

CONTENTS

Unit 1: Exploratory Data Analysis	1-30
1.1 EDA Fundamentals	
1.2 Understanding Data Science	
1.3 Significance of EDA	
1.4 Making Sense of Data	
1.5 Comparing EDA with Classical and Bayesian Analysis	
1.6 Software Tools for EDA	
1.7 Visual Aids for EDA	
1.8 Data Transformation Techniques – Merging Database, Reshaping and Pivoting	
1.9 Case Study: Attack for Tampering with Recommender Systems	
Unit 2: EDA Using Python	31-66
2.1 Data Manipulation using Pandas	
2.2 Pandas Objects	
2.3 Data Indexing and Selection	
2.4 Operating on Data	
2.5 Handling Missing Data	
2.6 Hierarchical Indexing	
2.7 Combining Datasets – Concat, Append, Merge and Join	
2.8 Aggregation and Grouping	
2.9 Pivot Tables	
2.10 Vectorized String Operations	
Unit 3: Univariate Analysis	67-79
3.1 Introduction to Single Variable: Distribution Variables	
3.2 Numerical Summaries of Level and Spread	
3.3 Scaling and Standardizing	
3.4 Inequality	
Unit 4: Bivariate Analysis	80-96
4.1 Relationships between Two Variables	
4.2 Percentage Tables	
4.3 Analysing Contingency Tables	
4.4 Handling Several Batches	
4.5 Scatterplots and Resistant Lines	

Unit 5: Multivariate and Time Series Analysis

97-127

- 5.1 Introducing a Third Variable
- 5.2 Causal Explanations
- 5.3 Three-Variable Contingency Tables and Beyond
- 5.4 Fundamentals of Time Series Analysis (TSA)
- 5.5 Characteristics of Time Series Data
- 5.6 Data Cleaning
- 5.7 Time Based Indexing
- 5.8 Visualizing
- 5.9 Grouping
- 5.10 Resampling