

Contents

About the Authors xiv

Preface xvii

Chapter 1 Introduction 1

1.1 Decision Making 4

1.2 Business Analytics Defined 5

1.3 A Categorization of Analytical Methods and Models 5

Descriptive Analytics 5

Predictive Analytics 6

Prescriptive Analytics 6

**Analytics in Action: Procter & Gamble Uses Business Analytics
to Redesign its Supply Chain 7**

1.4 Big Data 8

1.5 Business Analytics in Practice 9

Financial Analytics 9

Human Resource (HR) Analytics 10

Marketing Analytics 10

Health Care Analytics 10

Supply Chain Analytics 11

Analytics for Government and Nonprofits 11

Sports Analytics 12

Web Analytics 12

Summary 13

Glossary 13

Chapter 2 Descriptive Statistics 15

Analytics in Action: U.S. Census Bureau 16

2.1 Overview of Using Data: Definitions and Goals 16

2.2 Types of Data 17

Population and Sample Data 17

Quantitative and Categorical Data 18

Cross-Sectional and Time Series Data 18

Sources of Data 18

2.3 Modifying Data in Excel 21

Sorting and Filtering Data in Excel 21

Conditional Formatting of Data in Excel 23

2.4	Creating Distributions from Data	25
	Frequency Distributions for Categorical Data	25
	Relative Frequency and Percent Frequency Distributions	27
	Frequency Distributions for Quantitative Data	28
	Histograms	31
	Cumulative Distributions	34
2.5	Measures of Location	35
	Mean (Arithmetic Mean)	35
	Median	36
	Mode	37
	Geometric Mean	38
2.6	Measures of Variability	40
	Range	41
	Variance	41
	Standard Deviation	43
	Coefficient of Variation	44
2.7	Analyzing Distributions	44
	Percentiles	44
	Quartiles	45
	z-scores	46
	Empirical Rule	48
	Identifying Outliers	48
	Box Plots	49
2.8	Measures of Association Between Two Variables	51
	Scatter Charts	51
	Covariance	52
	Correlation Coefficient	55
	Summary	57
	Glossary	57
	Problems	58
	Case: Heavenly Chocolates Web Site Transactions	66
	Appendix: Creating Box Plots in XLMiner	67

Chapter 3 Data Visualization 70

Analytics in Action: Cincinnati Zoo & Botanical Garden 71

3.1 Overview of Data Visualization 73

Effective Design Techniques 73

3.2 Tables 75

Table Design Principles 77

Crosstabulation 79

PivotTables in Excel 80

3.3 Charts 85

Scatter Charts 85

Line Charts 87

	Bar Charts and Column Charts	90
	A Note on Pie Charts and 3-D Charts	93
	Bubble Charts	93
	Heat Maps	95
	Additional Charts for Multiple Variables	97
	PivotCharts in Excel	101
3.4	Advanced Data Visualization	102
	Advanced Charts	103
	Geographic Information Systems Charts	104
3.5	Data Dashboards	105
	Principles of Effective Data Dashboards	106
	Applications of Data Dashboards	106
	Summary	108
	Glossary	109
	Problems	110
	Case Problem: All-Time Movie Box Office Data	118
	Appendix: Creating a Scatter Chart Matrix and a Parallel Coordinates Plot with XLMiner	119

Chapter 4 Linear Regression 123

Analytics in Action: Alliance Data Systems 124

4.1 The Simple Linear Regression Model 125

Regression Model and Regression Equation 125

Estimated Regression Equation 126

4.2 Least Squares Method 127

Least Squares Estimates of the Regression Parameters 129

Using Excel's Chart Tools to Compute the Estimated Regression Equation 132

4.3 Assessing the Fit of the Simple Linear Regression Model 133

The Sums of Squares 134

The Coefficient of Determination 136

Using Excel's Chart Tools to Compute the Coefficient of Determination 137

4.4 The Multiple Regression Model 138

Regression Model and Regression Equation 138

Estimated Multiple Regression Equation 138

Least Squares Method and Multiple Regression 139

Butler Trucking Company and Multiple Regression 140

Using Excel's Regression Tool to Develop the Estimated Multiple Regression Equation 140

4.5 Inference and Regression 143

Conditions Necessary for Valid Inference in the Least Squares Regression Model 144

Testing for an Overall Regression Relationship 148

Testing Individual Regression Parameters	150
Addressing Nonsignificant Independent Variables	153
Multicollinearity	154
Inference and Very Large Samples	156
4.6 Categorical Independent Variables	161
Butler Trucking Company and Rush Hour	161
Interpreting the Parameters	162
More Complex Categorical Variables	164
4.7 Modeling Nonlinear Relationships	165
Quadratic Regression Models	167
Piecewise Linear Regression Models	170
Interaction Between Independent Variables	173
4.8 Model Fitting	177
Variable Selection Procedures	177
Overfitting	179
Summary	180
Glossary	180
Problems	182
Case Problem: Alumni Giving	197
Appendix: Using XLMiner for Regression	198

Chapter 5 Time Series Analysis and Forecasting 202

Analytics in Action: Forecasting Demand for a Broad Line of Office Products 203

5.1 Time Series Patterns	205
Horizontal Pattern	205
Trend Pattern	207
Seasonal Pattern	209
Trend and Seasonal Pattern	209
Cyclical Pattern	211
Identifying Time Series Patterns	212
5.2 Forecast Accuracy	212
5.3 Moving Averages and Exponential Smoothing	217
Moving Averages	217
Forecast Accuracy	221
Exponential Smoothing	221
Forecast Accuracy	224
5.4 Using Regression Analysis for Forecasting	226
Linear Trend Projection	226
Seasonality	228
Seasonality Without Trend	228
Seasonality with Trend	230
Using Regression Analysis as a Causal Forecasting Method	231

Combining Causal Variables with Trend and Seasonality
Effects 235

Considerations in Using Regression in Forecasting 235

5.5 Determining the Best Forecasting Model to Use 236

Summary 237

Glossary 237

Problems 238

Case Problem: Forecasting Food and Beverage Sales 246

Appendix: Using XLMiner for Forecasting 247

Chapter 6 Data Mining 251

**Analytics in Action: Online Retailers Using Predictive Analytics
to Cater to Customers 252**

6.1 Data Sampling 253

6.2 Data Preparation 254

Treatment of Missing Data 254

Identification of Outliers and Erroneous Data 254

Variable Representation 254

6.3 Unsupervised Learning 255

Cluster Analysis 256

Association Rules 265

6.4 Supervised Learning 269

Partitioning Data 269

Classification Accuracy 273

Prediction Accuracy 277

k-Nearest Neighbors 277

Classification and Regression Trees 283

Logistic Regression 299

Summary 308

Glossary 309

Problems 311

Case Problem: Grey Code Corporation 319

Chapter 7 Spreadsheet Models 320

**Analytics in Action: Procter and Gamble Sets Inventory Targets Using
Spreadsheet Models 321**

7.1 Building Good Spreadsheet Models 322

Influence Diagrams 322

Building a Mathematical Model 322

Spreadsheet Design and Implementing the Model
in a Spreadsheet 324

- 7.2 What-If Analysis 327**
 - Data Tables 327
 - Goal Seek 331
- 7.3 Some Useful Excel Functions for Modeling 332**
 - SUM and SUMPRODUCT 332
 - IF and COUNTIF 333
 - VLOOKUP 337
- 7.4 Auditing Spreadsheet Models 339**
 - Trace Precedents and Dependents 339
 - Show Formulas 340
 - Evaluate Formulas 340
 - Error Checking 341
 - Watch Window 342

Summary 343

Glossary 343

Problems 344

Case Problem: Retirement Plan 350

Chapter 8 Linear Optimization Models 352

Analytics in Action: Timber Harvesting Model at MeadWestvaco Corporation 353

8.1 A Simple Maximization Problem 354

Problem Formulation 355

Mathematical Model for the Par, Inc. Problem 357

8.2 Solving the Par, Inc. Problem 358

The Geometry of the Par, Inc. Problem 358

Solving Linear Programs with Excel Solver 360

8.3 A Simple Minimization Problem 364

Problem Formulation 364

Solution for the M&D Chemicals Problem 365

8.4 Special Cases of Linear Program Outcomes 367

Alternative Optimal Solutions 367

Infeasibility 368

Unbounded 370

8.5 Sensitivity Analysis 372

Interpreting Excel Solver Sensitivity Report 372

8.6 General Linear Programming Notation and More Examples 374

Investment Portfolio Selection 375

Transportation Planning 378

Advertising Campaign Planning 381

8.7 Generating an Alternative Optimal Solution for a Linear Program 386

Summary 388

Glossary 389

Problems 390

Case Problem: Investment Strategy 398

Appendix: Solving Linear Optimization Models Using Analytic Solver Platform 399

Chapter 9 Integer Linear Optimization Models 405

Analytics in Action: Optimizing the Transport of Oil Rig Crews 406

9.1 Types of Integer Linear Optimization Models 406

9.2 Eastborne Realty, An Example of Integer Optimization 407

The Geometry of Linear All-Integer Optimization 408

9.3 Solving Integer Optimization Problems with Excel Solver 410

A Cautionary Note About Sensitivity Analysis 414

9.4 Applications Involving Binary Variables 415

Capital Budgeting 415

Fixed Cost 416

Bank Location 420

Product Design and Market Share Optimization 424

9.5 Modeling Flexibility Provided by Binary Variables 426

Multiple-Choice and Mutually Exclusive Constraints 427

k out of n Alternatives Constraint 427

Conditional and Corequisite Constraints 427

9.6 Generating Alternatives in Binary Optimization 428

Summary 430

Glossary 430

Problems 431

Case Problem: Applecore Children's Clothing 441

Appendix: Solving Integer Linear Optimization Problems Using Analytic Solver Platform 442

Chapter 10 Nonlinear Optimization Models 448

Analytics in Action: Intercontinental Hotels Optimizes Retail Pricing 449

10.1 A Production Application: Par, Inc. Revisited 449

An Unconstrained Problem 450

A Constrained Problem 450

Solving Nonlinear Optimization Models Using Excel Solver 453

Sensitivity Analysis and Shadow Prices in Nonlinear Models 454

10.2 Local and Global Optima 455

Overcoming Local Optima with Excel Solver 457

10.3 A Location Problem 459

10.4 Markowitz Portfolio Model 461

10.5 Forecasting Adoption of a New Product 465

Summary 469

Glossary 470

Problems 470

Case Problem: Portfolio Optimization with Transaction Costs 477

Appendix: Solving Nonlinear Optimization Problems with Analytic Solver Platform 480

Chapter 11 Monte Carlo Simulation 485

Analytics in Action: Reducing Patient Infections in the ICU 486

11.1 What-If Analysis 487

The Sanotronics Problem 487

Base-Case Scenario 487

Worst-Case Scenario 488

Best-Case Scenario 488

11.2 Simulation Modeling with Native Excel Functions 488

Use of Probability Distributions to Represent

Random Variables 489

Generating Values for Random Variables with Excel 491

Executing Simulation Trials with Excel 495

Measuring and Analyzing Simulation Output 495

11.3 Simulation Modeling with Analytic Solver Platform 498

The Land Shark Problem 499

Spreadsheet Model for Land Shark 499

Generating Values for Land Shark's Random Variables 500

Tracking Output Measures for Land Shark 503

Executing Simulation Trials and Analyzing Output for Land Shark 504

The Zappos Problem 506

Spreadsheet Model for Zappos 507

Modeling Random Variables for Zappos 510

Tracking Output Measures for Zappos 515

Executing Simulation Trials and Analyzing Output for Zappos 517

11.4 Simulation Optimization 518

11.5 Simulation Considerations 524

Verification and Validation 524

Advantages and Disadvantages of Using Simulation 524

Summary 525

Glossary 526

Problems 527

Case Problem: Four Corners 536

**Appendix 11.1: Incorporating Dependence Between
Random Variables 537**

Appendix 11.2: Probability Distributions for Random Variables 545

Chapter 12 Decision Analysis 550

Analytics in Action: Phytopharm's New Product Research and Development 551

12.1 Problem Formulation 552

Payoff Tables 553

Decision Trees 553

12.2 Decision Analysis Without Probabilities 554

Optimistic Approach 554

Conservative Approach 555

Minimax Regret Approach 555

12.3 Decision Analysis with Probabilities 557

Expected Value Approach 557

Risk Analysis 559

Sensitivity Analysis 560

12.4 Decision Analysis with Sample Information 561

Expected Value of Sample Information 566

Expected Value of Perfect Information 567

12.5 Computing Branch Probabilities with Bayes' Theorem 568

12.6 Utility Theory 571

Utility and Decision Analysis 573

Utility Functions 577

Exponential Utility Function 580

Summary 581

Glossary 582

Problems 584

Case Problem: Property Purchase Strategy 595

Appendix: Using Analytic Solver Platform to Create Decision Trees 596

Appendix A Basics of Excel 609

Appendix B Data Management and Microsoft Access 621

Appendix C Answers to Even-Numbered Exercises (online)

References 659

Index 661