
Table of Contents

Preface.....	vii
1. Introduction.....	1
Telling Better Stories with Data	1
A Quick Look: Data Science for SAP Professionals	3
A Quick Look: SAP Basics for Data Scientists	6
Getting Data Out of SAP	8
Roles and Responsibilities	11
Summary	11
2. Data Science for SAP Professionals.....	13
Machine Learning	14
Supervised Machine Learning	15
Unsupervised Machine Learning	18
Semi-Supervised Machine Learning	21
Reinforcement Machine Learning	22
Neural Networks	26
Summary	43
3. SAP for Data Scientists.....	45
Getting Started with SAP	46
The ABAP Data Dictionary	49
Tables	50
Structures	53
Data Elements and Domains	54
Where-Used	58
ABAP QuickViewer	62
SE16 Export	68

OData Services	68
Core Data Services	80
Summary	91
4. Exploratory Data Analysis with R.....	93
The Four Phases of EDA	95
Phase 1: Collecting Our Data	96
Importing with R	104
Phase 2: Cleaning Our Data	107
Null Removal	107
Binary Indicators	107
Removing Extraneous Columns	108
Whitespace	108
Numbers	109
Phase 3: Analyzing Our Data	109
DataExplorer	110
Discrete Features	113
Continuous Features	117
Phase 4: Modeling Our Data	121
TensorFlow and Keras	122
Training and Testing Split	122
Shaping and One-Hot Encoding	123
Recipes	124
Preparing Data for the Neural Network	126
Results	130
Summary	132
5. Anomaly Detection with R and Python.....	133
Types of Anomalies	134
Tools in R	135
AnomalyDetection	135
Anomalize	136
Getting the Data	136
SAP ECC System	137
SAP NetWeaver Gateway	142
SQL Server	153
Finding Anomalies	174
PowerBI and R	174
PowerBI and Python	184
Summary	189

6. Predictive Analytics in R and Python.....	191
Predicting Sales in R	193
Step 1: Identify Data	193
Step 2: Gather Data	193
Step 3: Explore Data	194
Step 4: Model Data	195
Step 5: Evaluate Model	206
Predicting Sales in Python	210
Step 1: Identify Data	210
Step 2: Gather Data	210
Step 3: Explore Data	216
Step 4: Model Data	219
Step 5: Evaluate Model	220
Summary	222
7. Clustering and Segmentation in R.....	225
Understanding Clustering and Segmentation	226
RFM	227
Pareto Principle	228
k-Means	229
k-Medoid	230
Hierarchical Clustering	231
Time-Series Clustering	233
Step 1: Collecting the Data	233
Step 2: Cleaning the Data	234
Step 3: Analyzing the Data	240
Revisiting the Pareto Principle	240
Finding Optimal Clusters	241
k-Means Clustering	244
k-Medoid Clustering	249
Hierarchical Clustering	253
Manual RFM	255
Step 4: Report the Findings	258
R Markdown Code	261
R Markdown Knit	262
Summary	264
8. Association Rule Mining.....	267
Understanding Association Rule Mining	269
Support	269
Confidence	269
Lift	270

Apriori Algorithm	270
Operationalization Overview	270
Collecting the Data	271
Cleaning the Data	276
Analyzing the Data	277
Fiori	282
Summary	287
9. Natural Language Processing with the Google Cloud Natural Language API.	289
Understanding Natural Language Processing	290
Sentiment Analysis	290
Translation	292
Preparing the Cloud API	292
Collecting the Data	298
Analyzing the Data	301
Summary	303
10. Conclusion.	305
Original Mission	305
Recap	306
Chapter 1: Introduction	306
Chapter 2: Data Science for SAP Professionals	306
Chapter 3: SAP for Data Scientists	306
Chapter 4: Exploratory Data Analysis	307
Chapter 5: Anomaly Detection with R and Python	307
Chapter 6: Prediction with R	307
Chapter 7: Clustering and Segmentation in R	307
Chapter 8: Association Rule Mining	307
Chapter 9: Natural Language Processing with the Google Cloud Natural Language API	308
Tips and Recommendations	308
Be Creative	308
Be Practical	308
Enjoy the Ride	309
Stay in Touch	309
Index.	311