# Objective:

**Practical No. 3 NumPy, Pandas**

Write Python program to demonstrate use of built-in packages (e.g. NumPy, Pandas)

# Practical Significance

Though Python is simple to learn language but it also very strong with its features. As mentioned earlier Python supports various built in packages. Apart from built-in package user can also make their own packages i.e. User Defined Packages. **Numpy** is a general-purpose array-processing package. It provides a high-performance multidimensional array object, and tools for working with these arrays. This practical will allow students to write a code.

# Minimum Theoretical Background

NumPy, which stands for Numerical Python, is a library consisting of multidimensional array objects and a collection of routines for processing those arrays. Using NumPy, mathematical and logical operations on arrays can be performed.

Steps for Installing numpy in windows OS

* 1. goto Command prompt
  2. run command pip install numpy
  3. open IDLE Python Interpreter
  4. Check numpy is working or not

>>> import numpy

>>> import numpy as np

>>> a=np.array([10,20,30,40,50])

>>> print(a)

[10 20 30 40 50]

# Example:

>>> student=np.dtype([('name','S20'),('age','i1'),('marks','f4')])

>>> a=np.array([('Vijay',43,90),('Prasad',38,80)],dtype=student)

>>> print(a)

[('Vijay', 43, 90.) ('Prasad', 38, 80.)]

# Example:

>>> print(a)

[10 20 30 40 50 60]

>>> a.shape=(2,3)

>>> print(a) [[10 20 30]

[40 50 60]]

>>> a.shape=(3,2)

>>> print(a) [[10 20]

[30 40]

[50 60]]

# What is Pandas?

Pandas is a Python library used for working with data sets.

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

The name "Pandas" has a reference to both "Panel Data", and "Python Data Analysis" and was created by Wes McKinney in 2008.

# Why Use Pandas?

Pandas allows us to analyze big data and make conclusions based on statistical theories. Pandas can clean messy data sets, and make them readable and relevant.

Relevant data is very important in data science.

# Exercise

* 1. Write a Python program to create two matrices and perform addition, subtraction, multiplication and division operation on matrix.
  2. Write a NumPy program to generate six random integers between 10 and 30.
  3. Write a python code to read a csv file using pandas module and print the first five and last five lines of a file.
  4. Write a program code to compute summary statistics such as mean, median, mode, standard deviation and variance of the given dataset.
  5. Write a python code to read a csv file using pandas module and preprocess the dataset for null, duplicate values.