DAY – 1

29 July 2025

In Python, a library is a group of modules that contain functions, classes and methods to perform common tasks like data manipulation, math operations, web scraping and more. Python libraries make coding faster, cleaner and more efficient by providing ready-to-use solutions for different domains such as data science, web development, machine learning and automation.

Types of Python Library

Python libraries are divided into two main types:

1. Built-in Python Standard Library

It is a collection of modules that come bundled with every Python installation; we don't need to install anything separately. Most of these modules are written in C for better performance.

Examples of built-in modules:

- Math: Mathematical operations
- os: Interact with the operating system
- date time: Date and time operations
- Random: Generate random numbers
- json: Handles JSON data encoding and decoding.

2. External Python Libraries

External (third-party) libraries are not included with Python by default. You can install them easily using the pip package manager. Popular External Python Libraries:

- NumPy: It, short for Numerical Python, is the core library for numerical and scientific computing in Python. It provides powerful tools for creating and manipulating arrays, matrices and multidimensional data.

- Pandas: It is a data analysis and manipulation library built on top of NumPy. It introduces data structures like DataFrame and Series that make it easy to handle structured data efficiently.
- Matplotlib: It is a data visualization library that helps you create a wide variety of static, animated and interactive plots. It supports charts such as bar graphs, line charts, histograms and scatter plots.
- SciPy: It, short for Scientific Python, extends the functionality of NumPy by providing tools for advanced mathematical, scientific and engineering computations. It includes modules for optimization, integration, signal processing and linear algebra.
- Tensor Flow: It is an open-source machine learning and deep learning framework developed by Google. It allows developers to build and train neural networks for tasks such as image recognition, natural language processing and predictive modeling.
- Scikit-learn: It is a popular machine learning library built on top of NumPy and SciPy. It offers simple and efficient tools for classification, regression, clustering and dimensionality reduction.
- Scrapy: It is a web scraping and data extraction library designed to efficiently crawl websites and gather structured information. It allows developers to create spiders that automatically navigate web pages and collect data.
- TensorFlow: An open-source, end-to-end platform developed by Google for large-scale machine learning and deep learning tasks. It excels in high-performance numerical computation across various platforms (CPUs, GPUs, TPUs) and is ideal for large-scale production deployments. Visit the official TensorFlow website for details.
- PyTorch: Developed by Facebook's AI Research lab, PyTorch is known for its flexibility, dynamic computational graphs, and ease of use, making it popular in the research community for rapid prototyping, computer vision, and natural language processing (NLP).
- Keras: A high-level, user-friendly neural networks API that acts as an interface for backends like TensorFlow, Theano, or CNTK. It is designed for fast and easy experimentation with deep neural networks.
- XGBoost & LightGBM: High-performance, decision-tree-based ensemble libraries that use gradient boosting frameworks. They are known for their speed and accuracy in classification and regression tasks involving structured data.

- NLTK & spaCy: Libraries specifically for Natural Language Processing (NLP) tasks, such as tokenization, sentiment analysis, and semantic reasoning.