4-Month (16-Week) Data Science Curriculum

Module 1: Python + Statistics Foundations (Weeks 1–3)

Objective: Build strong programming + analytical base

Week 1 – Python Basics

- Variables, Data Types, Operators
- Conditional Statements, Loops
- Functions, Lambda
- Data Structures (List, Tuple, Dict, Set)

Week 2 – NumPy + Data Handling

- NumPy Arrays, Indexing, Slicing
- Vectorized Operations
- File Handling (CSV, JSON)

Week 3 – Statistics & Probability

- Mean, Median, Mode, Variance, Std Dev
- Probability Basics
- Distributions (Normal, Skewed)
- Sampling & Central Limit Theorem

Module 2: Data Analysis & Visualization (Weeks 4–6)

Objective: Real-world data cleaning + insight extraction

Week 4 – Pandas & Data Cleaning

- DataFrames, Filtering, Sorting
- Handling Missing Values
- Outliers Treatment
- Feature Understanding

Week 5 – Exploratory Data Analysis (EDA)

- Univariate & Bivariate Analysis
- Correlation Analysis
- Feature Relationships

Week 6 – Visualization

- Matplotlib (line, bar, scatter)
- Seaborn (heatmaps, pairplots)

- Storytelling with Data

Module 3: SQL for Data Analytics (Weeks 7–8)

Objective: Data extraction + backend analytics

Week 7 – SQL Basics

- SELECT, WHERE, ORDER BY
- JOINs (INNER, LEFT, RIGHT)
- GROUP BY, Aggregations

Week 8 – Advanced SQL

- Subqueries & CTEs
- Window Functions (RANK, ROW_NUMBER)
- Case Statements
- Feature Engineering using SQL

Module 4: Machine Learning (Core) (Weeks 9–11)

Objective: Model building + evaluation

Week 9 – ML Foundations

- ML Workflow (Train/Test Split)
- Data Preprocessing (Scaling, Encoding)
- Feature Selection

Week 10 – Regression

- Linear Regression
- Polynomial Regression
- Evaluation (MAE, MSE, RMSE, R^2)

Week 11 – Classification

- Logistic Regression
- KNN
- Decision Tree
- Metrics: Accuracy, Precision, Recall, F1

Module 5: Advanced ML + Unsupervised Learning (Weeks 12–13)

Objective: Industry-level modeling

Week 12 – Ensemble Learning

- Random Forest

- Gradient Boosting
- XGBoost
- Hyperparameter Tuning (GridSearchCV)

Week 13 – Unsupervised Learning

- Clustering (K-Means, DBSCAN)
- PCA (Dimensionality Reduction)
- Feature Engineering Basics

Module 6: BI Tools + Deployment (Weeks 14–15)

Objective: Business + visualization layer

Week 14 – Power BI / Tableau

- Dashboard Creation
- KPI Metrics
- Business Reporting

Week 15 – Deployment Basics

- Model Saving (pickle/joblib)
- Streamlit App
- Basic API Integration

Module 7: Capstone + Career Prep (Week 16)

Objective: End-to-end real-world readiness

Week 16 – Final Capstone

- Problem Statement
- Data Collection & Cleaning
- Model Building
- Dashboard Creation
- Deployment (Streamlit)