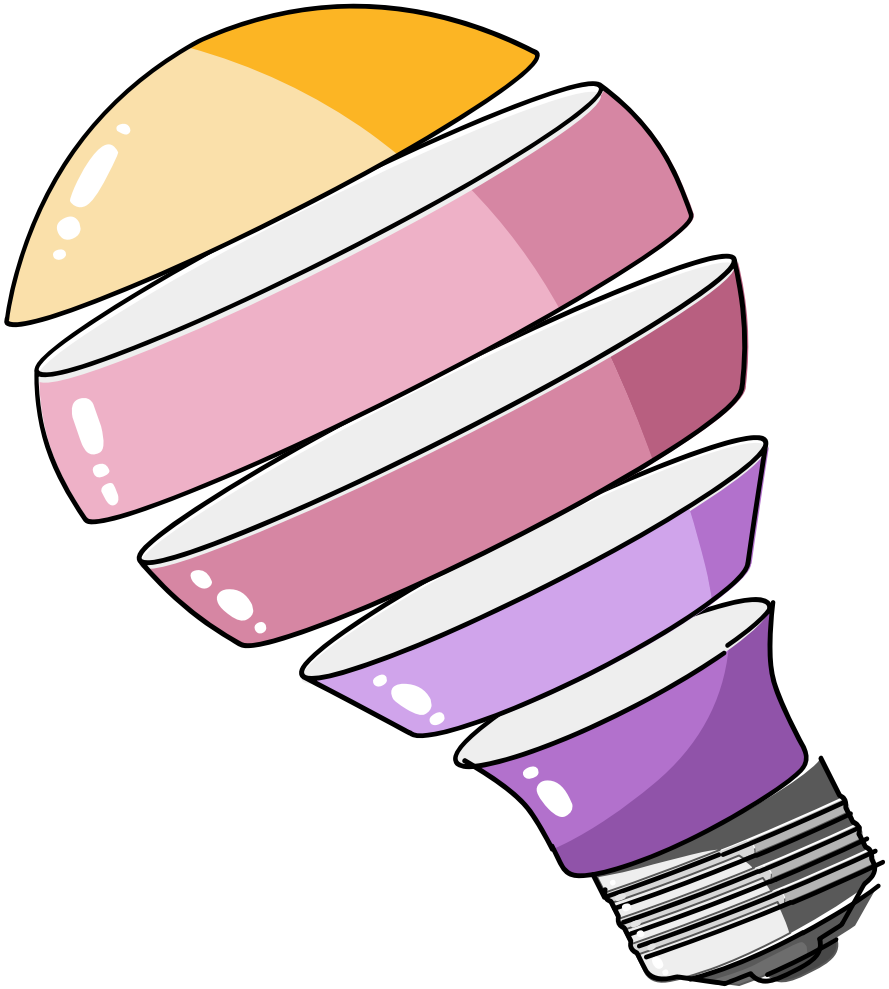
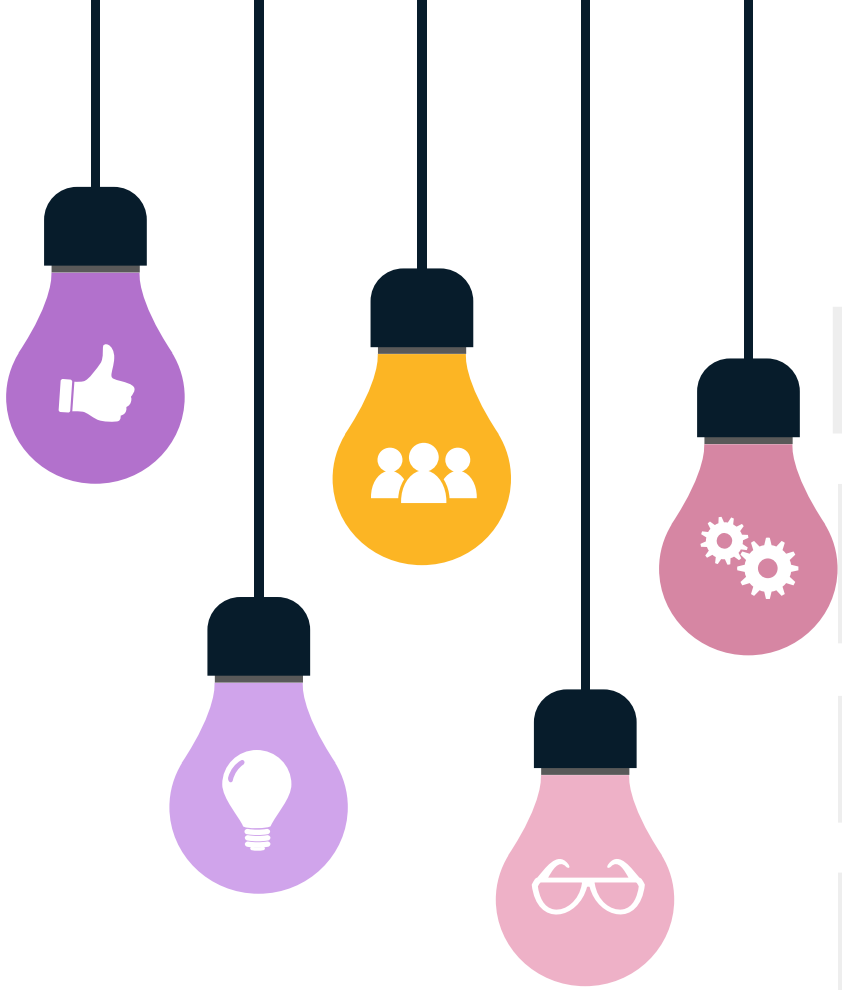


Python Libraries for ML

Anaconda installation



Steps for Linux

Download Anaconda
https://www.anaconda.com/products/individual

1

Download Essential Libraries
sudo apt-get install libgl1-mesa-glx libegl1-mesa libxrandr2 libxrandr2 libxss1 libxcursor1 libxcomposite1 libasound2 libxi6 libxtst6

2

Install Anaconda
bash ~/Downloads/Anaconda3-2020.02-Linux-x86_64.sh

3

Incase base is not visible

- source /home/neelup/anaconda3/bin/activate

4

Steps for Windows

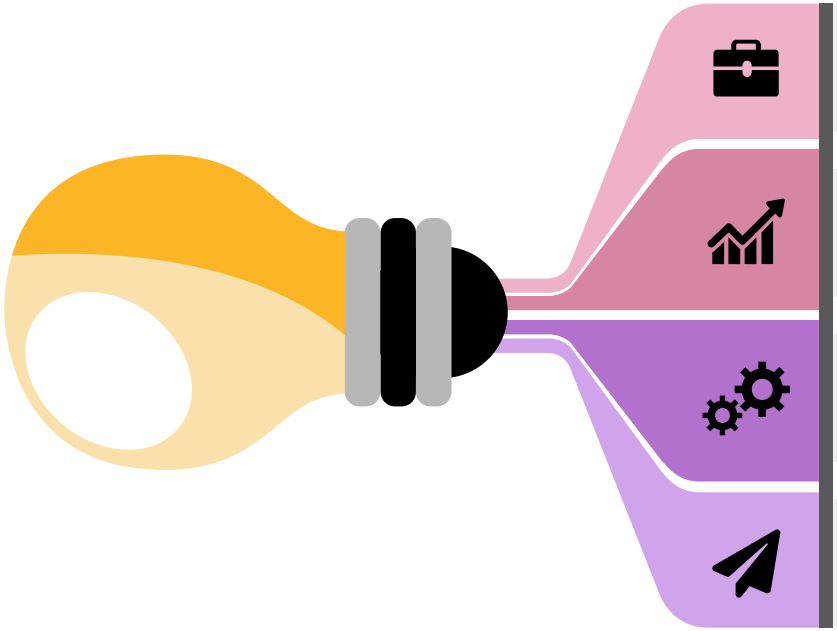
https://www.anaconda.com/products/individual	1
Double click on the Anaconda exe and click on “**Next**”	2
Check the box “**Add Anaconda3 to my PATH environment variable**”	3
Open the powershell and run **jupyter notebook** command	4

Numpy

Agenda for Numpy

- What is Numpy?
- Numpy vs List
- Numpy Operation
- Numpy Special functions

What is Numpy



Numpy is the core library for scientific computing in python. 1

It provides a high performance multi dimensional array object, and tools for working with these arrays. 2

Useful in Linear Algebra, Fourier Transform, and Random Number capabilities 3

Installation:- **pip install numpy** 4

Numpy vs List

We use python NumPy array instead of a list because of the below three reasons :-

Less Memory

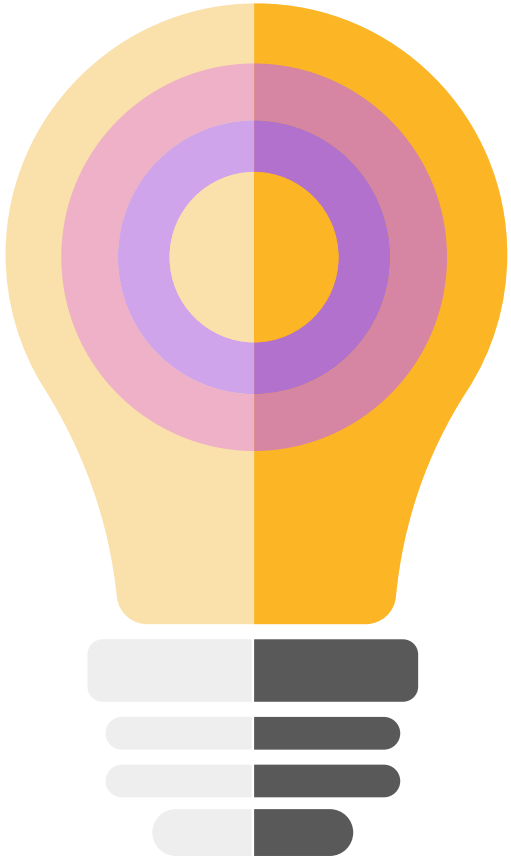
Numpy uses much less memory to store data

Fast

Numpy is around 5 to 100 times faster than the standard List

Convenient

Numpy is efficient and easy to use



Numpy Operations

- How to find the dimension of the array

ndim()

- How to find the size of each elements

itemsize()

- Find the data type of the elements

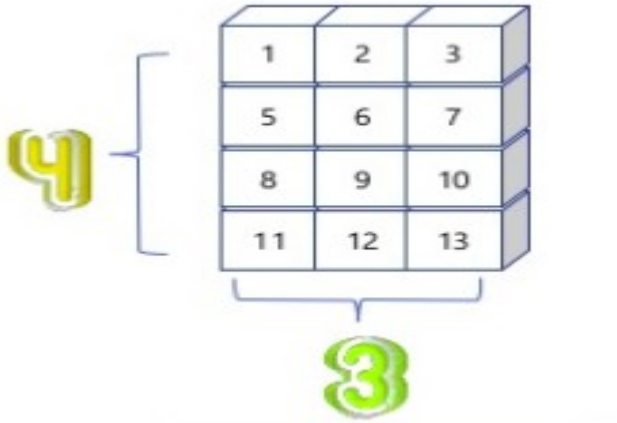
dtype()

Continuation..

- How to find out the number of elements in the array.

size()

- How to find out the shape of an array

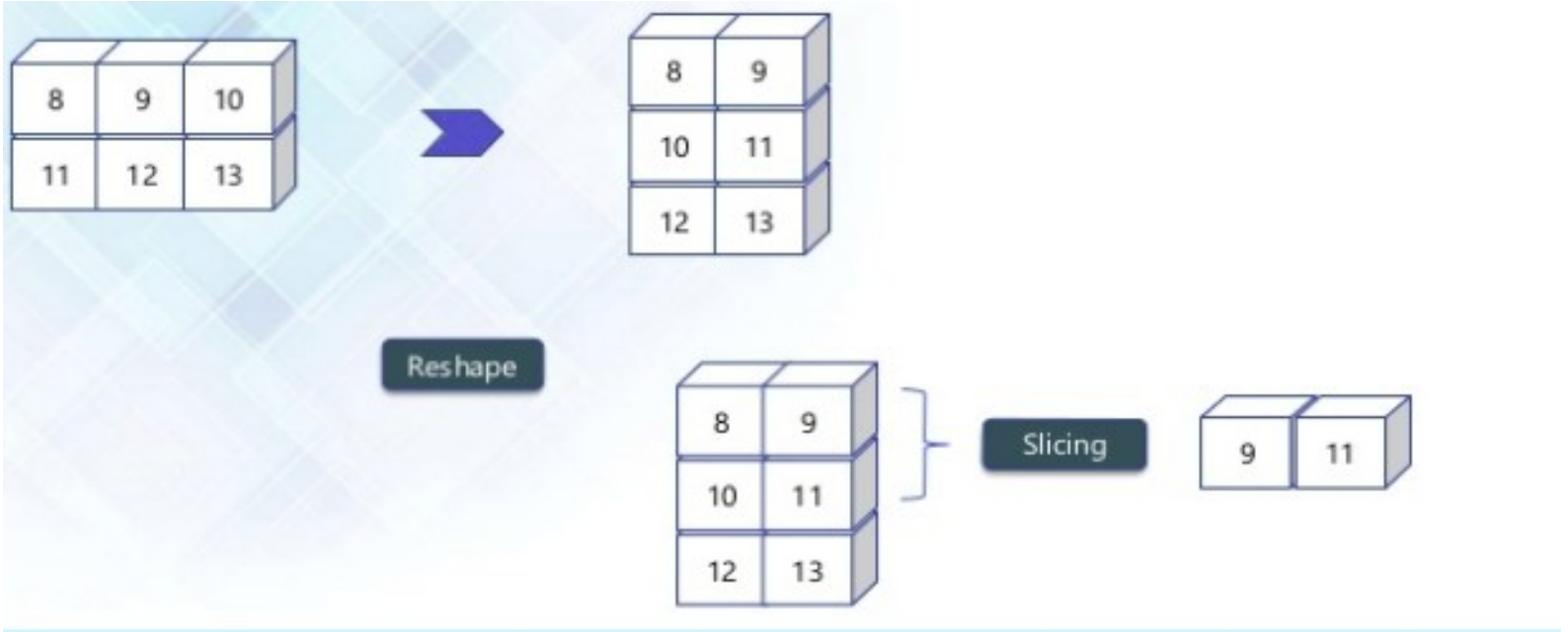


shape()

Continuation..

- Reshape and Slicing

reshape() for reshaping the numpy array

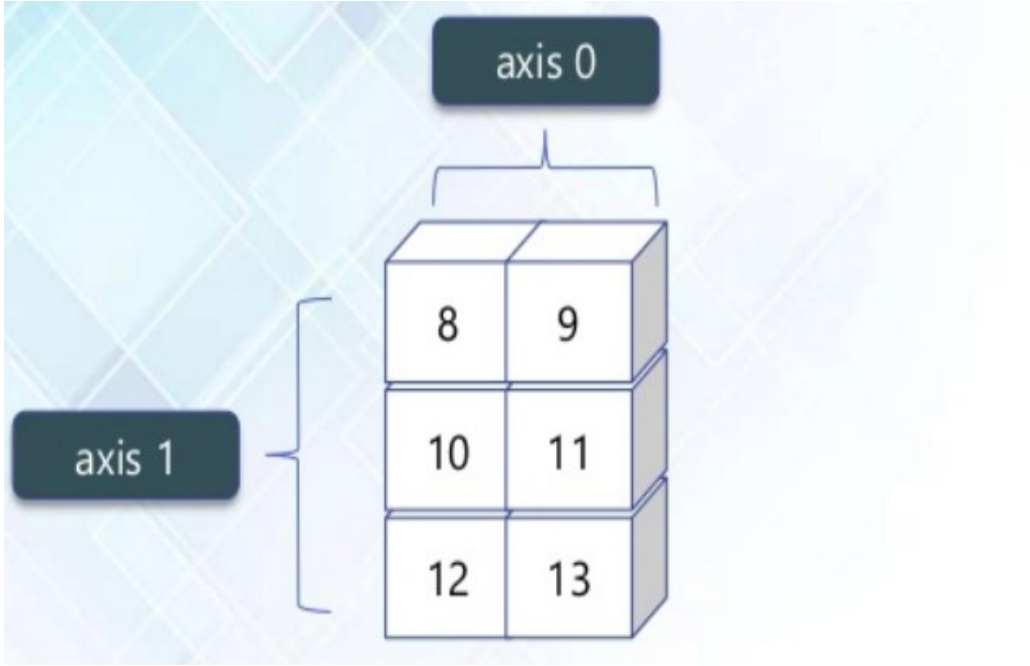


Continuation..

- Min, Max, Sum Operation

Continuation..

- Sum of axis 0 and axis 1

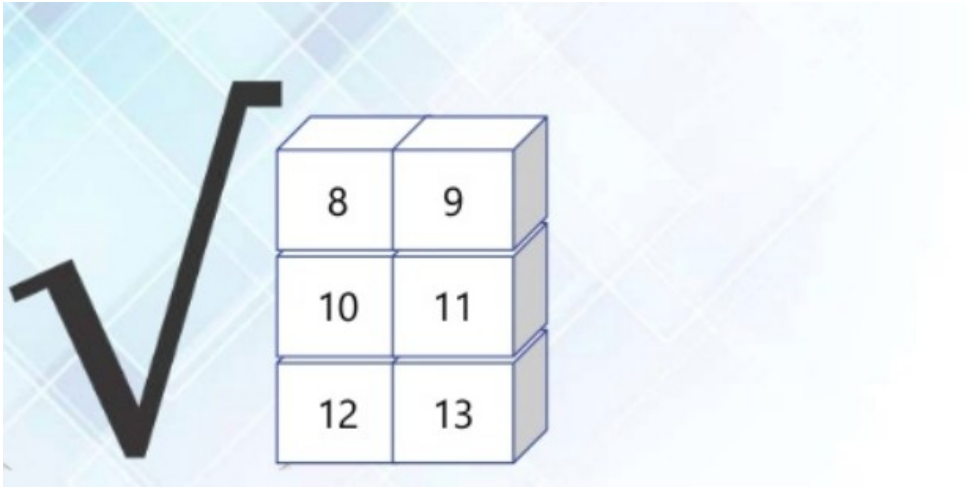


Sum of axis 0: [30, 33]

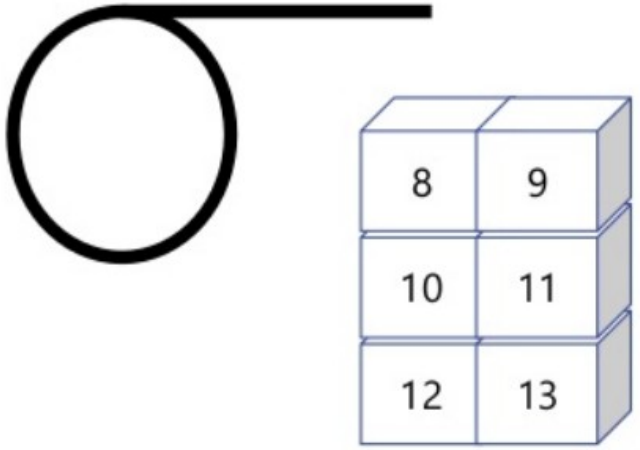
Sum of axis 1: [17, 21, 25]

Continuation..

- How to find out square root.

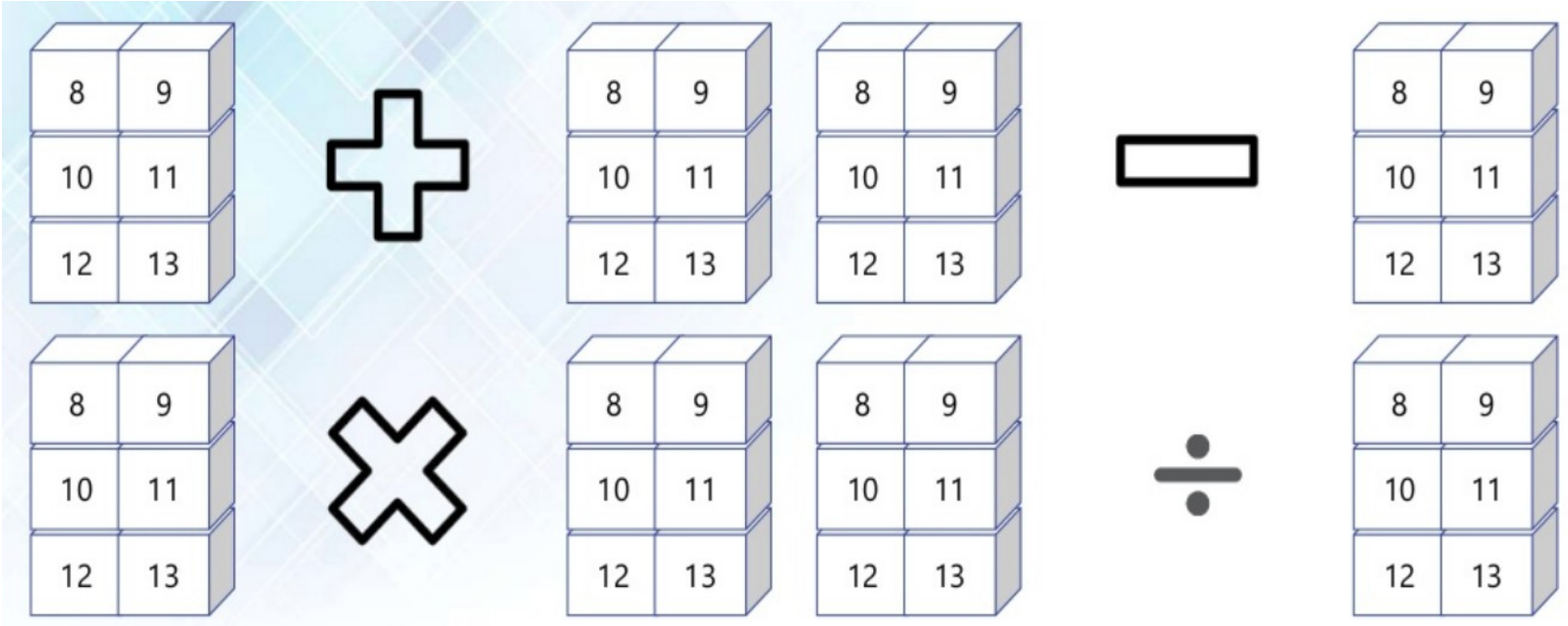


- How to find out standard deviation.



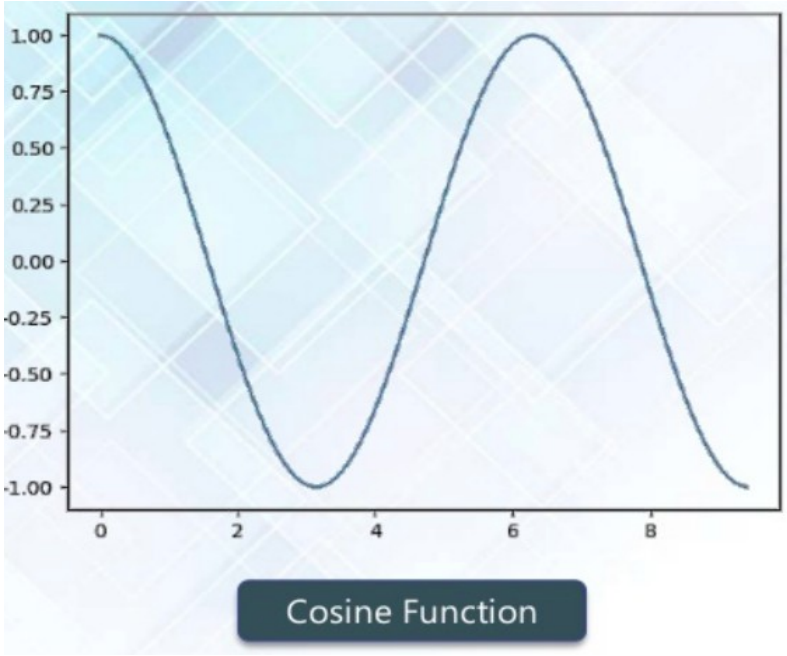
Continuation..

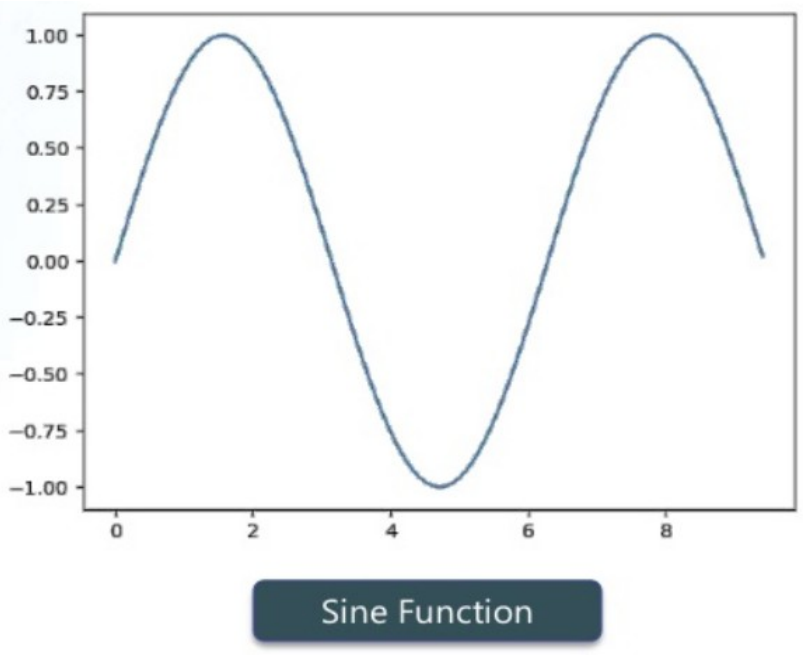
- Array addition, subtraction, multiplication and division



Continuation..

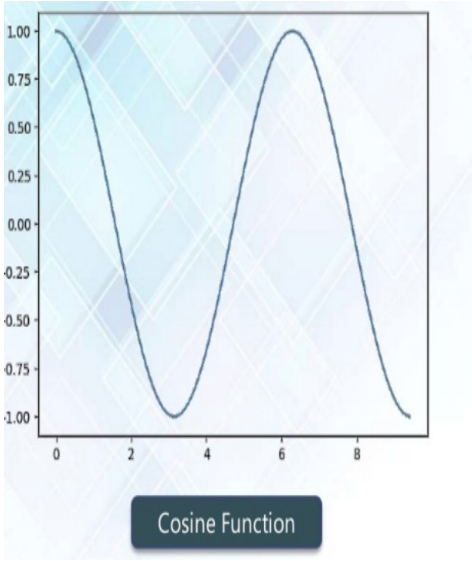
- Special Functions

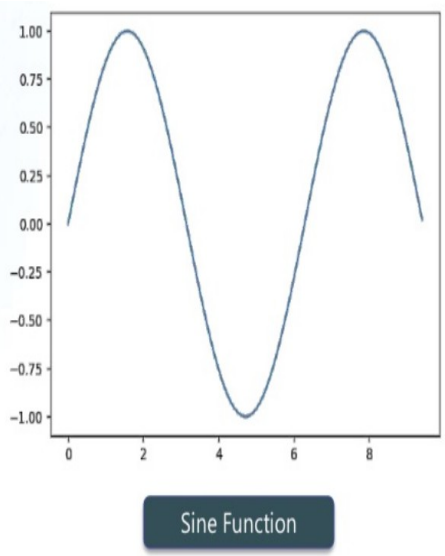




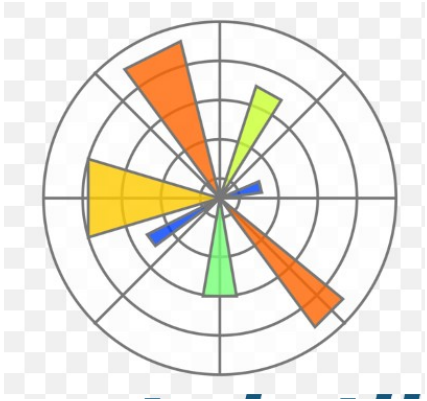
Continuation..

- Special Functions





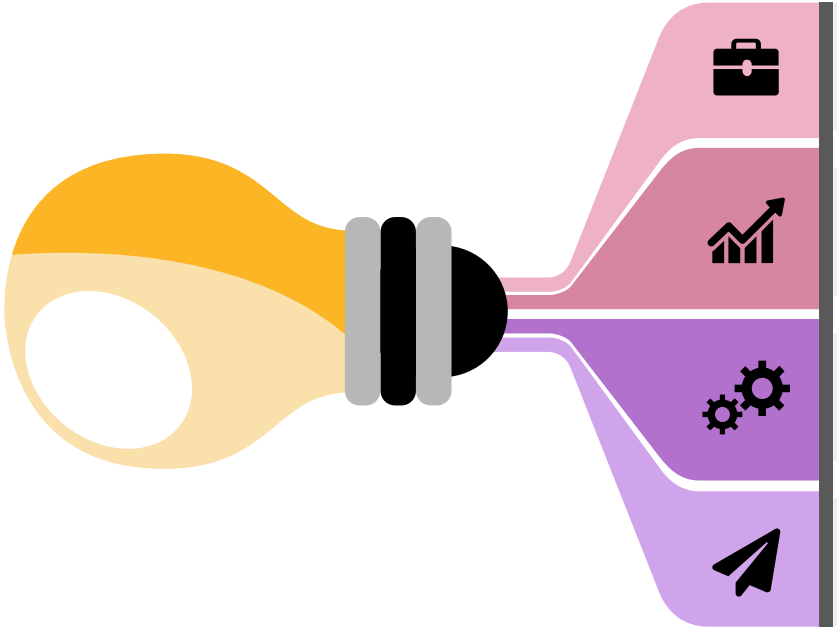
Matplotlib



Agenda for Matplotlib

- What is Matplotlib?
- Why it is required?
- Types of plots
- Getting started with matplotlib

What is Matplotlib?



Matplotlib is a cross-platform, data visualization and graphical plotting library for Python

1

It can be used in python scripts, shell, web application servers and other graphical user interface toolkits.

2

it is a robust, free and easy library for data visualization.

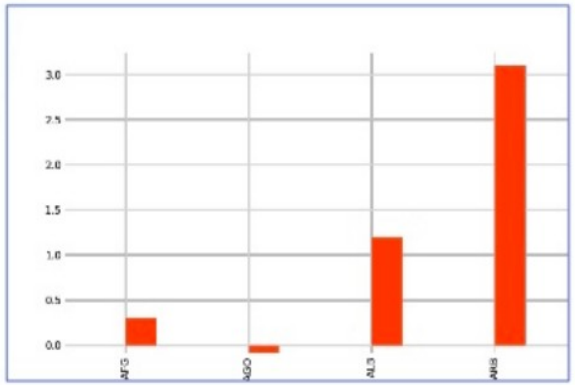
3

Installation:- **pip install matplotlib**

4

Continuation..

Country Name	Country Code	2010	2011	2012	2013	2014
Afghanistan	AFG	20.6	20.9	19.7	21.1	20.8
Angola	AGO	10.8	10.7	10.7	10.6	10.5
Albania	ALB	25.799999	27	28.3	28.7	29.2
Arab World	ARB	25.022214	28.11752	29.11321	29.33531	29.70457
United Arab Emirates	ARE	9.8000002	9.8	9.8	9.9	10
Argentina	ARG	19.5	18.8	18.4	19.7	21.3
Armenia	ARM	38.299999	38.7	35	32.5	35.1
Australia	AUS	11.4	11.4	11.7	12.2	13.1
Austria	AUT	8.8000002	8.2	8.7	9.1	9.2
Azerbaijan	AZE	14.6	14.5	14.3	13.4	13.6
Burundi	BDI	10.8	10.8	10.8	10.8	10.7
Belgium	BEL	22.5	18.6	19.7	23.1	23.6
Benin	BEN	2	2	2	1.8	1.7
Burkina Faso	BFA	5.1999998	5.3	5.2	5.2	5
Bangladesh	BGD	8.1999998	8.2	8.2	8.9	9.1
Bulgaria	BGR	22.9	25.2	28.2	29.7	25.9
Bahrain	BHR	10.2	11.4	10.5	10.6	10.9
Bahamas, The	BHS	36	27.2	30.4	30.8	30.1
Bosnia and Herzegovina	BIH	57.200001	57.1	61.7	57.4	57.5
Belarus	BLR	13.2	12.5	11.8	12	12
Belize	BLZ	20.9	24.3	26	22.4	22



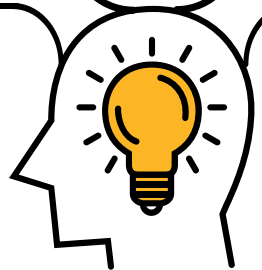
Why matplotlib is required ?

Matplotlib is effective when we integrate to use with other GUI toolkits.

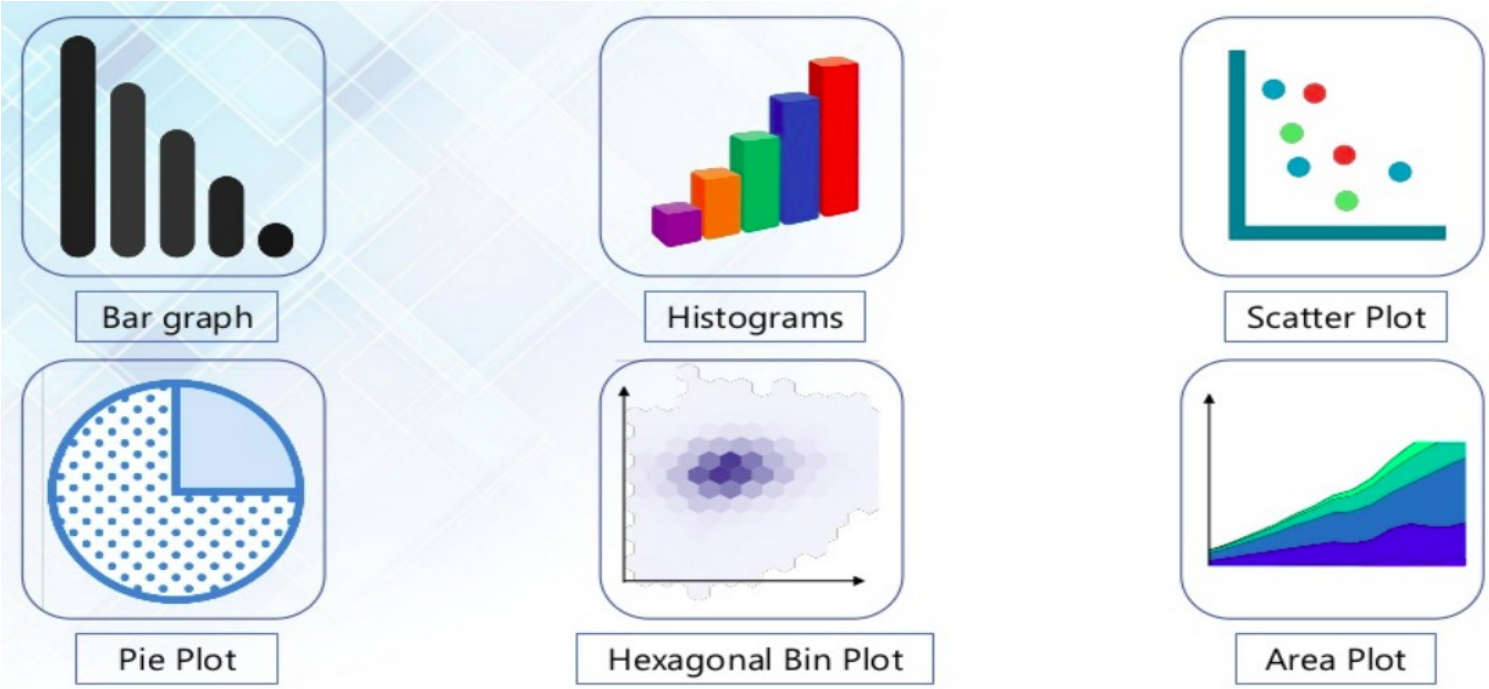
Matplotlib to ease the analysis of statistical data.

Matploitlib is a Python Library used for plotting, This python library provides an object-oriented APIs for integrating plots into applications.

Matplotlib is an effective replacement for the MatLab tool.It contains all the requirements for replacements.



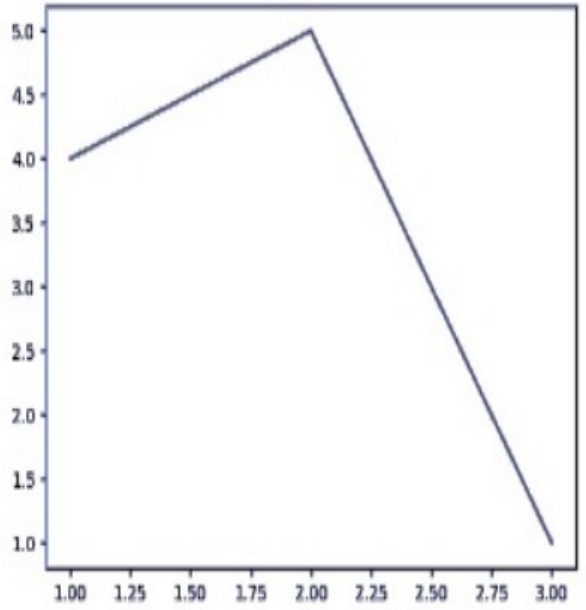
Types of Plots



Getting Started with Matplotlib

- How to create a simple graph.

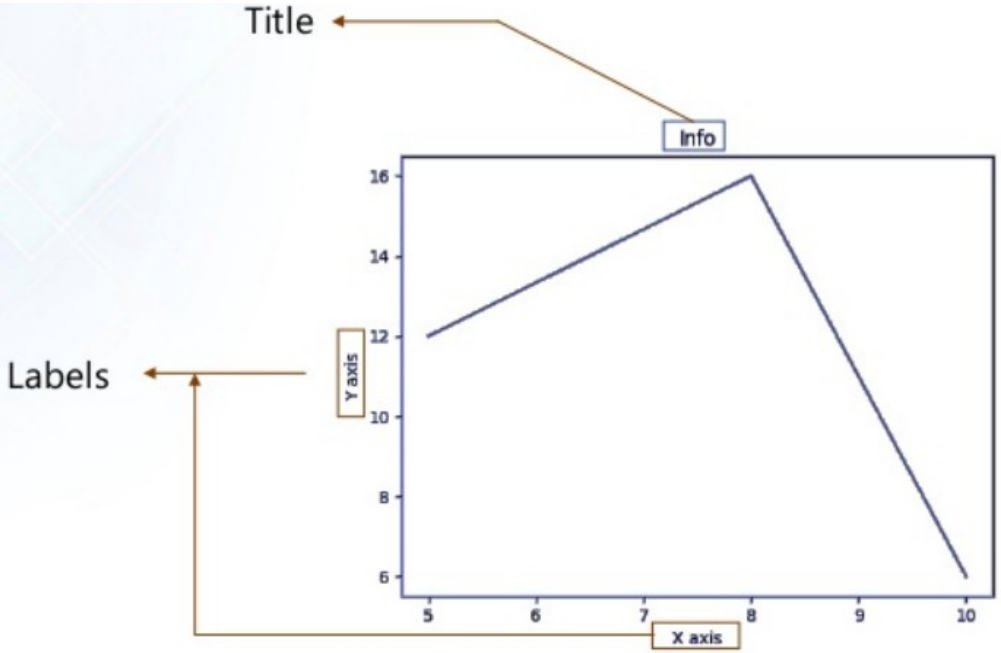
```
from matplotlib import pyplot as plt

#Plotting to our canvas

plt.plot([1,2,3],[4,5,1])

#Showing what we plotted

plt.show()
```



Continuation..

- Add title and label to the graph

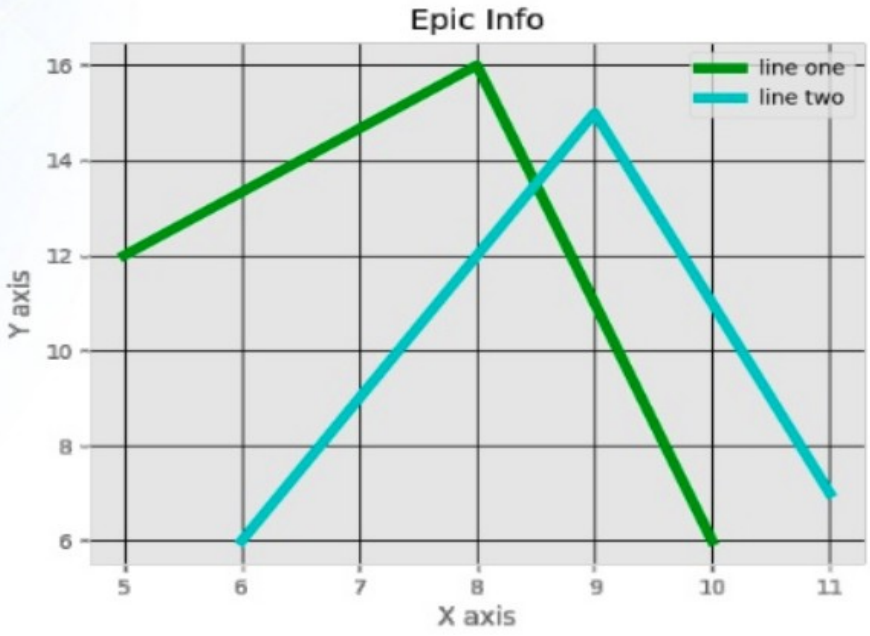
```
from matplotlib import pyplot as plt

x = [5,8,10]
y = [12,16,6]

plt.plot(x,y)

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

plt.show()
```



Continuation..

- Adding Style to the Graph

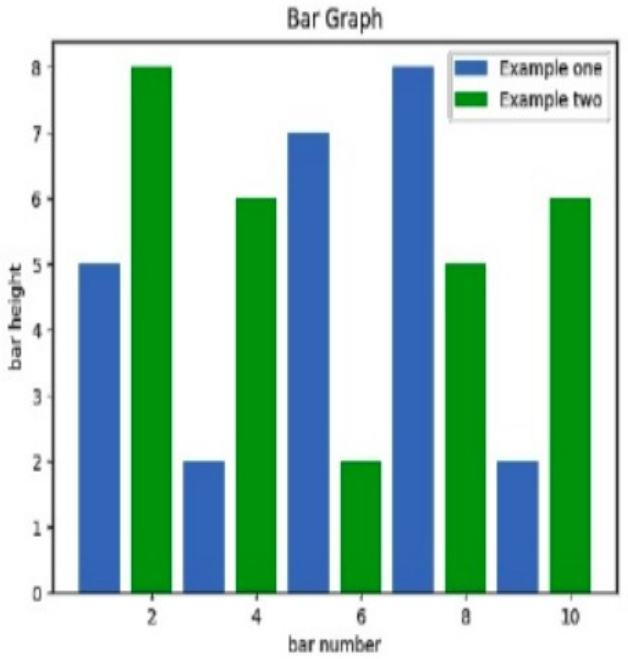
```
from matplotlib import pyplot as plt
from matplotlib import style

style.use('ggplot')

x = [5,8,10]
y = [12,16,6]

x2 = [6,9,11]
y2 = [6,15,7]

plt.plot(x,y,'g',label='line one', linewidth=5)
plt.plot(x2,y2,'c',label='line two',linewidth=5)

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

plt.legend()

plt.grid(True,color='k')

plt.show()
```



Continuation..

- Bar Graph

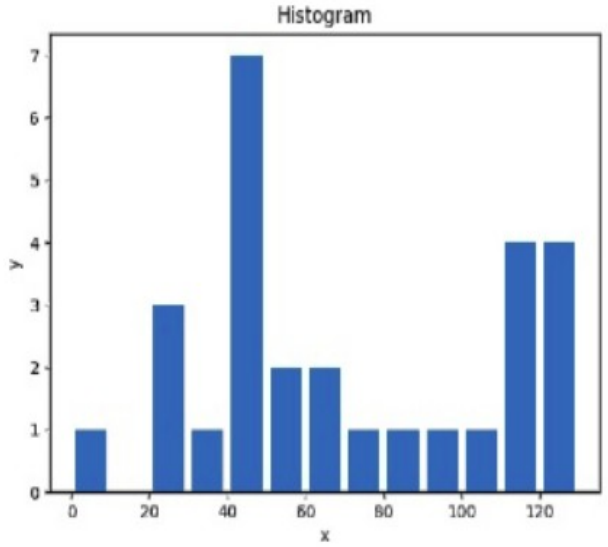
```
import matplotlib.pyplot as plt

plt.bar([1,3,5,7,9],[5,2,7,8,2], label="Example one")

plt.bar([2,4,6,8,10],[8,6,2,5,6], label="Example two", color='g')
plt.legend()
plt.xlabel('bar number')
plt.ylabel('bar height')

plt.title('Info')

plt.show()
```



Continuation..

- Histogram

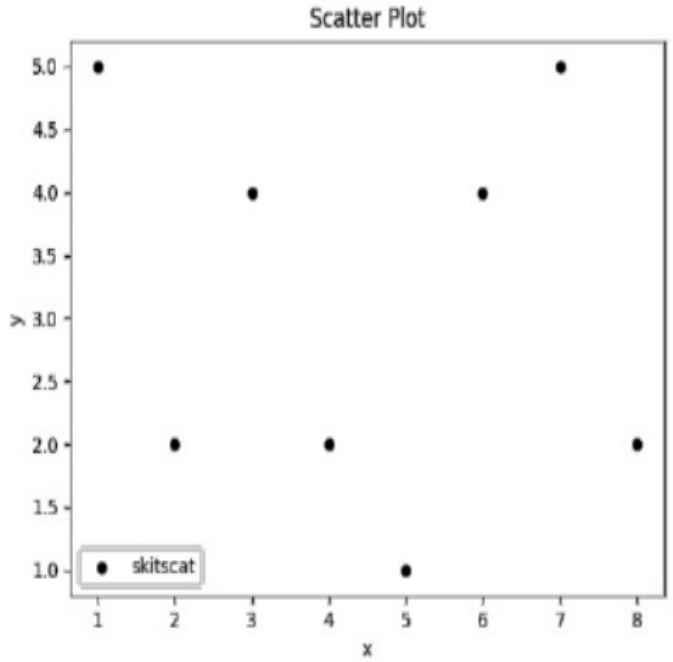
```
import matplotlib.pyplot as plt

population_ages =
[22,55,62,45,21,22,34,42,42,4,99,102,110,120,121,122,130,111,115,112,80,75,6
5,54,44,43,42,48]

bins = [0,10,20,30,40,50,60,70,80,90,100,110,120,130]

plt.hist(population_ages, bins, histtype='bar', rwidth=0.8)

plt.xlabel('x')
plt.ylabel('y')
plt.title('Histogram')
plt.legend()
plt.show()
```



Continuation..

- Scatter Plot

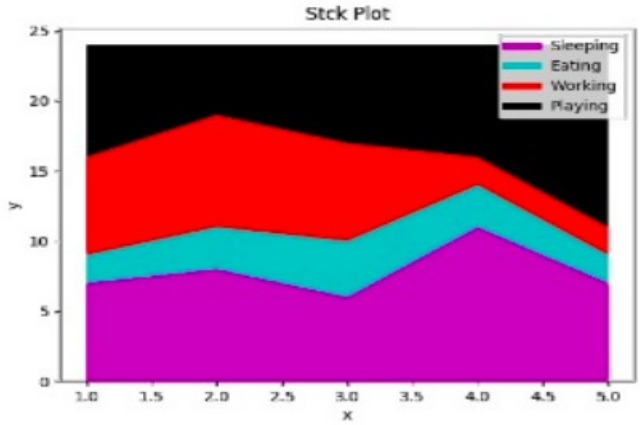
```
import matplotlib.pyplot as plt

x = [1,2,3,4,5,6,7,8]
y = [5,2,4,2,1,4,5,2]

plt.scatter(x,y, label='skitscat', color='k')

plt.xlabel('x')
plt.ylabel('y')
plt.title('Scatter Plot')
plt.legend()
plt.show()
```



Continuation..

- Stack Plot

```
import matplotlib.pyplot as plt

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

sleeping = [7,8,6,11,7]
eating =   [2,3,4,3,2]
working =  [7,8,7,2,2]
playing =  [8,5,7,8,13]

plt.plot([],[],color='m', label='Sleeping', linewidth=5)
plt.plot([],[],color='c', label='Eating', linewidth=5)
plt.plot([],[],color='r', label='Working', linewidth=5)
plt.plot([],[],color='k', label='Playing', linewidth=5)

plt.stackplot(days, sleeping,eating,working,playing, colors=['m','c','r','k'])

plt.xlabel('x')
plt.ylabel('y')
plt.title('Stck Plot')
plt.legend()
plt.show()
```



Continuation..

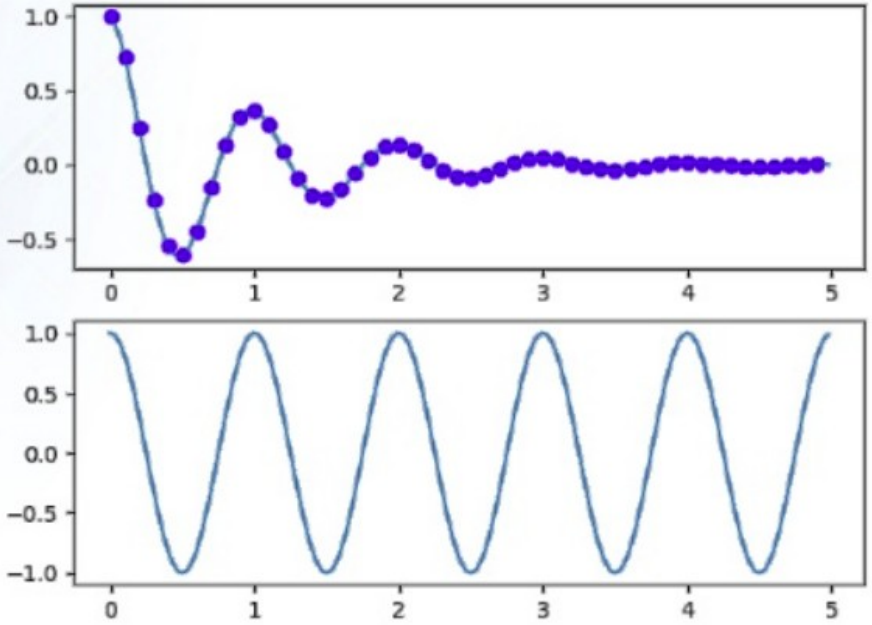
- Pie Chart

```
import matplotlib.pyplot as plt

slices = [7,2,2,13]
activities = ['sleeping','eating','working','playing']
cols = ['c','m','r','b']

plt.pie(slices,
        labels=activities,
        colors=cols,
        startangle=90,
        shadow= True,
        explode=(0,0.1,0,0),
        autopct='%1.1f%%')

plt.title('Pie Plot')
plt.show()
```

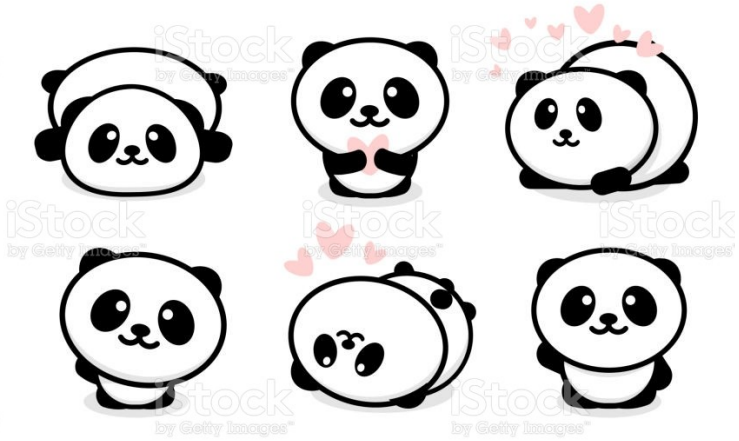
Continuation..

- Working with multiple charts

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

def f(t):
    return np.exp(-t) * np.cos(2*np.pi*t)

t1 = np.arange(0.0, 5.0, 0.1)
t2 = np.arange(0.0, 5.0, 0.02)

plt.subplot(211)
plt.plot(t1, f(t1), 'bo', t2, f(t2))

plt.subplot(212)
plt.plot(t2, np.cos(2*np.pi*t2))
plt.show()
```



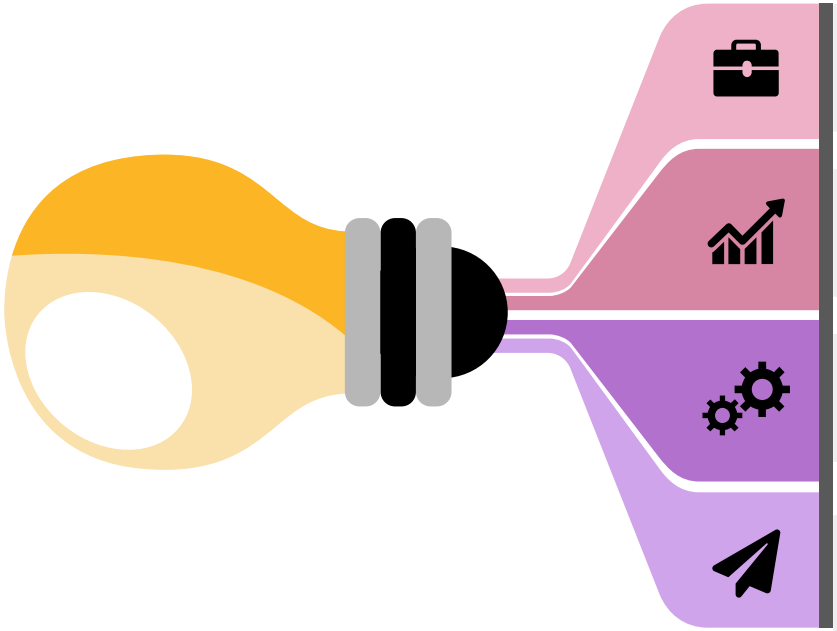
Pandas



Agenda for Pandas

- What is Pandas?
- Why Pandas is required ?
- Different Ways Of Creating Dataframe
- Basic Operations on Dataframe
- Handling Missing Data
- Filter pandas dataframe

What is Pandas ?



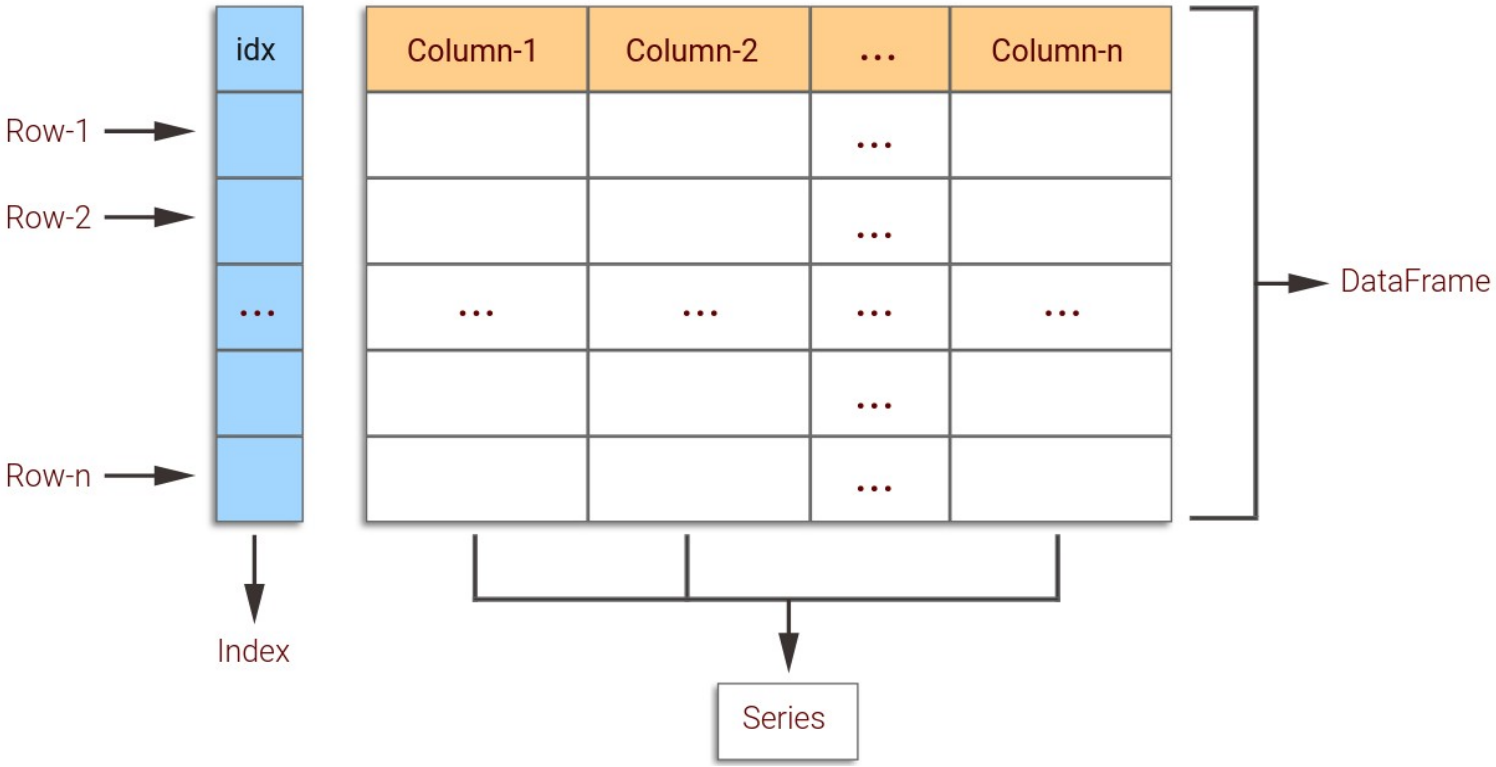
pandas is a fast, powerful, flexible and easy to use open source data analysis. 1

It is a manipulation tool, built on top of the Python programming language. 2

It has functions for analyzing, cleaning, exploring, and manipulating data. 3

Installation:- **pip install pandas** 4

Continuation..



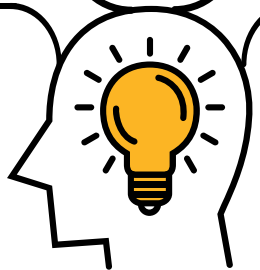
Why Pandas is required ?

Cross Platform and Easy to use

High-performance

Provide data analysis tools

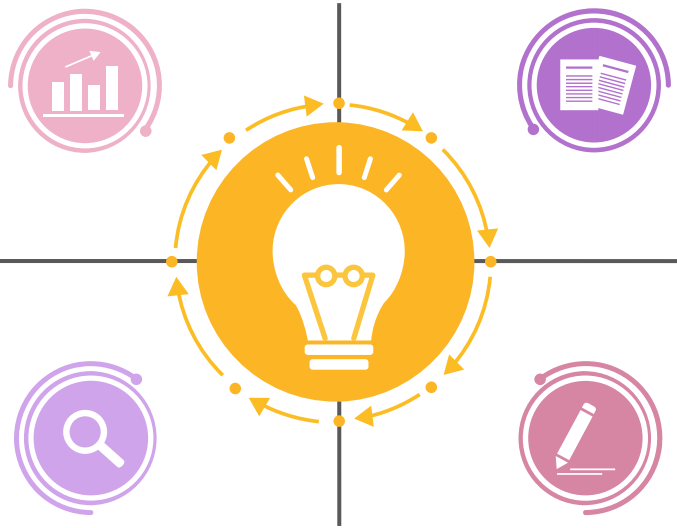
Provides data filtration.



Different Ways Of Creating Dataframe

Using CSV

Using EXCEL

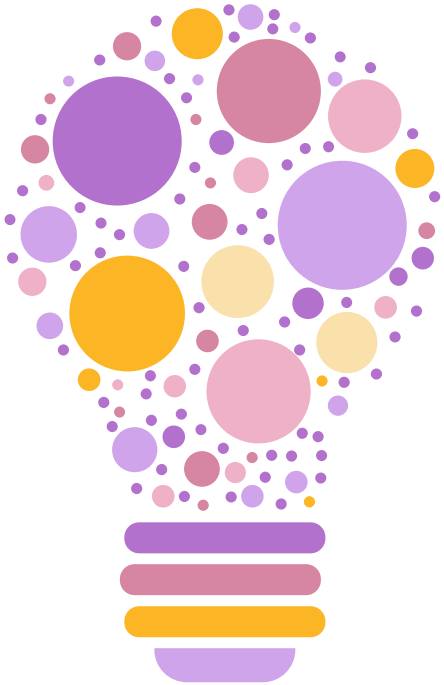


Using Dictionary

Using Tuple list

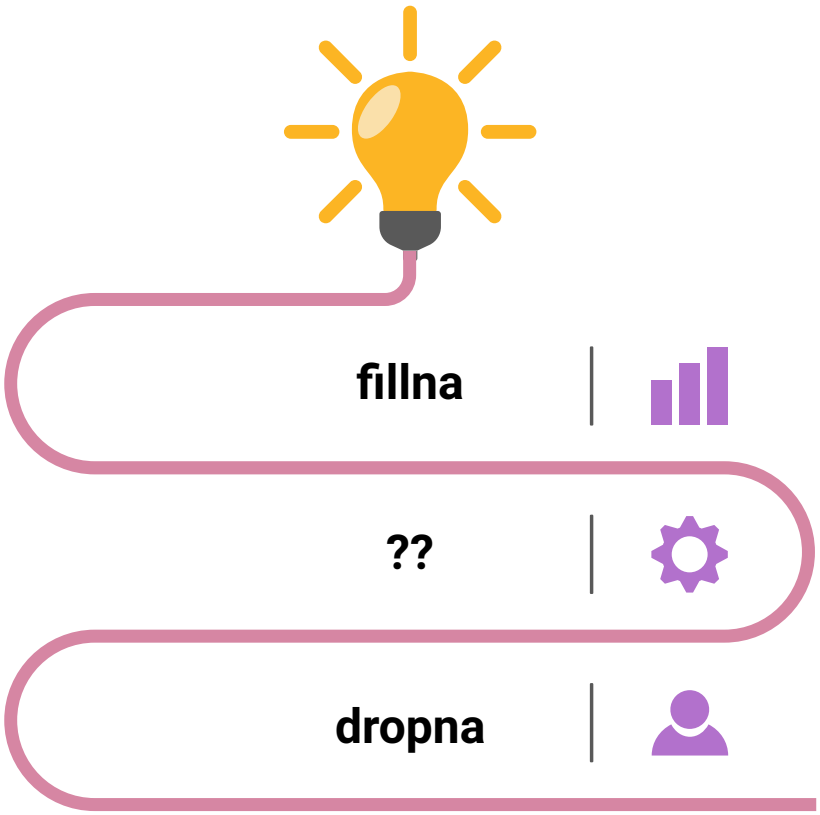
Basic Operations on Dataframe

- Shape, head and tail of dataframe
- Slice the rows of dataframe
- Extract one and multiple columns



- Min, max and std on dataframe
- Logical operations on dataframe
- Filter out data based on requirement

Handling Missing Data



This function is used to fill NA/NaN values using the specified method

Find out any method which used to handle the missing data

It allows the user to analyze and drop Rows/Columns with Null values in different ways

Thank you