Contents

Introduction ix

Part I Data Quality Defined 1

Chapter 1 Introductory Case Studies 3

Chapter 2 Definition and Scope of Data Quality for Analytics 21

Chapter 3 Data Availability 31

Chapter 4 Data Quantity 47

Chapter 5 Data Completeness 59

Chapter 6 Data Correctness 71

Chapter 7 Predictive Modeling 83

Chapter 8 Analytics for Data Quality 95

Chapter 9 Process Considerations for Data Quality 107

Part II Data Quality—Profiling and Improvement 121

Chapter 10 Profiling and Imputation of Missing Values 123

Chapter 11 Profiling and Replacement of Missing Data in a Time Series 141

Chapter 12 Data Quality Control across Related Tables 159

Chapter 13 Data Quality with Analytics 167

Chapter 14 Data Quality Profiling and Improvement with SAS Analytic Tools 181

Part III Consequences of Poor Data Quality—Simulation Studies . 199

Chapter 15 Introdution to Simulation Studies 201

Chapter 16 Simulating the Consequences of Poor Data Quality for Predictive Modeling 207

Chapter 17 Influence of Data Quality and Data Availability on Model Quality in Predictive Modeling 219

Chapter 18 Influence of Data Completeness on Model Quality in Predictive Modeling 231

Chapter 19 Influence of Data Correctness on Model Quality in Predictive Modeling 243

Svolba, Gerhard. Data Quality for Analytics Using SAS®. Copyright © 2012, SAS Institute Inc., Cary, North Carolina, USA. ALL RIGHTS RESERVED. For additional SAS resources, visit support.sas.com/publishing.

Chapter 20 Simulating the Consequences of Poor Data Quality in Time Series Forecasting 255
Chapter 21 Consequences of Data Quantity and Data Completeness in Time Series Forecasting 265
Chapter 22 Consequences of Random Disturbances in Time Series Data 273
Chapter 23 Consequences of Systematic Disturbances in Time Series Data 281
Appendix A: Macro Code 289
Appendix B: General SAS Content and Programs 301
Appendix C: Using SAS Enterprise Miner for Simulation Studies 305
Appendix D: Macro to Determine the Optimal Length of the Available Data History 311
Appendix E: A Short Overview on Data Structures and Analytic Data Preparation 319
References 327
Index 329

Svolba, Gerhard. Data Quality for Analytics Using SAS®. Copyright © 2012, SAS Institute Inc., Cary, North Carolina, USA. ALL RIGHTS RESERVED. For additional SAS resources, visit support.sas.com/publishing.