

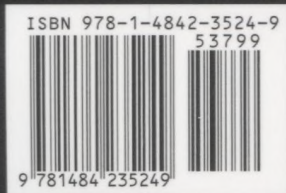
FOR PROFESSIONALS BY PROFESSIONALS®

Applied Analytics through Case Studies Using SAS and R

Examine business problems and use a practical analytical approach to solve them by implementing predictive models and machine learning techniques using SAS and R analytical language.

This book is ideal for those who are well-versed in writing code and have a basic understanding of statistics, but have limited experience in implementing predictive models and machine learning techniques for analyzing real world data. The most challenging part of solving industrial business problems is the practical and hands-on knowledge of building and deploying advanced predictive models and machine learning algorithms.

Applied Analytics through Case Studies Using SAS and R is your answer to solving these business problems by sharpening your analytical skills.



Shelve in:
Databases/General

User Level:
Intermediate–Advanced

SOURCE CODE ONLINE



Table of Contents

About the Author	xi
About the Contributor	xiii
About the Technical Reviewer	xv
Acknowledgments	xvii
Introduction	xix
Chapter 1: Data Analytics and Its Application in Various Industries	1
What Is Data Analytics?	2
Data Collection	3
Data Preparation	4
Data Analysis	4
Model Building	5
Results	5
Put into Use	5
Types of Analytics	6
Understanding Data and Its Types	7
What Is Big Data Analytics?	8
Big Data Analytics Challenges	10
Data Analytics and Big Data Tools	11
Role of Analytics in Various Industries	14
Who Are Analytical Competitors?	18
Key Models and Their Applications in Various Industries	18
Summary	21
References	21

Chapter 2: Banking Case Study	27
Applications of Analytics in the Banking Sector.....	28
Increasing Revenue by Cross-Selling and Up-Selling	29
Minimizing Customer Churn	30
Increase in Customer Acquisition	30
Predicting Bank-Loan Default.....	31
Predicting Fraudulent Activity.....	32
Case Study: Predicting Bank-Loan Defaults with Logistic Regression Model.....	34
Logistic Regression Equation	35
Odds	36
Logistic Regression Curve.....	37
Logistic Regression Assumptions.....	38
Logistic Regression Model Fitting and Evaluation.....	39
Statistical Test for Individual Independent Variable in Logistic	40
Regression Model.....	40
Predictive Value Validation in Logistic Regression Model.....	41
Logistic Regression Model Using R.....	46
About Data.....	47
Performing Data Exploration	47
Model Building and Interpretation of Full Data.....	52
Model Building and Interpretation of Training and Testing Data.....	56
Predictive Value Validation.....	61
Logistic Regression Model Using SAS	65
Model Building and Interpretation of Full Data.....	74
Summary.....	92
References.....	92
Chapter 3: Retail Case Study	97
Supply Chain in the Retail Industry.....	98
Types of Retail Stores	99

Role of Analytics in the Retail Sector	100
Customer Engagement	100
Supply Chain Optimization	101
Price Optimization	103
Space Optimization and Assortment Planning.....	103
Case Study: Sales Forecasting for Gen Retailers with SARIMA Model.....	105
Overview of ARIMA Model	107
Three Steps of ARIMA Modeling.....	111
Identification Stage	111
Estimation and Diagnostic Checking Stage.....	113
Forecasting Stage.....	114
Seasonal ARIMA Models or SARIMA.....	115
Evaluating Predictive Accuracy of Time Series Model	117
Seasonal ARIMA Model Using R	118
About Data.....	119
Performing Data Exploration for Time Series Data	119
Seasonal ARIMA Model Using SAS.....	133
Summary.....	158
References	159
Chapter 4: Telecommunication Case Study	161
Types of Telecommunications Networks.....	162
Role of Analytics in the Telecommunications Industry.....	163
Predicting Customer Churn	163
Network Analysis and Optimization	165
Fraud Detection and Prevention	166
Price Optimization	166
Case Study: Predicting Customer Churn with Decision Tree Model	168
Advantages and Limitations of the Decision Tree.....	169
Handling Missing Values in the Decision Tree	170

Handling Model Overfitting in Decision Tree.....	170
How the Decision Tree Works	171
Measures of Choosing the Best Split Criteria in Decision Tree.....	172
Decision Tree Model Using R.....	179
About Data.....	179
Performing Data Exploration	180
Splitting Data Set into Training and Testing.....	183
Model Building & Interpretation on Training and Testing Data.....	184
Decision Tree Model Using SAS	193
Model Building and Interpretation of Full Data.....	200
Model Building and Interpretation on Training and Testing Data	208
Summary.....	217
References	217
Chapter 5: Healthcare Case Study	221
Application of Analytics in the Healthcare Industry	224
Predicting the Outbreak of Disease and Preventative Management	225
Predicting the Readmission Rate of the Patients	225
Healthcare Fraud Detection	227
Improve Patient Outcomes & Lower Costs	228
Case Study: Predicting Probability of Malignant and Benign Breast Cancer with Random Forest Model	230
Working of Random Forest Algorithm.....	230
Random Forests Model Using R	238
Random Forests Model Using SAS	249
Summary.....	271
References	271
Chapter 6: Airline Case Study	277
Application of Analytics in the Airline Industry.....	280
Personalized Offers and Passenger Experience	281
Safer Flights	282

Airline Fraud Detection	283
Predicting Flight Delays.....	284
Case Study: Predicting Flight Delays with Multiple Linear Regression Model	286
Multiple Linear Regression Equation	287
Multiple Linear Regression Assumptions and Checking for Violation of Model Assumptions	287
Variables Selection in Multiple Linear Regression Model.....	290
Evaluating the Multiple Linear Regression Model	290
Multiple Linear Regression Model Using R	292
About Data.....	293
Performing Data Exploration	293
Model Building & Interpretation on Training and Testing Data.....	299
Multiple Linear Regression Model Using SAS	311
Summary.....	340
References.....	340
Chapter 7: FMCG Case Study	345
Application of Analytics in FMCG Industry	346
Customer Experience & Engagement.....	347
Sales and Marketing.....	347
Logistics Management	348
Markdown Optimization	349
Case Study: Customer Segmentation with RFM Model and K-means Clustering	350
Overview of RFM Model	351
Overview of K-means Clustering.....	355
RFM Model & K-means Clustering Using R.....	358
About Data.....	358
Performing Data Exploration	359
RFM Model & K-means Clustering Using SAS.....	376
Summary.....	393
References.....	394
Index.....	397