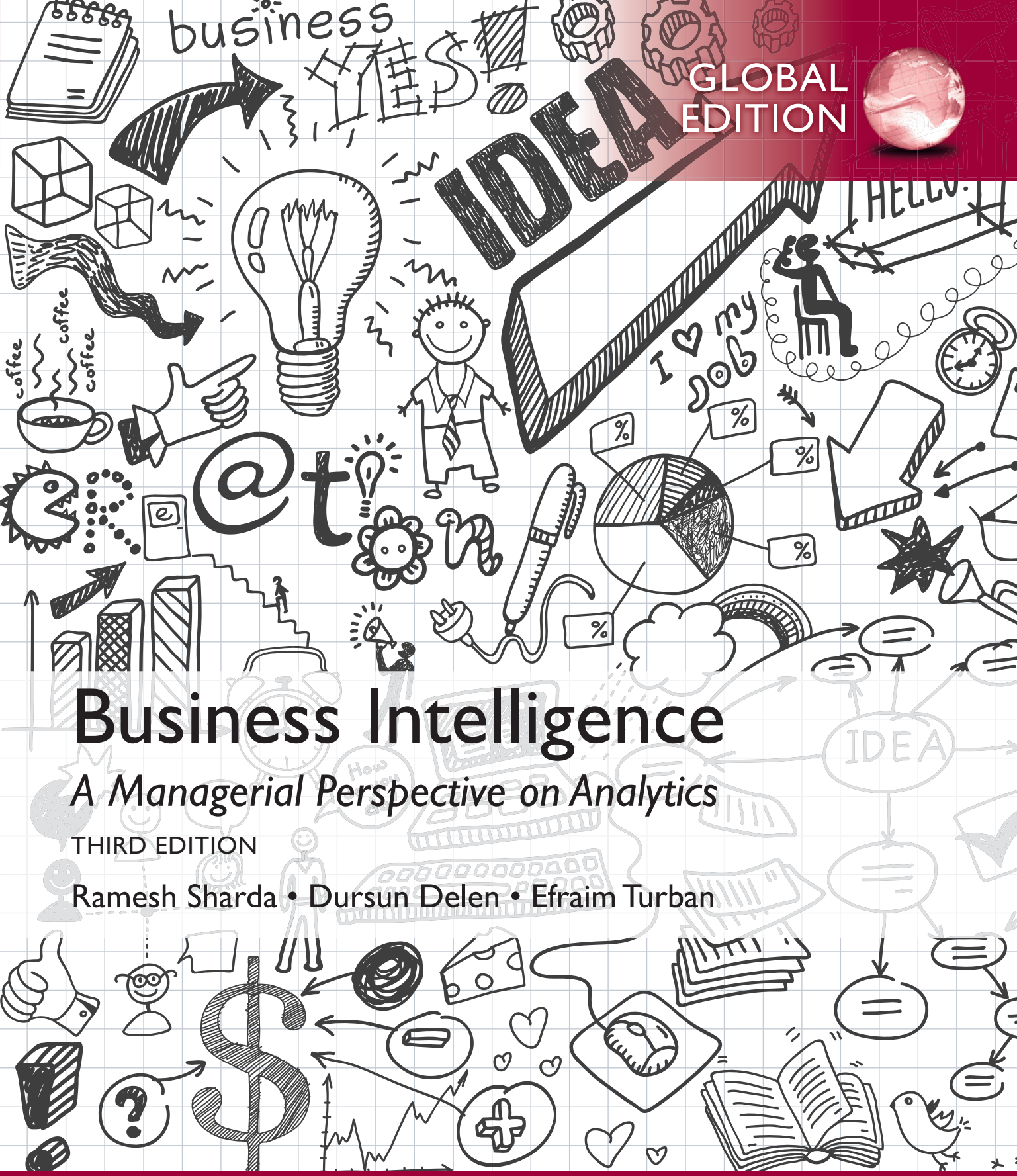


# Business Intelligence

## *A Managerial Perspective on Analytics*

THIRD EDITION

Ramesh Sharda • Dursun Delen • Efraim Turban

ALWAYS LEARNING

PEARSON

# Business Intelligence

THIRD EDITION

# BUSINESS INTELLIGENCE

## A MANAGERIAL PERSPECTIVE ON ANALYTICS

**Ramesh Sharda**
*Oklahoma State University*

**Dursun Delen**
*Oklahoma State University*

**Efraim Turban**
*University of Hawaii*

*With contributions by*

**J. E. Aronson**
*The University of Georgia*

**Ting-Peng Liang**
*National Sun Yat-sen University*

**David King**
*JDA Software Group, Inc.*

*Global Edition contributions by*

**Loveleen Gaur**
*Amity University, Noida*

PEARSON

Boston Columbus Indianapolis New York San Francisco Upper Saddle River
Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montréal Toronto
Delhi Mexico City São Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo

**Editor in Chief:** Stephanie Wall
**Acquisitions Editor:** Bob Horan
**Program Management Lead:** Ashley Santora
**Program Manager:** Denise Vaughn
**Marketing Manager:** Anne Fahlgren
**Project Management Lead:** Judy Leale
**Project Manager:** Tom Benfatti
**Publishing Operations Director, Global Edition:** Angshuman Chakraborty
**Publisher, Global Edition:** Laura Dent
**Manager - Publishing Operations, Global Edition:** Shokhi Shah Khandelwal
**Associate Print and Media Editor, Global Edition:** Anuprova Dey Chowdhuri
**Acquisitions Editor, Global Edition:** Sandhya Ghoshal
**Publishing Administrator, Global Edition:** Hema Mehta
**Project Editor, Global Edition:** Karthik Subramanian
**Operations Specialist:** Michelle Klein
**Creative Director:** Jayne Conte
**Cover Designer:** Suzanne Behnke
**Media Project Manager:** Lisa Rinaldi
**Full-Service Project Management:** George Jacob, Integra Software Solutions.
**Cover Photo:** Shutterstock/Art'nLera
**Cover Printer:** Courier Kendallville

Pearson Education Limited
Edinburgh Gate
Harlow
Essex CM20 2JE
England and Associated Companies throughout the world

Visit us on the World Wide Web at: www.pearsonglobaleditions.com

© Pearson Education Limited 2014

The rights of Ramesh Sharda, Dursun Delen, Efraim Turban, J. E. Aronson, Ting-Peng Liang, and David King to be identified as the authors of this work have been asserted by him in accordance with the Copyright, Designs and Patents Act 1988.

*Authorized adaptation from the United States edition, entitled* Business Intelligence: A Managerial Perspective on Analytics, Third Edition, ISBN 978-0-13-305105-6, *by Ramesh Sharda, Dursun Delen, Efraim Turban, J. E. Aronson, Ting-Peng Liang, and David King, published by Pearson Education © 2014.*

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without either the prior written permission of the publisher or a license permitting restricted copying in the United Kingdom issued by the Copyright Licensing Agency Ltd, Saffron House, 6–10 Kirby Street, London EC1N 8TS.

All trademarks used herein are the property of their respective owners. The use of any trademark in this text does not vest in the author or publisher any trademark ownership rights in such trademarks, nor does the use of such trademarks imply any affiliation with or endorsement of this book by such owners.

Microsoft and/or its respective suppliers make no representations about the suitability of the information contained in the documents and related graphics published as part of the services for any purpose. All such documents and related graphics are provided "as is" without warranty of any kind. Microsoft and/or its respective suppliers hereby disclaim all warranties and conditions with regard to this information, including all warranties and conditions of merchantability, whether express, implied or statutory, fitness for a particular purpose, title and non-infringement. In no event shall Microsoft and/or its respective suppliers be liable for any special, indirect or consequential damages or any damages whatsoever resulting from loss of use, data or profits, whether in an action of contract, negligence or other tortious action, arising out of or in connection with the use or performance of information available from the services.

The documents and related graphics contained herein could include technical inaccuracies or typographical errors. Changes are periodically added to the information herein. Microsoft and/or its respective suppliers may make improvements and/or changes in the product(s) and/or the program(s) described herein at any time. Partial screen shots may be viewed in full within the software version specified.

Microsoft® and Windows® are registered trademarks of the Microsoft Corporation in the U.S.A. and other countries. This book is not sponsored or endorsed by or affiliated with the Microsoft Corporation.

British Library Cataloguing-in-Publication Data A catalogue record for this book is available from the British Library

Typeset in Garamond by Integra Software Solutions

Printed and bound by Courier Kendalville in The United States of America

PEARSON

ISBN 10: 1-292-00487-8
ISBN 13: 978-1-292-00487-7

# BRIEF CONTENTS

*Preface 17*
*About the Authors 23*

**Chapter 1** An Overview of Business Intelligence, Analytics, and Decision Support 27
**Chapter 2** Data Warehousing 61
**Chapter 3** Business Reporting, Visual Analytics, and Business Performance Management 119
**Chapter 4** Data Mining 169
**Chapter 5** Text and Web Analytics 225
**Chapter 6** Big Data and Analytics 299
**Chapter 7** Business Analytics: Emerging Trends and Future Impacts 351

*Glossary 393*
*Index 401*

# CONTENTS

*Preface 17*
*About the Authors 23*

**Chapter 1 An Overview of Business Intelligence, Analytics, and Decision Support 27**

1.1 Opening Vignette: Magpie Sensing Employs Analytics to Manage a Vaccine Supply Chain Effectively and Safely 28
1.2 Changing Business Environments and Computerized Decision Support 30
The Business Pressures–Responses–Support Model 30
1.3 A Framework for Business Intelligence (BI) 32
Definitions of BI 32
A Brief History of BI 33
The Architecture of BI 34
The Origins and Drivers of BI 34
► **APPLICATION CASE 1.1 Sabre Helps Its Clients Through Dashboards and Analytics 35**
A Multimedia Exercise in Business Intelligence 36
1.4 Intelligence Creation, Use, and BI Governance 37
A Cyclical Process of Intelligence Creation and Use 37
Intelligence and Espionage 38
1.5 Transaction Processing Versus Analytic Processing 39
1.6 Successful BI Implementation 40
The Typical BI User Community 40
Appropriate Planning and Alignment with the Business Strategy 40
Real-Time, On-Demand BI Is Attainable 41
Developing or Acquiring BI Systems 42
Justification and Cost–Benefit Analysis 42
Security and Protection of Privacy 42
Integration of Systems and Applications 42
1.7 Analytics Overview 43
Descriptive Analytics 44
Predictive Analytics 44
► **APPLICATION CASE 1.2 Eliminating Inefficiencies at Seattle Children's Hospital 45**
► **APPLICATION CASE 1.3 Analysis at the Speed of Thought 46**
Prescriptive Analytics 46
► **APPLICATION CASE 1.4 *Moneyball*: Analytics in Sports and Movies 47**
► **APPLICATION CASE 1.5 Analyzing Athletic Injuries 48**
Analytics Applied to Different Domains 48

► APPLICATION CASE 1.6 **Industrial and Commercial Bank of China (ICBC) Employs Models to Reconfigure Its Branch Network 49**
Analytics or Data Science? 50
1.8 Brief Introduction to Big Data Analytics 51
What Is Big Data? 51
► APPLICATION CASE 1.7 **Gilt Groupe's Flash Sales Streamlined by Big Data Analytics** 52
1.9 Plan of the Book 53
1.10 Resources, Links, and the Teradata University Network Connection 55
Resources and Links 55
Vendors, Products, and Demos 55
Periodicals 55
The Teradata University Network Connection 55
The Book's Web Site 55
*Key Terms 56*
*Questions for Discussion 56*
*Exercises 56*
*End-of-Chapter Application Case 57*
*References 59*

**Chapter 2 Data Warehousing 61**
2.1 Opening Vignette: Isle of Capri Casinos Is Winning with Enterprise Data Warehouse 62
2.2 Data Warehousing Definitions and Concepts 64
What Is a Data Warehouse? 64
A Historical Perspective to Data Warehousing 64
Characteristics of Data Warehousing 66
Data Marts 67
Operational Data Stores 67
Enterprise Data Warehouses (EDW) 68
► APPLICATION CASE 2.1 **A Better Data Plan: Well-Established TELCOs Leverage Data Warehousing and Analytics to Stay on Top in a Competitive Industry 68**
Metadata 69
2.3 Data Warehousing Process Overview 70
► APPLICATION CASE 2.2 **Data Warehousing Helps MultiCare Save More Lives 71**
2.4 Data Warehousing Architectures 73
Alternative Data Warehousing Architectures 76
Which Architecture Is the Best? 79
2.5 Data Integration and the Extraction, Transformation, and Load (ETL) Processes 80
Data Integration 81
► APPLICATION CASE 2.3 **BP Lubricants Achieves BIGS Success 81**
Extraction, Transformation, and Load 83

2.6 Data Warehouse Development 85
► APPLICATION CASE 2.4 Things Go Better with Coke's Data Warehouse 86
Data Warehouse Development Approaches 88
► APPLICATION CASE 2.5 Starwood Hotels & Resorts Manages Hotel Profitability with Data Warehousing 89
Additional Data Warehouse Development Considerations 91
Representation of Data in Data Warehouse 92
Analysis of Data in Data Warehouse 93
OLAP Versus OLTP 93
OLAP Operations 94
2.7 Data Warehousing Implementation Issues 97
► APPLICATION CASE 2.6 EDW Helps Connect State Agencies in Michigan 99
Massive Data Warehouses and Scalability 100
2.8 Real-Time Data Warehousing 101
► APPLICATION CASE 2.7 Egg Plc Fries the Competition in Near Real Time 102
2.9 Data Warehouse Administration, Security Issues, and Future Trends 105
The Future of Data Warehousing 107
2.10 Resources, Links, and the Teradata University Network Connection 110
Resources and Links 110
Cases 110
Vendors, Products, and Demos 111
Periodicals 111
Additional References 111
The Teradata University Network (TUN) Connection 111
*Key Terms 112*
*Questions for Discussion 112*
*Exercises 113*
*End-of-Chapter Application Case 114*
*References 116*

**Chapter 3 Business Reporting, Visual Analytics, and Business Performance Management 119**

3.1 Opening Vignette: Self-Service Reporting Environment Saves Millions for Corporate Customers 120
3.2 Business Reporting Definitions and Concepts 123
What Is a Business Report? 124
► APPLICATION CASE 3.1 Delta Lloyd Group Ensures Accuracy and Efficiency in Financial Reporting 126
Components of Business Reporting Systems 127
► APPLICATION CASE 3.2 Flood of Paper Ends at FEMA 128

3.3 Data and Information Visualization 129
► **APPLICATION CASE 3.3 Tableau Saves Blastrac Thousands of Dollars with Simplified Information Sharing 130**
A Brief History of Data Visualization 131
► **APPLICATION CASE 3.4 TIBCO Spotfire Provides Dana-Farber Cancer Institute with Unprecedented Insight into Cancer Vaccine Clinical Trials 133**
3.4 Different Types of Charts and Graphs 134
Basic Charts and Graphs 134
Specialized Charts and Graphs 135
3.5 The Emergence of Data Visualization and Visual Analytics 138
Visual Analytics 140
High-Powered Visual Analytics Environments 140
3.6 Performance Dashboards 143
Dashboard Design 145
► **APPLICATION CASE 3.5 Dallas Cowboys Score Big with Tableau and Teknion 145**
► **APPLICATION CASE 3.6 Saudi Telecom Company Excels with Information Visualization 146**
What to Look For in a Dashboard 148
Best Practices in Dashboard Design 148
Benchmark Key Performance Indicators with Industry Standards 149
Wrap the Dashboard Metrics with Contextual Metadata 149
Validate the Dashboard Design by a Usability Specialist 149
Prioritize and Rank Alerts/Exceptions Streamed to the Dashboard 149
Enrich Dashboard with Business-User Comments 149
Present Information in Three Different Levels 149
Pick the Right Visual Construct Using Dashboard Design Principles 150
Provide for Guided Analytics 150
3.7 Business Performance Management 150
Closed-Loop BPM Cycle 150
► **APPLICATION CASE 3.7 IBM Cognos Express Helps Mace for Faster and Better Business Reporting 153**
3.8 Performance Measurement 154
Key Performance Indicator (KPI) 154
Performance Measurement System 155
3.9 Balanced Scorecards 156
The Four Perspectives 156
The Meaning of Balance in BSC 158
Dashboards Versus Scorecards 159
3.10 Six Sigma as a Performance Measurement System 159
The DMAIC Performance Model 160
Balanced Scorecard Versus Six Sigma 160
Effective Performance Measurement 160

► **APPLICATION CASE 3.8 Expedia.com's Customer Satisfaction Scorecard 162**
*Key Terms 164*
*Questions for Discussion 165*
*Exercises 165*
*End-of-Chapter Application Case 166*
*References* 168

**Chapter 4 Data Mining 169**

4.1 Opening Vignette: Cabela's Reels in More Customers with Advanced Analytics and Data Mining 170
4.2 Data Mining Concepts and Applications 172
Definitions, Characteristics, and Benefits 173
► **APPLICATION CASE 4.1 Smarter Insurance: Infinity P&C Improves Customer Service and Combats Fraud with Predictive Analytics 174**
How Data Mining Works 179
► **APPLICATION CASE 4.2 Harnessing Analytics to Combat Crime: Predictive Analytics Helps Memphis Police Department Pinpoint Crime and Focus Police Resources 179**
Data Mining Versus Statistics 183
4.3 Data Mining Applications 183
► **APPLICATION CASE 4.3 A Mine on Terrorist Funding 186**
4.4 Data Mining Process 187
Step 1: Business Understanding 187
Step 2: Data Understanding 188
Step 3: Data Preparation 188
Step 4: Model Building 190
Step 5: Testing and Evaluation 192
Step 6: Deployment 192
► **APPLICATION CASE 4.4 Data Mining in Cancer Research 193**
Other Data Mining Standardized Processes and Methodologies 194
4.5 Data Mining Methods 196
Classification 196
Estimating the True Accuracy of Classification Models 197
► **APPLICATION CASE 4.5 2degrees Gets a 1275 Percent Boost in Churn Identification 203**
Cluster Analysis for Data Mining 204
Association Rule Mining 206
4.6 Data Mining Software Tools 210
► **APPLICATION CASE 4.6 Data Mining Goes to Hollywood: Predicting Financial Success of Movies 213**
4.7 Data Mining Privacy Issues, Myths, and Blunders 216
Data Mining and Privacy Issues 216
► **APPLICATION CASE 4.7 Predicting Customer Buying Patterns—The Target Story 217**

Data Mining Myths and Blunders 218
*Key Terms 220*
*Questions for Discussion 220*
*Exercises 221*
*End-of-Chapter Application Case 223*
*References* 223

**Chapter 5 Text and Web Analytics 225**

5.1 Opening Vignette: Machine Versus Men on *Jeopardy!*: The Story of Watson 226

5.2 Text Analytics and Text Mining Overview 229

▶ **APPLICATION CASE 5.1 Text Mining for Patent Analysis 232**

5.3 Natural Language Processing 233

▶ **APPLICATION CASE 5.2 Text Mining Improves Hong Kong Government's Ability to Anticipate and Address Public Complaints 235**

5.4 Text Mining Applications 237

Marketing Applications 238

Security Applications 238

▶ **APPLICATION CASE 5.3 Mining for Lies 239**

Biomedical Applications 241

Academic Applications 242

▶ **APPLICATION CASE 5.4 Text mining and Sentiment Analysis Help Improve Customer Service Performance 243**

5.5 Text Mining Process 244

Task 1: Establish the Corpus 245

Task 2: Create the Term–Document Matrix 246

Task 3: Extract the Knowledge 248

▶ **APPLICATION CASE 5.5 Research Literature Survey with Text Mining 250**

5.6 Sentiment Analysis 253

▶ **APPLICATION CASE 5.6 Whirlpool Achieves Customer Loyalty and Product Success with Text Analytics 255**

Sentiment Analysis Applications 256

Sentiment Analysis Process 258

Methods for Polarity Identification 259

Using a Lexicon 260

Using a Collection of Training Documents 261

Identifying Semantic Orientation of Sentences and Phrases 261

Identifying Semantic Orientation of Document 261

5.7 Web Mining Overview 262

Web Content and Web Structure Mining 264

5.8 Search Engines 267

Anatomy of a Search Engine 267

Development Cycle 267

Web Crawler 268
Document Indexer 268
Response Cycle 269
Query Analyzer 269
Document Matcher/Ranker 269
Search Engine Optimization 270
Methods for Search Engine Optimization 271
► **APPLICATION CASE 5.7 Understanding Why Customers Abandon Shopping Carts Results in a $10 Million Sales Increase 272**
5.9 Web Usage Mining (Web Analytics) 274
Web Analytics Technologies 274
► **APPLICATION CASE 5.8 Allegro Boosts Online Click-Through Rates by 500 Percent with Web Analysis 275**
Web Analytics Metrics 277
Web Site Usability 277
Traffic Sources 278
Visitor Profiles 229
Conversion Statistics 280
5.10 Social Analytics 281
Social Network Analysis 282
Social Network Analysis Metrics 283
► **APPLICATION CASE 5.9 Social Network Analysis Helps Telecommunication Firms 283**
Connections 284
Distributions 285
Segmentation 285
Social Media Analytics 285
How Do People Use Social Media? 286
► **APPLICATION CASE 5.10 Measuring the Impact of Social Media at Lollapalooza 287**
Measuring the Social Media Impact 288
Best Practices in Social Media Analytics 289
► **APPLICATION CASE 5.11 eHarmony Uses Social Media to Help Take the Mystery Out of Online Dating 290**
*Key Terms 293*
*Questions for Discussion 293*
*Exercises 293*
*End-of-Chapter Application Case 294*
*References* 296

**Chapter 6 Big Data and Analytics 299**
6.1 Opening Vignette: Big Data Meets Big Science at CERN 300
6.2 Definition of Big Data 303
The Vs That Define Big Data 304
► **APPLICATION CASE 6.1 BigData Analytics Helps Luxottica Improvement its Marketing Effectiveness 307**

6.3 Fundamentals of Big Data Analytics 308
Business Problems Addressed by Big Data Analytics 311
► **APPLICATION CASE 6.2 Top 5 Investment Bank Achieves Single Source of the Truth 312**
6.4 Big Data Technologies 313
MapReduce 313
Why Use MapReduce? 315
Hadoop 315
How Does Hadoop Work? 315
Hadoop Technical Components 316
Hadoop: The Pros and Cons 317
NoSQL 319
► **APPLICATION CASE 6.3 eBay's Big Data Solution 320**
6.5 Data Scientist 321
Where Do Data Scientists Come From? 322
► **APPLICATION CASE 6.4 Big Data and Analytics in Politics 325**
6.6 Big Data and Data Warehousing 326
Use Cases for Hadoop 327
Use Cases for Data Warehousing 328
The Gray Areas (Any One of the Two Would Do the Job) 329
Coexistence of Hadoop and Data Warehouse 329
6.7 Big Data Vendors 331
► **APPLICATION CASE 6.5 Dublin City Council Is Leveraging Big Data to Reduce Traffic Congestion 333**
► **APPLICATION CASE 6.6 Creditreform Boosts Credit Rating Quality with Big Data Visual Analytics 337**
6.8 Big Data And Stream Analytics 338
Stream Analytics Versus Perpetual Analytics 339
Critical Event Processing 340
Data Stream Mining 341
6.9 Applications of Stream Analytics 341
e-Commerce 341
Telecommunications 342
► **APPLICATION CASE 6.7 Turning Machine-Generated Streaming Data into Valuable Business Insights 342**
Law Enforcement and Cyber Security 344
Power Industry 344
Financial Services 344
Health Sciences 344
Government 345
*Key Terms 346*
*Questions for Discussion 346*

*Exercises 346*
*End-of-Chapter Application Case 347*
*References 348*

**Chapter 7 Business Analytics: Emerging Trends and Future Impacts 351**

7.1 Opening Vignette: Oklahoma Gas and Electric Employs Analytics to Promote Smart Energy Use 352

7.2 Location-Based Analytics for Organizations 353

Geospatial Analytics 353

► **APPLICATION CASE 7.1 Great Clips Employs Spatial Analytics to Shave Time in Location Decisions 355**

Real-Time Location Intelligence 357

► **APPLICATION CASE 7.2 Quiznos Targets Customers for its Sandwiches 358**

7.3 Analytics Applications for Consumers 359

► **APPLICATION CASE 7.3 A Life Coach in Your Pocket 360**

7.4 Recommendation Engines 362

7.5 The Web 2.0 Revolution and Online Social Networking 363

Representative Characteristics of Web 2.0 364

Social Networking 364

A Definition and Basic Information 365

Implications of Business and Enterprise Social Networks 365

7.6 Cloud Computing and BI 366

Service-Oriented DSS 367

Data-as-a-Service (DaaS) 369

Information-as-a-Service (Information on Demand) (IaaS) 370

Analytics-as-a-Service (AaaS) 371

7.7 Impacts of Analytics In Organizations: An Overview 372

New Organizational Units 373

Restructuring Business Processes and Virtual Teams 373

Job Satisfaction 374

Job Stress and Anxiety 374

Analytics' Impact on Managers' Activities and Their Performance 374

7.8 Issues of Legality, Privacy, and Ethics 376

Legal Issues 376

Privacy 376

Recent Technology Issues in Privacy and Analytics 376

Ethics in Decision Making and Support 379

7.9 An Overview of the Analytics Ecosystem 379

Analytics Industry Clusters 380

Data Infrastructure Providers 380

Data Warehouse Industry 381

Middleware/BI Platform Industry 381

Data Aggregators/Distributors 382
Analytics-Focused Software Developers 382
Reporting/Analytics 382
Predictive Analytics 382
Prescriptive Analytics 383
Application Developers or System Integrators: Industry Specific or General 383
Analytics User Organizations 385
Analytics Industry Analysts and Influencers 385
Academic Providers and Certification Agencies 387
*Key Terms 389*
*Questions for Discussion 389*
*Exercises 389*
*End-of-Chapter Application Case 390*
*References 391*
*Glossary 393*
*Index 401*

# PREFACE

Analytics has become the technology driver of this decade. Companies such as IBM, Oracle, Microsoft, and others are creating new organizational units focused on analytics that help businesses become more effective and efficient in their operations. Decision makers are using more computerized tools to support their work. Even consumers are using analytics tools directly or indirectly to make decisions on routine activities such as shopping, healthcare, and entertainment. The field of Business intelligence (BI) is evolving rapidly to become more focused on innovative applications of data streams that were not even captured some time back, much less analyzed in any significant way. New applications turn up daily in healthcare, sports, entertainment, supply chain management, utilities, and virtually every industry imaginable.

In addition to traditional decision support applications, this edition expands the reader's understanding of the various types of analytics by providing examples, products, services, and exercises by discussing Web-related issues throughout the text. We highlight Web intelligence/Web analytics, which parallel BI/business analytics (BA) for e-commerce and other Web applications. The book is supported by a Web site (**www.pearsonglobaleditions.com/sharda**), and also by an independent site at **dssbibook.com**. We will also provide links to software tutorials through a special section of the Web site.

The purpose of this book is to introduce the reader to these technologies that are generally called analytics but have been known by other names. This book presents the fundamentals of the techniques and the manner in which these systems are constructed and used. We follow an EEE approach to introducing these topics: **Exposure, Experience, and Explore**. The book primarily provides **exposure** to various analytics techniques and their applications. The idea is that a student will be inspired to learn from how other organizations have employed analytics to make decisions or to gain a competitive edge. We believe that such **exposure** to what is being done with analytics and how it can be achieved is the key component of learning about analytics. In describing the techniques we also introduce specific software tools that can be used for developing such applications. The book is not limited to any one software tool, so the students can **experience** these techniques using any number of available software tools. Specific suggestions are given in each chapter, but the student and the professor are able to use this book with many different software tools. Our book's companion website will include specific software guides, but students can gain **experience** with these techniques in many different ways. Finally, we hope that this **exposure** and **experience** enables and motivates a reader to **explore** the potential of these techniques in their own domain. To facilitate such **exploration** we include exercises that direct them to Teradata University Network and other sites as well that include team-oriented exercises where appropriate. We will also highlight new and innovative applications that we learn about on the book's website.

Most of the specific improvements made in this third edition concentrate on three areas: reorganization, content update, and a sharper focus. Despite the many changes, we have preserved the comprehensiveness and user friendliness that have made the text a market leader. Finally, we present accurate and updated material that is not available in any other text. We next describe the changes in the third edition.

## WHAT'S NEW IN THE THIRD EDITION?

With the goal of improving the text, this edition marks a major reorganization of the text to reflect the focus on analytics. This edition is now organized around three major types of analytics. The new edition has many timely additions, and the dated content has been deleted. The following major specific changes have been made:

- ***New Organization.*** The book recognizes three types of analytics: descriptive, predictive, and prescriptive, a classification promoted by INFORMS. After introducing BI and analytics in Chapter 1, the book begins with an overview of Data Warehousing and data foundations in Chapter 2. Then, Chapter 3 covers descriptive or reporting analytics, specifically, visualization and business performance measurement. Chapter 4 covers predictive analytics. Chapter 5 extends the application of analytics to text, web, and social media. Chapter 6 introduces big data and analytics. The book concludes with Chapter 7, emerging trends and topics in business analytics including location intelligence, mobile computing, cloud-based analytics, and privacy/ethical considerations in analytics. This chapter also includes an overview of the analytics ecosystem to help the user explore all of the different ways one can participate and grow in the analytics environment. The discussion of analytics ecosystem recognizes prescriptive analytics as well.
- ***New chapters.*** The following chapters have been added:

  **Chapter 5, "*Web Analytics, Web Mining & Social Analytics*"** This chapter covers the poppular topics of web analytics and social media analytics. It is an almost entriely new chapter. [95% new material].

  **Chapter 6, "*Big Data & Analytics*"** This chapter introduces the hot topics of big data and analytics. It covers the basics of major components of big data techniques and charcteristics. It is also a new chapter. [99% new material].

  **Chapter 7, "*Business Analytics: Emerging Trends and Future Direction*"** This chapter examines several new phenomena that are already changing or are likely to change analytics. It includes coverage of geospatial in analytics, location-based analytics applications, consumer oriented analytical applications, mobile platforms, and cloud-based analytics. It also updates some coverage from the last edition on ethical and privacy considerations. It concludes with a major discussion of the analytics ecosystem. [90% new material].

- ***Streamlined coverage.*** We have made the book shorter by keeping the most commonly used content. We also mostly eliminated the preformatted online content. Instead, we will use a Web site to provide updated content and links on a regular basis. We also reduced the number of references in each chapter.
- ***Revamped author team.*** Building upon the excellent content that has been prepared by the authors of the previous editions (Turban, Sharda, Delen, and King,), this edition was revised primarily by Ramesh Sharda and Dursun Delen. Both Ramesh and Dursun have worked extensively in DSS and analytics and have industry as well as research experience.
- ***All new figures for PowerPoint.*** Although the figures in the print edition have been retained from previous editions and new figures added for the new content, all the figures have been redrawn in color and are available through the image library for use in PowerPoint presentations.
- ***A live update Web site.*** Adopters of the textbook will have access to a Web site that will include links to news stories, software, tutorials, and even YouTube videos related to topics covered in the book. This site will be accessible at **http://dssbibook.com**.

- ***Revised and updated content.*** Almost all of the chapters have new opening vignettes and closing cases that are based on recent stories and events. In addition, Application Cases throughout the book have been updated to include recent examples of applications of a specific technique/model. New Web site links have been added throughout the book. We also deleted many older product links and references. Finally, most chapters have new exercises, Internet assignments, and discussion questions throughout.

Specific changes made in previous chapters in the third edition are summarized next:

**Chapter 1 – An Overview of Business Intelligence, Analytics and Decision Support** Introduces the three types of analytics as proposed by INFORMS: descriptive, predictive, and prescriptive analytics. A noted earlier, this classification is used is guiding the complete reorganization of the book itself. It includes about 60% new material.

**Chapter 2 – Data Warehousing**

- 30% new material (including the cases)
- New opening case
- Mostly new cases throughout
- NEW - A historic perspective to DW – how did we come here
- Better coverage of multidimensional modeling (star schema and snow flake schema)
- An updated coverage on Future of Data Warehousing

**Chapter 3 – Business Reporting, Visual Analytics and Business Performance Management**

- 60% of the material new – especially in Visual Analytics and Reporting
- Most of the cases are new

**Chapter 4 – Data Mining**

- 25% of the material new
- Most of the cases are new

**Chapters 5–7** are almost entirely new chapters, as noted earlier.

We have retained many of the enhancements made in the last editions and updated the content. These are summarized next:

- ***Links to Teradata University Network (TUN).*** Most chapters include new links to TUN (**teradatauniversitynetwork.com**).
- ***Book title.*** As is already evident, the book's title and focus has changed substantially.
- ***Software support.*** The TUN Web site provides software support at no charge. It also provides links to free data mining and other software. In addition, the site provides exercises in the use of such software.

## THE SUPPLEMENT PACKAGE

A comprehensive and flexible technology-support package is available to enhance the teaching and learning experience. The following instructor and student supplements are available on the book's Web site, **www.pearsonglobaleditions.com/sharda**:

- ***Instructor's Manual.*** The Instructor's Manual includes learning objectives for the entire course and for each chapter, answers to the questions and exercises at the end of each chapter, and teaching suggestions (including instructions for

projects). The Instructor's Manual is available on the secure faculty section of **www.pearsonglobaleditions.com/sharda**.

- ***Test Item File and TestGen Software.*** The Test Item File is a comprehensive collection of true/false, multiple-choice, fill-in-the-blank, and essay questions. The questions are rated by difficulty level, and the answers are referenced by book page number. The Test Item File is available in Microsoft Word and in the computerized form of Prentice Hall TestGen. TestGen is a comprehensive suite of tools for testing and assessment. It allows instructors to easily create and distribute tests for their courses, either by printing and distributing through traditional methods or by online delivery via a local area network (LAN) server. TestGen features wizards that assist in moving through the program, and the software is backed with full technical support. Both the Test Item File and testgen software are available on the secure faculty section of **www.pearsonglobaleditions.com/sharda**.
- ***PowerPoint slides.*** PowerPoint slides are available that illuminate and build on key concepts in the text. Faculty can download the PowerPoint slides from **www.pearsonglobaleditions.com/sharda**.

## ACKNOWLEDGMENTS

Many individuals have provided suggestions and criticisms since the publication of the first edition of this book. Dozens of students participated in class testing of various chapters, software, and problems and assisted in collecting material. It is not possible to name everyone who participated in this project, but our thanks go to all of them. Certain individuals made significant contributions, and they deserve special recognition.

First, we appreciate the efforts of those individuals who provided formal reviews of the first through third editions (school affiliations as of the date of review):

Ann Aksut, Central Piedmont Community College
Bay Arinze, Drexel University
Ranjit Bose, University of New Mexico
Marty Crossland, MidAmerica Nazarene University
Kurt Engemann, Iona College
Badie Farah, Eastern Michigan University
Gary Farrar, Columbia College
Jerry Fjermestad, New Jersey Institute of Technology
Christie M. Fuller, Louisiana Tech University
Martin Grossman, Bridgewater State College
Jahangir Karimi, University of Colorado, Denver
Huei Lee, Eastern Michigan University
Natalie Nazarenko, SUNY Fredonia
Asil Oztekin, University of Massachusetts Lowell
Kala Chand Seal, Loyola Marymount University
Roger Wilson, Fairmont State University
Vincent Yu, Missouri University of Science and Technology
Fan Zhao, Florida Gulf Coast University

We also appreciate the efforts of those individuals who provided formal reviews of this text and our other DSS book—Decision Support and Business Intelligence Systems, 9th Edition, Prentice Hall, 2011.

Secondly, several individuals contributed material to the text or the supporting material. Susan Baxley and Dr. David Schrader of Teradata provides special help in identifying new TUN content for the book and arranging permissions for the same. Peter Horner, Editor of *OR/MS Today* allowed us to summarize new application stories from *OR/MS Today* and Analytics Magazine. We also thanks INFORMS for their permission to highlight content from Interfaces. Assistance provided by Natraj Ponna, Daniel Asamoah, Kartik Dasika, Clara Gregory, and Amy Wallace (all of Oklahoma State University) is gratefully acknowledged for this edition.

For the previous edition, we acknowledge the contributions of Dave King (JDA Software Group, Inc.). Other major contributors to the previous edition include J. Aronson (University of Georgia), who was our coauthor, contributing to the data warehousing chapter; Mike Goul (Arizona State University), whose contributions were included in Chapter 1; and T. P. Liang (National Sun Yet-Sen University, Taiwan), who contributed material on neural networks in the previous editions. Judy Lang collaborated with all of us, provided editing, and guided us during the entire project in the first edition.

Fourth, several vendors cooperated by providing development and/or demonstration software: Acxiom (Little Rock, AR), California Scientific Software (Nevada City, CA), Cary Harwin of Catalyst Development (Yucca Valley, CA), Demandtec (San Carlos, CA), DS Group, Inc. (Greenwich, CT), Gregory Piatetsky-Shapiro of KDNuggets.com, Gary Lynn of NeuroDimension Inc. (Gainesville, Florida), Palisade Software (Newfield, New York), Promised Land Technologies (New Haven, Connecticut), Salford Systems (La Jolla, California), Sense Networks (New York, New York), Gary Miner of Statsoft, Inc. (Tulsa, Oklahoma), Ward Systems Group, Inc. (Frederick, Maryland), Idea Fisher Systems, Inc. (Irving, California), and Wordtech Systems (Orinda, California).

Fifth, special thanks to the Teradata University Network and especially to Susan Baxley, Program Director; Hugh Watson, who started TUN; Michael Goul, Barb Wixom, and Mary Gros for their encouragement to tie this book with TUN and for providing useful material for the book.

Finally, the Prentice Hall team is to be commended: Executive Editor Bob Horan, who orchestrated this project; our editorial project manager Kelly Loftus, who kept us on the timeline; Kitty Jarrett, who copyedited the manuscript; and the production team, Karalyn Holland at Prentice Hall, and the staff at GGS Book Services, who transformed the manuscript into a book.

We would like to thank all these individuals and corporations. Without their help, the creation of this book would not have been possible.

*R.S.*

*D.D.*

*E.T.*

The publishers wish to thank Susmi Routray of IMT Ghaziabad for reviewing the content of the Global Edition.

---

Note that Web site URLs are dynamic. As this book went to press, we verified that all the cited Web sites were active and valid. Web sites to which we refer in the text sometimes change or are discontinued because companies change names, are bought or sold, merge, or fail. Sometimes Web sites are down for maintenance, repair, or redesign. Most organizations have dropped the initial "www" designation for their sites, but some still use it. If you have a problem connecting to a Web site that we mention, please be patient and simply run a Web search to try to identify the new site. Most times, the new site can be found quickly. We apologize in advance for this inconvenience.

# ABOUT THE AUTHORS

**Ramesh Sharda** (M.B.A., Ph.D., University of Wisconsin–Madison) is director of the Ph.D. in Business for Executives Program and Institute for Research in Information Systems (IRIS), ConocoPhillips Chair of Management of Technology, and a Regents Professor of Management Science and Information Systems in the Spears School of Business at Oklahoma State University (OSU). About 200 papers describing his research have been published in major journals, including *Operations Research*, *Management Science*, *Information Systems Research*, *Decision Support Systems*, and *Journal of MIS*. He cofounded the AIS SIG on Decision Support Systems and Knowledge Management (SIGDSS). Dr. Sharda serves on several editorial boards, including those of *INFORMS Journal on Computing*, *Decision Support Systems*, and *ACM Transactions on Management Information Systems*. He has authored and edited several textbooks and research books and serves as the co-editor of several book series (Integrated Series in Information Systems, Operations Research/Computer Science Interfaces, and Annals of Information Systems) with Springer. He is also currently serving as the executive director of the Teradata University Network. His current research interests are in decision support systems, business analytics, and technologies for managing information overload.

**Dursun Delen** (Ph.D., Oklahoma State University) is the Spears and Patterson Chairs in Business Analytics, Director of Research for the Center for Health Systems Innovation, and Professor of Management Science and Information Systems in the Spears School of Business at Oklahoma State University (OSU). Prior to his academic career, he worked for a privately owned research and consultancy company, Knowledge Based Systems Inc., in College Station, Texas, as a research scientist for five years, during which he led a number of decision support and other information systems–related research projects funded by federal agencies, such as DoD, NASA, NIST, and DOE. Dr. Delen's research has appeared in major journals including *Decision Support Systems, Communications of the ACM, Computers and Operations Research, Computers in Industry, Journal of Production Operations Management, Artificial Intelligence in Medicine, Expert Systems with Applications*, and *IEEE Wireless Communications*, among others. He recently published four textbooks: *Advanced Data Mining Techniques* with Springer, 2008; *Decision Support and Business Intelligence Systems* with Prentice Hall, 2010; *Business Intelligence: A Managerial Approach* with Prentice Hall, 2010; and *Practical Text Mining* with Elsevier, 2012. He is often invited to national and international conferences for keynote addresses on topics related to data/text mining, business intelligence, decision support systems, and knowledge management. He served as the general co-chair for the 4th International Conference on Network Computing and Advanced Information Management (September 2–4, 2008, in Soul, South Korea) and regularly chairs tracks and mini-tracks at various information systems conferences. He is the associate editor-in-chief for *International Journal of Experimental Algorithms*, associate editor for *International Journal of RF Technologies* and *Journal of Decision Analytics*, and is on the editorial boards of five other technical journals. His research and teaching interests are in data and text mining, business analytics, decision support systems, knowledge management, business intelligence, and enterprise modeling.

**Efraim Turban** (M.B.A., Ph.D., University of California, Berkeley) is a visiting scholar at the Pacific Institute for Information System Management, University of Hawaii. Prior to this, he was on the staff of several universities, including City University of Hong Kong; Lehigh University; Florida International University; California State University, Long

Beach; Eastern Illinois University; and the University of Southern California. Dr. Turban is the author of more than 100 refereed papers published in leading journals, such as *Management Science*, *MIS Quarterly*, and *Decision Support Systems*. He is also the author of 20 books, including *Electronic Commerce: A Managerial Perspective* and *Information Technology for Management*. He is also a consultant to major corporations worldwide. Dr. Turban's current areas of interest are Web-based decision support systems, social commerce, and collaborative decision making.

# Business Intelligence

## A Managerial Perspective on Analytics

This book deals with a collection of computer technologies that support managerial work—essentially, decision making. These technologies have had a profound impact on corporate strategy, performance, and competitiveness. Collectively, these technologies are called *business analytics* and *business intelligence*.

CHAPTER

# 1

# An Overview of Business Intelligence, Analytics, and Decision Support

**LEARNING OBJECTIVES**

- Understand today's turbulent business environment and describe how organizations survive and even excel in such an environment (solving problems and exploiting opportunities)
- Understand the need for computerized support of managerial decision making
- Describe the business intelligence (BI) methodology and concepts
- Understand the various types of analytics

The business environment (climate) is constantly changing, and it is becoming more and more complex. Organizations, private and public, are under pressures that force them to respond quickly to changing conditions and to be innovative in the way they operate. Such activities require organizations to be agile and to make frequent and quick strategic, tactical, and operational decisions, some of which are very complex. Making such decisions may require considerable amounts of relevant data, information, and knowledge. Processing these, in the framework of the needed decisions, must be done quickly, frequently in real time, and usually requires some computerized support.

This book is about using business analytics as computerized support for managerial decision making. It concentrates the theoretical and conceptual foundations of decision support, as well as on the commercial tools and techniques that are available. This introductory chapter provides more details of these topics as well as an overview of the book. This chapter has the following sections:


**1.1** Opening Vignette: Magpie Sensing Employs Analytics to Manage a Vaccine Supply Chain Effectively and Safely 28

**1.2** Changing Business Environments and Computerized Decision Support 30

**1.3** A Framework for Business Intelligence (BI) 32

**1.4** Intelligence Creation, Use, and BI Governance 37

**1.5** Transaction Processing Versus Analytic Processing 39
**1.6** Successful BI Implementation 40
**1.7** Analytics Overview 43
**1.8** Brief Introduction to Big Data Analytics 51
**1.9** Plan of the Book 53
**1.10** Resources, Links, and the Teradata University Network Connection 55

## 1.1 OPENING VIGNETTE: Magpie Sensing Employs Analytics to Manage a Vaccine Supply Chain Effectively and Safely

*Cold chain* in healthcare is defined as the temperature-controlled supply chain involving a system of transporting and storing vaccines and pharmaceutical drugs. It consists of three major components—transport and storage equipment, trained personnel, and efficient management procedures. The majority of the vaccines in the cold chain are typically maintained at a temperature of 35–46 degrees Fahrenheit (2–8 degrees Centigrade). Maintaining cold chain integrity is extremely important for healthcare product manufacturers.

Especially for vaccines, improper storage and handling practices that compromise vaccine viability prove to be a costly, time-consuming affair. Vaccines must be stored properly from manufacture until they are available for use. Any extreme temperatures of heat or cold will reduce the vaccine potency; such vaccines, if administered, might not yield effective results or could cause adverse effects.

Effectively maintaining the temperatures of the storage units throughout the healthcare supply chain in real time (i.e., beginning from the gathering of the resources, manufacturing, distribution, and dispensing of the products) is the most effective solution desired in the cold chain. The location-tagged real-time environmental data about the storage units help in monitoring the cold chain for the spoiled products. The chain of custody can be easily identified to assign product liability.

A study conducted by the Centers for Disease Control and Prevention (CDC) looked at the handling of cold chain vaccines by 45 healthcare providers around United States. The CDC reported that three-quarters of the providers experienced serious cold chain violations.

### A WAY TOWARD A POSSIBLE SOLUTION

Magpie Sensing, a start-up project under Ebers Smith and Douglas Associated LLC., provides a suite of cold chain monitoring and analysis technologies for the healthcare industry. It is a shippable, wireless temperature and humidity monitor that provides real-time, location-aware tracking of the cold chain products during the shipment. Magpie Sensing's solutions rely on rich analytics algorithms that leverage the data gathered from the monitoring devices to improve the efficiency of cold chain processes and predict cold storage problems before they occur. Magpie Sensing applies all three types of analytical techniques—descriptive, predictive, and prescriptive analytics—to turn the raw data returned from the monitoring devices into actionable recommendations and warnings.

The properties of the cold storage system, which include the set point of the storage system's thermostat, the typical range of temperature values in the storage system, and the duty cycle of the system's compressor, are monitored and reported in real time. This information helps trained personnel ensure that the storage unit is properly configured to

store a particular product. All temperature information is displayed on a dashboard that shows a graph of the temperature inside a specific storage unit.

Based on the information derived from the monitoring devices, Magpie's predictive analytic algorithms can determine the set point of the storage unit's thermostat and alert the system's users if the system is incorrectly configured, depending on the various types of products stored. Magpie's system also sends alerts about possible temperature violations based on the storage unit's average temperature and subsequent compressor cycle runs, which may drop the temperature below the freezing point. Magpie's predictive analytics further report possible human errors, such as failing to shut the storage unit doors or having an incomplete seal, by analyzing the temperature trend and alerting the users via Web interface, text message, or audible alert before the temperature bounds are actually violated. A compressor or a power failure also can be detected. The estimated time before the storage unit reaches an unsafe temperature also can be determined, which allows users to look for backup solutions, such as using dry ice to restore power or working on other actions to prevent product spoilage.

In addition to predictive analytics, Magpie Sensing's analytics systems can provide prescriptive recommendations for improving the cold storage processes and business decision making. Prescriptive analytics help users dial in the optimal temperature setting, which helps to achieve the correct balance between freezing and spoilage risk; this, in turn, provides a cushion-time to react to the situation before the products spoil. Its prescriptive analytics also gather useful meta-information on cold storage units, including the times of day that are busiest or periods where the system's doors are opened. This can be used to help formulate additional design training plans and institutional policies to ensure that the system is properly maintained and not overused.

Furthermore, prescriptive analytics can be used to guide equipment purchase decisions by constantly analyzing the performance of current storage units. Based on the storage system's efficiency, decisions on distributing the products across available storage units can be made based on the product's sensitivity.

Using Magpie Sensing's cold chain analytics, additional manufacturing time and expenditure can be eliminated by ensuring product safety throughout the supply chain, and effective products can be administered to the patients. Compliance with state and federal safety regulations can be better achieved through automatic data gathering and reporting about the products involved in the cold chain.

### QUESTIONS FOR THE OPENING VIGNETTE

1. What information is provided by the descriptive analytics employed at Magpie Sensing?
2. What type of support is provided by the predictive analytics employed at Magpie Sensing?
3. How do prescriptive analytics help in business decision making?
4. What are possible ways to report actionable information in real time to the concerned users of the system?
5. What other situations might need real-time monitoring applications?

### WHAT WE CAN LEARN FROM THIS VIGNETTE

This vignette illustrates how data from a business process can be used to generate insights at various levels. First, the graphical analysis of the data (termed *reporting analytics*) allows users to get a good feel for the situation. Then, additional analysis using **data mining** techniques can be used to estimate what future behavior would be

like. This is the domain of predictive analytics. Such analysis can then be taken to create specific recommendations for operators. This is an example of prescriptive analytics. Finally, this opening vignette also suggests that innovative applications of analytics can create new business ventures. Identifying opportunities for applications of analytics and assisting with decision making in specific domains is an emerging entrepreneurial opportunity.

*Sources:* **Magpiesensing.com**, "Magpie Sensing Cold Chain Analytics and Monitoring," **magpiesensing.com/wp-content/uploads/2013/01/ColdChainAnalyticsMagpieSensing-Whitepaper.pdf** (accessed July 2013); Centers for Disease Control and Prevention, Vaccine Storage and Handling, **www.cdc.gov/vaccines/pubs/pinkbook/vac-storage.html#storage** (accessed July 2013); and A. Zaleski, "Magpie Analytics System Tracks Cold-Chain Products to Keep Vaccines, Reagents Fresh," 2012, **technicallybaltimore.com/profiles/startups/magpie-analytics-system-tracks-cold-chain-products-to-keep-vaccines-reagents-fresh** (accessed February 2013).

## 1.2 CHANGING BUSINESS ENVIRONMENTS AND COMPUTERIZED DECISION SUPPORT

The opening vignette illustrates how an organization can employ analytics to develop reports on what is happening, predict what is likely to happen, and then also make decisions to make best use of the situation at hand. These steps require an organization to collect and analyze vast stores of data. Companies are moving aggressively to computerized support of their operations. To understand why companies are embracing computerized support, including business intelligence, we developed a model called the *Business Pressures–Responses–Support Model*, which is shown in Figure 1.1.

### The Business Pressures–Responses–Support Model

The Business Pressures–Responses–Support Model, as its name indicates, has three components: business pressures that result from today's business climate; responses (actions taken) by companies to counter the pressures (or to take advantage of the opportunities available in the environment); and computerized support that facilitates

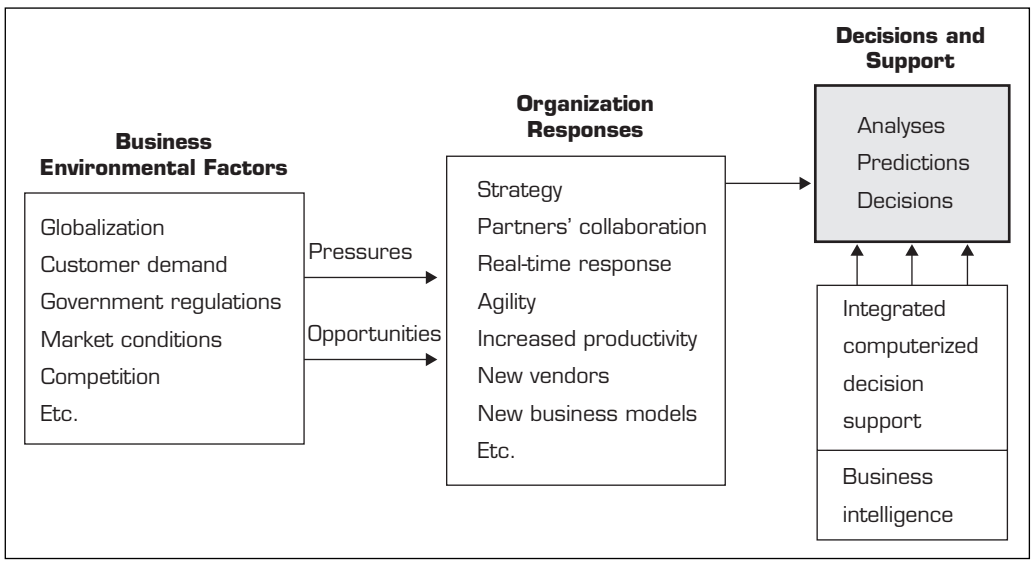


**FIGURE 1.1 The Business Pressures–Responses–Support Model.**