

Contents

Foreword	xvii
Preface	xix
About the Book	xix
Quality Principles Applied in This Book	xx
Organization of the Book	xxi
Who Should Read This Book?	xxiii
References	xxiii
Acknowledgments	xxv
PART I: DEFINE PHASE	1
Chapter 1: Introduction	3
Introduction	3
Data, Analytics, AI, and Business Performance	5
Data as a Business Asset or Liability	6
Data Governance, Data Management, and Data Quality	7
Leadership Commitment to Data Quality	10
Key Takeaways	12
Conclusion	13
References	13
Chapter 2: Business Data	17
Introduction	17
Data in Business	18
Telemetry Data	21
Purpose of Data in Business	22
Business Data Views	24
Key Characteristics of Business Data	31
Critical Data Elements (CDEs)	32

xiii

Key Takeaways	34
Conclusion	35
References	35
Chapter 3: Data Quality in Business	37
Introduction	37
Data Quality Dimensions	39
Context in Data Quality	51
Consequences and Costs of Poor Data Quality	52
Data Depreciation and Its Factors	54
Data in IT Systems	56
Data Quality and Trusted Information	59
Key Takeaways	60
Conclusion	61
References	62
PART II: ANALYZE PHASE	63
Chapter 4: Causes for Poor Data Quality	65
Introduction	65
Data Quality RCA Techniques	66
Typical Causes of Poor Data Quality	71
Key Takeaways	78
Conclusion	79
References	80
Chapter 5: Data Lifecycle and Lineage	81
Introduction	81
Business-Enabled DLC Stages	82
IT Business-Enabled DLC Stages	86
Data Lineage	88
Key Takeaways	90
Conclusion	90
References	91
Chapter 6: Profiling for Data Quality	93
Introduction	93
Criteria for Data Profiling	95
Data Profiling Techniques for Measures of Centrality	98
Data Profiling Techniques for Measures of Variation	100
Integrating Centrality and Variation KPIs	109
Key Takeaways	112

Conclusion	112
References	112

PART III: REALIZE PHASE **113**

Chapter 7: Reference Architecture for Data Quality **115**

Introduction	115
Options to Remediate Data Quality	116
DataOps	118
Data Product	120
Data Fabric and Data Mesh	123
Data Enrichment	126
Key Takeaways	131
Conclusion	132
References	132

Chapter 8: Best Practices to Realize Data Quality **133**

Introduction	133
Overview of Best Practices	134
BP 1: Identify the Business KPIs and the Ownership of These KPIs and the Pertinent Data	136
BP 2: Build and Improve the Data Culture and Literacy in the Organization	138
BP 3: Define the Current and Desired State of Data Quality	142
BP 4: Follow the Minimalistic Approach to Data Capture	145
BP 5: Select and Define the Data Attributes for Data Quality	148
BP 6: Capture and Manage Critical Data with Data Standards in MDM Systems	152
Key Takeaways	155
Conclusion	158
References	158

Chapter 9: Best Practices to Realize Data Quality **161**

Introduction	161
BP 7: Rationalize and Automate the Integration of Critical Data Elements	162
BP 8: Define the SoR and Securely Capture Transactional Data in the SoR/OLTP System	168
BP 9: Build and Manage Robust Data Integration Capabilities	173
BP 10: Distribute Data Sourcing and Insight Consumption	181
Key Takeaways	186
Conclusion	189
References	189

PART IV: SUSTAIN PHASE	191
Chapter 10: Data Governance	193
Introduction	193
Data Governance Principles	195
Data Governance Design Components	197
Implementing the Data Governance Program	202
Data Observability	203
Data Compliance – ISO 27001, SOC1, and SOC2	205
Key Takeaways	206
Conclusion	208
References	209
Chapter 11: Protecting Data	211
Introduction	211
Data Classification	212
Data Safety	216
Data Security	218
Key Takeaways	220
Conclusion	220
References	221
Chapter 12: Data Ethics	223
Introduction	223
Data Ethics	224
Importance of Data Ethics	224
Principles of Data Ethics	225
Model Drift in Data Ethics	226
Data Privacy	228
Managing Data Ethically	230
Key Takeaways	235
Conclusion	235
References	236
Appendix 1: Abbreviations and Acronyms	237
Appendix 2: Glossary	241
Appendix 3: Data Literacy Competencies	245
About the Author	249
Index	251