Introduction à la Data Science

Program: Introduction à la Data Science

Objective:

To provide foundational knowledge in data science, covering its principles, tools, and applications, enabling participants to analyze data and derive actionable insights.

Day 1: Introduction to Data Science

1. What is Data Science?
   - Definition and scope.
   - Key concepts: data, analytics, machine learning, AI.
   - The role of a data scientist.

2. Data Science Workflow
   - Problem definition.
   - Data collection and preprocessing.
   - Analysis and modeling.
   - Interpretation and communication.

3. Applications of Data Science
   - Real-world examples in healthcare, finance, e-commerce, etc.

4. Tools Overview
   - Popular tools: Python, R, Jupyter Notebook, SQL.
   - Introduction to libraries like pandas, NumPy, and Matplotlib.

Day 2: Data Manipulation and Exploration

1. Data Acquisition

- Importing data from CSV, Excel, and databases.
- APIs and web scraping basics.

2. Data Cleaning

- Handling missing data.
- Removing duplicates.
- Data transformation.

3. Data Exploration

- Descriptive statistics.
- Visualization tools (Matplotlib, Seaborn).

4. Hands-On

- Exploratory Data Analysis (EDA) on a sample dataset.

Day 3: Introduction to Machine Learning

1. What is Machine Learning?

- Types: supervised, unsupervised, and reinforcement learning.
- Basic ML terminology.

2. Supervised Learning

- Regression vs classification.
- Simple examples (e.g., linear regression).

3. Unsupervised Learning

- Clustering and dimensionality reduction.
- Examples with K-Means.

4. Hands-On

- Building a basic ML model in Python.

Day 4: Data Science in Practice

1. Data Science Projects

- Structuring a project: defining objectives and delivering insights.
- Challenges and ethical considerations in data science.

2. Case Study

- Walkthrough of a data science project from start to finish.

3. Advanced Topics (Optional)

- Introduction to deep learning.
- Time-series analysis or Natural Language Processing (NLP).

4. Wrap-Up

- Recap of key concepts.
- Q&A and next steps in data science learning.

Prerequisites:

- Basic programming knowledge (preferably Python).
- Familiarity with statistics and linear algebra is helpful but not mandatory.

Deliverables:

- Hands-on Python notebooks with examples.

- Resources for further learning (books, online courses, datasets).