**Module2: Introduction to Python**

Theoretical Assignments:

1. **Explain Python’s Role in Data Science.**
2. **Easy to Learn and Use :**

* **Readable syntax** makes it accessible for non-programmers (e.g., statisticians, analysts).
* Encourages quick development and experimentation.

1. **Rich Ecosystem of Libraries :**

Python has a robust collection of open-source libraries specifically designed for data science:

**Data Manipulation & Analysis**

* **Pandas** – Data structures and operations for manipulating structured data.
* **NumPy** – Powerful n-dimensional arrays and numerical operations.

**Visualization**

* **Matplotlib** and **Seaborn** – Plotting and statistical visualization.
* **Plotly** – Interactive charts and dashboards.

**Machine Learning & Modeling**

* **Scikit-learn** – Classical ML algorithms (regression, classification, clustering).
* **TensorFlow** and **PyTorch** – Deep learning frameworks.
* **XGBoost**, **LightGBM** – Gradient boosting models for structured data.

**Data Collection & Processing**

* **BeautifulSoup**, **Scrapy** – Web scraping.
* **Requests**, **Selenium** – API calls and browser automation.

1. **Integration and Deployment :**

* Easily integrates with **databases**, **big data platforms** (e.g., Spark via PySpark), and **cloud services**.
* Supports deployment of models via **Flask**, **FastAPI**, or **streamlit**.

1. **Strong Community and Support :**

* Massive community contributes to continuous development.
* Extensive documentation, tutorials, and forums.

1. **Versatility Across the Data Science Workflow :**

From data cleaning and exploration to modeling, evaluation, and deployment, Python supports **the entire data science lifecycle**.