Comprehensive Data Science Syllabus

1. Python Programming

- Variables, Data Types, and Operators
- Control Flow (if, loops)
- Functions and Modules
- File Handling
- Object-Oriented Programming

2. Python Libraries for Data Science

- NumPy: Arrays, broadcasting, linear algebra
- Pandas: DataFrames, data cleaning, manipulation
- Matplotlib & Seaborn: Visualization
- Scikit-learn: ML tools and preprocessing
- TensorFlow/Keras: Deep Learning frameworks

3. Statistics for Data Science

- Descriptive Statistics: Mean, Median, Mode, Standard Deviation
- Probability Theory
- Distributions (Normal, Binomial, etc.)
- Inferential Statistics and Hypothesis Testing

4. Machine Learning

- Supervised Learning: Linear & Logistic Regression, Decision Trees, Random Forest, SVM
- Unsupervised Learning: Clustering (K-Means, Hierarchical), PCA
- Model Evaluation Metrics (Accuracy, Precision, Recall, F1-Score)
- Cross Validation and Hyperparameter Tuning

5. Deep Learning

- Neural Network Basics

Comprehensive Data Science Syllabus

- Activation Functions, Optimizers
- CNNs (Convolutional Neural Networks)
- RNNs (Recurrent Neural Networks), LSTM
- Model Training with TensorFlow/Keras

6. MySQL for Data Science

- Database Basics
- Writing SQL Queries (SELECT, INSERT, UPDATE, DELETE)
- Joins, Aggregations, Subqueries
- Data Filtering and Sorting
- Using SQL for Data Analysis

7. Artificial Intelligence (AI)

- Introduction to AI and its Applications
- AI vs ML vs DL
- Search Algorithms
- Expert Systems
- NLP basics

8. Generative AI (GenAI)

- Understanding GenAI
- Applications of GenAI (Chatbots, Content Creation)
- Working with Large Language Models (LLMs)
- Prompt Engineering Basics
- Ethics and Bias in GenAI

9. Capstone Project

- Real-world Dataset
- Data Collection to Model Building

Comprehensive Data Science Syllabus

- Visualization and Insight Generation

- Final Report and Presentation