Too Big to Ignore

The Business Case for Big Data

Phil Simon

WILEY
Contents

List of Tables and Figures xv

Preface xvii

Acknowledgments xxiii

Introduction This Ain't Your Father's Data 1
   Better Car Insurance through Data 2
   Potholes and General Road Hazards 5
   Recruiting and Retention 8
   How Big is Big? The Size of Big Data 10
   Why Now? Explaining the Big Data Revolution 12
   Central Thesis of Book 22
   Plan of Attack 24
   Who Should Read This Book? 25
   Summary 25
   Notes 26

Chapter 1 Data 101 and the Data Deluge 29
   The Beginnings: Structured Data 30
   Structure This! Web 2.0 and the Arrival of Big Data 33
   The Composition of Data: Then and Now 39
   The Current State of the Data Union 41
   The Enterprise and the Brave New Big Data World 43
   Summary 46
   Notes 47

Chapter 2 Demystifying Big Data 49
   Characteristics of Big Data 50
   The Anti-Definition: What Big Data Is Not 71
   Summary 72
   Notes 72

Chapter 3 The Elements of Persuasion: Big Data Techniques 77
   The Big Overview 79
Statistical Techniques and Methods 80
Data Visualization 84
Automation 88
Semantics 93
Big Data and the Gang of Four 98
Predictive Analytics 100
Limitations of Big Data 105
Summary 106
Notes 107

Chapter 4 Big Data Solutions 111
Projects, Applications, and Platforms 114
Other Data Storage Solutions 121
Websites, Start-ups, and Web Services 128
Hardware Considerations 133
The Art and Science of Predictive Analytics 136
Summary 137
Notes 137

Chapter 5 Case Studies: The Big Rewards of Big Data 141
Quantcast: A Small Big Data Company 141
Explorys: The Human Case for Big Data 147
NASA: How Contests, Gamification, and Open Innovation Enable Big Data 152
Summary 158
Notes 158

Chapter 6 Taking the Big Plunge 161
Before Starting 161
Starting the Journey 165
Avoiding the Big Pitfalls 174
Summary 181
Notes 181

Chapter 7 Big Data: Big Issues and Big Problems 183
Privacy: Big Data = Big Brother? 184
Big Security Concerns 188
Big, Pragmatic Issues 189
Summary 195
Notes 196

Chapter 8 Looking Forward: The Future of Big Data 197
Predicting Pregnancy 198
Big Data Is Here to Stay 200
Big Data Will Evolve 201
Projects and Movements 203
Big Data Will Only Get Bigger...and Smarter 205
The Internet of Things: The Move from Active to Passive Data Generation 206
Big Data: No Longer a Big Luxury 211
Stasis Is Not an Option 212
Summary 213
Notes 214

Final Thoughts 217

Selected Bibliography 221

About the Author 223

Index 225