DATA ANALYTICS *for* ACCOUNTING

McGraw Hill

RICHARDSON   TEETER   TERRELL

Data Analytics for Accounting

SECOND EDITION

Vernon J. Richardson
University of Arkansas,
Baruch College

Ryan A. Teeter
University of Pittsburgh

Katie L. Terrell
University of Arkansas

DATA ANALYTICS FOR ACCOUNTING

All credits appearing on page or at the end of the book are considered to be an extension of the copyright page.

The Internet addresses listed in the text were accurate at the time of publication. The inclusion of a website does not indicate an endorsement by the authors or McGraw-Hill Education, and McGraw-Hill Education does not guarantee the accuracy of the information presented at these sites.

mheducation.com/highered

Dedications

My wonderful daughter, Melissa, for your constant love, encouragement and support.

—Vern Richardson

My wife, Erin, and children, Sylvia and Theodore.

—Ryan Teeter

To my co-author, friend, and colleague, Vernon Richardson. Thank you for inviting me to be on this textbook journey. And thank you for your guidance and patience—I'm thrilled to be a part of your team!

—Katie Terrell

Preface

Data Analytics is changing the business world—data simply surrounds us! So much data is available to businesses about each of us—how we shop, what we read, what we buy, what music we listen to, where we travel, whom we trust, where we invest our time and money, etc. Accountants can create value by addressing fundamental business and accounting questions using data analytics.

All accountants must develop data analytic skills to address the needs of the profession in the future. *Data Analytics for Accounting, 2e* recognizes that accountants don't need to become data scientists —they may never need to build a data repository or do the real hard-core Data Analytics or learn how to program a computer to do machine learning. However, there are seven skills that analytic-minded accountants must have to be prepared for a data-filled world, including:

An analytics mindset—recognize when and how Data Analytics can address accounting questions.

Data scrubbing and data preparation—comprehend the process needed to extract (query), clean and prepare the data before analysis.

Data quality—recognize what is meant by data quality, be it completeness, reliability, or validity.

Descriptive data analysis—perform basic analysis to understand the quality of the underlying data and their ability to address the business question.

Data analysis through data manipulation—demonstrate ability to sort, rearrange, merge, and reconfigure data in a manner that allows enhanced analysis.

Problem solving through statistical data analysis—identify and implement an approach that will use statistical data analysis to draw conclusions and make recommendations on a timely basis.

Data visualization and data reporting—report results of analysis in an accessible way to each varied decision maker and his or her specific needs.

Consistent with these skills, it's important to recognize that Data Analytics is a process. The process begins by identifying business questions that can be addressed with data, extracting and testing the data, refining our testing, and finally, communicating those findings to management. *Data Analytics for Accounting,* 2e describes this process by relying on an established data analytics model called the IMPACT cycle[1]
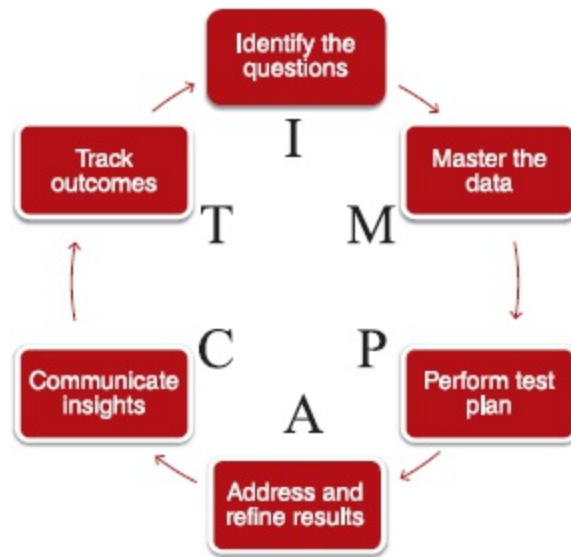
Identify the question.

Master the data.

Perform test plan.

Address and refine results.

Communicate insights.

Track outcomes

Adapted from *Win with Advanced Business Analytics: Creating Business Value from Your Data*, by Jean Paul Isson and Jesse S. Harriott.

The IMPACT cycle is described in the first four chapters and then the process is illustrated in audit, managerial accounting, financial accounting and tax in Chapters 5-9, adding an all-new tax chapter to *Data Analytics for Accounting,* 2e. In response to instructor feedback, *Data Analytics for Accounting,* 2e now also includes two new project chapters, giving students a chance to practice the full IMPACT model with multiple labs that build on each other.

Data Analytics for Accounting, 2e emphasizes hands-on practice. Students are provided with hands-on instruction (e.g., click-by-click instructions, screenshots, etc.) on datasets within the chapter; within the end-of-chapter materials; and in the labs at the end of each chapter. Throughout the text, students identify questions, extract and download data, perform testing, and then communicate the results of that testing.

The use of real-world data is highlighted by using data from **LendingClub**, **College Scorecard**, **Dillard's**, the **State of Oklahoma**, as well as other data from our labs. In particular, we emphasize the rich data from **Dillard's** sales transactions that we use in more than fifteen of the labs throughout the text (including Chapter 11).

Data Analytics for Accounting, 2e also emphasizes the various data analysis tools students will use throughout the rest of their career —Microsoft Excel, Microsoft Access (including SQL), Tableau (free student license), IDEA (free student license), and Weka (free student license). Using multiple tools allows students to learn which tool is best suited for the necessary data analysis, data visualization, and communication of the insights gained—for example, which tool is easiest for internal controls testing, which is best for analysis or querying (using SQL) big datasets, which is best for data visualizations, and so on.

[1]Jean Paul Isson and Jesse S. Harriott, *Win with Advanced Business Analytics: Creating Business Value from Your Data* (Hoboken, NJ: Wiley, 2013).

About the Authors

Vernon J. Richardson

Vernon J. Richardson is a Distinguished Professor of Accounting and the G. William Glezen Chair in the Sam M. Walton College of Business at the University of Arkansas and a Visiting Professor at Baruch College. He received his BS, Master of Accountancy, and MBA from Brigham Young University and a PhD in accounting from the University of Illinois at Urbana–Champaign. He has taught students at the University of Arkansas, Baruch College, University of Illinois, Brigham Young University, Aarhus University, and University of Kansas and internationally at the China Europe International Business School (Shanghai), Xi'an Jiaotong Liverpool University, and the University of Technology Sydney.

Dr. Richardson is a member of the American Accounting Association. He has served as president of the American Accounting Association Information Systems section. He previously served as an editor of *The Accounting Review* and is currently an editor at *Accounting Horizons.* He has published articles in *The Accounting Review, Journal of Information Systems, Journal of Accounting and Economics, Contemporary Accounting Research, MIS Quarterly, International Journal of Accounting Information Systems, Journal of Management Information Systems, Journal of Operations*

Management, and *Journal of Marketing.* Dr. Richardson is also the author of McGraw-Hill's *Accounting Information Systems* textbook.
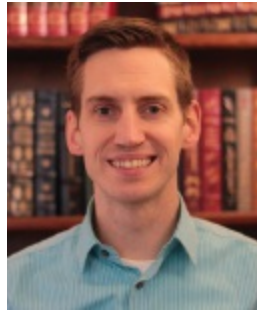
Ryan A. Teeter

Ryan A. Teeter is a Clinical Assistant Professor of Accounting in the Katz Graduate School of Business at the University of Pittsburgh. He teaches accounting information systems, auditing, and accounting data analytics. Prior to receiving his PhD in accounting information systems from Rutgers University, he worked at Google in Mountain View, California. He has since worked with internal audit organizations at Siemens, Procter & Gamble, Alcoa/Arconic, and FedEx, helping to develop robotic process automation programs and data analytic solutions.

Dr. Teeter is a member of the American Accounting Association and has published articles in the *Journal of Strategic Technologies in Accounting* and *Issues in Accounting Education.* He has received grant funding for data analytics research from PwC.

Katie L. Terrell

Katie L. Terrell is an instructor in the Sam M. Walton College of Business at the University of Arkansas. She received her BA degrees in English literature and in the Spanish language from the University of Central Arkansas and her MBA from the University of Arkansas. She expects a doctoral degree by 2020. She has taught students at the University of Arkansas; Soochow University (Suzhou, China); the University College Dublin (Ireland); and Duoc UC, a branch of the Catholic University of Chile (Vina del Mar, Chile).

She is a member of the American Accounting Association and has published a *Statement on Management Accounting* for the Institute of Management Accountants on managing organizational change in operational change initiatives. Terrell was named the 2019 Business Professional of the Year (Education) by the national Beta Alpha Psi organization. She has recently been recognized for her innovative teaching by being the recipient of the Mark Chain/FSA Teaching Award for innovative graduate-level accounting teaching practices in 2016. She has worked with Tyson Foods, where she held various information system roles, focusing on business analysis, project management for ERP implementations and upgrades, and organizational change management.

Acknowledgments

Our sincere thanks to all who helped us on this project.

Our biggest thanks to the awesome team at McGraw-Hill Education, including Steve Schuetz, Tim Vertovec, Allie Kukla, Fran Simon, Kevin Moran, and Sarah Wood.

Our thanks also to each of the following:

The Walton College Enterprise Team (Paul Cronan, Ron Freeze, Michael Gibbs, Michael Martz, Tanya Russell) for their work helping us get access to the Dillard's data.

Shane Lunceford from LendingClub for helping gain access to LendingClub data.

Marcia Watson, University of North Carolina–Charlotte; Ryan Baxter, Boise State University; Antoinette Smith, Florida International University; and Lorrie Metzger, University at Buffalo for their accuracy check and review of the manuscript.

In addition, the following reviewers and classroom testers who provided ideas and insights for this edition. We appreciate their contributions.

Amelia Annette Baldwin

University of South Alabama

Dereck Barr-Pulliam

University of Wisconsin–Madison

Heather Carrasco

Texas Tech University

Elizabeth Felski

State University of New York at Geneseo

Chris C. Hsu

York College, City University of New York

Venkataraman Iyer

University of North Carolina at Greensboro

Andrea S. Kelton

Middle Tennessee State University

Brandon Lock

Baruch College, CUNY

Sharon M. Lightner

National University

Margarita Maria Lenk

Colorado State University

Partha Mohapatra

California State University, Sacramento

Uday Murthy

University of South Florida

Kalana Malimage

University of Wisconsin–Whitewater

Bonnie Morris

Duquesne University

Kathy Nesper

University at Buffalo

Ali Saeedi

University of Minnesota Crookston

Karen Schuele

John Carroll University

Drew Sellers

Kent State University

Joe Shangguan

Robert Morris University

Vincent J. Shea

St. John's University

Marcia Watson

University of North Carolina at Charlotte

Liu Yang

Southeast Missouri State University

Zhongxia Ye

University of Texas, San Antonio

Qiongyao (Yao) Zhang

Robert Morris University

Vernon Richardson
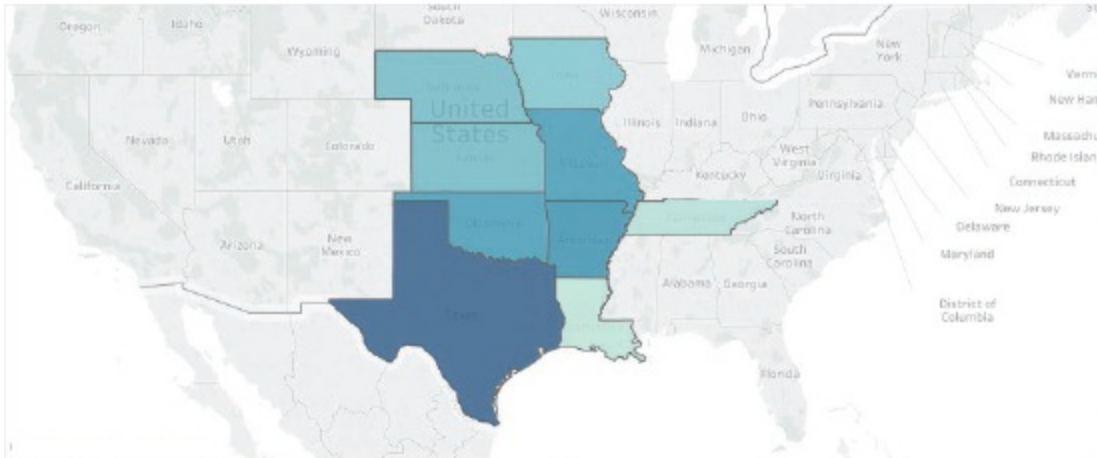
Ryan Teeter

Katie Terrell

Key Features

Emphasis on Skills: Working through the IMPACT cycle framework, students will learn problem assessment, data preparation, data analysis, data visualization, control contesting, and more.

Emphasis on Hands-On Practice: Students will be provided hands-on learning (click-by-click instructions with screenshots) on datasets within each chapter, within the end-of-chapter materials, and in the labs and comprehensive cases.

Emphasis on Datasets: To illustrate data analysis techniques and skills, multiple practice datasets (audit, financial, and managerial data) will be used in every chapter. Students gain real-world experience working with data from **LendingClub, Dillard's, College Scorecard,** the **State of Oklahoma,** as well as financial statement data (via XBRL) from *Fortune* 100 companies.

Emphasis on Tools: Students will learn how to conduct data analysis using Excel Access (including SQL), Tableau (free student license), IDEA (free student license), and Weka (free student license). Students will compare and contrast the different tools to determine which are best suited for basic data analysis and data visualization, which are easiest for internal controls testing, which are best for SQL queries, and so on.

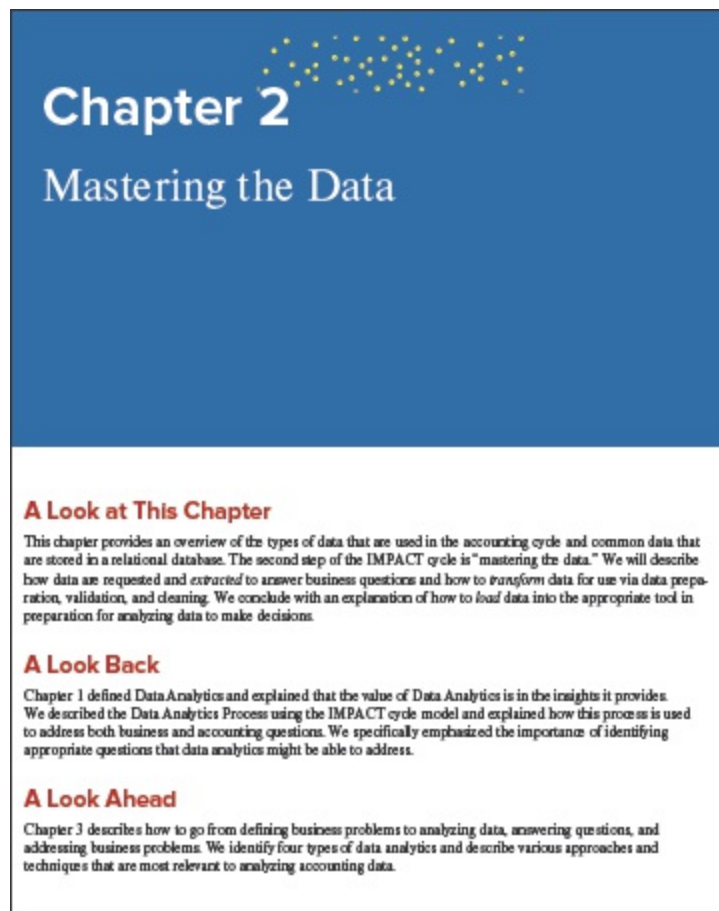## Total Products Sold by State

Total Products Sold

Product Description

Imperial Stout	258.0
IPA	235.0
Stout	190.0
Wheat	119.0
Pale Ale	116.0
Imperial IPA	61.0

Total Products Sold by Year

Sales Order Date / Product Description

2019
Imperial Stout	92.0
IPA	64.0
Stout	72.0
Wheat	33.0
Pale Ale	35.0
Imperial IPA	19.0

2020
Imperial Stout	166.0
IPA	171.0
Stout	118.0
Wheat	86.0
Pale Ale	81.0
Imperial IPA	42.0

Main Text Features

Chapter Maps

These maps provide a guide of what we're going to cover in the chapter as well as a guide of what we've just learned and what's coming next.

Chapter 2

Mastering the Data

A Look at This Chapter

This chapter provides an overview of the types of data that are used in the accounting cycle and common data that are stored in a relational database. The second step of the IMPACT cycle is "mastering the data." We will describe how data are requested and *extracted* to answer business questions and how to *transform* data for use via data preparation, validation, and cleaning. We conclude with an explanation of how to *load* data into the appropriate tool in preparation for analyzing data to make decisions.

A Look Back

Chapter 1 defined Data Analytics and explained that the value of Data Analytics is in the insights it provides. We described the Data Analytics Process using the IMPACT cycle model and explained how this process is used to address both business and accounting questions. We specifically emphasized the importance of identifying appropriate questions that data analytics might be able to address.

A Look Ahead

Chapter 3 describes how to go from defining business problems to analyzing data, answering questions, and addressing business problems. We identify four types of data analytics and describe various approaches and techniques that are most relevant to analyzing accounting data.

Chapter-Opening Vignettes

Because companies are facing the new and exciting opportunities with

their use of Data Analytics to help with accounting and business decisions, we detail what they're doing and why in our chapter-opening vignettes.

Learning Objectives

We feature learning objectives at the beginning of each chapter. Having these learning objectives provides students with an overview of the concepts to be taught in the chapter and the labs.

Progress Checks

Periodic progress check questions are posed to the students throughout each chapter. These checks provoke the student to stop and consider the concepts presented.

End-of-Chapter Materials

Answers to Progress Checks

Allow students to evaluate if they are on track with their understanding of the materials presented in the chapter.

> **✓ ANSWERS TO PROGRESS CHECKS**
>
> 1. The unique identifier of the Supplier table is [Supplier ID], and the unique I Purchase Order table is [PO No.]. The Purchase Order table contains the t
> 2. The foreign key attributes in the Purchase Order table that do not relate in the view are EmployeeID and CashDisbursementID. These attributes to the Employee table (so that we can tell which employee was respor Purchase Order) and the Cash Disbursement table (so that we can tell Orders have been paid for yet, and if so, on which check). The Employee a complete listing of each Employee, as well containing the details about (for example, phone number, address, etc.). The Cash Disbursement tak listing of the payments the company has made.

Multiple Choice Questions

Quickly assess student's knowledge of chapter content.

> **Multiple Choice Questions**
>
> 1. Mastering the data can also be described via the ETL process. Th stands for:
>   a. extract, total, and load data.
>   b. enter, transform, and load data.
>   c. extract, transform, and load data.
>   d. enter, total, and load data.
> 2. Which of the following describes part of the goal of the ETL process:
>   a. identify which approach to data analytics should be used.
>   b. load the data into a relational database for storage.
>   c. communicate the results and insights found through the analysis.

Discussion Questions

Provide questions for group discussion.

Discussion Questions

1. The advantages of a relational database include limiting the amount of ┄
   that are stored in a database. Why is this an important advantage? Wha┄
   when redundant data are stored?
2. The advantages of a relational database include integrating business p┄
   is it preferable to integrate business processes in one information syst┄
   store different business process data in separate, isolated databases?
3. Even though it is preferable to store data in a relational database, stori┄
   separate tables can make data analysis cumbersome. Describe three rea┄
   the trouble to store data in a relational database.
4. Among the advantages of using a relational database is enforcing busine┄
   on your understanding of how the structure of a relational database help┄
   redundancy and other advantages, how does the primary key/foreign k┄

Problems

Challenge the student's ability to see relationships in the learning objectives by employing higher-level thinking and analytical skills.

Problems

The following problems correspond to the **College Scorecard** data. You sho┄
answer each question by just looking at the data dictionary included in App┄
you would like to use the raw data, feel free to do so (CollegeScorecard_Raw┄

1. Which attributes from the College Scorecard data would you need to c┄
   attendance across types of institutions (public, private nonprofit, or priva┄
2. Which attributes from the College Scorecard data would you need to┄
   scores across types of institutions (public, private nonprofit, or private fo┄
3. Which attributes from the College Scorecard data would you need to co┄
   diversity across types of institutions (public, private nonprofit, or private ┄
4. If you were conducting a data analysis in order to compare the percenta┄
   who receive federal loans at universities above and below the media┄
   data across all institutions┄ analysis would requi┄ al steps┄

Labs

Give students hands-on experience working with different types of data and the tools used to analyze them. Students will conduct data analysis using Excel, Access (including SQL), Tableau, IDEA, XBRL,

and Weka.

Lab 2-1 Create a Request for Data Extra...

One of the biggest challenges you face with data a... may have the best questions in the world, but if there... hypothesis, you will have difficulty providing value... which the IT workers may be reluctant to share data... data, the wrong data, or completely ignore your reque... look for creative ways to find insight with an incompl...

Company summary

Sláinte is a fictional brewery that has recently gone th... ferent products. The brewery has only recently expand... state to nine states, and now its business has begun sta...

Comprehensive Cases

Use a real-life Big Data set based on **Dillard's** actual company data. This dataset allows students to build their skills and test their conclusions across concepts covered in each chapter. The Comprehensive Cases can be followed continuously from the first chapter or picked up at any later point in the book; enough information is provided to ensure students can get right to work.

Lab 2-8 Comprehensive Case: Dillard's S...
Connecting Excel to a SQL Data...

Company summary

Dillard's is a department store with approximately 330... in Little Rock, Arkansas. You can learn more about **Di**... (Ticker symbol = DDS) and the Wikipedia site for DI... Dillard II is an accounting grad of the University of...

Data Analytics for Accounting, 2e Content Updates

General Updates for the 2nd Edition

Added additional End-of-Chapter Multiple Choice Questions and Problems throughout the text.

Significantly revised many End-of-Chapter Problems for availability and auto-grading within Connect.

Revised and added many new Discussion Questions in most chapters.

Chapter by Chapter Updates

Specific chapter changes for *Data Analytics for Accounting,* 2nd Edition, are as follows:

Chapter 1

Updated the opening vignette and statistics on Alibaba sales and use of e-commerce.

Updated the statistics and screenshots for Lending Club Analysis.

Revised Connect questions for problems and labs.

Chapter 2

Improved and clarified the discussion of relational databases, including updated figures.

Expanded the discussion of different RDBMS (Access, SQLite, and SQL Server).

Improved discussion of Excel and SQL. The brief introduction to how to use SQL now has its own place in a dedicated appendix at the end of the text, and it has been vastly expanded to teach beginners how to write queries.

Expanded the discussion on data quality.

Added a brief discussion of ETL v. ELT.

Improved labs for clarity and a better learning experience, particularly Labs 2-1, 2-2, and 2-4.

Chapter 3

Reorganized chapter structure to follow the descriptive, diagnostic, predictive, and prescriptive approaches to Data Analytics.

New exhibits and examples to illustrate analytics approaches.

Removed previous edition flowchart for model selection.

Additional explanation and examples of each of the methods and approaches.

Improved labs for clarity.

Chapter 4

Updated the opening vignette.

Improved the discussion on the differences between qualitative and quantitative data and the discussion of the normal distribution.

Improved and clarified how to select a visualization based on the four chart types (qualitative vs. quantitative and declarative vs. exploratory).

Updated the discussion on the Gartner Quadrant to take into account Gartner's January 2019 analysis of BI tools (focusing on

Excel and Tableau).

Extended the discussion on written and spoken communication.

Added a lab to work with visualizing data and creating dashboards in Power BI to interactively compare the tool with Tableau.

Chapter 5

Expanded discussion on the modern data environment.

Included additional examples of the Audit Data Standard.

Improved and clarified content to match the focus on descriptive, diagnostic, predictive, and prescriptive analytics.

New labs (5-1 and 5-2) that have students transform data using a common data model.

Improved existing labs.

Chapter 6

Clarified chapter content to match the focus on descriptive, diagnostic, predictive, and prescriptive analytics.

Improved labs.

Chapter 7

Clarified chapter content and provided additional new exhibits and examples, such as variance analysis.

Improved labs.

Chapter 8

Reorganized chapter content to focus on financial statement analysis using descriptive, diagnostic, predictive, and prescriptive approaches.

Added new content on common size and ratio analysis.

Improved discussion of XBRL data.

Improved XBRL dataset (in Lab 8-4), accessible via Microsoft Access and included options to do analysis in Excel.

Chapter 9

All-new chapter on tax analytics, including examples of tax data, tax analysis, tax planning, and tax visualizations.

Chapter 10

All-new basic project chapter that explores the order-to-cash and procure-to-pay cycles from different user perspectives.

Chapter 11

All-new advanced project chapter, estimating sales returns at **Dillard's** with three question sets highlighting descriptive and exploratory analysis, hypothesis testing, and predictive analytics.

Appendixes

Several all-new appendixes have been added to ease the lab experience and introduce tools used or mentioned throughout the text:

Appendix A: Basic Statistics Tutorial.

Appendix B: Accessing the Excel Data Analysis Toolpak.

Appendix C: Excel (Formatting, Sorting, Filtering, and PivotTables).

Appendix D: SQL Part 1. This tutorial introduces the SQL language for extracting data and explains the following SQL syntax: SELECT, FROM, INNER JOIN, ON, WHERE, GROUP BY, HAVING, ORDER BY.

Appendix E: SQLite. We have added SQLite files as an option for each lab that uses Microsoft Access. This lab explains how to download SQLite and how to use the tool.

Appendix F: Power Query. This appendix contains a short tutorial on transforming data using Power Query. How to access data files on the University of Arkansas' remote desktop is also discussed.

Appendix G: Tableau.

Appendix H: SQL Part 2: On the heels of learning Tableau, students learn about more complex joins—LEFT and RIGHT.

Appendix I: Power BI.

Appendix J: Dillard's ER Diagram.

Appendix K: Data Dictionaries.