1.NumPy

NumPy, short for Numerical Python, is a scientific computing library in Python that provides efficient array operations and mathematical computation capabilities. Its core component is the ndarray (N-dimensional array), which supports vectorized operations. Compared to native Python lists, it is faster and more memory-efficient.

After installing NumPy, we can retrieve specific elements using commands like print(arr[X]), or modify elements in a matrix with expressions like matrix[X][Y] = value.

We can also create special arrays using:

Zeros=np.zeros((3,3)) (create a 3\*3 matrix filled with zeros)ones=np.ones((2))(create a 2\*2 matrix filled with one ) Once matrices are created,we can perform various operations such as np.dot np.linalg.inv

2. Pandas

Pandas, short for Python Data Analysis Library, is a powerful library designed for handling tabular data. It is widely used for data analysis, cleaning, and transformation.

Pandas provides two main data structures:

• Series: A one-dimensional labeled array, similar to a NumPy array.

• DataFrame: A two-dimensional labeled data structure, similar to an Excel spreadsheet.

Using Pandas, we can read and write data with commands like:df=pd.read\_csv df=pd.to\_csv

To explore data,we can use print(df.head())(first 5 rows) print (df.tail())(last five rows) df.info(data information) df.describe()(statistical summary)

Pandas also supports:

• Handling missing values

• Removing duplicates

• Replacing values

Importantly, Pandas can be seamlessly integrated with Matplotlib for data visualization, making it very convenient for exploratory data analysis.

3. Difference Between NumPy and Pandas

Apart from their different core data structures and application scenarios, NumPy is generally faster than Pandas because Pandas is built on top of NumPy.

• If you’re dealing with arrays, matrices, and scientific computations, NumPy is the better choice.

• If you’re working with structured data like CSV or Excel files for cleaning or analysis, Pandas is the go-to tool.

Summary:

• Use NumPy for mathematical and numerical computations.

• Use Pandas for data analysis and manipulation.

Although their purposes differ, both are essential tools in data science and highly recommended for beginners to master.