Task 4: Basics of NumPy and Pandas

What is NumPy

NumPy (Numerical Python) is a powerful open-source Python library used for scientific computing. It provides support for large multi-dimensional arrays and matrices, along with a collection of high-level mathematical functions to operate on these arrays efficiently.

Why NumPy is better than traditional methods like lists

- Operations are vectorised, making them significantly faster than looping over Python lists.
- Memory efficient — stores data in contiguous blocks with fixed types.
- Supports broadcasting, enabling arithmetic between arrays of different shapes.
- Includes a vast set of built-in functions (e.g., mean, std, dot product) optimised in C.

Applications of NumPy

- Linear algebra, Fourier transforms, and numerical integration.
- Data pre-processing and statistical analysis.
- Used as the foundation for libraries like Pandas, Scikit-learn, and TensorFlow.
  Simulation and modeling in physics, finance, and engineering.

What is Pandas

Pandas is a Python library for data manipulation and analysis. It makes it easy to handle structured data, similar to how tables are used in Excel or SQL.

Task 4: Basics of NumPy and Pandas

How it is better than traditional methods

- Offers label-based indexing, making data selection intuitive and powerful.
- Simplifies data cleaning, transformation, and aggregation.
- Integrates well with NumPy and Matplotlib, forming a complete data pipeline.
- Efficient handling of missing or duplicate data.

Use of Pandas in real-world applications

- Financial analysis and stock market predictions.
- Cleaning and analyzing large-scale business data in Excel-like workflows.
- Backend processing of data for machine learning models.
- Real-time data dashboards and reporting tools.

What is Matplotlib

Matplotlib is a comprehensive 2D plotting library for Python that enables the creation of static, interactive, and animated visualizations. It is often used alongside NumPy and Pandas for visual data analysis.

How it is better than traditional methods

- Allows for highly customizable plots, from basic line graphs to complex heatmaps.
- Can export visuals in high-quality formats (PNG, SVG, PDF).
- Works seamlessly with NumPy arrays and Pandas DataFrames.
- Offers both low-level control (via pyplot) and high-level plotting APIs (e.g., Seaborn).

Task 4: Basics of NumPy and Pandas

Use in real-world applications

- Data visualization in research papers and reports.
- Dashboard visual components in web apps.
  Visual analytics in finance, healthcare, and machine learning.
- Exploratory data analysis (EDA) for decision making.