NumPy

- Stands for Numerical Python
- It is one of the fundamental packages for mathematical, logical, and statistical operations with Python
- It contains
    - Powerful N-dimensional array object, called ndarray
    - Large set of functions for creating, manipulating, and transforming ndarrays
- ndarrays can only contain data of a single datatype
- Useful in linear algebra, vector calculus, random number capabilities, etc

Pandas

- Pandas is one of the fundamental packages for analysis and manipulation of tabular data
- Offers two major data structures - series & dataframe
- We can think of a pandas dataframe like an excel spreadsheet that is storing some data in rows and columns.
- A pandas dataframe is made up of several pandas series
- Each column of a dataframe is a series.
- Pandas dataframes can contain data of multiple datatypes

Common NumPy Functions

Function	Description
np.array()	To create an array
np.arange()	Return evenly spaced values within a given interval
np.linspace()	Return evenly spaced numbers over a specified interval
np.zeros()	To create an array of zeros
np.ones()	To create an array of ones
np.transpose()	Permute array dimensions

Common NumPy Functions

Function	Description
np.random.rand()	To create an array of specified shape filled with random values
np.random.randint()	Return random integers from low (inclusive) to high (exclusive)
np.random.randn()	Return a sample (or samples) from the “standard normal” distribution.
np.concatenate()	Concatenate two arrays
np.save()	Save an array to a binary file in .npy format.
np.savez()	Save several arrays into a single file in uncompressed .npz format.

Common Pandas Functions

Function	Description
pd.read_csv()	Read a comma-separated values (csv) file into DataFrame
df.loc[]	Access a group of rows and columns by label(s)
df.iloc[]	Purely integer-location based indexing for selection by position
df.drop()	Drop specified labels from rows or columns
pd.concat()	To concatenate two pandas objects
pd.merge()	To merge the pandas dataframes
df.groupby()	To split, apply or combine the data structures

Common Pandas Functions

Function	Description
df.value_counts()	To get count of some attributes
df.unique()	To get unique values
df.dtype	To get the data types
df.shape	To get the shape (number or rows and columns)
df.head()	To get the top rows
df.tail()	To get the last rows
df.describe()	To get the quick statistic summary

Common Python Libraries for Data Science

Library	Use
NumPy	Handling multi-dimensional arrays
Scipy	Scientific computation package
Matplotlib, Seaborn	Data visualisation
Pandas	Handling tabular data
Scikit-learn	Machine learning