

Data Science Course Syllabus

Starter Level (Beginner)

- Introduction to Data Science

- * What is Data Science?
- * Applications and use-cases
- * Data Science lifecycle
- * Roles in Data Science

- Python for Data Science

- * Python basics
- * Functions, loops, comprehensions
- * Libraries: NumPy, Pandas
- * File handling

- Basic Statistics & Mathematics

- * Descriptive statistics
- * Probability basics
- * Data distributions
- * Intro to linear algebra

- Data Cleaning and Exploration

- * Handling missing values
- * Outlier detection
- * EDA techniques

- Data Visualization

- * Matplotlib & Seaborn
- * Histogram, bar, box, scatter plots
- * Plotting multiple variables

Intermediate Level

- Intermediate Python & Pandas

- * Lambda functions
- * GroupBy, Merge, Join
- * Date/time data
- * Regex and text processing

- Probability and Statistics (Intermediate)

- * Combinatorics
- * Bayes' Theorem
- * Central Limit Theorem
- * Hypothesis testing

- Intro to Machine Learning

- * Supervised vs. unsupervised learning
- * Regression
- * Classification
- * Clustering

- Model Evaluation Techniques

- * Train-test split
- * Cross-validation
- * Accuracy, precision, recall
- * Confusion matrix
- * ROC-AUC

- Working with Real Datasets

- * APIs and Web Scraping
- * Intro to SQL
- * Simple data projects

Advanced Level

- Advanced Machine Learning

- * Random Forests, Gradient Boosting
- * Feature engineering
- * Hyperparameter tuning
- * Pipelines

- Natural Language Processing (NLP)

- * Text preprocessing
- * TF-IDF, Word2Vec
- * Sentiment analysis
- * NLP libraries

- Time Series Analysis

- * Stationarity
- * ARIMA/SARIMA
- * Forecasting

- Deep Learning (Intro)

- * Neural networks basics
- * TensorFlow/Keras or PyTorch
- * CNNs
- * RNNs

- Big Data Tools

- * Hadoop & MapReduce
- * Apache Spark
- * Cloud platforms intro

Expert Level

- Deep Learning (Advanced)

- * Transfer learning
- * Transformers
- * GANs
- * Fine-tuning pre-trained models

- Model Deployment

- * Flask/FastAPI
- * Streamlit/Gradio
- * Docker
- * CI/CD

- MLOps & Model Monitoring

- * ML lifecycle
- * Model versioning
- * Data drift handling

- AI Ethics and Interpretability

- * Explainable AI
- * Bias and fairness
- * Privacy and data protection

- Capstone Projects

- * End-to-end project
- * Domain-specific applications