth=5)
plt.plot(x,y4,'y',label='KTM',linewidth=5)
plt.title('bike details in line plot')
plt.ylabel('Distance in kms')
plt.xlabel('Days')
plt.legend()

```

<matplotlib.legend.Legend at 0x7f10553be910>



## ▼ 2 Bar Plot

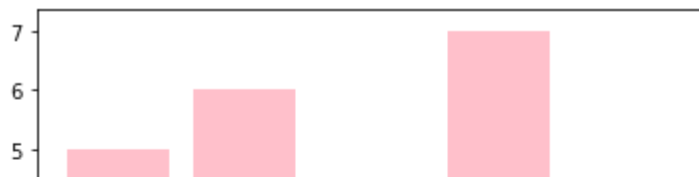
### ▼ 2.1 Plot a Bar Chart

```

x_values = [5,6,3,7,2]
y_values = ["A", "B", "C", "D", "E"]

plt.bar(y_values,x_values, color = "pink")
plt.show()

```



## ▼ 2.2 Multiple Bar Plot



# Example 1

```
plt.bar([0.25,1.25,2.25,3.25,4.25],[50,40,70,80,20],
label="Enfield",width=.5)
```

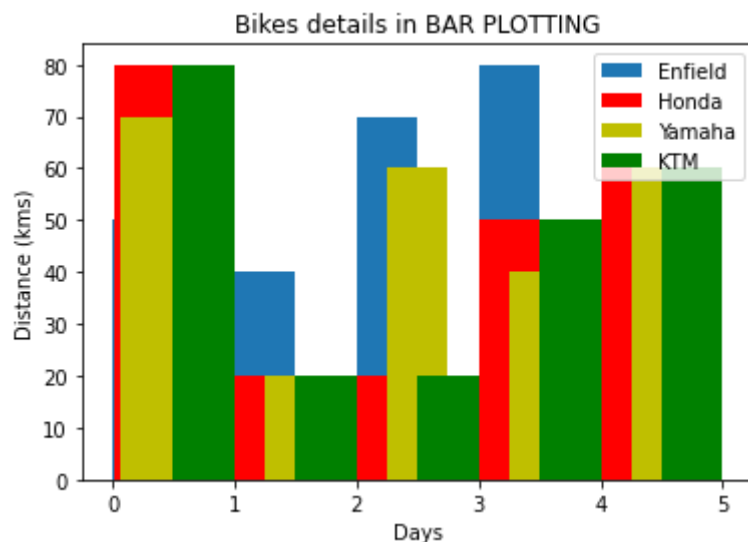
```
plt.bar([0.26,1.25,2.25,3.25,4.25],[80,20,20,50,60],
label="Honda", color='r',width=.5)
```

```
plt.bar([0.31,1.5,2.5,3.5,4.5],[70,20,60,40,60],
label="Yamaha", color='y',width=.5)
```

```
plt.bar([.75,1.75,2.75,3.75,4.75],[80,20,20,50,60],
label="KTM", color='g',width=.5)
```

```
plt.legend()
plt.xlabel('Days')
plt.ylabel('Distance (kms)')
plt.title('Bikes details in BAR PLOTTING')
```

Text(0.5, 1.0, 'Bikes details in BAR PLOTTING')



#Example 2

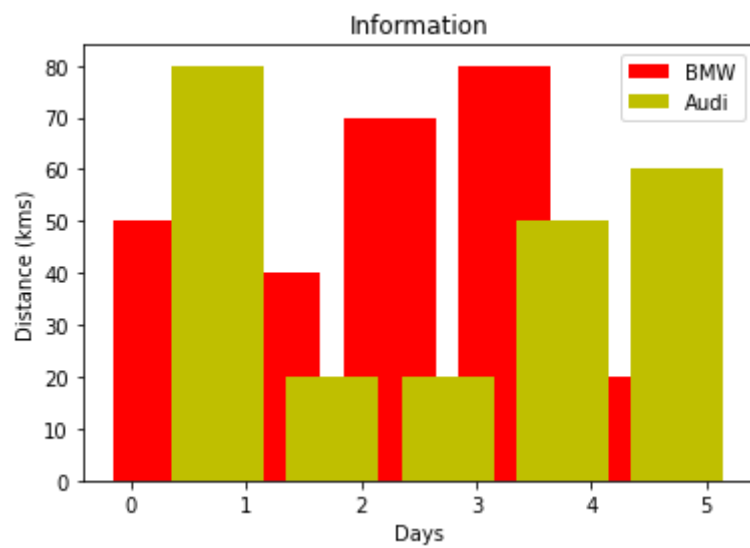
```
#from matplotlib import pyplot as plt
```

```
plt.bar([0.25,1.25,2.25,3.25,4.25],[50,40,70,80,20], label="BMW", color='r')
```

```
plt.bar([.75,1.75,2.75,3.75,4.75],[80,20,20,50,60], label="Audi", color='y')
```

```
plt.legend()
plt.xlabel('Days')
plt.ylabel('Distance (kms)')
```

```
plt.title('Information')
plt.show()
```

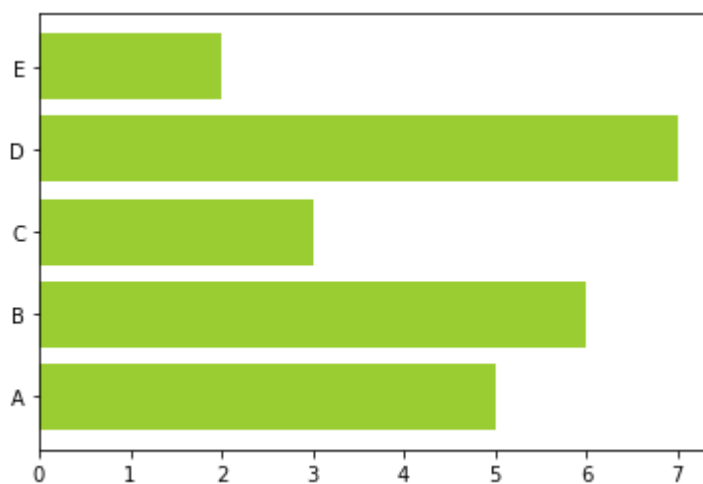


## ▼ 2.3 Horizontal Bar Plot

```
x_values = [5,6,3,7,2]
y_val = ["A", "B", "C", "D", "E"]

plt.barh(y_val,x_values, color ="yellowgreen")

# Adding an "h" after bar will flip the graph
plt.show()
```



## ▼ 3 Scatter Plot

### ▼ 3.1 Scatter plot without Details

```

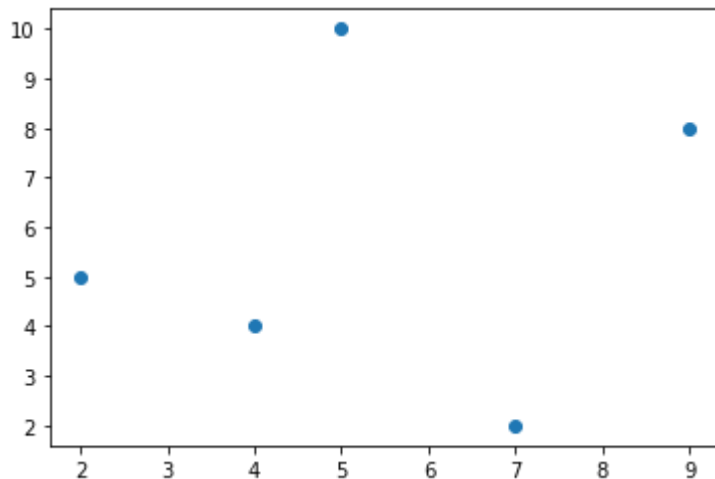
x = [5, 2, 9, 4, 7]      # x-axis values

y = [10, 5, 8, 4, 2]      # Y-axis values

plt.scatter(x, y)          # Function to plot scatter

plt.show()                # function to show the plot

```



## ▼ 3.2 Multiple Scatter Plot

```

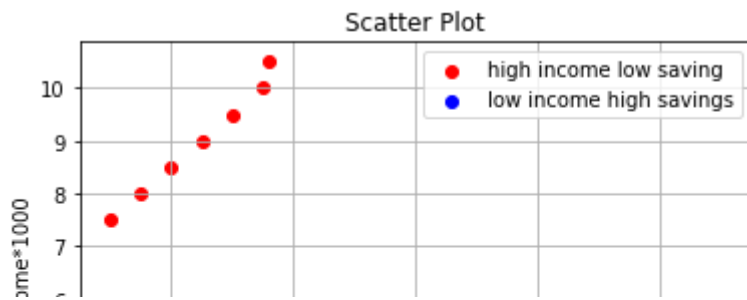
x = [1,1.5,2,2.5,3,3.5,3.6]
y = [7.5,8,8.5,9,9.5,10,10.5]
x1=[8,8.5,9,9.5,10,10.5,11]
y1=[3,3.5,3.7,4,4.5,5,5.2]

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()
plt.grid()

```



## ▼ 4 pie plot

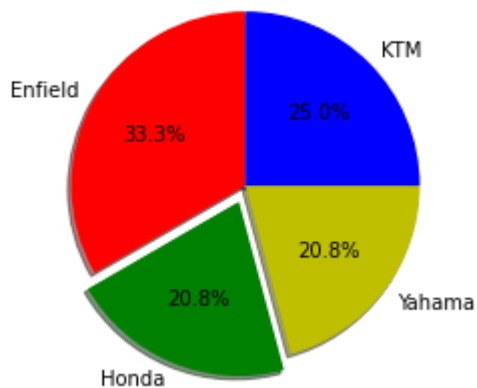


```
days = [1,2,3,4,5]
Enfield =[50,40,70,80,20]
Honda = [80,20,20,50,60]
Yahama =[70,20,60,40,60]
KTM = [80,20,20,50,60]
slices = [8,5,5,6]

activities = ['Enfield','Honda','Yahama','KTM']
cols = ['r','g','y','b']
plt.pie(slices, labels=activities, colors=cols, startangle=90, shadow= True,explode=(0,0.1

plt.title('Bike details in Pie Plot')
```

```
Text(0.5, 1.0, 'Bike details in Pie Plot')
Bike details in Pie Plot
```



## ▼ 5 Matplotlib Plots on NBA DataSet

```
import pandas as pd
import matplotlib.pyplot as plt
from google.colab import files
uploaded = files.upload()
df = pd.read_csv('nba.csv')
df
```

Choose Files nba.csv

- **nba.csv**(n/a) - 32823 bytes, last modified: 8/23/2020 - 100% done  
Saving nba.csv to nba.csv

	Name	Team	Number	Position	Age	Height	Weight	Colle
0	Avery Bradley	Boston Celtics	0.0	PG	25.0	6-2	180.0	Tex
1	Jae Crowder	Boston Celtics	99.0	SF	25.0	6-6	235.0	Marque
2	John Holland	Boston Celtics	30.0	SG	27.0	6-5	205.0	Boston Univers
3	R.J. Hunter	Boston Celtics	28.0	SG	22.0	6-5	185.0	Georgia Sta
4	Jonas Jerebko	Boston Celtics	8.0	PF	29.0	6-10	231.0	Ni
...	...	...	...	...	...	...	...	...
453	Shelvin Mack	Utah Jazz	8.0	PG	26.0	6-3	203.0	But
454	Raul Neto	Utah Jazz	25.0	PG	24.0	6-1	179.0	Ni
455	Tibor Pleiss	Utah Jazz	21.0	C	26.0	7-3	256.0	Ni
456	Jeff Withey	Utah Jazz	24.0	C	26.0	7-0	231.0	Kans
457	NaN	NaN	NaN	NaN	NaN	NaN	NaN	Ni

458 rows × 9 columns

## ▼ 5.1 Bar Graph

```
#Example 1
y_values = df['Salary']
x_values = df['Age']

plt.xlabel('Age')
plt.ylabel('Salary (in millions)')

#To plot a bar graph plt.bar() command is used
#This plots a bar graph between Age and Salaries of NBA players
plt.bar(x_values,y_values,color = "purple")
plt.show()
```

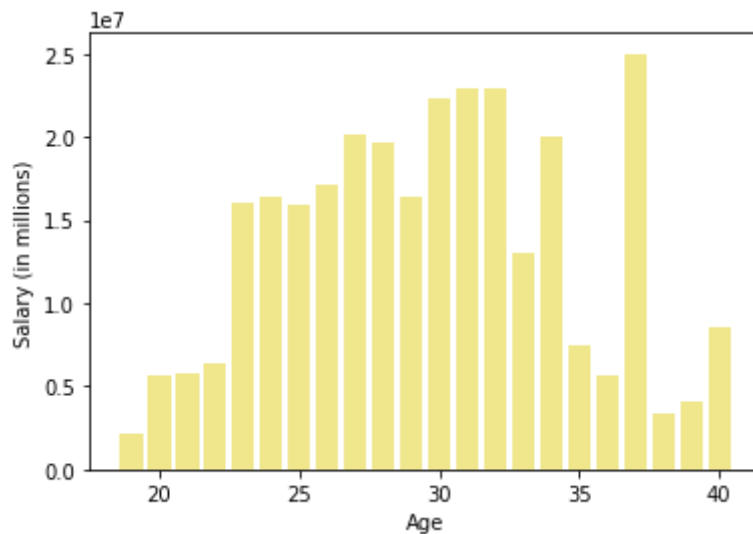




```
# Example 2
y_values = df['Salary']
x_values = df['Age']

plt.xlabel('Age')
plt.ylabel('Salary (in millions)')

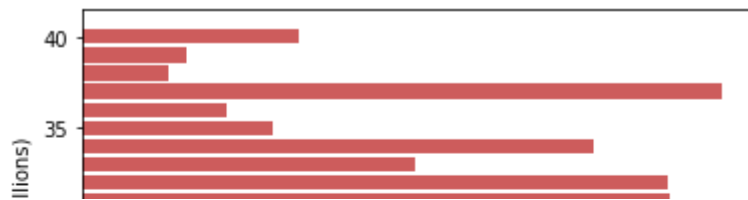
# Making changes in the color field changes the colour of the graph
plt.bar(x_values,y_values,color = "khaki")
plt.show()
```



## ▼ 5.2 Horizontal Bar Graph

```
y_values = df['Salary']
x_values = df['Age']

plt.xlabel('Age')
plt.ylabel('Salary (in millions)')
plt.barh(x_values,y_values,color = "indianred")
plt.show()
```



## ▼ 5.3 Scatter Plot

5.3.1 Scatter Plot

#Example 1

```
y_value=df['Salary']
```

```
x_values=df['Age']
```

#To display a scatter plot we use plt.scatter() command

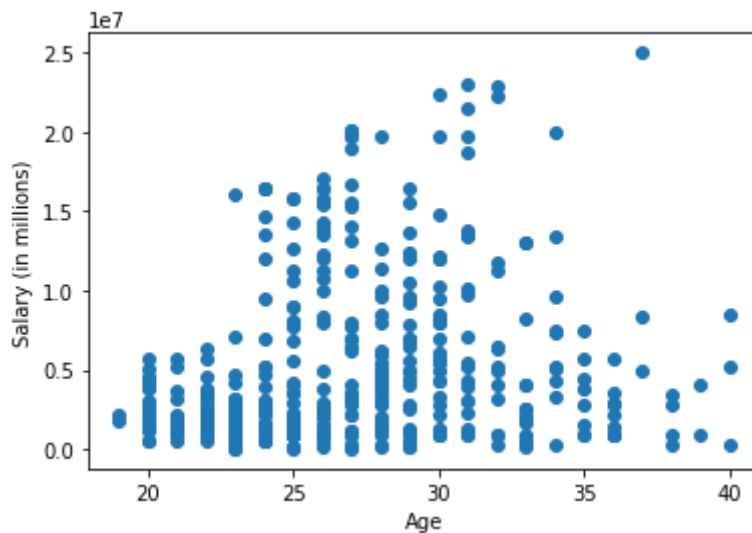
```
plt.xlabel('Age')
```

```
plt.ylabel('Salary (in millions)')
```

#This displays a scatter plot between Age and Salaries of NBA players

```
plt.scatter(x_values,y_values)
```

<matplotlib.collections.PathCollection at 0x7f1055188650>



#Example 2

```
y_value=df['Salary']
```

```
x_values=df['Age']
```

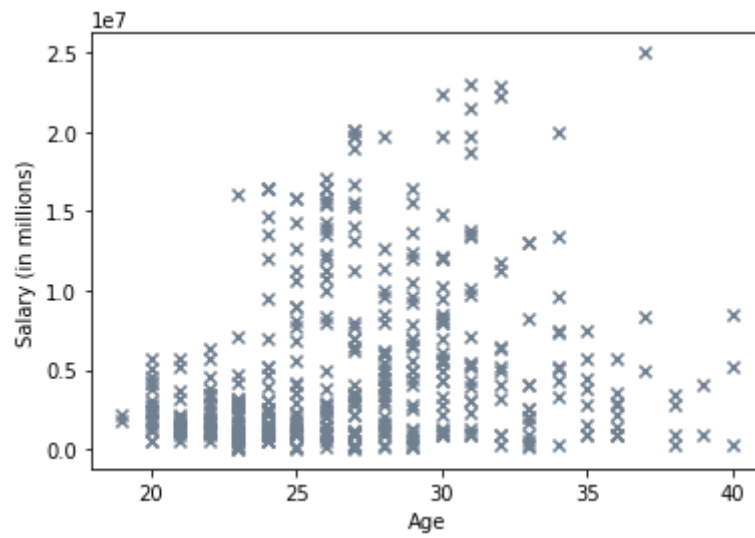
```
plt.xlabel('Age')
```

```
plt.ylabel('Salary (in millions)')
```

#We can also change the type and colors of scatter plot as shown

```
plt.scatter(x_values,y_values, color = "slategrey", marker = "x")
```

<matplotlib.collections.PathCollection at 0x7f1055360c10>



✓ 0s completed at 5:06 PM

● ✕