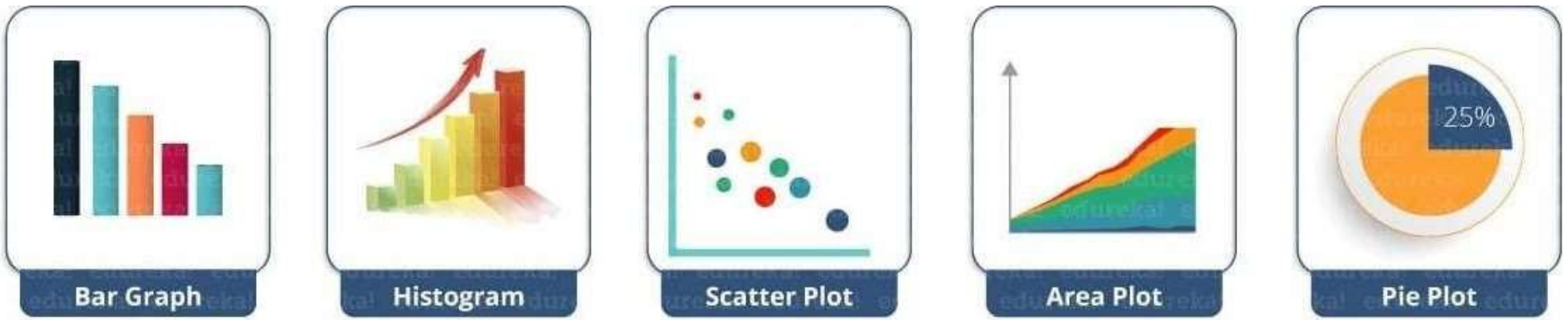
I

Topic: DATA VISUALIZATION -NOTES

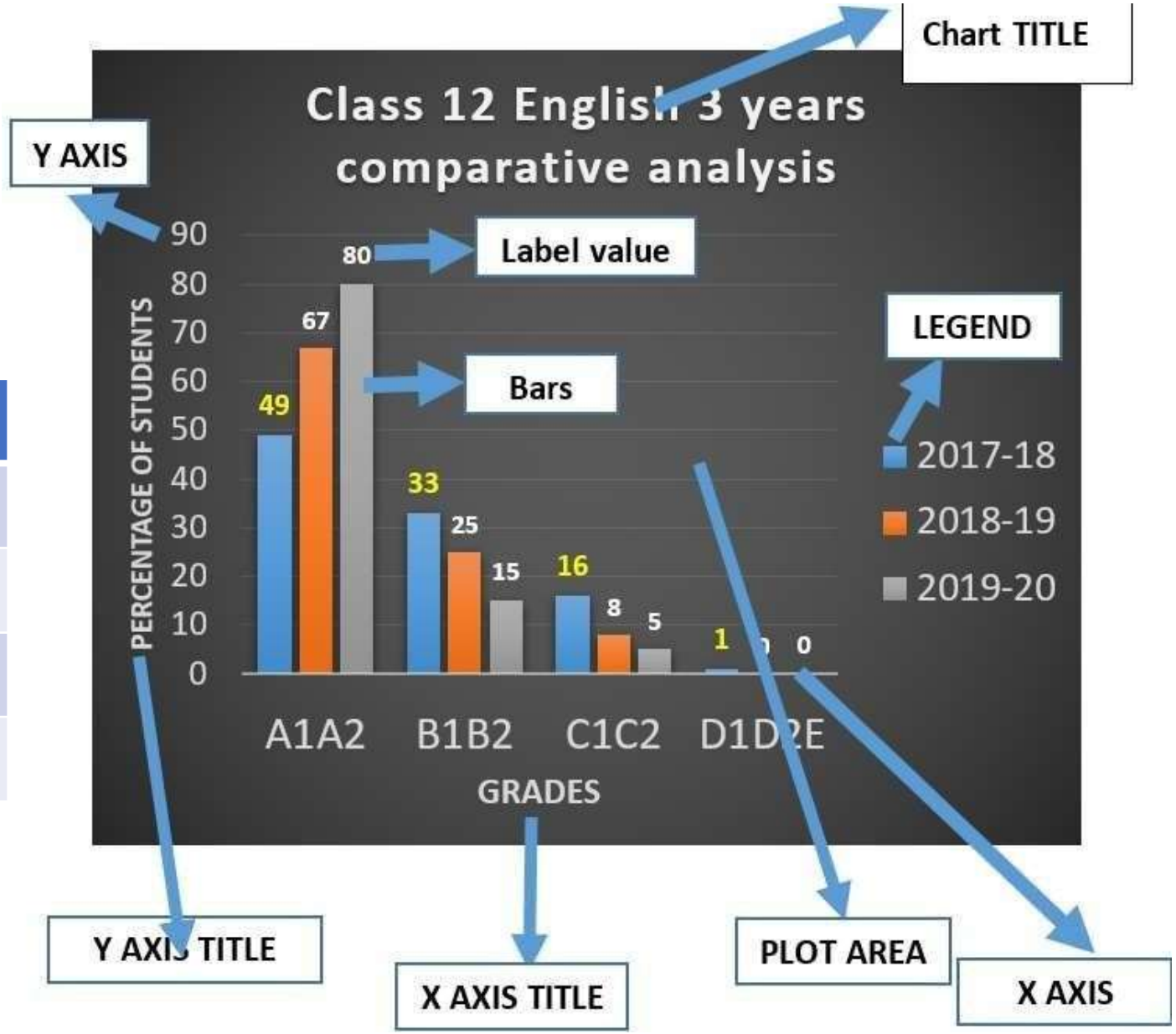


DATA VISUALIZATION in PANDAS

- ***Data visualization*** basically refers to the **graphical or visual representation of information** and data using visual elements like charts, graphs, maps, etc.
- Charts are often used to ease understanding of large quantities of data and the relationships between parts of the data. Charts can usually be read more quickly than the raw data.

VARIOUS COMPONENTS OF CHART

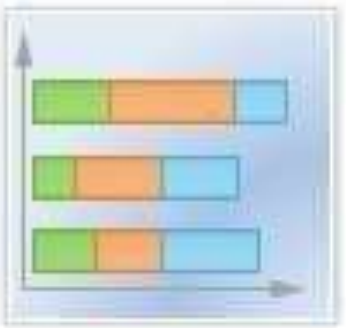
ENGLISH				
	A1A2	B1B2	C1C2	D1D2E
2017-18	49	33	16	1
2018-19	67	25	8	0
2019-20	80	15	5	0



Type of charts:

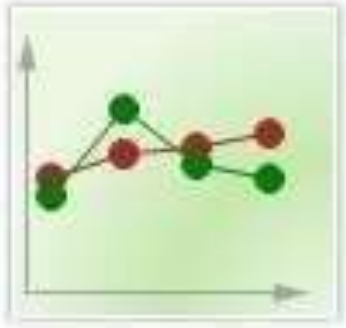
- LINE CHART
- BAR CHART
- HISTOGRAM
- PIE CHART
- FREQUENCY POLYGON
- BOX PLOT
- SCATTER PLOT


Column

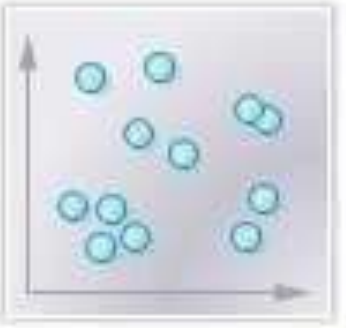

Bar

Pie

Doughnut


Line


Area

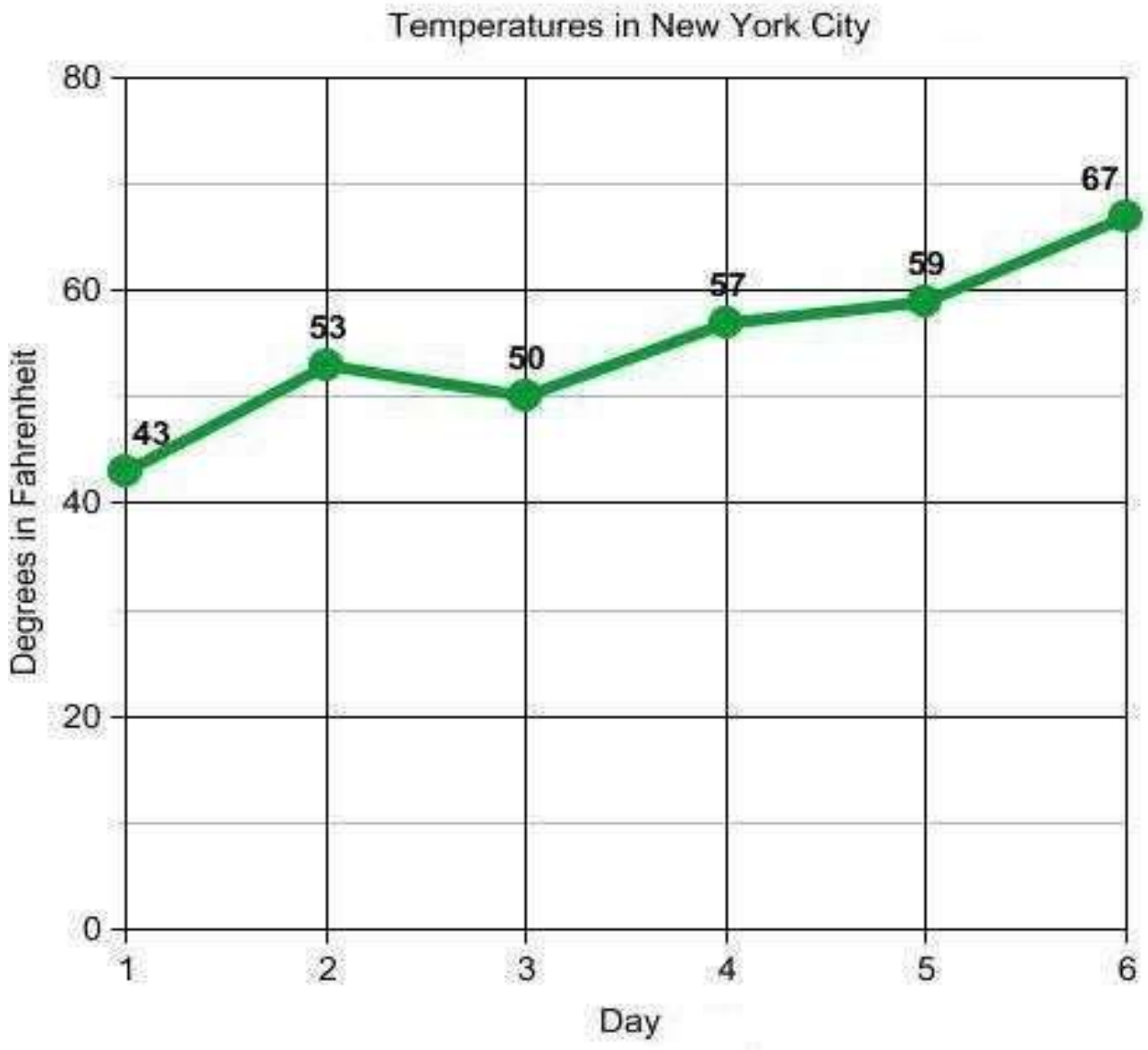

Scatter


Spider

LINE CHART

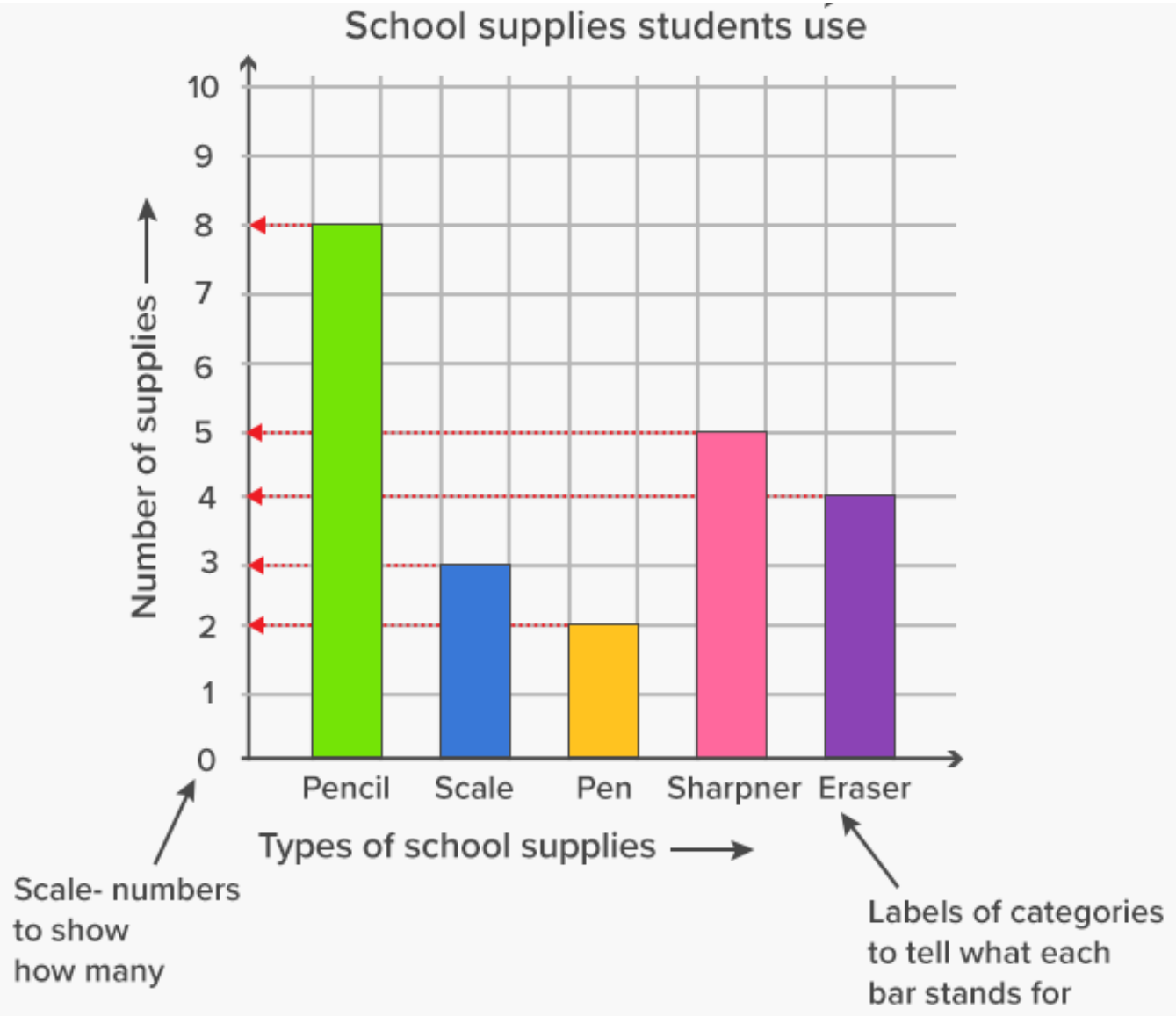
- A **LINE CHART** or line graph is a type of chart which displays information as a series of data points called 'markers' connected by straight line segments. Line graphs are usually used to find relationship between two data sets on different axis.

In the example given below, we are trying to show a relationship between every day temperatures in F in the New York city.



BAR CHART

1. A **BAR CHART** or bar graph is a chart or graph that presents categorical data with rectangular bars with heights or lengths proportional to the values that they represent. The bars can be plotted vertically or horizontally.

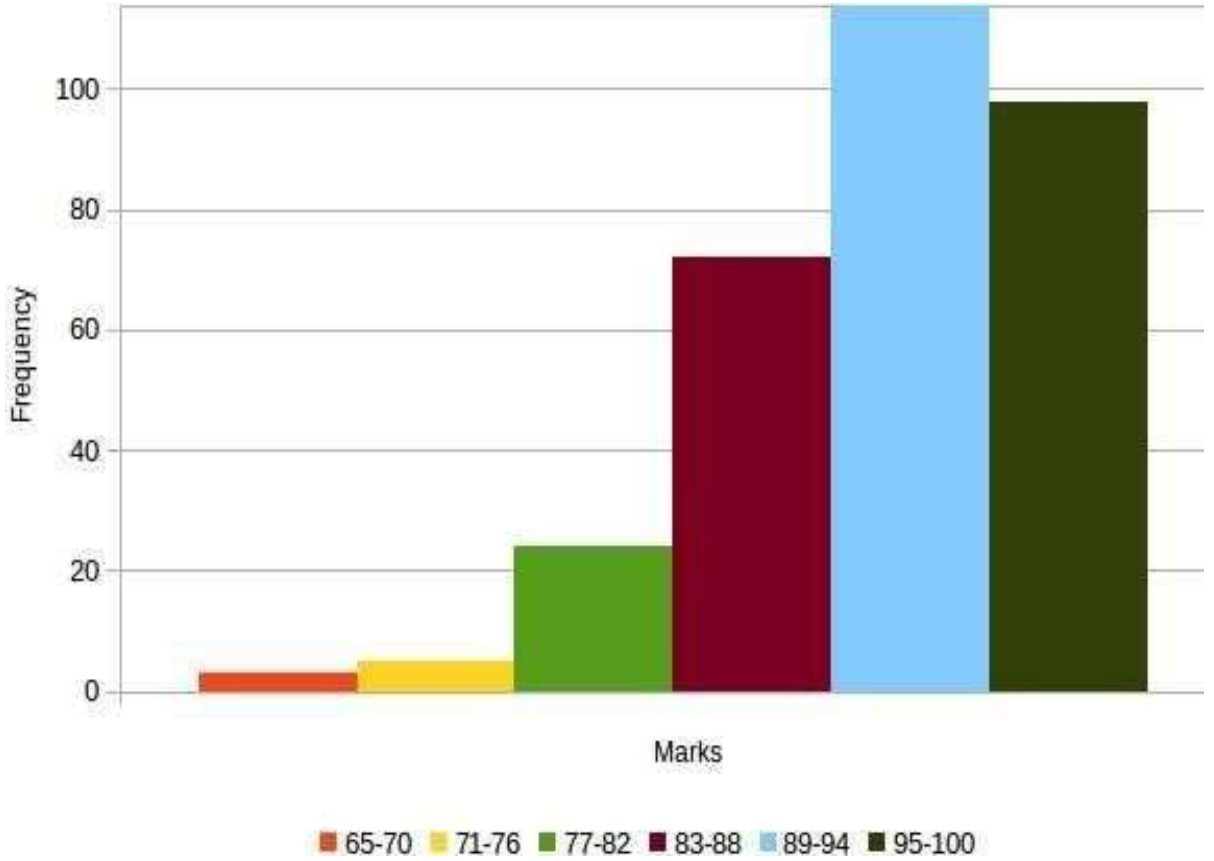


HISTOGRAM

It represents the frequency on the vertical axis. The horizontal axis is another dimension. It is similar to a Bar Chart, but a histogram groups numbers into ranges. The height of each bar shows how many fall into each range, which can be called as **frequency.**

Usually, it has bins, where every bin has a minimum and maximum value.

Range of Marks (x axis)	No of students who have scored in this range also called as Frequency (y axis)
65-70	2
71-76	5
77-82	22
83-88	70
89-94	150
95-100	98



HOW TO START MAKING GRAPHS IN PANDAS?

Matplotlib is the whole Python package/library used to create 2D graphs and plots by using Python scripts.

Pyplot is a module in matplotlib which supports a very wide variety of graphs and plots. It helps save images in several output formats (PNG, PS and others).

How to install matplotlib in IDLE

Go to command prompt where your pip installed and write the command.

pip install matplotlib

CUSTOMIZING PLOTS

List of Pyplot functions to customize plots

1. grid - Configure the grid lines.

   Ex: plt.grid(True): shows grid line

2. title - Set a title for the axes.

   Ex: plt.title(“Title name”)

3. Adding labels:

a) xlabel - Set the label for the x-axis.

b) ylabel - Set the label for the y-axis.

plt.xlabel(“ x axis label name”)

plt.ylabel(“y axis label”)

4. legend - Place a legend on the axes.

   Ex: plt.legend()

5. savefig - Save the current figure.

   Ex: plt.savefig(“ path&filename”)

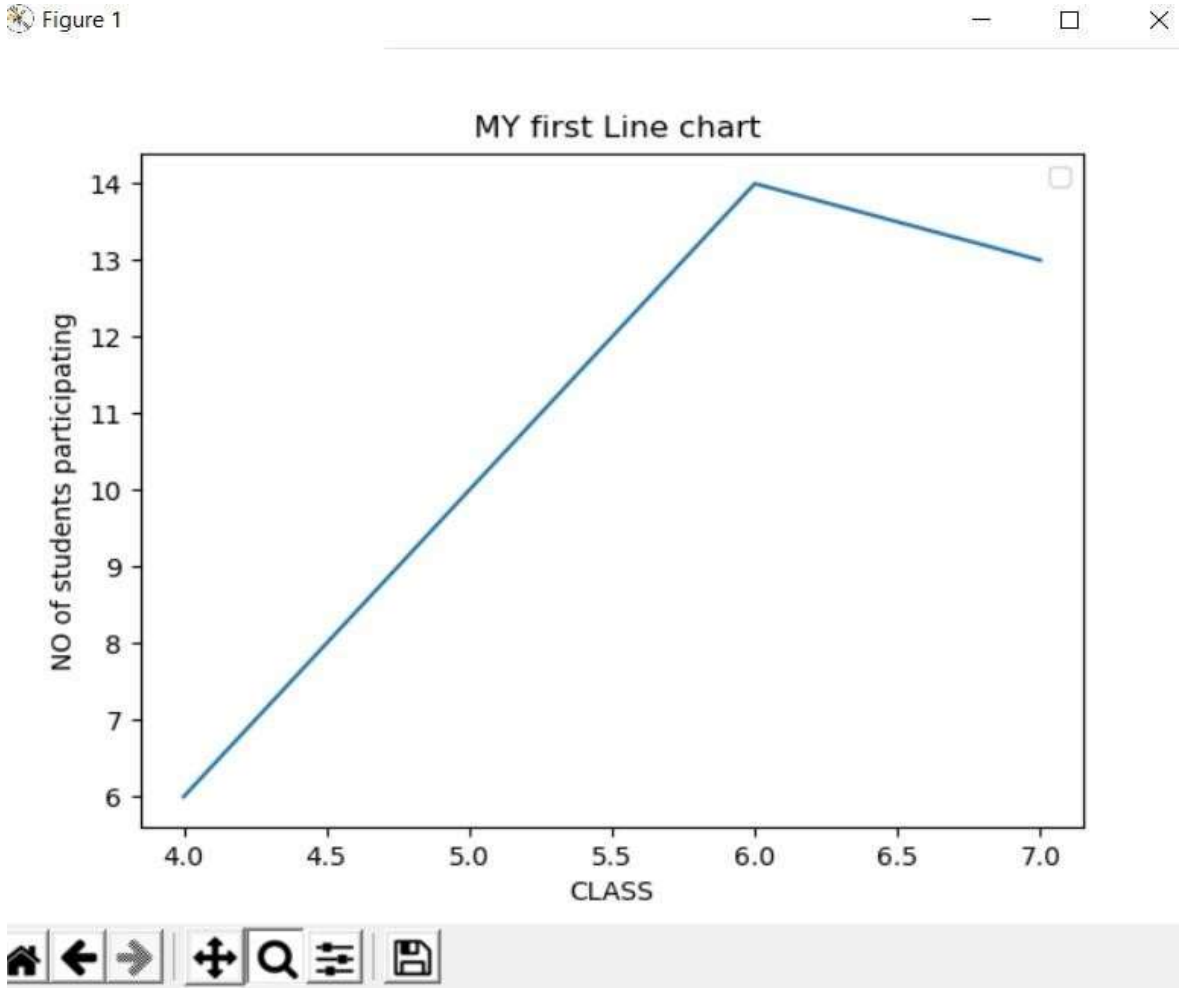
6. show - Display the chart.

   Ex: plt.show()

LINE PLOT example1:

Question: Draw a line graph where the x axis shows the individual classes and the y axis shows the number of students participating in an interhouse event.

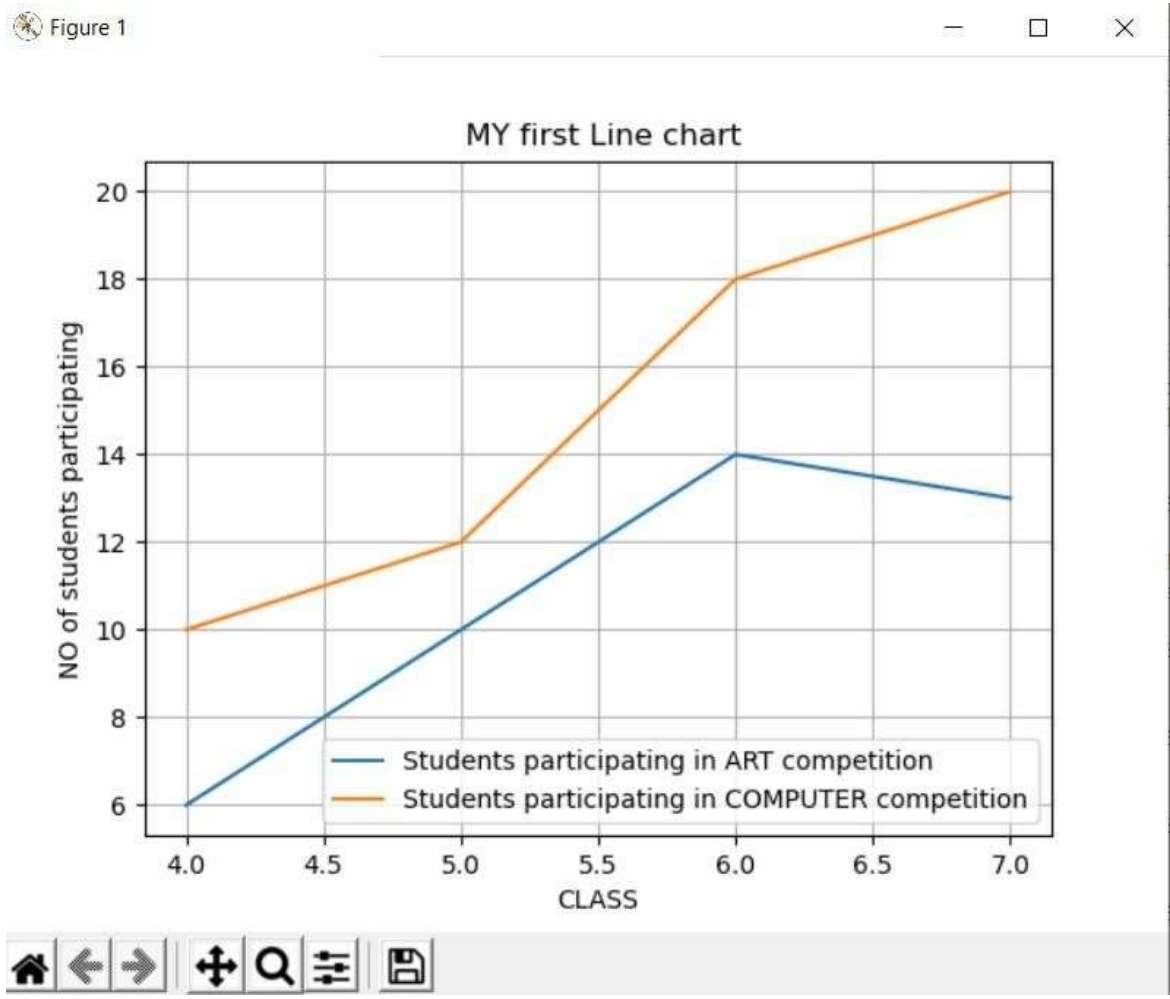
```
import matplotlib.pyplot as plt
x=[4,5,6,7]
y=[6,10,14,13]
plt.plot(x,y)
plt.title(" MY first Line chart")
plt.xlabel("CLASS")
plt.ylabel("NO of students participating")
plt.show()
```



LINE PLOT example2:

Question: Draw two line graph where the x axis shows the individual classes and the y axis shows the number of students participating in ART/COMPUTER interhouse event.

```
import matplotlib.pyplot as plt
x=[4,5,6,7]
y=[6,10,14,13]
z=[10,12,18,20]
plt.plot(x,y,label="Students participating inART competition")
plt.plot(x,z,label="Students participating in COMPUTER
competition")
plt.legend()
plt.title(" MY first Line chart")
plt.xlabel("CLASS")
plt.ylabel("NO of students participating")
plt.grid(True)
plt.show()
```



CHANGING COLOURS OF THE LINES:

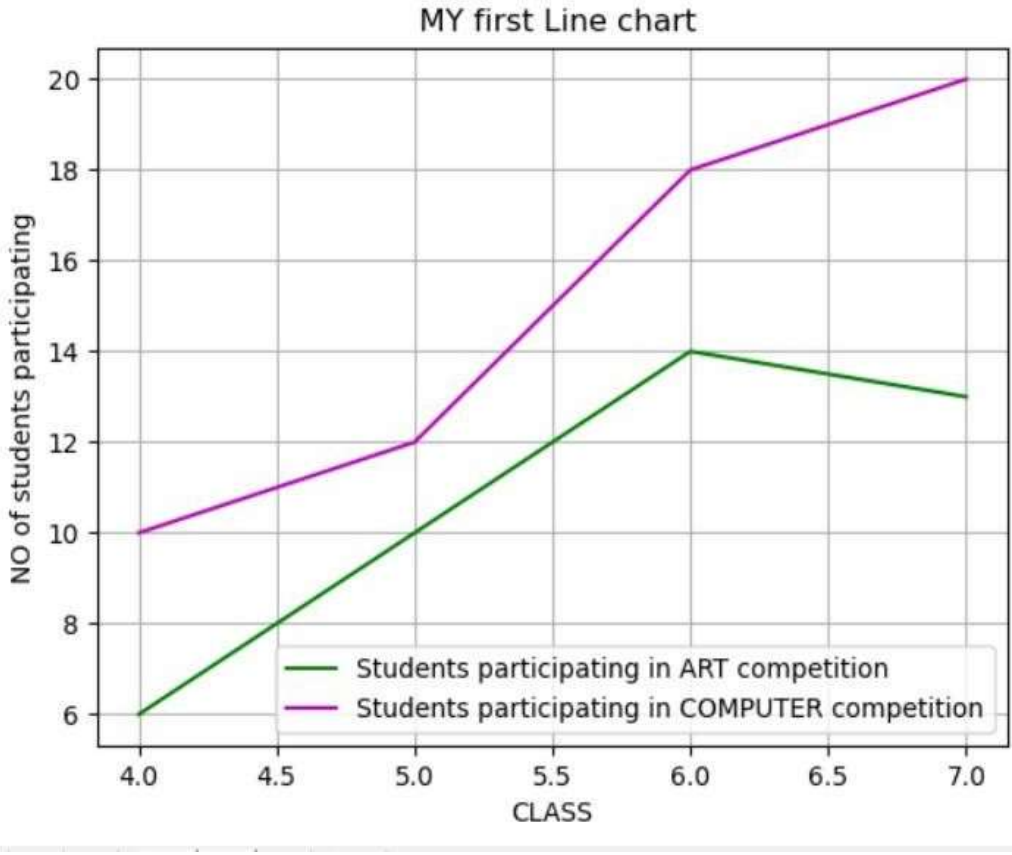
Table of the abbreviations used to select colors:

Colour abbreviation	Colour name
b	blue
c	cyan
g	green
k	black
m	magenta
r	red
w	white
y	yellow

LINE PLOT Example3:

Question: Draw two line graph where the x axis shows the individual classes and the y axis shows the number of students participating in ART/COMPUTER interhouse event. Use different colours for the plots.

```
import matplotlib.pyplot as plt
x=[4,5,6,7]
y=[6,10,14,13]
z=[10,12,18,20]
plt.plot(x,y,'g',label="Students participating in ART competition")
plt.plot(x,z,'m',label="Students participating in COMPUTER
competition")
plt.legend()
plt.title(" MY first Line chart")
plt.xlabel("CLASS")
plt.ylabel("NO of students participating")
plt.grid(True)
plt.show()
```



Some MORE VARIATIONS which we can add with the following functions:

a) **plot()**

- **linestyle**- set linestyle to any of '-‘ for solid line style, '--‘ for dashed, '-.‘ , ':‘ for dotted line
- **linewidth**- sets the width of the line
- **color-** sets the color of the line

b) **xlabel(), ylabel():**

- **fontsize:** can specify the size of the font for the labels
- **color:** can specify the color of the font used to display the x and y labels.

Some MORE VARIATIONS which we can add with the following functions:

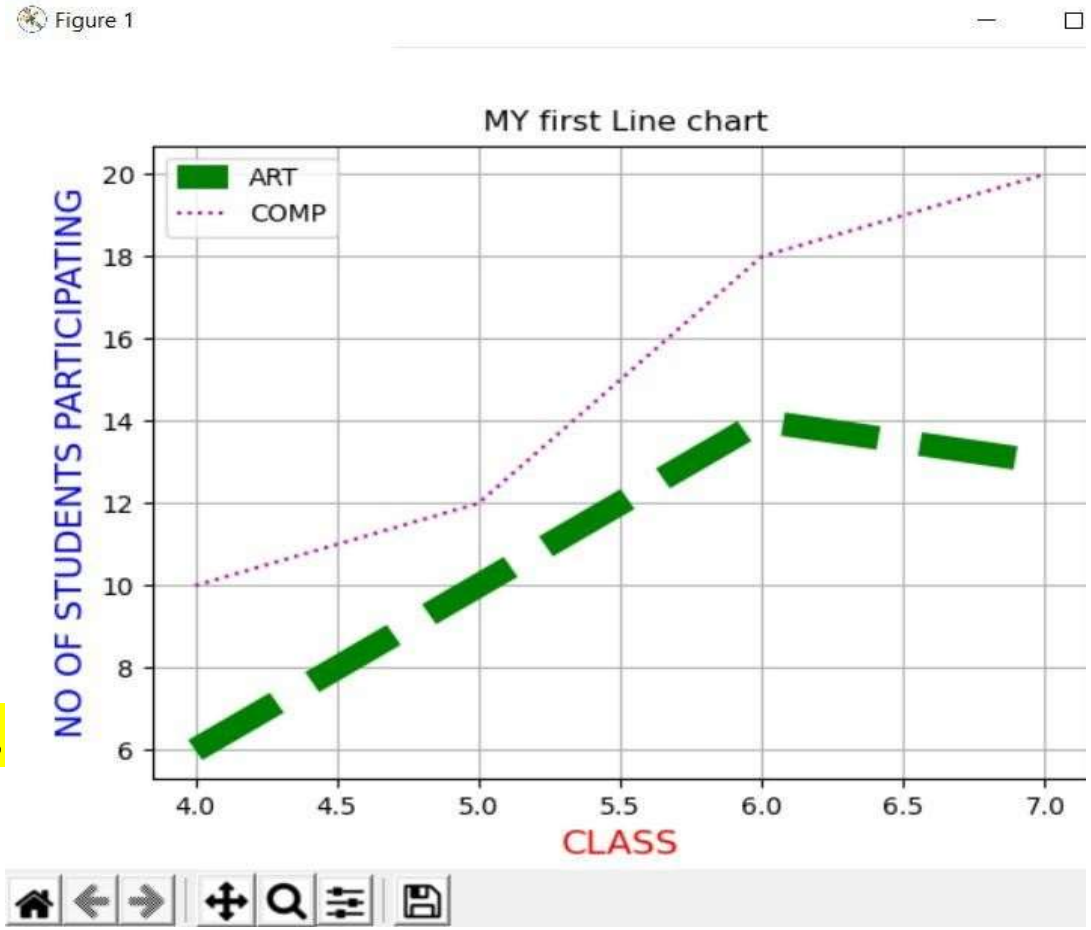
c) legend()

- **loc-** specify the location of the legend as “upper left” or “upper right”, “lower left”,”lower right”
- To give different names other than the labels mentioned in the plot(). (“**new names** s eparated by commas in their own half circle bracket”)
  - Ex: plt.legend((„Art“,“Comp“))

LINE PLOT Example4:

Question: Draw two line graph where the x axis shows the individual classes and the y axis shows the number of students participating in ART/COMPUTER interhouse event. Use different colors for the x and y labels. Specify the position and text for the legend. Alsochange the style and width of the lines.

```
import matplotlib.pyplot as plt
x=[4,5,6,7]
y=[6,10,14,13]
z=[10,12,18,20]
plt.plot(x,y,'g',linestyle="--", linewidth=10)
plt.plot(x,z,'m',linestyle="dotted" )
plt.legend(('ART','COMP'), loc='upper left')
plt.title(" MY first Line chart")
plt.xlabel("CLASS", fontsize="14",color="red")
plt.ylabel("NO OF STUDENTS PARTICIPATING", fontsize="14", color="blue")
plt.grid(True)
plt.show()
```



Creating multiple views in the same chart

Using subplot(): allows u to draw two or more graphs on the same area.

Ex: subplot(2,1,1)

2 represents number of graphs

The Last 1 represent the first one out of two

WE can adjust the horizontal and vertical space in corresponding to the 2 plots

subplots_adjust (hspace=0.4, wspace=0.4)

- **hspace** -> Horizontal space between 2 plots
- **vspace** --. Vertical space

LINE PLOT Example5:

Question: Draw two line graph where the x axis shows the individual classes and the y axis shows the number of students participating in ART/COMPUTER interhouse event, in two different plots

```
•import matplotlib.pyplot as plt
x=[4,5,6,7]
y=[6,10,14,13]
z=[10,12,18,20]
plt.subplot(2,1,1)
plt.plot(x,y,'g',linestyle="--", linewidth=5)
plt.title(" Students participating for ART")
plt.xlabel("CLASS", fontsize="8",color="red")
plt.ylabel("NO OF STUDENTS PARTICIPATING", fontsize="8", color="blue")
plt.grid(True)
plt.show()

plt.subplot(2,1,2)
plt.plot(x,z,'m',linestyle="dotted")
plt.title(" Students participating for COMPUTERS")
plt.xlabel("CLASS", fontsize="8",color="red")
plt.ylabel("NO OF STUDENTS PARTICIPATING", fontsize="8", color="blue")
plt.grid(True)
plt.show()
```