

# Data Science & Machine Learning:

## Lab Cycle 3

### Mathplotlib

**Demonstrate creating various types of charts and plots using functions in matplotlib library**

1. Sarah bought a new car in 2001 for \$24,000. The dollar value of her car changed each year as shown in the table below.

#### **Value of Sarah's Car**

**Year Value**

**2001 \$24,000**

**2002 \$22,500**

**2003 \$19,700**

**2004 \$17,500**

**2005 \$14,500**

**2006 \$10,000**

**2007 \$ 5,800**

Represent the following information using a line graph with following style properties

- **X- axis - Year**
- **Y -axis - Car Value**
- **title -Value Depreciation (left Aligned)**
- **Line Style dashdot and Line-color should be red**
- **point using \* symbol with green color and size 20**

Subplot() provides multiple plots in one figure.

#### **Code:**

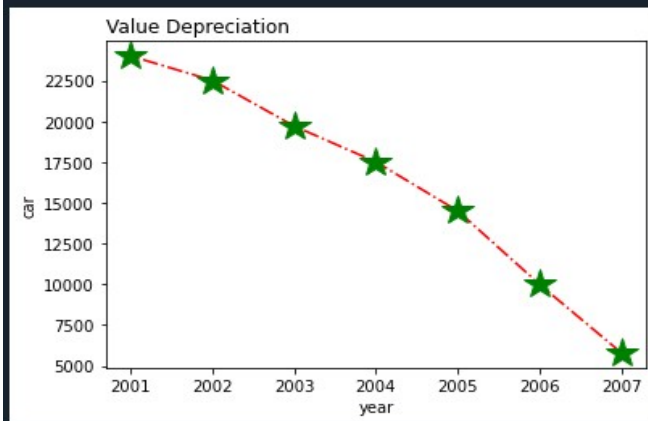
```
import matplotlib.pyplot as plt
import numpy as np
```

```
xaxis = np.array([2007, 2006, 2005, 2004, 2003, 2002, 2001])
yaxis = np.array([5800, 10000, 14500, 17500, 19700, 22500, 24000])
```

```
plt.plot(xaxis, yaxis, linestyle="dashdot", c='r', marker= '*', ms='20',
mfc='green', mec='green')
plt.title("Value Depreciation", loc='left')
plt.xlabel("year")
plt.ylabel("car")
plt.show()
```

### Output:

```
In [2]: runfile('/home/sjctet/Documents/Sandra-20MCA049/DSML/Cycle3/
p1FunctionsInMathplotlib.py', wdir='/home/sjctet/Documents/Sandra-20MCA049/DSML/Cycle3')
```



2. Following table gives the daily sales of the following items in a shop

Day	Mon	Tues	Wed	Thurs	Fri
Drinks	300	450	150	400	650
Food	400	500	350	300	500

Use subplot function to draw the line graphs with **grids(color as blue and line style dotted)** for the above information as 2 separate graphs in two rows

a) Properties for the Graph 1:

- X label- Days of week
- Y label-Sale of Drinks
- Title-Sales Data1 (right aligned)
- Line -dotted with cyan color
- Points- hexagon shape with color magenta and outline black

b) Properties for the Graph 2:

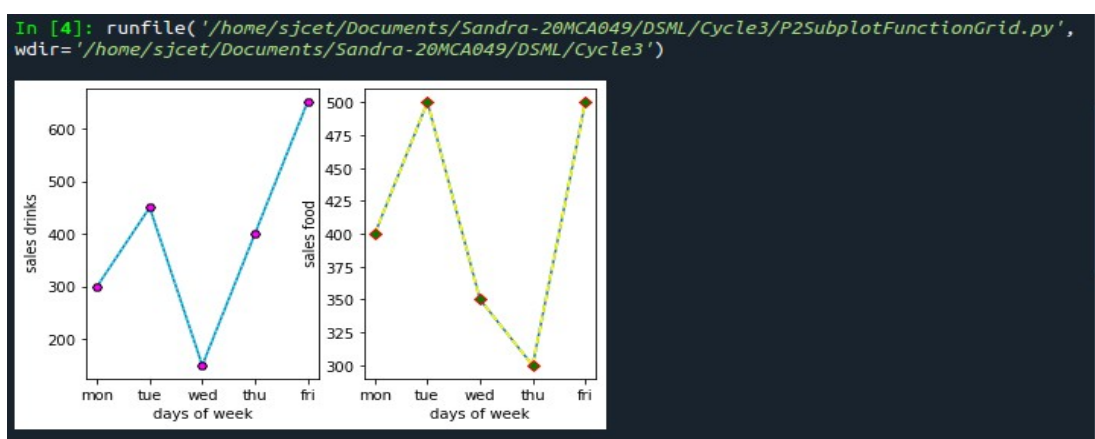
- X label- Days of Week
- Y label-Sale of Food
- Title-Sales Data2 ( center aligned)
- Line -dashed with yellow color

- Points- diamond shape with color green and outline red

### Code:

```
import matplotlib.pyplot as plt
import numpy as np
x = np.array(["mon", "tue", "wed", "thu", "fri"])
y = np.array([300, 450, 150, 400, 650])
plt.subplot(1, 2, 1)
plt.plot(x, y)
plt.plot(x, y, linestyle =
'dotted', color='cyan', marker='H', mec='black', mfc='magenta')
plt.xlabel("days of week")
plt.ylabel(" sales drinks")
#plot2
x = np.array(["mon", "tue", "wed", "thu", "fri"])
y = np.array([400, 500, 350, 300, 500])
plt.subplot(1, 2, 2)
plt.plot(x, y)
plt.plot(x, y, linestyle =
'dashed', color='yellow', marker='D', mec='red', mfc='green')
plt.xlabel("days of week")
plt.ylabel(" sales food")
plt.show()
```

### Output:



3. Create scatter plot for the below data:(use Scatter function)

Product	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Affordable Segment	173	153	195	147	120	144	148	109	174	130	172	131
Luxury Segment	189	189	105	112	173	109	151	197	174	145	177	161
Super Luxury Segment	185	185	126	134	196	153	112	133	200	145	167	110

Create scatter plot for each Segment with following properties within one graph

- X Label- Months of Year with font size 18
- Y-Label- Sales of Segments
- Title -Sales Data
- Color for Affordable segment- pink
- Color for Luxury Segment- Yellow
- Color for Super luxury segment-blue

**Code:**

```
import matplotlib.pyplot as plt
import numpy as np
plt.title("Sales Data")
plt.xlabel("Months of Year")
plt.ylabel("Sale of Food")

x =
np.array(['Jan','Feb','Mar','Apr','May','Jun','Jul','Aug','Sep','Oct','Nov','Dec'])

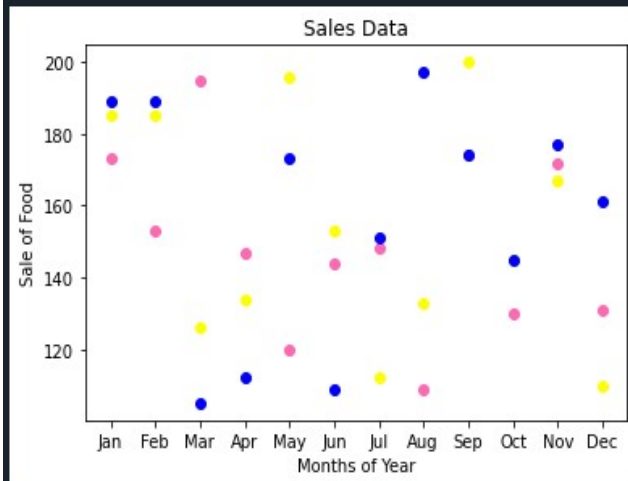
y1 = np.array([173,153,195,147,120,144,148,109,174,130,172,131])
plt.scatter(x,y1, color = 'hotpink')

y2 = np.array([185,185,126,134,196,153,112,133,200,145,167,110])
plt.scatter(x, y2, color = 'yellow')

y3 = np.array([189,189,105,112,173,109,151,197,174,145,177,161])
plt.scatter(x, y3, color = 'blue')
plt.show()
```

## Output:

```
In [5]: runfile('/home/sjct/Documents/Sandra-20MCA049/DSML/Cycle3/P3ScatterPlot.py', wdir='/home/sjct/Documents/Sandra-20MCA049/DSML/Cycle3')
```



4. Display the above data using multiline plot  
( 3 different lines in same graph)
  - Display the description of the graph in upper right corner(**use legend()**)
  - Use different colors and line styles for 3 different lines

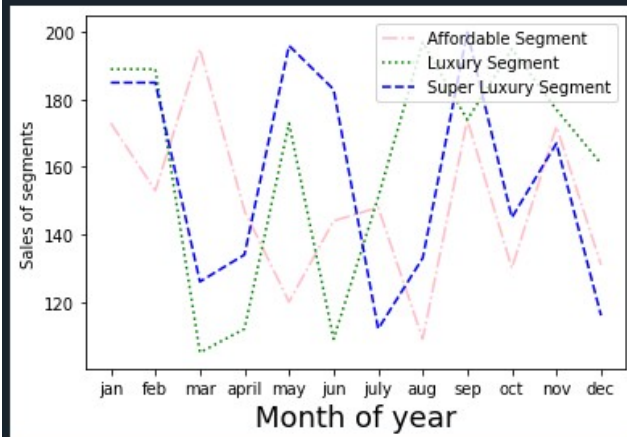
### Code:

```
import matplotlib.pyplot as plt
import numpy as np

x = np.array(['jan','feb','mar','april','may','jun','july','aug','sep','oct','nov','dec'])
y1 = np.array([173,153,195,147,120,144,148,109,174,130,172,131])
y2 = np.array([189,189,105,112,173,109,151,197,174,195,177,161])
y3 = np.array([185,185,126,134,196,183,112,133,200,145,167,116])
font1={'size':18}
plt.plot(x,y1,color='pink' ,label = "Affordable Segment" , linestyle="-.")
plt.plot(x,y2,color='green', label = "Luxury Segment" ,linestyle=":")
plt.plot(x,y3,color='blue',label = "Super Luxury Segment", linestyle="--")
plt.xlabel("Month of year",fontdict=font1)
plt.ylabel("Sales of segments")
plt.legend()
plt.show()
```

### Output:

```
In [3]: runfile('/home/sjcet/Documents/Sandra-20MCA049/DSML/Cycle3/P4.py', wdir='/home/sjcet/Documents/Sandra-20MCA049/DSML/Cycle3')
```



5. 100 students were asked what their primary mode of transport for getting to school was. The results of this survey are recorded in the table below. Construct a bar graph representing this information.

### Code:

### Output:

Mode of transport	Frequency
Walking	29
Cycling	15
Car	35
Bus	18
Train	3

Create a bar graph with `import matplotlib.pyplot as plt`

- X axis -mode of Transport and Y axis 'frequency'
- Provide appropriate labels and title
- Width .1, color green

### Code:

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

```
import numpy as np
```

```
x = np.array(["walking", "cycling", "car", "bus", "train"])
```

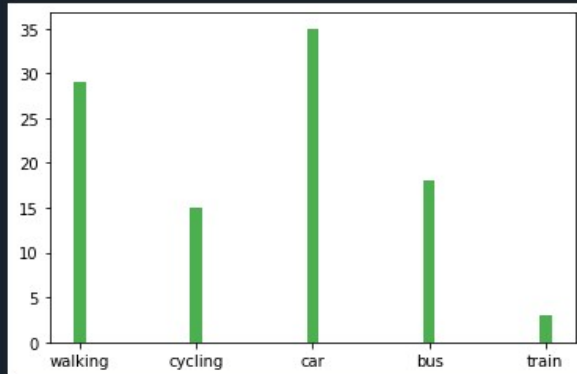
```
y = np.array([29,15,35,18,3])
```

```
plt.bar(x,y, color="#4CAF50",width=0.1)
```

```
plt.show()
```

### Output:

```
In [10]: runfile('/home/sjctet/Documents/Sandra-20MCA049/DSML/Cycle3/P5StudentBarGraph.py',  
wdir='/home/sjctet/Documents/Sandra-20MCA049/DSML/Cycle3')
```



6. We are provided with the height of 30 cherry trees.  
The height of the trees (in inches): 61, 63, 64, 66, 68, 69, 71, 71.5, 72, 72.5, 73, 73.5, 74, 74.5, 76, 76.2, 76.5, 77, 77.5, 78, 78.5, 79, 79.2, 80, 81, 82, 83, 84, 85, 87. Create a histogram with a bin size of 5

### Code:

```
import matplotlib.pyplot as plt  
height = [61, 63, 64, 66, 68, 69, 71, 71.5,  
72, 72.5, 73, 73.5, 74, 74.5, 76, 76.2, 76.5, 77, 77.5, 78, 78.5, 79,  
79.2, 80, 81, 82, 83, 84, 85, 87]  
plt.hist(height, edgecolor="red", bins=5)  
plt.show()
```

### Output:

```
In [4]: runfile('/home/sjctet/Documents/Sandra-20MCA049/DSML/Cycle3/P6Histogram.py', wdir='/  
home/sjctet/Documents/Sandra-20MCA049/DSML/Cycle3')
```

