# Data Science Syllabus

**Chapter 1: Introduction to Python for Data Science**

* Introduction to Data Science & its applications
* Setting up the environment (Anaconda, Jupyter Notebook)
* Python basics (data types, loops, functions)
* Introduction to NumPy for numerical computing
* Hands-on exercises

**Chapter 2: Data Handling with Pandas**

* Introduction to Pandas for data manipulation
* Reading data from CSV, Excel, and databases
* Data cleaning: handling missing values, filtering, sorting
* Data transformation and aggregation (groupby, etc)
* Hands-on project: Cleaning a real-world dataset

**Chapter 3: Exploratory Data Analysis (EDA)**

* Descriptive statistics (mean, median, mode, variance)
* Visualizing data with Matplotlib & Seaborn
* Boxplots, histograms, scatter plots, pair plots
* Detecting outliers and correlations
* Hands-on project: Exploratory analysis of a dataset

**Chapter 4: Data Preprocessing & Feature Engineering**

* Handling missing data, duplicates, and outliers
* Encoding categorical variables (OneHot, Label Encoding)
* Feature scaling (StandardScaler, MinMaxScaler)
* Feature selection techniques
* Hands-on case study: Preparing data for ML models

**Chapter 5: Introduction to Machine Learning with Scikit-Learn**

* Understanding Supervised vs. Unsupervised Learning
* Introduction to Scikit-Learn library
* Building a simple Linear Regression model
* Train-test split, evaluation metrics (MAE, RMSE)
* Hands-on project: Predicting house prices

**Chapter 6: Classification Algorithms**

* Logistic Regression, Decision Trees, Random Forest
* Evaluating models: Precision, Recall, F1-score, ROC curve
* Hyperparameter tuning with GridSearchCV
* Hands-on project: Classifying email spam

**Chapter 7: Clustering & Dimensionality Reduction**

* K-Means clustering
* PCA (Principal Component Analysis) for dimensionality reduction
* Hands-on project: Customer segmentation using K-Means

**Chapter 8: Natural Language Processing (NLP)**

* Text preprocessing (Tokenization, Stemming, Lemmatization)
* TF-IDF, Word Embeddings
* Sentiment Analysis with NLTK & Scikit-Learn
* Hands-on project: Sentiment analysis on Twitter data

**Chapter 9: Mini project on NLP/ML**

* Development of NLP/ML Model
* Development of Front End

**Chapter 10: Mini project on NLP/ML Continuation**

* Testing
* Demonstration
* Final assessment