

GLOBAL
EDITION



Business Intelligence, Analytics, and Data Science

A Managerial Perspective

FOURTH EDITION

Ramesh Sharda • Dursun Delen • Efraim Turban



FOURTH EDITION

GLOBAL EDITION

BUSINESS INTELLIGENCE, ANALYTICS, AND DATA SCIENCE:

A Managerial Perspective

Ramesh Sharda

Oklahoma State University

Dursun Delen

Oklahoma State University

Efraim Turban

University of Hawaii

With contributions to previous editions by

J. E. Aronson

The University of Georgia

Ting-Peng Liang

National Sun Yat-sen University

David King

JDA Software Group, Inc.



Pearson

Harlow, England • London • New York • Boston • San Francisco • Toronto • Sydney • Dubai • Singapore • Hong Kong
Tokyo • Seoul • Taipei • New Delhi • Cape Town • Sao Paulo • Mexico City • Madrid • Amsterdam • Munich • Paris • Milan

Business Intelligence: A Managerial Approach, Global Edition

Table of Contents

Cover

Title Page

Copyright Page

Brief Contents

Contents

Preface

Acknowledgments

About the Authors

Chapter 1: An Overview of Business Intelligence, Analytics, and Data
Science

1.1. Opening Vignette: Sports AnalyticsAn Exciting Frontier for Learning and
Understanding Applications of Analytics

1.2. Changing Business Environments and Evolving Needs for Decision
Support and Analytics

1.3. Evolution of Computerized Decision Support to Analytics/Data Science

1.4. A Framework for Business Intelligence

Definitions of BI

A Brief History of BI

The Architecture of BI

The Origins and Drivers of BI

Application Case 1.1: Sabre Helps Its Clients Through Dashboards and Analytics

A Multimedia Exercise in Business Intelligence

Transaction Processing versus Analytic Processing

Appropriate Planning and Alignment with the Business Strategy

Real-Time, On-Demand BI Is Attainable

Table of Contents

Developing or Acquiring BI Systems

Justification and CostBenefit Analysis

Security and Protection of Privacy

Integration of Systems and Applications

1.5. Analytics Overview

Descriptive Analytics

Application Case 1.2: Silvaris Increases Business with Visual Analysis and Real-Time Reporting Capabilities

Application Case 1.3: Siemens Reduces Cost with the Use of Data Visualization

Predictive Analytics

Application Case 1.4: Analyzing Athletic Injuries

Prescriptive Analytics

Analytics Applied to Different Domains

Application Case 1.5: A Specialty Steel Bar Company Uses Analytics to Determine Available-to-Promise Dates

Analytics or Data Science?

1.6. Analytics Examples in Selected Domains

Analytics Applications in HealthcareHumana Examples

Analytics in the Retail Value Chain

1.7. A Brief Introduction to Big Data Analytics

What Is Big Data?

Application Case 1.6: CenterPoint Energy Uses Real-Time Big Data Analytics to Improve Customer Service

1.8. An Overview of the Analytics Ecosystem

Data Generation Infrastructure Providers

Data Management Infrastructure Providers

Data Warehouse Providers

Middleware Providers

Data Service Providers

Analytics-Focused Software Developers

Application Developers: Industry Specific or General

Analytics Industry Analysts and Influencers

Academic Institutions and Certification Agencies

Table of Contents

Regulators and Policy Makers

Analytics User Organizations

1.9. Plan of the Book

1.10. Resources, Links, and the Teradata University Network Connection

Resources and Links

Vendors, Products, and Demos

Periodicals

The Teradata University Network Connection

The Books Web Site

Chapter Highlights

Key Terms

Questions for Discussion

Exercises

References

Chapter 2: Descriptive Analytics I: Nature of Data, Statistical Modeling, and Visualization

2.1. Opening Vignette: SiriusXM Attracts and Engages a New Generation of Radio Consumers with Data-Driven Marketing

2.2. The Nature of Data

2.3. A Simple Taxonomy of Data

Application Case 2.1: Medical Device Company Ensures Product Quality While Saving Money

2.4. The Art and Science of Data Preprocessing

Application Case 2.2: Improving Student Retention with Data-Driven Analytics

2.5. Statistical Modeling for Business Analytics

Descriptive Statistics for Descriptive Analytics

Measures of Centrality Tendency (May Also Be Called Measures of Location or Centrality)

Arithmetic Mean

Median

Mode

Measures of Dispersion (May Also Be Called Measures of Spread Decentrality)

Table of Contents

Range

Variance

Standard Deviation

Mean Absolute Deviation

Quartiles and Interquartile Range

Box-and-Whiskers Plot

The Shape of a Distribution

Application Case 2.3: Town of Cary Uses Analytics to Analyze Data from Sensors,
Assess Demand, and Detect Problems

2.6 Regression Modeling for Inferential Statistics

How Do We Develop the Linear Regression Model?

How Do We Know If the Model Is Good Enough?

What Are the Most Important Assumptions in Linear Regression?

Logistic Regression

Application Case 2.4: Predicting NCAA Bowl Game Outcomes

Time Series Forecasting

2.7. Business Reporting

Application Case 2.5: Flood of Paper Ends at FEMA

2.8. Data Visualization

A Brief History of Data Visualization

Application Case 2.6: Macfarlan Smith Improves Operational Performance Insight with
Tableau Online

2.9. Different Types of Charts and Graphs

Basic Charts and Graphs

Specialized Charts and Graphs

Which Chart or Graph Should You Use?

2.10. The Emergence of Visual Analytics

Visual Analytics

High-Powered Visual Analytics Environments

2.11. Information Dashboards

Application Case 2.7: Dallas Cowboys Score Big with Tableau and Teknion

Dashboard Design

Application Case 2.8: Visual Analytics Helps Energy Supplier Make Better

Table of Contents

Connections

What to Look for in a Dashboard

Best Practices in Dashboard Design

Benchmark Key Performance Indicators with Industry Standards

Wrap the Dashboard Metrics with Contextual Metadata

Validate the Dashboard Design by a Usability Specialist

Prioritize and Rank Alerts/Exceptions Streamed to the Dashboard

Enrich the Dashboard with Business-User Comments

Present Information in Three Different Levels

Pick the Right Visual Construct Using Dashboard Design Principles

Provide for Guided Analytics

Chapter Highlights

Key Terms

Questions for Discussion

Exercises

References

Chapter 3: Descriptive Analytics II: Business Intelligence and Data Warehousing

3.1. Opening Vignette: Targeting Tax Fraud with Business Intelligence and Data Warehousing

3.2. Business Intelligence and Data Warehousing

What Is a Data Warehouse?

A Historical Perspective to Data Warehousing

Characteristics of Data Warehousing

Data Marts

Operational Data Stores

Enterprise Data Warehouses (EDW)

Metadata

Application Case 3.1: A Better Data Plan: Well- Established TELCOs Leverage Data Warehousing and Analytics to Stay on Top in a Competitive Industry

3.3. Data Warehousing Process

3.4. Data Warehousing Architectures

Table of Contents

Alternative Data Warehousing Architectures

Which Architecture Is the Best?

3.5. Data Integration and the Extraction, Transformation, and Load (ETL) Processes

Data Integration

Application Case 3.2: BP Lubricants Achieves BIGS Success

Extraction, Transformation, and Load

3.6. Data Warehouse Development

Application Case 3.3: Use of Teradata Analytics for SAP Solutions Accelerates Big
Data Delivery

Data Warehouse Development Approaches

Additional Data Warehouse Development Considerations

Representation of Data in Data Warehouse

Analysis of Data in Data Warehouse

OLAP versus OLTP

OLAP Operations

3.7. Data Warehousing Implementation Issues

Massive Data Warehouses and Scalability

Application Case 3.4: EDW Helps Connect State Agencies in Michigan

3.8. Data Warehouse Administration, Security Issues, and Future Trends

The Future of Data Warehousing

3.9. Business Performance Management

Closed-Loop BPM Cycle

Application Case 3.5: AARP Transforms Its BI Infrastructure and Achieves a 347% ROI
in Three Years

3.10. Performance Measurement

Key Performance Indicator (KPI)

Performance Measurement System

3.11. Balanced Scorecards

The Four Perspectives

The Meaning of Balance in BSC

3.12. Six Sigma as a Performance Measurement System

The DMAIC Performance Model

Balanced Scorecard versus Six Sigma

Table of Contents

Effective Performance Measurement

Application Case 3.6: Expedia.com's Customer Satisfaction Scorecard

Chapter Highlights

Key Terms

Questions for Discussion

Exercises

References

Chapter 4: Predictive Analytics I: Data Mining Process, Methods, and Algorithms

4.1. Opening Vignette: Miami-Dade Police Department Is Using Predictive Analytics to Foresee and Fight Crime

4.2. Data Mining Concepts and Applications

Application Case 4.1: Visa Is Enhancing the Customer Experience While Reducing Fraud with Predictive Analytics and Data Mining

Definitions, Characteristics, and Benefits

How Data Mining Works

Application Case 4.2: Dell Is Staying Agile and Effective with Analytics in the 21st Century

Data Mining versus Statistics

4.3. Data Mining Applications

Application Case 4.3: Bank Speeds Time to Market with Advanced Analytics

4.4. Data Mining Process

Step 1: Business Understanding

Step 2: Data Understanding

Step 3: Data Preparation

Step 4: Model Building

Application Case 4.4: Data Mining Helps in Cancer Research

Step 5: Testing and Evaluation

Step 6: Deployment

Other Data Mining Standardized Processes and Methodologies

4.5. Data Mining Methods

Classification

Estimating the True Accuracy of Classification Models

Table of Contents

Application Case 4.5: Influence Health Uses Advanced Predictive Analytics to Focus
on the Factors That Really Influence Peoples Healthcare Decisions

Cluster Analysis for Data Mining

Association Rule Mining

4.6. Data Mining Software Tools

Application Case 4.6: Data Mining Goes to Hollywood: Predicting Financial Success
of Movies

4.7. Data Mining Privacy Issues, Myths, and Blunders

Application Case 4.7: Predicting Customer Buying PatternsThe Target Story

Data Mining Myths and Blunders

Chapter Highlights

Key Terms

Questions for Discussion

Exercises

References

Chapter 5: Predictive Analytics II: Text, Web, and Social Media Analytics

5.1. Opening Vignette: Machine versus Men on Jeopardy!: The Story of Watson

5.2. Text Analytics and Text Mining Overview

Application Case 5.1: Insurance Group Strengthens Risk Management with Text Mining
Solution

5.3. Natural Language Processing (NLP)

Application Case 5.2: AMC Networks Is Using Analytics to Capture New Viewers,
Predict Ratings, and Add Value for Advertisers in a Multichannel World

5.4. Text Mining Applications

Marketing Applications

Security Applications

Application Case 5.3: Mining for Lies

Biomedical Applications

Academic Applications

Application Case 5.4: Bringing the Customer into the Quality Equation: Lenovo Uses Analytics to

Table of Contents

Rethink Its Redesign

5.5. Text Mining Process

Task 1: Establish the Corpus

Task 2: Create the TermDocument Matrix

Task 3: Extract the Knowledge

Application Case 5.5: Research Literature Survey with Text Mining

5.6. Sentiment Analysis

Application Case 5.6: Creating a Unique Digital Experience to Capture the Moments
That Matter at Wimbledon

Sentiment Analysis Applications

Sentiment Analysis Process

Methods for Polarity Identification

Using a Lexicon

Using a Collection of Training Documents

Identifying Semantic Orientation of Sentences and Phrases

Identifying Semantic Orientation of Documents

5.7. Web Mining Overview

Web Content and Web Structure Mining

5.8. Search Engines

Anatomy of a Search Engine

1. Development Cycle

2. Response Cycle

Search Engine Optimization

Methods for Search Engine Optimization

Application Case 5.7: Understanding Why Customers Abandon Shopping Carts Results in
a \$10 Million Sales Increase

5.9. Web Usage Mining (Web Analytics)

Web Analytics Technologies

Web Analytics Metrics

Web Site Usability

Traffic Sources

Visitor Profiles

Conversion Statistics

Table of Contents

5.10. Social Analytics

Social Network Analysis

Social Network Analysis Metrics

Application Case 5.8: Titos Vodka Establishes Brand Loyalty with an Authentic Social Strategy

Connections

Distributions

Segmentation

Social Media Analytics

How Do People Use Social Media?

Measuring the Social Media Impact

Best Practices in Social Media Analytics

Chapter Highlights

Key Terms

Questions for Discussion

Exercises

References

Chapter 6: Prescriptive Analytics: Optimization and Simulation

6.1. Opening Vignette: School District of Philadelphia Uses Prescriptive Analytics to Find Optimal Solution for Awarding Bus Route Contracts

6.2. Model-Based Decision Making

Prescriptive Analytics Model Examples

Application Case 6.1: Optimal Transport for ExxonMobil Downstream through a DSS

Identification of the Problem and Environmental Analysis

Model Categories

Application Case 6.2: Ingram Micro Uses Business Intelligence Applications to Make Pricing Decisions

6.3. Structure of Mathematical Models for Decision Support

The Components of Decision Support Mathematical Models

The Structure of Mathematical Models

6.4. Certainty, Uncertainty, and Risk

Decision Making under Certainty

Decision Making under Uncertainty

Table of Contents

Decision Making under Risk (Risk Analysis)

6.5. Decision Modeling with Spreadsheets

Application Case 6.3: Primary Schools in Slovenia Use Interactive and Automated Scheduling Systems to Produce Quality Timetables

Application Case 6.4: Spreadsheet Helps Optimize Production Planning in Chilean Swine Companies

Application Case 6.5: Metro Meals on Wheels Treasure Valley Uses Excel to Find Optimal Delivery Routes

6.6 Mathematical Programming Optimization

Application Case 6.6: Mixed-Integer Programming Model Helps the University of Tennessee Medical Center with Scheduling Physicians

Linear Programming Model

Modeling in LP: An Example

Implementation

6.7. Multiple Goals, Sensitivity Analysis, What-If Analysis, and Goal Seeking

Multiple Goals

Sensitivity Analysis

What-If Analysis

Goal Seeking

6.8. Decision Analysis with Decision Tables and Decision Trees

Decision Tables

Decision Trees

6.9. Introduction to Simulation

Major Characteristics of Simulation

Application Case 6.7: Syngenta Uses Monte Carlo Simulation Models to Increase Soybean Crop Production

Advantages of Simulation

Disadvantages of Simulation

The Methodology of Simulation

Simulation Types

Monte Carlo Simulation

Discrete Event Simulation

Application Case 6.8: Cosan Improves Its Renewable Energy Supply Chain Using

Table of Contents

Simulation

6.10. Visual Interactive Simulation

Conventional Simulation Inadequacies

Visual Interactive Simulation

Visual Interactive Models and DSS

Simulation Software

Application Case 6.9: Improving Job-Shop Scheduling Decisions through RFID: A

Simulation-Based Assessment

Chapter Highlights

Key Terms

Questions for Discussion

Exercises

References

Chapter 7: Big Data Concepts and Tools

7.1. Opening Vignette: Analyzing Customer Churn in a Telecom Company Using Big Data Methods

7.2. Definition of Big Data

The Vs That Define Big Data

Application Case 7.1: Alternative Data for Market Analysis or Forecasts

7.3. Fundamentals of Big Data Analytics

Business Problems Addressed by Big Data Analytics

Application Case 7.2: Top Five Investment Bank Achieves Single Source of the Truth

7.4. Big Data Technologies

MapReduce

Why Use MapReduce?

Hadoop

How Does Hadoop Work?

Hadoop Technical Components

Hadoop: The Pros and Cons

NoSQL

Application Case 7.3: eBays Big Data Solution

Application Case 7.4: Understanding Quality and Reliability of Healthcare Support Information on

Table of Contents

Twitter

7.5. Big Data and Data Warehousing

Use Cases for Hadoop

Use Cases for Data Warehousing

The Gray Areas (Any One of the Two Would Do the Job)

Coexistence of Hadoop and Data Warehouse

7.6. Big Data Vendors and Platforms

IBM InfoSphere BigInsights

Application Case 7.5: Using Social Media for Nowcasting the Flu Activity

Teradata Aster

Application Case 7.6: Analyzing Disease Patterns from an Electronic Medical Records

Data Warehouse

7.7. Big Data and Stream Analytics

Stream Analytics versus Perpetual Analytics

Critical Event Processing

Data Stream Mining

7.8. Applications of Stream Analytics

e-Commerce

Telecommunications

Application Case 7.7: Salesforce Is Using Streaming Data to Enhance Customer Value

Law Enforcement and Cybersecurity

Power Industry

Financial Services

Health Sciences

Government

Chapter Highlights

Key Terms

Questions for Discussion

Exercises

References

Chapter 8: Future Trends, Privacy and Managerial Considerations in Analytics

Table of Contents

8.1. Opening Vignette: Analysis of Sensor Data Helps Siemens Avoid Train Failures

8.2. Internet of Things

Application Case 8.1: SilverHook Powerboats Uses Real-Time Data Analysis to Inform Racers and Fans

Application Case 8.2: Rockwell Automation Monitors Expensive Oil and Gas Exploration Assets

IoT Technology Infrastructure

RFID Sensors

Fog Computing

IoT Platforms

Application Case 8.3: Pitney Bowes Collaborates with General Electric IoT Platform to Optimize Production

IoT Start-Up Ecosystem

Managerial Considerations in the Internet of Things

8.3. Cloud Computing and Business Analytics

Data as a Service (DaaS)

Software as a Service (SaaS)

Platform as a Service (PaaS)

Infrastructure as a Service (IaaS)

Essential Technologies for Cloud Computing

Cloud Deployment Models

Major Cloud Platform Providers in Analytics

Analytics as a Service (AaaS)

Representative Analytics as a Service Offerings

Illustrative Analytics Applications Employing the Cloud Infrastructure

MD Anderson Cancer Center Utilizes Cognitive Computing Capabilities of IBM Watson to Give Better Treatment to Cancer Patients

Public School Education in Tacoma, Washington, Uses Microsoft Azure Machine Learning to Predict School Dropouts

Dartmouth-Hitchcock Medical Center Provides Personalized Proactive Healthcare Using Microsoft Cortana Analytics Suite

Mankind Pharma Uses IBM Cloud Infrastructure to Reduce Application Implementation Time by 98%

Gulf Air Uses Big Data to Get Deeper Customer Insight

Table of Contents

Chime Enhances Customer Experience Using Snowflake

8.4. Location-Based Analytics for Organizations

Geospatial Analytics

Application Case 8.4: Indian Police Departments Use Geospatial Analytics to Fight Crime

Application Case 8.5: Starbucks Exploits GIS and Analytics to Grow Worldwide

Real-Time Location Intelligence

Application Case 8.6: Quiznos Targets Customers for Its Sandwiches

Analytics Applications for Consumers

8.5. Issues of Legality, Privacy, and Ethics

Legal Issues

Privacy

Collecting Information about Individuals

Mobile User Privacy

Homeland Security and Individual Privacy

Recent Technology Issues in Privacy and Analytics

Who Owns Our Private Data?

Ethics in Decision Making and Support

8.6. Impacts of Analytics in Organizations: An Overview

New Organizational Units

Redesign of an Organization through the Use of Analytics

Analytics Impact on Managers Activities, Performance, and Job Satisfaction

Industrial Restructuring

Automations Impact on Jobs

Unintended Effects of Analytics

8.7. Data Scientist as a Profession

Where Do Data Scientists Come From?

Chapter Highlights

Key Terms

Questions for Discussion

Exercises

References

Table of Contents

Glossary

Index

Back Cover