

---

# Table of Contents

---

<b>1. Principal Component Analysis (PCA)</b>	<b>11</b>
1.1 Data to be Processed and Stages of the Analysis	11
1.1.1 Observation-driven Approach	12
1.1.2 Variable-driven Approach	16
1.2 Rationale – Criteria for Choosing the Number of Axes	18
1.2.1 Coverage Percentage Criterion	18
1.2.2 Kaiser's Criterion	19
1.2.3 Cattell's Criterion	19
1.3 Results Evaluation	20
1.3.1 Scores	20
1.3.2 Quality of Observations Representation	20
1.3.3 Contribution of Individuals to the Variance of Axes	21
1.3.4 Communality	21
1.3.5 Factor Loadings	21
1.4 Graphical Representations	22
1.4.1 Representation of Individuals (Observations)	22
1.4.2 Representation of Variables	22
1.5 Additional Variables and Observations	23
1.6 Reconstitution of Initial Data	23
1.6.1 Reconstitution of the Initial Data Table	23
1.6.2 Reconstitution of the Correlation Matrix	24
1.7 Unstandardised Principal Component Analysis	24
1.8 Weighted Principal Component Analysis	25
1.9 PCA – Case Study	25
1.9.1 Data Set	26
1.9.2 Python Implementation	27
1.9.3 Results and Discussion	37
<b>2. Exploratory Factor Analysis (EFA)</b>	<b>47</b>
2.1 Data to be Processed	47
2.2 Model Assumptions	48
2.3 Estimation of Factors Existence	49
2.4 Estimation of Model Parameters. Factor Extraction	50
2.5 Method of Maximum Likelihood	51
2.6 Method of Least-squares	51
2.7 Method of Principal Components	52
2.8 Estimate the Number of Factors. Bartlett's Test	53
2.9 Factor Rotation	54
2.10 EFA – Case Study	55
2.10.1 Data Set	55
2.10.2 Python Implementation	57
2.10.3 Results and Discussion	64

<b>3. Canonical Correlation Analysis (CCA)</b>	<b>69</b>
3.1 Data to be Processed	69
3.2 Stages of Analysis	70
3.3 Canonical Factors	72
3.4 Relations Between Factors	74
3.5 Explained Variance and Informational Redundancy	75
3.6 Standardisation of Canonical Factors	76
3.7 Relevance of Canonical Roots. Bartlett's Chi-square Test	76
3.8 Generalised Canonical Correlation Analysis (GCCA)	77
3.9 CCA – Case Study	79
3.9.1 Data Sets	79
3.9.2 Python Implementation	82
3.9.3 Results and Discussion	87
<b>4. Factorial Correspondence Analysis (FCA)</b>	<b>91</b>
4.1 Canonical Analysis of Complete Disjunctive Tables	91
4.2 Principal Component Analysis of Profile Tables	94
4.3 Principal Component Analysis of Inertia Matrix	96
4.4 Multiple Correspondences Analysis (MCA)	98
4.5 Principal Component Analysis of Standard Deviation Matrix	100
4.6 MCA – Case Study	101
4.6.1 Data Set	101
4.6.2 Python Implementation	106
4.6.3 Results and Discussion	111
<b>5. Hierarchical Cluster Analysis (HCA)</b>	<b>115</b>
5.1 Hierarchical Algorithms	116
5.2 Hierarchical Grouping Methods	117
5.3 Distances for Proximity Calculation	120
5.4 Choosing the Optimal Number of Clusters	121
5.5 HCA – Case Study	124
5.5.1 Data Set	124
5.5.2 Python Implementation	126
5.5.3 Results and Discussion	132
<b>6. Discriminant Analysis (DA)</b>	<b>138</b>
6.1 Data Organisation and Notation	138
6.2 Indicators of Variability and Dispersion	139
6.3 Model Meaning. Statistical Tests	141
6.4 Linear Discriminant Analysis (LDA) – Fisher Functions	143
6.5 Discriminant Analysis – Special Case of Canonical Correlation Analysis	146
6.6 Classification Functions	147
6.7 Bayesian Discrimination	149
6.7.1 Total Probability and Bayesian Probability	149
6.7.2 Bayesian Classifier	149
6.7.3 Nonparametric Estimation of Conditional Probabilities. Histogram Method	150
6.7.4 Parametric Methods	151
6.8 Discriminant Analysis – Case Study	152
6.8.1 Data Set	152
6.8.2 Python Implementation	153
6.8.3 Results and Discussion	161
<b>Bibliography</b>	<b>167</b>
<b>Appendix 1</b>	<b>169</b>
<b>Appendix 2</b>	<b>176</b>