



Data Science with Python Syllabus

1. Introduction to Data Science
 - What is Data Science?
 - Applications in Corporate & Research
 - Python for Data Science Overview
2. Python Refresher for Data Science
 - Data Types, Operators, Control Flow
 - Functions and Modules
 - File Handling and Exceptions
3. Data Handling with NumPy
 - Arrays and Vectorized Computations
 - Indexing, Slicing, Reshaping
 - Mathematical and Statistical Operations
4. Data Manipulation with Pandas
 - Series and DataFrames
 - Importing and Cleaning Data
 - Grouping, Aggregation, and Merging
 - Handling Missing Data
5. Data Visualization
 - Matplotlib Basics
 - Seaborn for Statistical Visualization
 - Plotly for Interactive Dashboards
6. Exploratory Data Analysis (EDA)
 - Understanding Data Distributions
 - Correlation and Feature Relationships
 - Outlier Detection
7. Statistics for Data Science
 - Descriptive and Inferential Statistics
 - Probability Distributions
 - Hypothesis Testing & p-values
8. Machine Learning Basics
 - Supervised vs Unsupervised Learning
 - Regression, Classification, Clustering
 - Model Training and Evaluation Metrics



9. Supervised Learning Algorithms

- Linear & Logistic Regression
- Decision Trees and Random Forests
- Support Vector Machines (SVM)
- Gradient Boosting (XGBoost, LightGBM)

10. Unsupervised Learning Algorithms

- K-Means Clustering
- Hierarchical Clustering
- Dimensionality Reduction (PCA, t-SNE)

11. Feature Engineering

- Handling Missing Values
- Feature Scaling and Transformation
- Encoding Categorical Variables
- Feature Selection Techniques

12. Deep Learning with Python (Intro)

- Neural Networks Basics
- TensorFlow / Keras Overview
- Building a Simple Neural Network

13. Model Deployment

- Saving & Loading Models (Pickle, Joblib)
- Flask/FastAPI for Model Deployment
- Streamlit for Interactive Dashboards

14. Capstone Project

- End-to-End Data Science Project
- Data Collection → Cleaning → Modeling → Deployment