Data Science

Chapter 1: Introduction to Data Science

- What is data science
- Data science lifecycle
- Difference between data analytics & data science
- Applications across industries
- Career paths & roles

Chapter 2: Python Programming for Data Science

- Python environment setup
- Python syntax & data types
- Control flow & functions
- Working with libraries
- Writing clean, reusable code

Chapter 3: Statistics & Probability

- Descriptive statistics
- Probability fundamentals
- Data distributions
- Hypothesis testing
- Statistical inference

Chapter 4: Data Collection & Data Wrangling

- Data sources (CSV, APIs, databases)
- Data cleaning techniques
- Handling missing & noisy data
- Feature transformation
- Data validation

Chapter 5: Exploratory Data Analysis (EDA)

- Data summarization
- Visualization techniques
- Correlation analysis

- Outlier detection
- Insights generation

Chapter 6: Data Visualization Tools

- Matplotlib fundamentals
- Seaborn basics
- Plotly overview
- Storytelling with data
- Visualization best practices

Chapter 7: Machine Learning Fundamentals

- Introduction to machine learning
- Supervised vs unsupervised learning
- Model training workflow
- Bias–variance tradeoff
- Evaluation metrics

Chapter 8: Supervised Learning Algorithms

- Linear & logistic regression
- Decision trees
- Random forests
- K-Nearest Neighbors
- Model evaluation

Chapter 9: Unsupervised Learning Algorithms

- Clustering concepts
- K-Means clustering
- Hierarchical clustering
- Dimensionality reduction
- Principal Component Analysis (PCA)

Chapter 10: Feature Engineering & Model Optimization

- Feature selection
- Feature scaling
- Hyperparameter tuning
- Cross-validation
- Model optimization techniques

Chapter 11: Introduction to Deep Learning

- Neural network basics
- Activation functions
- Deep learning frameworks overview
- Use cases of deep learning
- Model training basics

Chapter 12: Big Data & Data Science Tools

- Introduction to big data
- SQL for data science
- Cloud platforms overview
- Data pipelines basics
- Version control with Git

Chapter 13: Ethics, Governance & Deployment

- Ethical AI & bias
- Data privacy & security
- Model deployment basics
- Monitoring & maintenance
- Responsible AI practices

Chapter 14: Capstone Project & Industry Readiness

- End-to-end data science project
- Problem formulation
- Model building & evaluation
- Final presentation
- Resume & interview preparation