**Module2: Introduction to Python**

Theoretical Assignments:

1. Explain Python’s Role in Data Science

Ans:

* **What is Python?**
* Python is a programming language that is easy to read and write.
* Python is one of the most popular programming languages.
* Python is the most used language in Data Science  
  Because it is easy, has powerful tools, and supports every step of the data science process.
* **What is Data Science?**

Data Science = Data + Tools + Thinking → Insights + Decisions

**Data Science** is the process of using data to understand things, solve problems, and make smart decisions.

It combines:

* Math and statistics
* Programming (like Python)
* Knowledge of the topic (business, healthcare, etc.)
* **What Do Data Scientists Do?**

They follow these main steps:

1. **Collect Data**

Get data from files, websites, databases, sensors, etc.

1. **Clean the Data**

Remove mistakes, missing values, or duplicates

1. **Analyse the Data**

Find patterns, trends, or insights

1. **Visualize the Data**

Create graphs or charts to understand and show the data

1. **Build Models**

Use machine learning to make predictions or automate tasks

1. **Make Decisions**

Use the results to help businesses or people make better choices

* **Why is Data Science Important?**

Every day, a huge amount of data is created (from phones, websites, apps, etc.).  
Data Science helps us make sense of it — like finding patterns, predicting trends, and improving products or services.

* **Why Python is Used in Data Science**

**1. Easy to Use**

* Python has simple and clear code.
* You can do a lot with just a few lines of code.

**2. Powerful Libraries**

Python has many built-in tools (called **libraries**) for data science.  
Some important ones are:

|  |  |
| --- | --- |
| Purpose | Python Library |
| Math & Arrays | NumPy |
| Data Handling | Pandas |
| Data Visualization | Matplotlib, Seaborn |
| Machine Learning | Scikit-learn |
| Deep Learning | TensorFlow, PyTorch |

**3. Great for Visualization**

Python helps you **create charts and graphs** to understand the data better.

**4. Used in Every Step**

Python can be used for:

* Collecting data (e.g., from websites or files)
* Cleaning messy data
* Analysing and finding patterns
* Building and testing models
* Making predictions

**5. Works Well with Other Tools**

Python can connect with:

* Databases (like MySQL)
* Excel
* Big data tools (like Spark)
* Websites and apps