



Preface	4
Prerequisites	4
1. The Basics	6
2. Data Visualization	7
2.1 Object Hierarchy in Matplotlib	7
2.2 Seaborn Styles	8
2.3 Plots	10
2.3.1 Distribution Plots	10
2.4 Continuous & Categorical Variables Relationships	14
2.5 Continuous Variables Relationships	18
2.6 Statistical models	19
2.7 Heatmaps	20
3. Exploratory Data Analysis	21
3.1 Univariate Analysis	21
3.1.1 Remove unnecessary features	22
3.1.2 Handling Types of Features	22
3.2 Distribution	23
3.2.1 Measures of Center	25
3.2.2 Measures of Spread	26
3.3 Detecting and Handling Outliers	27
3.4 Correlation Matrix	30
4. Feature Engineering	32
4.1 Imputation	32
4.2 Categorical Encoding	35
4.2.1 Label Encoding	36
4.2.2 One-Hot Encoding	37
4.2.3 Count Encoding	37
4.2.4 Target Encoding	38
4.2.5 Leave One Out Target Encoding	40
4.3 Binning	41
4.4 Scaling	42
4.4.1 Standard Scaling	43
4.4.2 Min-Max Scaling (Normalization)	44
4.4.3 Quantile Transformation	45
4.5 Log Transform	47
4.6 Feature Selection	48
4.7 Feature Grouping	49
4.8 Feature Split	50