Data Science Course Syllabus

This course is designed to help IT and Non-IT students become job-ready Data Scientists. It covers Python programming, statistics, machine learning, deep learning, big data, cloud deployment, and real-world projects. The focus is on hands-on learning, industry-relevant tools, and interview preparation.

🕐**Course Duration: 10 Months** 🎯**Class Duration**: 2hrs (Mon to Fry)

📌 **Prerequisites: No prior coding experience required (Basic IT knowledge is beneficial)**

📖 **Course Modules**

📌 **Module 1: Introduction to Data Science**

- What is Data Science? Applications of Data Science in Different Industries, Understanding Data Science Roles (Data Scientist, Analyst, Engineer), Data Science Workflow & Tools, Understanding the Data Science Lifecycle, How Data Science is Used in Business Decision-Making, Career Opportunities in Data Science

📌 **Module 2: Basic Computer & Mathematics Skills**

- Understanding Computers, Software, and Operating Systems, Introduction to Spreadsheets (Excel/Google Sheets), Basic Mathematics for Data Science

- Arithmetic Operations, Algebra & Functions, Logarithms & Exponents, Linear Equations

📌 **Module 3: Python & R Language for Data Science**

- **Python Basics:** Variables, Data Types, Operators, Control Statements & Loops, Functions, Modules & File Handling

- Object-Oriented Programming in Python, Working with Libraries: NumPy, Pandas, Matplotlib, Seaborn

- **R Programming Basics (Optional):** Introduction to R Language, Data Types & Data Structures in R, R Libraries: dplyr, ggplot2, tidyr, Basic Data Manipulation in R

📌 **Module 4: Data Handling & Excel for Data Science**: Importing & Cleaning Data, Handling Missing Values & Duplicates, Data Transformation & Feature Engineering, Data Visualization with Matplotlib & Seaborn

4.1 Data Handling in Excel: Data Entry & Cleaning in Excel, Data Formatting & Conditional Formatting, Sorting, Filtering & Pivot Tables, Basic Formulas & Functions (SUM, AVERAGE, COUNT, IF, VLOOKUP)

4.2 Introduction to Data Handling with SQL: What is SQL? Basic SQL Queries (SELECT, WHERE, ORDER BY, GROUP BY), Joins & Subqueries, Using SQL for Data Analysis

4.3 Working with Data in Python: Loading Datasets (CSV, Excel, JSON, SQL), Data Cleaning (Handling Missing Values, Duplicates, Outliers), Data Transformation (Normalization, Encoding, Feature Scaling), Exploratory Data Analysis (EDA)

4.4 Data Wrangling with Pandas: Data Aggregation & Grouping, Merging, Joining, & Concatenating Datasets, Pivot Tables & Cross-tabulation

📌 **Module 5: Data Wrangling & Exploration:** Importing & Cleaning Data, Handling Missing Values & Duplicates, Data Transformation & Feature Engineering, Data Visualization with Matplotlib & Seaborn

📌 **Module 6: Statistics & Probability for Data Science:** Importance of Statistics in Data Science, Descriptive Statistics (Mean, Median, Mode, Variance, Standard Deviation), Probability Basics (Probability Distributions, Random Variables), Inferential Statistics (Hypothesis Testing, p-value, Confidence Intervals), Correlation & Regression Basics**,** Descriptive & Inferential Statistics, Measures of Central Tendency & Dispersion, Probability Theory & Distributions, Hypothesis Testing & A/B Testing

📌 **Module 7: Machine Learning Fundamentals:** Introduction to Machine Learning, Supervised vs. Unsupervised Learning, Regression Models (Linear, Logistic), Classification Algorithms (KNN, SVM, Decision Trees, Random Forest), Clustering Algorithms (K-Means, DBSCAN), Model Evaluation Metrics

📌 **Module 8: Machine Learning Basics (No Coding Required Initially):**What is Machine Learning? Supervised vs. Unsupervised Learning, Understanding Regression & Classification, Real-Life Applications of Machine Learning, Introduction to AutoML (Automated Machine Learning)

📌 **Module 9: Advanced Machine Learning:** Feature Selection & Engineering, Hyperparameter Tuning (Grid Search, Random Search), Ensemble Learning (Bagging, Boosting, XGBoost), Handling Imbalanced Data, Model Deployment with Flask & FastAPI

📌 **Module 10: Deep Learning with TensorFlow & Keras:** Understanding Neural Networks, Real-World Applications (Image & Speech Recognition), Basics of TensorFlow & Keras, Introduction to Neural Networks, Building Deep Learning Models with TensorFlow & Keras, Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Natural Language Processing (NLP) with spaCy & NLTK

📌 **Module 11: Big Data & Cloud Computing:** Introduction to Big Data & Apache Spark, Working with Cloud Platforms (AWS, Google Cloud, Azure), Deploying Machine Learning Models in Cloud

📌 **Module 12: Final Project & Interview Preparation**

- Full Data Science Project (E-Commerce, Healthcare, Finance, etc.), Resume & LinkedIn Profile Optimization, Mock Interviews & HR Round Preparation, Freelancing & Job Search Strategies

💡 **Tools & Technologies Covered**

☑ **Programming: Python, SQL**
☑ **Data Visualization: Matplotlib, Seaborn, Plotly**
☑ **Machine Learning: Scikit-Learn, TensorFlow, Keras**
☑ **Big Data: Apache Spark, Hadoop**
☑ **Cloud & Deployment: AWS, Google Cloud, Azure, Flask, FastAPI**

🚀 **Final Outcome**

By the end of this course, students will be:
☑ **Proficient in Data Science & Machine Learning**
☑ **Able to Build & Deploy ML Models**
☑ **Job-Ready for Data Scientist & Analyst Roles**
☑ **Confident in Technical Interviews & Freelancing**