# Contents

 Preface ............................................................... ix  
 About the Author ................................................... xv  
 Acknowledgments ................................................... xvii  

 Chapter 1 • Introduction to the Central Textbook Example ...................... 1  
   Introduction .................................................... 1  
   The Company .................................................. 2  
   Current Research Needs of the Company .................................. 2  
   Your Brief for the Case Example ...................................... 5  
   Extended Analytical Skills Needed in the Project ...................... 6  

 Chapter 2 • Introduction to the Statistics Process ................................ 9  
   Introductory Case: Big Data in the Airline Industry ................... 9  
   Introduction to the Statistics Process .................................. 11  
   Step 1: Your Needs & Requirements .................................... 12  
   Step 2: Getting Data ........................................... 13  
   Step 3: Extracting Statistics from the Data ............................ 15  
   Step 4: Understanding & Decision Making ............................. 17  
   Summary: Challenges in the Statistics Process ........................ 17  
   Advice to the Statistically Terrified ................................. 18  

 Chapter 3 • Introduction to Data ...................................... 21  
   Introductory Case: Royal FrieslandCampina .......................... 21  
   Brief Introduction to Samples, Populations & Data .................. 23  
   Basic Characteristics of Variables .................................... 27  

 Chapter 4 • Data Collection & Capture ..................................... 33  
   Introduction .................................................... 33  
   Correct Sampling .............................................. 34  
   Choose Constructs and Variable Measurements ........................ 35  
   Initial Data Capture: Which Package? ................................ 43  
   Dealing with Data Once It Has Been Captured ....................... 43
Database & Data Analysis Software .......................... 48
Some Complications in Datasets ............................. 48

Chapter 5 • Introduction to SAS® ........................................ 51
 Introductory Vignette: SAS On Top of the Analytics World ........................................ 51
 Brief Introduction to SAS ........................................... 52
 Introduction to the Textbook Materials ....................... 53
 Getting Started with SAS 9 or SAS Studio ................. 53

Chapter 6 • Basics of SAS Programs, Data Manipulation, Analysis & Reporting .......... 69
 Introduction ......................................................... 70
 The Running Data Example ...................................... 70
 The Pre-Analysis Data Cleaning & Preparation Steps ...... 72
 Overview of the Three Big Tasks in Business Statistics .... 73
 Basic Introduction to SAS Programming ..................... 73
 Major Task #1: Data Manipulation in SAS ................... 77
 Major Task #2: Data Analysis .................................... 83
 Major Task #3: SAS Reporting through Output Formats .... 84
 The Visual Programmer Mode in SAS Studio ................. 86
 Conclusion ......................................................... 88

Chapter 7 • Descriptive Statistics: Understand your Data ........................................ 89
 Introductory Case: 2007 AngloGold Ashanti Look Ahead ........................................ 90
 Introduction ......................................................... 91
 End Outcome of a Descriptive Statistics Analysis .......... 91
 Getting Descriptive Statistics in SAS ......................... 92
 Statistics Measuring Centrality ................................. 94
 Basic Statistics Assessing Variable Spread ................. 97
 Assessing Shape of a Variable’s Distribution ............... 99
 Conclusion on Descriptive Statistics ........................... 104
 Appendix A to Chapter 7: Basic Normality Statistics .... 104

Chapter 8 • Basics of Associating Variables ........................................ 109
 Introduction ......................................................... 109
Contents

What is Statistical Association? 110
Association Does Not Mean Causation 110
Overview of Associations for Different Variable Types 111
Relating Continuous or Ordinal Data:
  Correlation & Covariance 112
  Relating Categorical Variables 119

Chapter 9 • Using Basic Statistics to Check & Fix Data 123
  Introduction 123
  Inappropriate Data Points 124
  Dealing Practically with Missing Data 126
  Checking Centrality & Spread 127
  Strange Variable Distributions 128
  Dealing Practically with Multi-Item Scales 128

Chapter 10 • Introduction to Graphing in SAS 135
  Introduction 135
  Major Graphing Procedures in SAS 136
  The PROC SGPLOT Routine in SAS 138
  Multiple Plots Simultaneously through
    PROC SGPANEL 143
  Business Dashboards through PROC GKPI 143
  Geographical Mapping Using PROC GMAP 145
  PROC SGSCATTER for Multiple Scatterplots 146
  Conclusion on SAS Graphing 147

Chapter 11 • The Statistics Process: Fitting Models to Data 149
  Introduction 149
  Look for Patterns in the Data (Fit) 151
  Step 3: Interpret the Pattern 164
  Summary of the Statistics Process 168

Chapter 12 • Key Concepts: Size & Accuracy 171
  Illustrative Case: Pharmaceuticals I –
    AstraZeneca’s Crestor 172
  Introduction 173
Issue # 1: Size of a Statistic ................................. 173
Issue # 2: Accuracy of Statistics ............................. 177
The Aspects of Inaccuracy .................................... 179
Putting Statistical Size and Accuracy Together ........ 200
Conclusion ....................................................... 202
Appendix A to Chapter 12: More on Accuracy (optional) 203

Chapter 13 • Introduction to Linear Regression .................. 211
Illustrative Case: West Point ................................. 212
Introduction ..................................................... 213
The Core Textbook Case Example for Chapter 13 ....... 213
Introduction to Linear Regression ......................... 215
A Pictorial Walk through Regression ...................... 217
Implementing Multiple Regression in SAS ............... 226
Step 1: Collect, Capture and Clean Data .................. 227
Step 2: Run an Initial Regression Analysis ............... 231
Step 3: Assess Fit and Apply Remedies If Necessary .... 233
Step 4: Interpret the Regression Slopes ................... 257
Step 5: Reporting a Multiple Regression Result .......... 265
Other Statistical Forms ....................................... 266
Conclusion ....................................................... 267

Chapter 14 • Categories Explaining a Continuous Variable:
Comparing Two Means ........................................ 269
Introduction to Comparison of Categories ................ 270
Features of the Continuous Variable to Compare Across Categories ........................ 270
Two Types of Categories to Compare ....................... 271
Numbers of Categories to Compare: Two vs. More than Two ................................. 272
Data Assumptions and Alternatives when Comparing Categories ................................. 273
Comparing Two Means: T-Tests .............................. 275
Big Data ................................................................. 326
Data Warehousing ............................................... 330
Machine Learning & Algorithms ......................... 335
Simulation in Business Situations ...................... 336
Bayesian Statistics .............................................. 340
Conclusion ........................................................... 342

Chapter 19 • Bibliography ........................................... 343
Books and Articles .............................................. 343

Index ........................................................................ 351