**Experiment 1**

**Objective:**

To study about numpy, pandas and matplotlib libraries in python.

**Theory**

In Python programming, libraries such as NumPy, Pandas, and Matplotlib are fundamental for data manipulation, analysis, and visualization. These libraries streamline complex operations, enabling data scientists, analysts, and developers to handle large datasets, perform numerical computations, and visualize data effectively.

**NumPy (Numerical Python)** is a powerful library primarily used for numerical and matrix computations. It introduces the ndarray object, an N-dimensional array for efficiently storing and manipulating large arrays of homogeneous data. NumPy arrays are faster and more memory-efficient than traditional Python lists due to their fixed size and storage of elements in contiguous memory. This efficiency allows for rapid mathematical computations and operations across entire arrays without the need for explicit loops. NumPy also provides a range of mathematical functions, including linear algebra, Fourier transforms, and random number generation. The broadcasting feature in NumPy enables arithmetic operations on arrays of different shapes, making it flexible for various data manipulation tasks. NumPy is often used in fields like machine learning, scientific computing, and engineering, where heavy numerical computation is required.

**Pandas** is another essential library in Python, designed specifically for data manipulation and analysis. It provides two primary data structures: Series (1D) and DataFrame (2D), which are built on top of NumPy arrays. A DataFrame can hold heterogeneous data types across columns, making it ideal for handling and analyzing structured data. Pandas offers a variety of functions for data cleaning, filtering, grouping, merging, and aggregation. It supports handling missing data, which is a common issue in real-world datasets. Pandas also has tools for time series analysis, making it valuable for financial and temporal data. The ability to handle large datasets in-memory, apply various operations on data frames, and reshape data makes Pandas a powerful tool for any data-driven task. Through Pandas, data can be imported from numerous file formats, such as CSV, Excel, and SQL databases, facilitating easy data integration and analysis.

**Matplotlib** is a popular Python library for data visualization. Its primary goal is to provide an easy way to generate visual representations of data, making it easier to understand complex patterns and relationships. Matplotlib offers an extensive range of plotting options, including line plots, scatter plots, bar charts, histograms, and more. The library’s core, pyplot, provides a MATLAB-like interface for creating interactive and customizable plots. It allows users to control various aspects of a plot, such as labels, colors, styles, and legends, enabling high levels of customization. Matplotlib is often combined with Pandas, as Pandas’ built-in plotting functions use Matplotlib as the backend, simplifying the visualization of DataFrames directly. Visualizations are essential in data analysis, as they provide insights and highlight trends that may not be evident through raw numbers alone.

Together, NumPy, Pandas, and Matplotlib form the foundation of Python’s data science stack. NumPy handles numerical computations, Pandas manages data manipulation, and Matplotlib provides visualization tools, making it easier for data scientists to analyze and interpret data.

**Result**

As a result of this Experiment, we successfully wrote and executed the program to study about numpy, pandas and matplotlib libraries in python.

**Learning Outcomes**

Understand and utilize Python libraries NumPy, Pandas, and Matplotlib for numerical operations, data manipulation, and data visualization in data science tasks.