Taming the Big Data Tidal Wave

Finding Opportunities in Huge Data Streams with Advanced Analytics

Bill Franks
Contents

Foreword xiii
Preface xvii
Acknowledgments xxv

PART ONE THE RISE OF BIG DATA ........................................... 1

Chapter 1 What Is Big Data and Why Does It Matter? .......... 3
  What Is Big Data? 4
  Is the "Big" Part or the "Data" Part More Important? 5
  How Is Big Data Different? 7
  How Is Big Data More of the Same? 9
  Risks of Big Data 10
  Why You Need to Tame Big Data 12
  The Structure of Big Data 14
  Exploring Big Data 16
  Most Big Data Doesn't Matter 17
  Filtering Big Data Effectively 20
  Mixing Big Data with Traditional Data 21
  The Need for Standards 22
  Today's Big Data Is Not Tomorrow's Big Data 24
  Wrap-Up 26
  Notes 27
Chapter 2  Web Data: The Original Big Data .................................. 29
  Web Data Overview  30
  What Web Data Reveals  36
  Web Data in Action  42
  Wrap-Up  50
  Note  51

Chapter 3  A Cross-Section of Big Data Sources and the Value They Hold ................................................. 53
  Auto Insurance: The Value of Telematics Data  54
  Multiple Industries: The Value of Text Data  57
  Multiple Industries: The Value of Time and Location Data  60
  Retail and Manufacturing: The Value of Radio Frequency Identification Data  64
  Utilities: The Value of Smart-Grid Data  68
  Gaming: The Value of Casino Chip Tracking Data  71
  Industrial Engines and Equipment: The Value of Sensor Data  73
  Video Games: The Value of Telemetry Data  76
  Telecommunications and Other Industries: The Value of Social Network Data  78
  Wrap-Up  82

PART TWO  TAMING BIG DATA: THE TECHNOLOGIES, PROCESSES, AND METHODS ........................................... 85

Chapter 4  The Evolution of Analytic Scalability .................... 87
  A History of Scalability  88
  The Convergence of the Analytic and Data Environments  90
  Massively Parallel Processing Systems  93
  Cloud Computing  102
  Grid Computing  109
Chapter 5  The Evolution of Analytic Processes .......................... 121

The Analytic Sandbox 122
What Is an Analytic Data Set? 133
Enterprise Analytic Data Sets 137
Embedded Scoring 145
Wrap-Up 151

Chapter 6  The Evolution of Analytic Tools and Methods ................................. 153

The Evolution of Analytic Methods 154
The Evolution of Analytic Tools 163
Wrap-Up 175
Notes 176

PART THREE  TAMING BIG DATA: THE PEOPLE AND APPROACHES .......................... 177

Chapter 7  What Makes a Great Analysis? ........................................ 179

Analysis versus Reporting 179
Analysis: Make It G.R.E.A.T.! 184
Core Analytics versus Advanced Analytics 186
Listen to Your Analysis 188
Framing the Problem Correctly 189
Statistical Significance versus Business Importance 191
Samples versus Populations 195
Making Inferences versus Computing Statistics 198
Wrap-Up 200
Chapter 8  What Makes a Great Analytic Professional? ..... 201

Who Is the Analytic Professional? 202
The Common Misconceptions about Analytic Professionals 203
Every Great Analytic Professional Is an Exception 204
The Often Underrated Traits of a Great Analytic Professional 208
Is Analytics Certification Needed, or Is It Noise? 222
Wrap-Up 224

Chapter 9  What Makes a Great Analytics Team? ............... 227

All Industries Are Not Created Equal 228
Just Get Started! 230
There’s a Talent Crunch out There 231
Team Structures 232
Keeping a Great Team’s Skills Up 237
Who Should Be Doing Advanced Analytics? 241
Why Can’t IT and Analytic Professionals Get Along? 245
Wrap-Up 247
Notes 248

PART FOUR  BRINGING IT TOGETHER: THE ANALYTICS CULTURE .......................................................... 249

Chapter 10 Enabling Analytic Innovation ...................... 251

Businesses Need More Innovation 252
Traditional Approaches Hamper Innovation 253
Defining Analytic Innovation 255
Iterative Approaches to Analytic Innovation 256
Consider a Change in Perspective 257
Are You Ready for an Analytic Innovation Center? 259
Wrap-Up 269
Note 270
Chapter 11 Creating a Culture of Innovation and Discovery

Setting the Stage 272
Overview of the Key Principles 274
Wrap-Up 290
Notes 291

Conclusion: Think Bigger! 293

About the Author 295

Index 297