

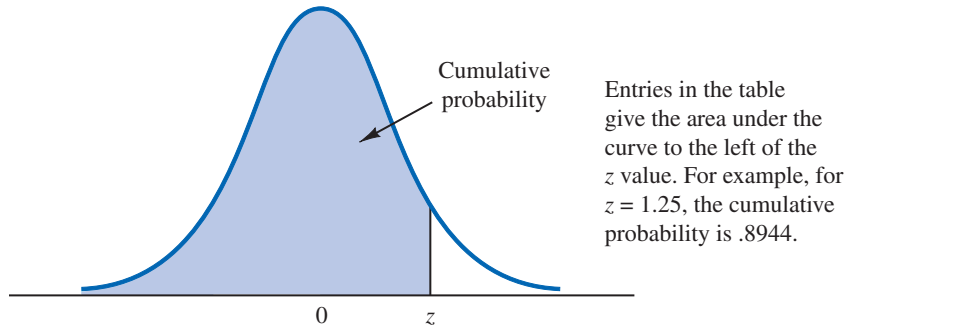
ANDERSON • SWEENEY • WILLIAMS

# Modern Business Statistics

With Microsoft® Office Excel® • 5e

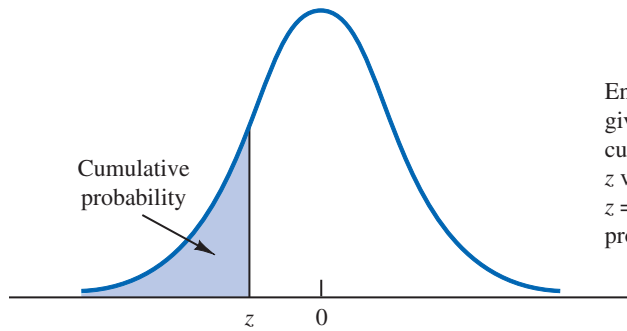


## CUMULATIVE PROBABILITIES FOR THE STANDARD NORMAL DISTRIBUTION



$z$	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

## CUMULATIVE PROBABILITIES FOR THE STANDARD NORMAL DISTRIBUTION



Entries in this table give the area under the curve to the left of the  $z$  value. For example, for  $z = -.85$ , the cumulative probability is .1977.

$z$	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990



It's time to give **YOU** something fresh.

W E L C O M E T O...



# MindTap™

Designed to support the way **YOU** learn and *where* you learn, MindTap is well beyond an eBook, assignments or digital supplement, a resource center website, a course delivery platform, or a Learning Management System.

**MindTap is the first in a new category—  
The Personal Learning Experience.**



See for yourself.  
View a MindTap demo at  
[www.cengage.com/mindtap](http://www.cengage.com/mindtap).



**“This thing is awesome. It’s very, very easy to use. Your book is right there and has easy tabs to go through. There are lots of extra functions on the side such as highlighting ... note taking ... it even reads to you!”**

**Chris Baggett**  
Student, Arkansas Tech University



**7 OUT OF 8**

MindTap students would recommend MindTap to their friends.

13M-AG0596



**MODERN  
BUSINESS  
STATISTICS  
WITH MICROSOFT®  
OFFICE EXCEL®, 5e**



# MODERN BUSINESS STATISTICS

## WITH MICROSOFT® OFFICE EXCEL®, 5e

**David R. Anderson**  
University of Cincinnati

**Dennis J. Sweeney**  
University of Cincinnati

**Thomas A. Williams**  
Rochester Institute of Technology



---

Australia • Brazil • Japan • Korea • Mexico • Singapore • Spain • United Kingdom • United States



This is an electronic version of the print textbook. Due to electronic rights restrictions, some third party content may be suppressed. Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. The publisher reserves the right to remove content from this title at any time if subsequent rights restrictions require it. For valuable information on pricing, previous editions, changes to current editions, and alternate formats, please visit [www.cengage.com/highered](http://www.cengage.com/highered) to search by ISBN#, author, title, or keyword for materials in your areas of interest.

**Modern Business Statistics with  
Microsoft Excel, Fifth Edition**

David R. Anderson, Dennis J. Sweeney,  
Thomas A. Williams

Senior Vice President, Global Product  
Manager, Higher Education:  
Jack W. Calhoun

Vice President, General Manager, Social  
Science & Qualitative Business:  
Erin Joyner

Product Director: Joe Sabatino

Product Manager: Aaron Arnsparger

Content Developer: Maggie Kubale

Product Assistant: Brad Sullender

Content Project Manager:  
Cliff Kallemeyn

Media Developer: Chris Valentine

Manufacturing Planner:  
Ron Montgomery

Marketing Director: Natalie King

Market Manager: Heather Mooney

Production Service: MPS Limited

Sr. Art Director: Stacy Jenkins Shirley

Internal Design: Michael Stratton/  
Chris Miller Design

Cover Design: Craig Ramsdell

Cover Image: ©iofoto/Veer

Sr. Rights Acquisitions Specialist:  
Amber Hosea

©2015, 2011 Cengage Learning

WCN: 02-200-203

ALL RIGHTS RESERVED. No part of this work covered by the copyright herein may be reproduced, transmitted, stored, or used in any form or by any means graphic, electronic, or mechanical, including but not limited to photocopying, recording, scanning, digitizing, taping, web distribution, information networks, or information storage and retrieval systems, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without the prior written permission of the publisher.

For product information and technology assistance, contact us at  
**Cengage Learning Customer & Sales Support, 1-800-354-9706**

For permission to use material from this text or product,  
submit all requests online at **www.cengage.com/permissions**

Further permissions questions can be emailed to  
**permissionrequest@cengage.com**

Microsoft Excel® is a registered trademark of Microsoft Corporation.  
© 2015 Microsoft.

Library of Congress Control Number: 2013938890

ISBN 13: 978-1-285-43330-1

ISBN 10: 1-285-43330-0

**Cengage Learning**

200 First Stamford Place, 4th Floor

Stamford, CT 06902

USA

Cengage Learning is a leading provider of customized learning solutions with office locations around the globe, including Singapore, the United Kingdom, Australia, Mexico, Brazil, and Japan. Locate your local office at: **www.cengage.com/global**

Cengage Learning products are represented in Canada by  
Nelson Education, Ltd.

To learn more about Cengage Learning Solutions, visit  
**www.cengage.com**

Purchase any of our products at your local college store or at our  
preferred online store **www.cengagebrain.com**

Printed in Canada

1 2 3 4 5 6 7 17 16 15 14 13

# Brief Contents

<b>Preface</b>	<b>xxiii</b>
<b>About the Authors</b>	<b>xxx</b>
<b>Chapter 1</b>	<b>Data and Statistics 1</b>
<b>Chapter 2</b>	<b>Descriptive Statistics: Tabular and Graphical Displays 36</b>
<b>Chapter 3</b>	<b>Descriptive Statistics: Numerical Measures 106</b>
<b>Chapter 4</b>	<b>Introduction to Probability 180</b>
<b>Chapter 5</b>	<b>Discrete Probability Distributions 226</b>
<b>Chapter 6</b>	<b>Continuous Probability Distributions 280</b>
<b>Chapter 7</b>	<b>Sampling and Sampling Distributions 312</b>
<b>Chapter 8</b>	<b>Interval Estimation 353</b>
<b>Chapter 9</b>	<b>Hypothesis Tests 393</b>
<b>Chapter 10</b>	<b>Inference About Means and Proportions with Two Populations 440</b>
<b>Chapter 11</b>	<b>Inferences About Population Variances 489</b>
<b>Chapter 12</b>	<b>Tests of Goodness of Fit, Independence, and Multiple Proportions 517</b>
<b>Chapter 13</b>	<b>Experimental Design and Analysis of Variance 551</b>
<b>Chapter 14</b>	<b>Simple Linear Regression 605</b>
<b>Chapter 15</b>	<b>Multiple Regression 692</b>
<b>Chapter 16</b>	<b>Regression Analysis: Model Building 745</b>
<b>Chapter 17</b>	<b>Time Series Analysis and Forecasting 791</b>
<b>Chapter 18</b>	<b>Nonparametric Methods 860</b>
<b>Chapter 19</b>	<b>Statistical Methods for Quality Control 903</b>
<b>Chapter 20</b>	<b>Decision Analysis On Website</b>
<b>Chapter 21</b>	<b>Sample Survey On Website</b>
<b>Appendix A</b>	<b>References and Bibliography 938</b>
<b>Appendix B</b>	<b>Tables 940</b>
<b>Appendix C</b>	<b>Summation Notation 951</b>
<b>Appendix D</b>	<b>Self-Test Solutions and Answers to Even-Numbered Exercises 953</b>
<b>Appendix E</b>	<b>Microsoft Excel 2013 and Tools for Statistical Analysis 1008</b>
<b>Index</b>	<b>1017</b>



# Contents

**Preface xxiii**  
**About the Authors xxx**

## **Chapter 1 Data and Statistics 1**

**Statistics in Practice: *Bloomberg Businessweek* 2**

### **1.1 Applications in Business and Economics 3**

Accounting 3  
Finance 4  
Marketing 4  
Production 4  
Economics 4  
Information Systems 5

### **1.2 Data 5**

Elements, Variables, and Observations 5  
Scales of Measurement 7  
Categorical and Quantitative Data 8  
Cross-Sectional and Time Series Data 8

### **1.3 Data Sources 11**

Existing Sources 11  
Statistical Studies 12  
Data Acquisition Errors 14

### **1.4 Descriptive Statistics 14**

### **1.5 Statistical Inference 16**

### **1.6 Statistical Analysis Using Microsoft Excel 18**

Data Sets and Excel Worksheets 18  
Using Excel for Statistical Analysis 20

### **1.7 Data Mining 21**

### **1.8 Ethical Guidelines for Statistical Practice 22**

**Summary 24**

**Glossary 24**

**Supplementary Exercises 25**

**Appendix An Introduction to StatTools 32**

## **Chapter 2 Descriptive Statistics: Tabular and Graphical Displays 36**

**Statistics in Practice: Colgate-Palmolive Company 37**

### **2.1 Summarizing Data for a Categorical Variable 38**

	Frequency Distribution	38
	Relative Frequency and Percent Frequency Distributions	39
	Using Excel to Construct a Frequency Distribution, a Relative Frequency Distribution, and a Percent Frequency Distribution	40
	Bar Charts and Pie Charts	41
	Using Excel to Construct a Bar Chart and a Pie Chart	43
<b>2.2</b>	<b>Summarizing Data for a Quantitative Variable</b>	<b>49</b>
	Frequency Distribution	49
	Relative Frequency and Percent Frequency Distributions	50
	Using Excel to Construct a Frequency Distribution	51
	Dot Plot	52
	Histogram	53
	Using Excel's Recommended Charts Tool to Construct a Histogram	55
	Cumulative Distributions	57
	Stem-and-Leaf Display	58
<b>2.3</b>	<b>Summarizing Data for Two Variables Using Tables</b>	<b>66</b>
	Crosstabulation	67
	Using Excel's PivotTable Tool to Construct a Crosstabulation	69
	Simpson's Paradox	71
<b>2.4</b>	<b>Summarizing Data for Two Variables Using Graphical Displays</b>	<b>78</b>
	Scatter Diagram and Trendline	78
	Using Excel to Construct a Scatter Diagram and a Trendline	79
	Side-by-Side and Stacked Bar Charts	82
	Using Excel's Recommended Charts Tool to Construct Side-by-Side and Stacked Bar Charts	83
<b>2.5</b>	<b>Data Visualization: Best Practices in Creating Effective Graphical Displays</b>	<b>88</b>
	Creating Effective Graphical Displays	88
	Choosing the Type of Graphical Display	90
	Data Dashboards	90
	Data Visualization in Practice: Cincinnati Zoo and Botanical Garden	92
	<b>Summary</b>	<b>95</b>
	<b>Glossary</b>	<b>96</b>
	<b>Key Formulas</b>	<b>97</b>
	<b>Supplementary Exercises</b>	<b>97</b>
	<b>Case Problem 1 Pelican Stores</b>	<b>102</b>
	<b>Case Problem 2 Motion Picture Industry</b>	<b>104</b>
	<b>Appendix Using StatTools for Tabular and Graphical Presentations</b>	<b>105</b>

**Chapter 3 Descriptive Statistics: Numerical Measures 106****Statistics in Practice: Small Fry Design 107****3.1 Measures of Location 108**

Mean 108

Median 110

Mode 111

Using Excel to Compute the Mean, Median, and Mode 112

Weighted Mean 113

Geometric Mean 114

Using Excel to Compute the Geometric Mean 116

Percentiles 117

Quartiles 118

Using Excel to Compute Percentiles and Quartiles 118

**3.2 Measures of Variability 125**

Range 126

Interquartile Range 126

Variance 127

Standard Deviation 129

Using Excel to Compute the Sample Variance  
and Sample Standard Deviation 129

Coefficient of Variation 130

Using Excel's Descriptive Statistics Tool 131

**3.3 Measures of Distribution Shape, Relative Location,  
and Detecting Outliers 135**

Distribution Shape 135

z-Scores 137

Chebyshev's Theorem 138

Empirical Rule 139

Detecting Outliers 140

**3.4 Five-Number Summaries and Box Plots 143**

Five-Number Summary 143

Box Plot 144

Comparative Analysis Using Box Plots 144

**3.5 Measures of Association Between Two Variables 148**

Covariance 149

Interpretation of the Covariance 150

Correlation Coefficient 153

Interpretation of the Correlation Coefficient 154

Using Excel to Compute the Sample Covariance  
and Sample Correlation Coefficient 155**3.6 Data Dashboards: Adding Numerical Measures to  
Improve Effectiveness 159****Summary 163****Glossary 163**

<b>Key Formulas</b>	<b>165</b>
<b>Supplementary Exercises</b>	<b>166</b>
<b>Case Problem 1 Pelican Stores</b>	<b>172</b>
<b>Case Problem 2 Motion Picture Industry</b>	<b>173</b>
<b>Case Problem 3 Business Schools of Asia-Pacific</b>	<b>174</b>
<b>Case Problem 4 Heavenly Chocolates Website Transactions</b>	<b>174</b>
<b>Case Problem 5 African Elephant Populations</b>	<b>177</b>
<b>Appendix Descriptive Statistics Using StatTools</b>	<b>178</b>

## **Chapter 4 Introduction to Probability 180**

<b>Statistics in Practice: National Aeronautics And Space Administration</b>	<b>181</b>
<b>4.1 Experiments, Counting Rules, and Assigning Probabilities</b>	<b>182</b>
Counting Rules, Combinations, and Permutations	183
Assigning Probabilities	187
Probabilities for the KP&L Project	189
<b>4.2 Events and Their Probabilities</b>	<b>192</b>
<b>4.3 Some Basic Relationships of Probability</b>	<b>196</b>
Complement of an Event	196
Addition Law	197
<b>4.4 Conditional Probability</b>	<b>203</b>
Independent Events	206
Multiplication Law	206
<b>4.5 Bayes' Theorem</b>	<b>211</b>
Tabular Approach	214
<b>Summary</b>	<b>217</b>
<b>Glossary</b>	<b>217</b>
<b>Key Formulas</b>	<b>218</b>
<b>Supplementary Exercises</b>	<b>219</b>
<b>Case Problem Hamilton County Judges</b>	<b>223</b>

## **Chapter 5 Discrete Probability Distributions 226**

<b>Statistics in Practice: Citibank</b>	<b>227</b>
<b>5.1 Random Variables</b>	<b>228</b>
Discrete Random Variables	228
Continuous Random Variables	228
<b>5.2 Developing Discrete Probability Distributions</b>	<b>231</b>
<b>5.3 Expected Value and Variance</b>	<b>236</b>
Expected Value	236
Variance	236
Using Excel to Compute the Expected Value, Variance, and Standard Deviation	238



<b>5.4</b>	<b>Bivariate Distributions, Covariance, and Financial Portfolios</b>	<b>242</b>
	A Bivariate Empirical Discrete Probability Distribution	243
	Financial Applications	246
	Summary	249
<b>5.5</b>	<b>Binomial Probability Distribution</b>	<b>252</b>
	A Binomial Experiment	252
	Martin Clothing Store Problem	254
	Using Excel to Compute Binomial Probabilities	258
	Expected Value and Variance for the Binomial Distribution	260
<b>5.6</b>	<b>Poisson Probability Distribution</b>	<b>263</b>
	An Example Involving Time Intervals	264
	An Example Involving Length or Distance Intervals	265
	Using Excel to Compute Poisson Probabilities	265
<b>5.7</b>	<b>Hypergeometric Probability Distribution</b>	<b>268</b>
	Using Excel to Compute Hypergeometric Probabilities	270
	<b>Summary</b>	<b>273</b>
	<b>Glossary</b>	<b>273</b>
	<b>Key Formulas</b>	<b>274</b>
	<b>Supplementary Exercises</b>	<b>276</b>
	<b>Chapter 6</b>	<b>Continuous Probability Distributions</b>
		<b>280</b>
	<b>Statistics in Practice: Procter &amp; Gamble</b>	<b>281</b>
<b>6.1</b>	<b>Uniform Probability Distribution</b>	<b>282</b>
	Area as a Measure of Probability	283
<b>6.2</b>	<b>Normal Probability Distribution</b>	<b>286</b>
	Normal Curve	286
	Standard Normal Probability Distribution	288
	Computing Probabilities for Any Normal Probability Distribution	293
	Grear Tire Company Problem	294
	Using Excel to Compute Normal Probabilities	295
<b>6.3</b>	<b>Exponential Probability Distribution</b>	<b>301</b>
	Computing Probabilities for the Exponential Distribution	302
	Relationship Between the Poisson and Exponential Distributions	303
	Using Excel to Compute Exponential Probabilities	303
	<b>Summary</b>	<b>306</b>
	<b>Glossary</b>	<b>306</b>
	<b>Key Formulas</b>	<b>307</b>
	<b>Supplementary Exercises</b>	<b>307</b>
	<b>Case Problem</b>	<b>Specialty Toys</b>
		<b>310</b>
	<b>Chapter 7</b>	<b>Sampling and Sampling Distributions</b>
		<b>312</b>
	<b>Statistics in Practice: MeadWestvaco Corporation</b>	<b>313</b>
<b>7.1</b>	<b>The Electronics Associates Sampling Problem</b>	<b>314</b>

<b>7.2</b>	<b>Selecting a Sample</b>	<b>315</b>
	Sampling from a Finite Population	315
	Sampling from an Infinite Population	318
<b>7.3</b>	<b>Point Estimation</b>	<b>322</b>
	Practical Advice	324
<b>7.4</b>	<b>Introduction to Sampling Distributions</b>	<b>326</b>
<b>7.5</b>	<b>Sampling Distribution of <math>\bar{x}</math></b>	<b>329</b>
	Expected Value of $\bar{x}$	329
	Standard Deviation of $\bar{x}$	330
	Form of the Sampling Distribution of $\bar{x}$	331
	Sampling Distribution of $\bar{x}$ for the EAI Problem	333
	Practical Value of the Sampling Distribution of $\bar{x}$	333
	Relationship Between the Sample Size and the Sampling Distribution of $\bar{x}$	335
<b>7.6</b>	<b>Sampling Distribution of <math>\bar{p}</math></b>	<b>339</b>
	Expected Value of $\bar{p}$	339
	Standard Deviation of $\bar{p}$	340
	Form of the Sampling Distribution of $\bar{p}$	341
	Practical Value of the Sampling Distribution of $\bar{p}$	342
<b>7.7</b>	<b>Other Sampling Methods</b>	<b>345</b>
	Stratified Random Sampling	345
	Cluster Sampling	345
	Systematic Sampling	346
	Convenience Sampling	346
	Judgment Sampling	347
	<b>Summary</b>	<b>347</b>
	<b>Glossary</b>	<b>348</b>
	<b>Key Formulas</b>	<b>349</b>
	<b>Supplementary Exercises</b>	<b>349</b>
	<b>Appendix Random Sampling with StatTools</b>	<b>351</b>

## **Chapter 8 Interval Estimation 353**

### **Statistics in Practice: Food Lion 354**

#### **8.1 Population Mean: $\sigma$ Known 355**

    Margin of Error and the Interval Estimate 355

    Using Excel 359

    Practical Advice 360

#### **8.2 Population Mean: $\sigma$ Unknown 362**

    Margin of Error and the Interval Estimate 364

    Using Excel 366

    Practical Advice 368

    Using a Small Sample 368

    Summary of Interval Estimation Procedures 369

<b>8.3</b>	<b>Determining the Sample Size</b>	<b>373</b>
<b>8.4</b>	<b>Population Proportion</b>	<b>376</b>
	Using Excel	377
	Determining the Sample Size	379
	<b>Summary</b>	<b>383</b>
	<b>Glossary</b>	<b>383</b>
	<b>Key Formulas</b>	<b>384</b>
	<b>Supplementary Exercises</b>	<b>384</b>
	<b>Case Problem 1</b> <i>Young Professional Magazine</i>	<b>387</b>
	<b>Case Problem 2</b> <i>Gulf Real Estate Properties</i>	<b>388</b>
	<b>Case Problem 3</b> <i>Metropolitan Research, Inc.</i>	<b>390</b>
	<b>Appendix</b> <i>Interval Estimation with StatTools</i>	<b>391</b>

## **Chapter 9** Hypothesis Tests **393**

### **Statistics in Practice: John Morrell & Company** 394

#### **9.1** Developing Null and Alternative Hypotheses 395

    The Alternative Hypothesis as a Research Hypothesis 395

    The Null Hypothesis as an Assumption to Be Challenged 396

    Summary of Forms for Null and Alternative Hypotheses 397

#### **9.2** Type I and Type II Errors 399

#### **9.3** Population Mean: $\sigma$ Known 401

    One-Tailed Test 401

    Two-Tailed Test 407

    Using Excel 410

    Summary and Practical Advice 412

    Relationship Between Interval Estimation and Hypothesis Testing 413

#### **9.4** Population Mean: $\sigma$ Unknown 417

    One-Tailed Test 418

    Two-Tailed Test 419

    Using Excel 420

    Summary and Practical Advice 422

#### **9.5** Population Proportion 425

    Using Excel 428

    Summary 429

### **Summary** 432

### **Glossary** 432

### **Key Formulas** 433

### **Supplementary Exercises** 433

#### **Case Problem 1** *Quality Associates, Inc.* 436

#### **Case Problem 2** *Ethical Behavior of Business Students at Bayview University* 437

### **Appendix** *Hypothesis Testing with StatTools* 439

## **Chapter 10 Inference About Means and Proportions with Two Populations 440**

### **Statistics in Practice: U.S. Food and Drug Administration 441**

#### **10.1 Inferences About the Difference Between Two Population Means: $\sigma_1$ and $\sigma_2$ Known 442**

Interval Estimation of  $\mu_1 - \mu_2$  442

Using Excel to Construct a Confidence Interval 444

Hypothesis Tests About  $\mu_1 - \mu_2$  446

Using Excel to Conduct a Hypothesis Test 448

Practical Advice 449

#### **10.2 Inferences About the Difference Between Two Population Means: $\sigma_1$ and $\sigma_2$ Unknown 452**

Interval Estimation of  $\mu_1 - \mu_2$  452

Using Excel to Construct a Confidence Interval 454

Hypothesis Tests About  $\mu_1 - \mu_2$  456

Using Excel to Conduct a Hypothesis Test 458

Practical Advice 460

#### **10.3 Inferences About the Difference Between Two Population Means: Matched Samples 463**

Using Excel to Conduct a Hypothesis Test 466

#### **10.4 Inferences About the Difference Between Two Population Proportions 471**

Interval Estimation of  $p_1 - p_2$  471

Using Excel to Construct a Confidence Interval 473

Hypothesis Tests About  $p_1 - p_2$  474

Using Excel to Conduct a Hypothesis Test 476

#### **Summary 480**

#### **Glossary 481**

#### **Key Formulas 481**

#### **Supplementary Exercises 482**

#### **Case Problem Par, Inc. 485**

#### **Appendix Inferences About Two Populations Using StatTools 486**

## **Chapter 11 Inferences About Population Variances 489**

### **Statistics in Practice: U.S. Government Accountability Office 490**

#### **11.1 Inferences About a Population Variance 491**

Interval Estimation 491

Using Excel to Construct a Confidence Interval 495

Hypothesis Testing 496

Using Excel to Conduct a Hypothesis Test 499

#### **11.2 Inferences About Two Population Variances 503**

Using Excel to Conduct a Hypothesis Test 507

**Summary 512**

**Key Formulas 512**

**Supplementary Exercises 512**

**Case Problem Air Force Training Program 514**

**Appendix Inferences About Population Variances with StatTools 515**

## **Chapter 12 Tests of Goodness of Fit, Independence, and Multiple Proportions 517**

**Statistics in Practice: United Way 518**

**12.1 Goodness of Fit Test 519**

Multinomial Probability Distribution 519

Using Excel to Conduct a Goodness of Fit Test 523

**12.2 Test of Independence 525**

Using Excel to Conduct a Test of Independence 529

**12.3 Testing for Equality of Three or More Population Proportions 534**

A Multiple Comparison Procedure 538

Using Excel to Conduct a Test of Multiple Proportions 540

**Summary 544**

**Glossary 545**

**Key Formulas 545**

**Supplementary Exercises 545**

**Case Problem A Bipartisan Agenda for Change 548**

**Appendix Tests of Independence and Multiple Proportions  
Using StatTools 549**

## **Chapter 13 Experimental Design and Analysis of Variance 551**

**Statistics in Practice: Burke Marketing Services, Inc. 552**

**13.1 An Introduction to Experimental Design and  
Analysis of Variance 553**

Data Collection 554

Assumptions for Analysis of Variance 555

Analysis of Variance: A Conceptual Overview 555

**13.2 Analysis of Variance and the Completely Randomized Design 558**

Between-Treatments Estimate of Population Variance 559

Within-Treatments Estimate of Population Variance 560

Comparing the Variance Estimates: The  $F$  Test 560

ANOVA Table 562

Using Excel 564

Testing for the Equality of  $k$  Population Means:

An Observational Study 565

**13.3 Multiple Comparison Procedures 570**

Fisher's LSD 570

Type I Error Rates 573

<b>13.4</b>	<b>Randomized Block Design</b>	<b>576</b>
	Air Traffic Controller Stress Test	577
	ANOVA Procedure	578
	Computations and Conclusions	579
	Using Excel	581
<b>13.5</b>	<b>Factorial Experiment</b>	<b>585</b>
	ANOVA Procedure	586
	Computations and Conclusions	587
	Using Excel	590
	<b>Summary</b>	<b>594</b>
	<b>Glossary</b>	<b>595</b>
	<b>Key Formulas</b>	<b>595</b>
	<b>Supplementary Exercises</b>	<b>597</b>
	<b>Case Problem 1</b>	<b>Wentworth Medical Center 602</b>
	<b>Case Problem 2</b>	<b>Compensation for Sales Professionals 603</b>
	<b>Appendix</b>	<b>Analysis of a Completely Randomized Design Using StatTools 603</b>

## **Chapter 14** Simple Linear Regression 605

### **Statistics in Practice: Alliance Data Systems 606**

<b>14.1</b>	<b>Simple Linear Regression Model</b>	<b>607</b>
	Regression Model and Regression Equation	607
	Estimated Regression Equation	608
<b>14.2</b>	<b>Least Squares Method</b>	<b>610</b>
	Using Excel to Construct a Scatter Diagram, Display the Estimated Regression Line, and Display the Estimated Regression Equation	614
<b>14.3</b>	<b>Coefficient of Determination</b>	<b>622</b>
	Using Excel to Compute the Coefficient of Determination	626
	Correlation Coefficient	626
<b>14.4</b>	<b>Model Assumptions</b>	<b>630</b>
<b>14.5</b>	<b>Testing for Significance</b>	<b>632</b>
	Estimate of $\sigma^2$	632
	$t$ Test	633
	Confidence Interval for $\beta_1$	635
	$F$ Test	636
	Some Cautions About the Interpretation of Significance Tests	638
<b>14.6</b>	<b>Using the Estimated Regression Equation for Estimation and Prediction</b>	<b>641</b>
	Interval Estimation	642
	Confidence Interval for the Mean Value of $y$	643
	Prediction Interval for an Individual Value of $y$	644

<b>14.7</b>	<b>Excel's Regression Tool</b>	<b>649</b>
	Using Excel's Regression Tool for the Armand's Pizza Parlors Example	649
	Interpretation of Estimated Regression Equation Output	650
	Interpretation of ANOVA Output	651
	Interpretation of Regression Statistics Output	652
	Using StatTools to Compute Prediction Intervals	652
<b>14.8</b>	<b>Residual Analysis: Validating Model Assumptions</b>	<b>656</b>
	Residual Plot Against $x$	657
	Residual Plot Against $\hat{y}$	658
	Standardized Residuals	660
	Using Excel to Construct a Residual Plot	662
	Normal Probability Plot	665
<b>14.9</b>	<b>Outliers and Influential Observations</b>	<b>668</b>
	Detecting Outliers	669
	Detecting Influential Observations	670
	<b>Summary</b>	<b>675</b>
	<b>Glossary</b>	<b>675</b>
	<b>Key Formulas</b>	<b>676</b>
	<b>Supplementary Exercises</b>	<b>679</b>
	<b>Case Problem 1</b> Measuring Stock Market Risk	<b>684</b>
	<b>Case Problem 2</b> U.S. Department of Transportation	<b>685</b>
	<b>Case Problem 3</b> Selecting a Point-and-Shoot Digital Camera	<b>686</b>
	<b>Case Problem 4</b> Finding the Best Car Value	<b>687</b>
	<b>Appendix 14.1</b> Calculus-Based Derivation of Least Squares Formulas	<b>688</b>
	<b>Appendix 14.2</b> A Test for Significance Using Correlation	<b>689</b>
	<b>Appendix 14.3</b> Regression Analysis Using StatTools	<b>690</b>

## **Chapter 15** Multiple Regression **692**

### **Statistics in Practice: International Paper** **693**

<b>15.1</b>	<b>Multiple Regression Model</b>	<b>694</b>
	Regression Model and Regression Equation	694
	Estimated Multiple Regression Equation	694
<b>15.2</b>	<b>Least Squares Method</b>	<b>695</b>
	An Example: Butler Trucking Company	696
	Using Excel's Regression Tool to Develop the Estimated Multiple Regression Equation	699
	Note on Interpretation of Coefficients	700
<b>15.3</b>	<b>Multiple Coefficient of Determination</b>	<b>706</b>
<b>15.4</b>	<b>Model Assumptions</b>	<b>709</b>
<b>15.5</b>	<b>Testing for Significance</b>	<b>710</b>
	$F$ Test	711

	<i>t</i> Test	713
	Multicollinearity	714
<b>15.6</b>	<b>Using the Estimated Regression Equation for Estimation and Prediction</b>	<b>717</b>
<b>15.7</b>	<b>Categorical Independent Variables</b>	<b>720</b>
	An Example: Johnson Filtration, Inc.	720
	Interpreting the Parameters	723
	More Complex Categorical Variables	724
<b>15.8</b>	<b>Residual Analysis</b>	<b>728</b>
	Residual Plot Against $\hat{y}$	728
	Standardized Residual Plot Against $\hat{y}$	730
	<b>Summary</b>	<b>732</b>
	<b>Glossary</b>	<b>733</b>
	<b>Key Formulas</b>	<b>734</b>
	<b>Supplementary Exercises</b>	<b>735</b>
	<b>Case Problem 1 Consumer Research, Inc.</b>	<b>740</b>
	<b>Case Problem 2 Predicting Winnings for NASCAR Drivers</b>	<b>741</b>
	<b>Case Problem 3 Finding the Best Car Value</b>	<b>743</b>
	<b>Appendix Multiple Regression Analysis Using StatTools</b>	<b>744</b>
<b>Chapter 16</b>	<b>Regression Analysis: Model Building</b>	<b>745</b>
	<b>Statistics in Practice: Monsanto Company</b>	<b>746</b>
<b>16.1</b>	<b>General Linear Model</b>	<b>747</b>
	Modeling Curvilinear Relationships	747
	Interaction	749
	Transformations Involving the Dependent Variable	753
	Nonlinear Models That Are Intrinsically Linear	756
<b>16.2</b>	<b>Determining When to Add or Delete Variables</b>	<b>760</b>
	General Case	762
<b>16.3</b>	<b>Analysis of a Larger Problem</b>	<b>766</b>
<b>16.4</b>	<b>Variable Selection Procedures</b>	<b>770</b>
	Stepwise Regression	770
	Forward Selection	771
	Backward Elimination	772
	Best-Subsets Regression	772
<b>16.5</b>	<b>Multiple Regression Approach to Experimental Design</b>	<b>775</b>
<b>16.6</b>	<b>Autocorrelation and the Durbin-Watson Test</b>	<b>780</b>
	<b>Summary</b>	<b>784</b>
	<b>Glossary</b>	<b>784</b>
	<b>Key Formulas</b>	<b>784</b>
	<b>Supplementary Exercises</b>	<b>785</b>
	<b>Case Problem 1 Analysis of LPGA Tour Statistics</b>	<b>788</b>



<b>Case Problem 2</b>	<b>Rating Wines from the Piedmont Region of Italy</b>	<b>789</b>
<b>Appendix</b>	<b>Variable Selection Procedures Using StatTools</b>	<b>790</b>
<b>Chapter 17</b>	<b>Time Series Analysis and Forecasting</b>	<b>791</b>
<b>Statistics in Practice:</b>	<b>Nevada Occupational Health Clinic</b>	<b>792</b>
<b>17.1</b>	<b>Time Series Patterns</b>	<b>793</b>
	Horizontal Pattern	793
	Trend Pattern	795
	Seasonal Pattern	795
	Trend and Seasonal Pattern	796
	Cyclical Pattern	797
	Using Excel's Chart Tools to Construct a Time Series Plot	799
	Selecting a Forecasting Method	799
<b>17.2</b>	<b>Forecast Accuracy</b>	<b>800</b>
<b>17.3</b>	<b>Moving Averages and Exponential Smoothing</b>	<b>805</b>
	Moving Averages	805
	Using Excel's Moving Average Tool	807
	Weighted Moving Averages	808
	Exponential Smoothing	809
	Using Excel's Exponential Smoothing Tool	812
<b>17.4</b>	<b>Trend Projection</b>	<b>817</b>
	Linear Trend Regression	817
	Using Excel's Regression Tool to Compute a Linear Trend Equation	821
	Nonlinear Trend Regression	822
	Using Excel's Regression Tool to Compute a Quadratic Trend Equation	822
	Using Excel's Chart Tools for Trend Projection	824
<b>17.5</b>	<b>Seasonality and Trend</b>	<b>830</b>
	Seasonality Without Trend	830
	Seasonality and Trend	832
	Models Based on Monthly Data	835
<b>17.6</b>	<b>Time Series Decomposition</b>	<b>839</b>
	Calculating the Seasonal Indexes	841
	Deseasonalizing the Time Series	844
	Using the Deseasonalized Time Series to Identify Trend	845
	Seasonal Adjustments	846
	Models Based on Monthly Data	847
	Cyclical Component	847
	<b>Summary</b>	<b>849</b>
	<b>Glossary</b>	<b>850</b>
	<b>Key Formulas</b>	<b>851</b>
	<b>Supplementary Exercises</b>	<b>852</b>

**Case Problem 1 Forecasting Food and Beverage Sales 856**

**Case Problem 2 Forecasting Lost Sales 857**

**Appendix Forecasting Using StatTools 858**

## **Chapter 18 Nonparametric Methods 860**

**Statistics in Practice: West Shell Realtors 861**

**18.1 Sign Test 862**

Hypothesis Test About a Population Median 862

Hypothesis Test with Matched Samples 866

Using Excel 868

**18.2 Wilcoxon Signed-Rank Test 870**

**18.3 Mann-Whitney-Wilcoxon Test 876**

**18.4 Kruskal-Wallis Test 886**

**18.5 Rank Correlation 891**

Using Excel 893

**Summary 896**

**Glossary 897**

**Key Formulas 898**

**Supplementary Exercises 898**

**Appendix Nonparametric Methods with StatTools 901**

## **Chapter 19 Statistical Methods for Quality Control 903**

**Statistics in Practice: Dow Chemical Company 904**

**19.1 Philosophies and Frameworks 905**

Malcolm Baldrige National Quality Award 906

ISO 9000 906

Six Sigma 906

Quality in the Service Sector 909

**19.2 Statistical Process Control 909**

Control Charts 910

$\bar{x}$  Chart: Process Mean and Standard Deviation Known 911

$\bar{x}$  Chart: Process Mean and Standard Deviation Unknown 913

$R$  Chart 916

$p$  Chart 918

$np$  Chart 920

Interpretation of Control Charts 920

**19.3 Acceptance Sampling 923**

KALI, Inc.: An Example of Acceptance Sampling 924

Computing the Probability of Accepting a Lot 925

Selecting an Acceptance Sampling Plan 928

Multiple Sampling Plans 930

**Summary 931**

<b>Glossary</b>	<b>931</b>
<b>Key Formulas</b>	<b>932</b>
<b>Supplementary Exercises</b>	<b>933</b>
<b>Appendix Control Charts Using StatTools</b>	<b>935</b>

## **Chapter 20** Decision Analysis On Website

<b>Statistics in Practice: Ohio Edison Company</b>	<b>20-2</b>
<b>20.1 Problem Formulation</b>	<b>20-3</b>
Payoff Tables	20-4
Decision Trees	20-4
<b>20.2 Decision Making with Probabilities</b>	<b>20-5</b>
Expected Value Approach	20-5
Expected Value of Perfect Information	20-7
<b>20.3 Decision Analysis with Sample Information</b>	<b>20-13</b>
Decision Tree	20-14
Decision Strategy	20-16
Expected Value of Sample Information	20-18
<b>20.4 Computing Branch Probabilities Using Bayes' Theorem</b>	<b>20-24</b>
<b>Summary</b>	<b>20-29</b>
<b>Glossary</b>	<b>20-29</b>
<b>Key Formulas</b>	<b>20-30</b>
<b>Supplementary Exercises</b>	<b>20-30</b>
<b>Case Problem Lawsuit Defense Strategy</b>	<b>20-33</b>
<b>Appendix An Introduction to PrecisionTree</b>	<b>20-34</b>

## **Chapter 21** Sample Survey On Website

<b>Statistics in Practice: Duke Energy</b>	<b>21-2</b>
<b>21.1 Terminology Used in Sample Surveys</b>	<b>21-2</b>
<b>21.2 Types of Surveys and Sampling Methods</b>	<b>21-3</b>
<b>21.3 Survey Errors</b>	<b>21-5</b>
Nonsampling Error	21-5
Sampling Error	21-5
<b>21.4 Simple Random Sampling</b>	<b>21-6</b>
Population Mean	21-6
Population Total	21-7
Population Proportion	21-8
Using Excel for Simple Random Sampling	21-9
Determining the Sample Size	21-11
<b>21.5 Stratified Simple Random Sampling</b>	<b>21-14</b>
Population Mean	21-14
Using Excel: Population Mean	21-16
Population Total	21-17
Population Proportion	21-18

	Using Excel: Population Proportion	21-20
	Determining the Sample Size	21-20
<b>21.6</b>	<b>Cluster Sampling</b>	<b>21-24</b>
	Population Mean	21-26
	Population Total	21-28
	Population Proportion	21-29
	Using Excel for Cluster Sampling	21-30
	Determining the Sample Size	21-30
<b>21.7</b>	<b>Systematic Sampling</b>	<b>21-32</b>
	<b>Summary</b>	<b>21-33</b>
	<b>Glossary</b>	<b>21-33</b>
	<b>Key Formulas</b>	<b>21-34</b>
	<b>Supplementary Exercises</b>	<b>21-38</b>
	<b>Appendix A</b>	<b>References and Bibliography 938</b>
	<b>Appendix B</b>	<b>Tables 940</b>
	<b>Appendix C</b>	<b>Summation Notation 951</b>
	<b>Appendix D</b>	<b>Self-Test Solutions and Answers to Even-Numbered Exercises 953</b>
	<b>Appendix E</b>	<b>Microsoft Excel 2013 and Tools for Statistical Analysis 1008</b>
	<b>Index</b>	<b>1017</b>

# Preface

The purpose of *Modern Business Statistics with Microsoft Office® Excel®* is to give students, primarily those in the fields of business administration and economics, a conceptual introduction to the field of statistics and its many applications. The text is applications oriented and written with the needs of the nonmathematician in mind; the mathematical prerequisite is knowledge of algebra.

Applications of data analysis and statistical methodology are an integral part of the organization and presentation of the text material. The discussion and development of each technique is presented in an applications setting, with the statistical results providing insights to decisions and solutions to applied problems.

Although the book is applications oriented, we have taken care to provide sound methodological development and to use notation that is generally accepted for the topic being covered. Hence, students will find that this text provides good preparation for the study of more advanced statistical material. A bibliography to guide further study is included as an appendix.

## Use of Microsoft Excel for Statistical Analysis

*Modern Business Statistics with Microsoft Excel* is first and foremost a statistics textbook that emphasizes statistical concepts and applications. But since most practical problems are too large to be solved using hand calculations, some type of statistical software package is required to solve these problems. There are several excellent statistical packages available today. However, because most students and potential employers value spreadsheet experience, many schools now use a spreadsheet package in their statistics courses. Microsoft Excel is the most widely used spreadsheet package in business as well as in colleges and universities. We have written *Modern Business Statistics with Microsoft Excel* especially for statistics courses in which Microsoft Excel is used as the software package.

Excel has been integrated within each of the chapters and plays an integral part in providing an application orientation. Although we assume that readers using this text are familiar with Excel basics such as selecting cells, entering formulas, copying, and so on, we do not assume that readers are familiar with Excel 2013 or Excel's tools for statistical analysis. As a result, we have included Appendix E, which provides an introduction to Excel 2013 and tools for statistical analysis.

Throughout the text the discussion of using Excel to perform a statistical procedure appears in a subsection immediately following the discussion of the statistical procedure. We believe that this style enables us to fully integrate the use of Excel throughout the text, but still maintain the primary emphasis on the statistical methodology being discussed. In each of these subsections, we use a standard format for using Excel for statistical analysis. There are four primary tasks: Enter/Access Data, Enter Functions and Formulas, Apply Tools, and Editing Options. We believe a consistent framework for applying Excel helps users to focus on the statistical methodology without getting bogged down in the details of using Excel.

In presenting worksheet figures we often use a nested approach in which the worksheet shown in the background of the figure displays the formulas and the worksheet shown in the foreground shows the values computed using the formulas. Different colors and shades of colors are used to differentiate worksheet cells containing data, highlight cells containing

Excel functions and formulas, and highlight material printed by Excel as a result of using one or more data analysis tools.

## Use of StatTools

StatTools is a commercial Excel add-in which we and Palisade Corporation have made available to adopters of this text for free. StatTools extends the range of statistical and graphical options for Excel users. In an appendix to Chapter 1 we show how to download and install StatTools. Most chapters also include an appendix that shows the steps required to accomplish a statistical procedure using StatTools.

We have been very careful to make the use of StatTools completely optional. Users who want to teach using the standard tools available in Excel 2013 can continue to do so. But users who want additional statistical capabilities not available in Excel 2013 now have access to an industry standard statistics add-in that students will be able to continue to use in the workplace.

## Changes in the Fifth Edition

We appreciate the acceptance and positive response to the previous editions of *Modern Business Statistics with Microsoft Excel*. Accordingly, in making modifications for this new edition, we have maintained the presentation style and readability of those editions. The significant changes in the new edition are summarized here.

- **Microsoft Excel 2013.** Step-by-step instructions and screen captures show how to use the latest version of Excel to implement statistical procedures.
- **Revised Chapter 2.** We have significantly revised Chapter 2 to incorporate new tools available with Excel 2013 and new material on data visualization. We now show how Excel's Recommended PivotTables tool can be used to construct a frequency distribution for a categorical variable and how Excel's Recommended Charts tool can be used to construct a histogram for a quantitative variable. Also, Chapter 2 has been reorganized to include new material on side-by-side and stacked bar charts, including showing how to use Excel's Recommended Charts tool to construct both types of charts. A new section has been added on data visualization, data dashboards, and best practices in creating effective visual displays.
- **Revised Chapter 3.** Chapter 3 now includes coverage of the weighted mean and geometric mean in the section on measures of location. The geometric mean has many applications in the computation of growth rates for financial assets, annual percentage rates, and so on. We have also completely rewritten the material on percentiles and quartiles, including a new procedure for computing percentiles that provides results consistent with Excel's PERCENTILE.EXC function. Chapter 3 also includes a new section on data dashboards and how summary statistics can be incorporated to enhance their effectiveness.
- **Revised Chapter 5.** The introductory material in this chapter has been revised to improve the explanation of the role of probability distributions and to show how the material on assigning probabilities in Chapter 4 can be used to develop discrete probability distributions. We point out that the empirical discrete probability distribution is developed by using the relative frequency method to assign probabilities. At the request of many users, we have added a new section (Section 5.4) that covers bivariate discrete distributions and financial applications.
- **Test of Goodness of Fit, Independence, and Multiple Proportions—Chapter 12.** This chapter has undergone a major revision. In Section 12.1 we show how the chi-square goodness of fit test can be used to determine if a frequency distribution

developed from categorical data is a good fit to a hypothesized probability distribution. In Section 12.2 we show how the chi-square test of independence is used to determine if two categorical variables sampled from one population are independent. Section 12.3 is a new section that describes how sample data from three or more populations can be used to determine if the population proportions are equal. For each test we show how Excel's Recommended PivotTables tool and Excel's CHISQ.TEST function can be used to compute the chi-square test statistic and corresponding  $p$ -value.

- **New Case Problems.** We have added 9 new case problems to this edition. The new case problems appear in the chapters on descriptive statistics and regression analysis. The 30 case problems in the text provide students with the opportunity to analyze somewhat larger data sets and prepare managerial reports based on the results of their analysis.
- **New Statistics in Practice Applications.** Each chapter begins with a Statistics in Practice vignette that describes an application of the statistical methodology in the chapter. New to this edition is a Statistics in Practice for Chapter 2 describing the use of data dashboards and data visualization at the Cincinnati Zoo. We have also added a new Statistics in Practice to Chapter 4 describing how a NASA team used probability to assist the rescue of 33 Chilean miners trapped by a cave-in.
- **New Examples and Exercises Based on Real Data.** We have added approximately 225 new examples and exercises based on real data and recently referenced sources of statistical information. Using data obtained from various data collection organizations and other sources, such as *The Wall Street Journal*, *USA Today*, *Fortune*, and *Barron's*, we have drawn upon actual studies to develop explanations and to create exercises that demonstrate many uses of statistics in business and economics. We believe the use of real data helps generate more student interest in the material and enables the student to learn about both the statistical methodology and its application.

## Features and Pedagogy

Authors Anderson, Sweeney, and Williams have continued many of the features that appeared in previous editions. Important ones for students are noted here.

### Methods Exercises and Applications Exercises

The end-of-section exercises are split into two parts, Methods and Applications. The Methods exercises require students to use the formulas and make the necessary computations. The Applications exercises require students to use the chapter material in real-world situations. Thus, students first focus on the computational “nuts and bolts” and then move on to the subtleties of statistical application and interpretation.

### Self-Test Exercises

Certain exercises are identified as “Self-Test Exercises.” Completely worked-out solutions for these exercises are provided in Appendix D. Students can attempt the Self-Test Exercises and immediately check the solution to evaluate their understanding of the concepts presented in the chapter.

### Margin Annotations and Notes and Comments

Margin annotations that highlight key points and provide additional insights for the student are a key feature of this text. These annotations, which appear in the margins, are designed to provide emphasis and enhance understanding of the terms and concepts being presented in the text.

At the end of many sections, we provide Notes and Comments designed to give the student additional insights about the statistical methodology and its application. Notes and Comments include warnings about or limitations of the methodology, recommendations for application, brief descriptions of additional technical considerations, and other matters.

## Data Files Accompany the Text

Over 200 Excel data files are available on the website that accompanies the text. Webfile logos are used in the text to identify the data sets that are available on the website. Data sets for all case problems as well as data sets for larger exercises are included.

## Acknowledgments

A special thanks goes to our associates from business and industry who supplied the Statistics in Practice features. We recognize them individually by a credit line in each of the articles. We are also indebted to our product manager Aaron Arnsperger, our content developer Margaret Kubale, our content project manager Cliff Kallemeyn, and others at Cengage Learning for their editorial counsel and support during the preparation of this text.

We would like to acknowledge the work of our reviewers who provided comments and suggestions of ways to continue to improve our text. Thanks to

Jamal Abdul-Hafidh University of Missouri– St. Louis	William H. Bleuel Pepperdine University	John R. Carpenter Cornerstone University
Chris Adalikwu Concordia College	Gary Bliss Florida State University– Panama City	Jasmine Chang Georgia State University
Eugene Allevato Woodbury University	Leslie M. Bobb New York Institute of Technology	Si Chen Murray State University
Solomon Antony Murray State University	Michelle Boddy Baker College	Alan S. Chesen Wright State University
Ardavan Asef-Vaziri California State University, Northridge	Thomas W. Bolland Ohio University	Michael Cicero Highline Community College
S. Scott Bailey Troy University	Derrick S. Boone, Sr. Wake Forest University	Robert Collins Marquette University
Robert J. Banis University of Missouri– St. Louis	Lawrence Bos Cornerstone University	Ping Deng Maryville University
Wayne Bedford University of West Alabama	Alan Brokaw Michigan Tech University	Sarvanan Devaraj Notre Dame University
Enoch K. Beraho South Carolina State University	Nancy Brooks University of Vermont	Terry Dielman Texas Christian University
Timothy M. Bergquist Northwest Christian College	Yvonne Brown Pima Community College	Cassandra DiRienzo Elon University
Darl Bien University of Denver	Robert Burgess Georgia Tech	Anne Drougas Dominican University
	Von L. Burton Athens State University	Jianjun Du University of Houston, Victoria



John N. Dyer Georgia Southern University	Erick Hofacker University of Wisconsin– River Falls	John Lawrence California State University, Fullerton
Hossein Eftekari University of Wisconsin– River Falls	Amy C. Hooper Gettysburg College	Tenpao Lee Niagara University
Mohammed A. El-Saidi Ferris State University	Paul Hudec Milwaukee School of Engineering	Daniel Light Northwest State College
Robert M. Escudero Pepperdine University	Alan Humphrey University of Rhode Island	Robert Lindsey College of Charleston
Allessandra Faggian Ohio State University	Wade Jackson University of Memphis	B. Lucas A&M College
Nicolas Farnum California State University, Fullerton	Timmy James Northwest Shoals Community College	Michael Machiorlatti City College of San Francisco
Abe Feinberg California State University, Northridge	Eugene Jones The Ohio State University	Malik B. Malik University of Maryland Eastern Shore
Maggie Williams Flint Northeast State Tech Community College	Naser Kamleh Wallace Community College	Lee McClain Western Washington University
Alfonso Flores-Lagunes University of Arizona	Mark P. Karscig Central Missouri State University	Timothy E. McDaniel Buena Vista University
James Flynn Cleveland State University	Howard Kittleson Riverland Community College	Kim I. Melton North Georgia College & State University
Alan F. Foltz Drury University	Kenneth Klassen California State University, Northridge	Brian Metz Cabrini College
Ronald L. Friesen Bluffton College	Eileen Quinn Knight St. Xavier University, Chicago	John M. Miller Sam Houston State University
Richard Gebhart University of Tulsa	Bharat Kolluri University of Hartford	Patricia A. Mullins University of Wisconsin– Madison
Paul Gentine Bethany College	Joseph Kosler Indiana University of Pennsylvania	Jack Muryn University of Wisconsin, Washington County
Deborah J. Gougeon University of Scranton	David A. Kravitz George Mason University	Muhammad Mustafa South Carolina State University
Jeffrey Gropp DePauw University	Laura Kuhl University of Phoenix, Cleveland Campus	Anthony Narsing Macon State College
V. Daniel Guide Duquesne University	June Lapidus Roosevelt University	Kenneth F. O’Brien Farmingdale State College
Aravind Narasipur Chennai Business School		Ceyhun Ozgur Valparaiso University
Rhonda Hensley North Carolina A&T University		

Michael Parzen Emory University	Probir Roy University of Missouri, Kansas City	Peter Wibawa Sutanto Prairie View A&M University
Barry Pasternack California State University, Fullerton	Randall K. Russell Yavapai College	Lee Tangedahk University of Montana
Lynne Pastor Carnegie Mellon University	Alan Safer California State University, Long Beach	Sudhir Thakur California State University, Sacramento
Ranjna Patel Bethune-Cookman College	David Satava University of Houston, Victoria	Alexander Thomson Schoolcraft College
Tremaine Pimperl Faulkner State Community College	Richard W. Schrader Bellarmine University	Suzanne Tilleman Montana State University Northern
Jennifer M. Platania Elon University	Larry Seifert Webster University	Daniel Tschopp Daemen College (NY)
Von Roderick Plessner Northwest State Community College	John Seydel Arkansas State University	Sushila Umashankar University of Arizona
Glenn Potts University of Wisconsin– River Falls	Jim Shi Robinson College of Business Georgia State University	Jack Vaughn University of Texas, El Paso
Irene Powell Grinnell College	Philip Shaw Fairfield University	Dave Vinson Pellissippi State Community College
Leonard Presby William Paterson University	Robert Simoneau Keene State College	John Vogt Newman University
W. N. Pruitt South Carolina State University	Harvey A. Singer George Mason University	Geoffrey L. Wallace University of Wisconsin, Madison
Narseeyappa Rajanikanth Mississippi Valley State University	Donald R. Smith Monmouth University	Michael Wiemann Metro Community Colleges
Elizabeth L. Rankin Centenary College of Louisiana	Toni M. Somers Wayne State University	Charles Wilf Duquesne University
Tim Raynor Albertus Magnus College	Clifford Sowell Berea College	John Wiorowski University of Texas, Dallas
Carolyn Renier Pellissippi State Tech Community College	Keith Starcher Indiana Wesleyan University	Joyce A. Zadzilka Morrisville State College
Ronny Richardson Southern Polytechnic State University	William Stein Texas A&M University	Guoqiang Peter Zhang Georgia Southern University
Leonard E. Ross California State University, Pomona	Jason Stine Troy University	Zhe George Zhang Western Washington University
	William Struning Seton Hall University	Deborah G. Ziegler Hannibal-LaGrange College
	Timothy S. Sullivan Southern Illinois University, Edwardsville	

We would like to recognize the following individuals who have helped us in the past and continue to influence our writing.

Glen Archibald University of Mississippi	David W. Cravens Texas Christian University	Paul Guy California State University, Chico
Mike Bourke Houston Baptist University	Robert Carver Stonehill College	Alan Humphrey University of Rhode Island
Peter Bryant University of Colorado	Tom Dahlstrom Eastern College	Ann Hussein Philadelphia College of Textiles and Science
Terri L. Byczkowski University of Cincinnati	Ronald Ehresman Baldwin-Wallace College	Ben Isselhardt Rochester Institute of Technology
Ying Chien University of Scranton	Michael Ford Rochester Institute of Technology	Jeffery Jarrett University of Rhode Island
Robert Cochran University of Wyoming	Phil Fry Boise State University	
Murray Côté University of Florida		

# About the Authors

**David R. Anderson.** David R. Anderson is Professor Emeritus of Quantitative Analysis in the College of Business Administration at the University of Cincinnati. Born in Grand Forks, North Dakota, he earned his B.S., M.S., and Ph.D. degrees from Purdue University. Professor Anderson has served as Head of the Department of Quantitative Analysis and Operations Management and as Associate Dean of the College of Business Administration at the University of Cincinnati. In addition, he was the coordinator of the College's first Executive Program.

At the University of Cincinnati, Professor Anderson has taught introductory statistics for business students as well as graduate-level courses in regression analysis, multivariate analysis, and management science. He has also taught statistical courses at the Department of Labor in Washington, D.C. He has been honored with nominations and awards for excellence in teaching and excellence in service to student organizations.

Professor Anderson has coauthored 10 textbooks in the areas of statistics, management science, linear programming, and production and operations management. He is an active consultant in the field of sampling and statistical methods.

**Dennis J. Sweeney.** Dennis J. Sweeney is Professor Emeritus of Quantitative Analysis and Founder of the Center for Productivity Improvement at the University of Cincinnati. Born in Des Moines, Iowa, he earned a B.S.B.A. degree from Drake University and his M.B.A. and D.B.A. degrees from Indiana University, where he was an NDEA Fellow. Professor Sweeney has worked in the management science group at Procter & Gamble and spent a year as a visiting professor at Duke University. Professor Sweeney served as Head of the Department of Quantitative Analysis and as Associate Dean of the College of Business Administration at the University of Cincinnati.

Professor Sweeney has published more than 30 articles and monographs in the area of management science and statistics. The National Science Foundation, IBM, Procter & Gamble, Federated Department Stores, Kroger, and Duke Energy have funded his research, which has been published in *Management Science*, *Operations Research*, *Mathematical Programming*, *Decision Sciences*, and other journals.

Professor Sweeney has coauthored 10 textbooks in the areas of statistics, management science, linear programming, and production and operations management.

**Thomas A. Williams.** Thomas A. Williams is Professor Emeritus of Management Science in the College of Business at Rochester Institute of Technology. Born in Elmira, New York, he earned his B.S. degree at Clarkson University. He did his graduate work at Rensselaer Polytechnic Institute, where he received his M.S. and Ph.D. degrees.

Before joining the College of Business at RIT, Professor Williams served for seven years as a faculty member in the College of Business Administration at the University of Cincinnati, where he developed the undergraduate program in Information Systems and then served as its coordinator. At RIT he was the first chairman of the Decision Sciences Department. He has taught courses in management science and statistics, as well as graduate courses in regression and decision analysis.

Professor Williams is the coauthor of 11 textbooks in the areas of management science, statistics, production and operations management, and mathematics. He has been a consultant for numerous Fortune 500 companies and has worked on projects ranging from the use of data analysis to the development of large-scale regression models.

# CHAPTER 1

## Data and Statistics

---

### CONTENTS

#### STATISTICS IN PRACTICE:

#### *BLOOMBERG BUSINESSWEEK*

#### **1.1** APPLICATIONS IN BUSINESS AND ECONOMICS

Accounting

Finance

Marketing

Production

Economics

Information Systems

#### **1.2** DATA

Elements, Variables, and

Observations

Scales of Measurement

Categorical and Quantitative Data

Cross-Sectional and Time

Series Data

#### **1.3** DATA SOURCES

Existing Sources

Statistical Studies

Data Acquisition Errors

#### **1.4** DESCRIPTIVE STATISTICS

#### **1.5** STATISTICAL INFERENCE

#### **1.6** STATISTICAL ANALYSIS

USING MICROSOFT EXCEL

Data Sets and Excel Worksheets

Using Excel for Statistical

Analysis

#### **1.7** DATA MINING

#### **1.8** ETHICAL GUIDELINES FOR STATISTICAL PRACTICE

**STATISTICS** *in* **PRACTICE**
**BLOOMBERG BUSINESSWEEK\***

NEW YORK, NEW YORK

With a global circulation of more than 1 million, *Bloomberg Businessweek* is one of the most widely read business magazines in the world. Bloomberg's 1700 reporters in 145 service bureaus around the world enable *Bloomberg Businessweek* to deliver a variety of articles of interest to the global business and economic community. Along with feature articles on current topics, the magazine contains articles on international business, economic analysis, information processing, and science and technology. Information in the feature articles and the regular sections helps readers stay abreast of current developments and assess the impact of those developments on business and economic conditions.

Most issues of *Bloomberg Businessweek*, formerly *BusinessWeek*, provide an in-depth report on a topic of current interest. Often, the in-depth reports contain statistical facts and summaries that help the reader understand the business and economic information. Examples of articles and reports include the impact of businesses moving important work to cloud computing, the crisis facing the U.S. Postal Service, and why the debt crisis is even worse than we think. In addition, *Bloomberg Businessweek* provides a variety of statistics about the state of the economy, including production indexes, stock prices, mutual funds, and interest rates.

*Bloomberg Businessweek* also uses statistics and statistical information in managing its own business. For example, an annual survey of subscribers helps the company learn about subscriber demographics, reading habits, likely purchases, lifestyles, and so on. *Bloomberg Businessweek* managers use statistical summaries from the survey to provide better services to subscribers and advertisers. One recent North American subscriber



*Bloomberg Businessweek* uses statistical facts and summaries in many of its articles. © Photoshot.

survey indicated that 90% of *Bloomberg Businessweek* subscribers use a personal computer at home and that 64% of *Bloomberg Businessweek* subscribers are involved with computer purchases at work. Such statistics alert *Bloomberg Businessweek* managers to subscriber interest in articles about new developments in computers. The results of the subscriber survey are also made available to potential advertisers. The high percentage of subscribers using personal computers at home and the high percentage of subscribers involved with computer purchases at work would be an incentive for a computer manufacturer to consider advertising in *Bloomberg Businessweek*.

In this chapter, we discuss the types of data available for statistical analysis and describe how the data are obtained. We introduce descriptive statistics and statistical inference as ways of converting data into meaningful and easily interpreted statistical information.

\*The authors are indebted to Charlene Trentham, Research Manager, for providing this Statistics in Practice.

Frequently, we see the following types of statements in newspapers and magazines:

- In the first nine months of last year, Turkish Airlines' profit increased to about \$482 million on sales of \$6.2 billion (*Fortune*, February 25, 2013).
- Women account for 14.1% of the executive officers in Fortune 500 companies (*The Wall Street Journal*, April 30, 2012).

- The average annual cost for a college education is \$17,100 for public, in-state universities and \$38,600 for private universities (*Money Magazine*, March 2012).
- A Yahoo! Finance survey reported 51% of workers say the key to getting ahead is internal politics, whereas 27% say the key to getting ahead is hard work (*USA Today*, September 29, 2012).
- The California State Teachers' Retirement System has \$154.3 billion under management (*Bloomberg Businessweek*, January 21–January 27, 2013).
- At a Sotheby's art auction held on February 5, 2013, Pablo Picasso's painting *Woman Sitting Near a Window* sold for \$45 million (*The Wall Street Journal*, February 15, 2013).
- Over the past three months, the industry average for sales incentives per vehicle by GM, Chrysler, Ford, Toyota, and Honda was \$2336 (*The Wall Street Journal*, February 14, 2013).

The numerical facts in the preceding statements—\$482 million, \$6.2 billion, 14.1%, \$17,100, \$38,600, 51%, 27%, \$154.3 billion, \$45 million, \$2336—are called **statistics**. In this usage, the term statistics refers to numerical facts such as averages, medians, percentages, and maximums that help us understand a variety of business and economic situations. However, as you will see, the field, or subject, of statistics involves much more than numerical facts. In a broader sense, statistics is the art and science of collecting, analyzing, presenting, and interpreting data. Particularly in business and economics, the information provided by collecting, analyzing, presenting, and interpreting data gives managers and decision makers a better understanding of the business and economic environment and thus enables them to make more informed and better decisions. In this text, we emphasize the use of statistics for business and economic decision making.

Chapter 1 begins with some illustrations of the applications of statistics in business and economics. In Section 1.2 we define the term *data* and introduce the concept of a data set. This section also introduces key terms such as *variables* and *observations*, discusses the difference between quantitative and categorical data, and illustrates the uses of cross-sectional and time series data. Section 1.3 discusses how data can be obtained from existing sources or through survey and experimental studies designed to obtain new data. The important role that the Internet now plays in obtaining data is also highlighted. The uses of data in developing descriptive statistics and in making statistical inferences are described in Sections 1.4 and 1.5. The last three sections of Chapter 1 provide the role of the computer in statistical analysis, an introduction to data mining, and a discussion of ethical guidelines for statistical practice. A chapter-ending appendix includes an introduction to the add-in StatTools which can be used to extend the statistical options for users of Microsoft Excel.

## 1.1

# Applications in Business and Economics

In today's global business and economic environment, anyone can access vast amounts of statistical information. The most successful managers and decision makers understand the information and know how to use it effectively. In this section, we provide examples that illustrate some of the uses of statistics in business and economics.

## Accounting

Public accounting firms use statistical sampling procedures when conducting audits for their clients. For instance, suppose an accounting firm wants to determine whether the amount of accounts receivable shown on a client's balance sheet fairly represents the actual amount of accounts receivable. Usually the large number of individual accounts receivable

makes reviewing and validating every account too time-consuming and expensive. As common practice in such situations, the audit staff selects a subset of the accounts called a sample. After reviewing the accuracy of the sampled accounts, the auditors draw a conclusion as to whether the accounts receivable amount shown on the client's balance sheet is acceptable.

## Finance

Financial analysts use a variety of statistical information to guide their investment recommendations. In the case of stocks, analysts review financial data such as price/earnings ratios and dividend yields. By comparing the information for an individual stock with information about the stock market averages, an analyst can begin to draw a conclusion as to whether the stock is a good investment. For example, *The Wall Street Journal* (March 19, 2012) reported that the average dividend yield for the S&P 500 companies was 2.2%. Microsoft showed a dividend yield of 2.42%. In this case, the statistical information on dividend yield indicates a higher dividend yield for Microsoft than the average dividend yield for the S&P 500 companies. This and other information about Microsoft would help the analyst make an informed buy, sell, or hold recommendation for Microsoft stock.

## Marketing

Electronic scanners at retail checkout counters collect data for a variety of marketing research applications. For example, data suppliers such as ACNielsen and Information Resources, Inc., purchase point-of-sale scanner data from grocery stores, process the data, and then sell statistical summaries of the data to manufacturers. Manufacturers spend hundreds of thousands of dollars per product category to obtain this type of scanner data. Manufacturers also purchase data and statistical summaries on promotional activities such as special pricing and the use of in-store displays. Brand managers can review the scanner statistics and the promotional activity statistics to gain a better understanding of the relationship between promotional activities and sales. Such analyses often prove helpful in establishing future marketing strategies for the various products.

## Production

Today's emphasis on quality makes quality control an important application of statistics in production. A variety of statistical quality control charts are used to monitor the output of a production process. In particular, an  $x$ -bar chart can be used to monitor the average output. Suppose, for example, that a machine fills containers with 12 ounces of a soft drink. Periodically, a production worker selects a sample of containers and computes the average number of ounces in the sample. This average, or  $x$ -bar value, is plotted on an  $x$ -bar chart. A plotted value above the chart's upper control limit indicates overfilling, and a plotted value below the chart's lower control limit indicates underfilling. The process is termed "in control" and allowed to continue as long as the plotted  $x$ -bar values fall between the chart's upper and lower control limits. Properly interpreted, an  $x$ -bar chart can help determine when adjustments are necessary to correct a production process.

## Economics

Economists frequently provide forecasts about the future of the economy or some aspect of it. They use a variety of statistical information in making such forecasts. For instance, in forecasting inflation rates, economists use statistical information on such indicators as the Producer Price Index, the unemployment rate, and manufacturing capacity utilization. Often these statistical indicators are entered into computerized forecasting models that predict inflation rates.



## Information Systems

Information systems administrators are responsible for the day-to-day operation of an organization's computer networks. A variety of statistical information helps administrators assess the performance of computer networks, including local area networks (LANs), wide area networks (WANs), network segments, intranets, and other data communication systems. Statistics such as the mean number of users on the system, the proportion of time any component of the system is down, and the proportion of bandwidth utilized at various times of the day are examples of statistical information that help the system administrator better understand and manage the computer network.

Applications of statistics such as those described in this section are an integral part of this text. Such examples provide an overview of the breadth of statistical applications. To supplement these examples, practitioners in the fields of business and economics provided chapter-opening Statistics in Practice articles that introduce the material covered in each chapter. The Statistics in Practice applications show the importance of statistics in a wide variety of business and economic situations.

### 1.2

## Data

**Data** are the facts and figures collected, analyzed, and summarized for presentation and interpretation. All the data collected in a particular study are referred to as the **data set** for the study. Table 1.1 shows a data set containing information for 60 nations that participate in the World Trade Organization. The World Trade Organization encourages the free flow of international trade and provides a forum for resolving trade dispute.

### Elements, Variables, and Observations

**Elements** are the entities on which data are collected. Each nation listed in Table 1.1 is an element with the nation or element name shown in the first column. With 60 nations, the data set contains 60 elements.

A **variable** is a characteristic of interest for the elements. The data set in Table 1.1 includes the following five variables:

- **WTO Status:** The nation's membership status in the World Trade Organization; this can be either as a member or an observer.
- **Per Capita GDP (\$):** The total market value (\$) of all goods and services produced by the nation divided by the number of people in the nation; this is commonly used to compare economic productivity of the nations.
- **Trade Deficit (\$1000s):** The difference between the total dollar value of the nation's imports and the total dollar value of the nation's exports.
- **Fitch Rating:** The nation's sovereign credit rating as appraised by the Fitch Group<sup>1</sup>; the credit ratings range from a high of AAA to a low of F and can be modified by + or -.
- **Fitch Outlook:** An indication of the direction the credit rating is likely to move over the upcoming two years; the outlook can be negative, stable, or positive.

Measurements collected on each variable for every element in a study provide the data. The set of measurements obtained for a particular element is called an **observation**. Referring to Table 1.1, we see that the first observation contains the following measurements:

---

<sup>1</sup>The Fitch Group is one of three nationally recognized statistical rating organizations designated by the U.S. Securities and Exchange Commission. The other two are Standard and Poor's and Moody's investor service.

**TABLE 1.1** DATA SET FOR 60 NATIONS IN THE WORLD TRADE ORGANIZATION

Nation	WTO Status	Per Capita GDP (\$)	Trade Deficit (\$1000s)	Fitch Rating	Fitch Outlook
Armenia	Member	5,400	2,673,359	BB-	Stable
Australia	Member	40,800	-33,304,157	AAA	Stable
Austria	Member	41,700	12,796,558	AAA	Stable
Azerbaijan	Observer	5,400	-16,747,320	BBB-	Positive
Bahrain	Member	27,300	3,102,665	BBB	Stable
Belgium	Member	37,600	-14,930,833	AA+	Negative
Brazil	Member	11,600	-29,796,166	BBB	Stable
Bulgaria	Member	13,500	4,049,237	BBB-	Positive
Canada	Member	40,300	-1,611,380	AAA	Stable
Cape Verde	Member	4,000	874,459	B+	Stable
Chile	Member	16,100	-14,558,218	A+	Stable
China	Member	8,400	-156,705,311	A+	Stable
Colombia	Member	10,100	-1,561,199	BBB-	Stable
Costa Rica	Member	11,500	5,807,509	BB+	Stable
Croatia	Member	18,300	8,108,103	BBB-	Negative
Cyprus	Member	29,100	6,623,337	BBB	Negative
Czech Republic	Member	25,900	-10,749,467	A+	Positive
Denmark	Member	40,200	-15,057,343	AAA	Stable
Ecuador	Member	8,300	1,993,819	B-	Stable
Egypt	Member	6,500	28,486,933	BB	Negative
El Salvador	Member	7,600	5,019,363	BB	Stable
Estonia	Member	20,200	802,234	A+	Stable
France	Member	35,000	118,841,542	AAA	Stable
Georgia	Member	5,400	4,398,153	B+	Positive
Germany	Member	37,900	-213,367,685	AAA	Stable
Hungary	Member	19,600	-9,421,301	BBB-	Negative
Iceland	Member	38,000	-504,939	BB+	Stable
Ireland	Member	39,500	-59,093,323	BBB+	Negative
Israel	Member	31,000	6,722,291	A	Stable
Italy	Member	30,100	33,568,668	A+	Negative
Japan	Member	34,300	31,675,424	AA	Negative
Kazakhstan	Observer	13,000	-33,220,437	BBB	Positive
Kenya	Member	1,700	9,174,198	B+	Stable
Latvia	Member	15,400	2,448,053	BBB-	Positive
Lebanon	Observer	15,600	13,715,550	B	Stable
Lithuania	Member	18,700	3,359,641	BBB	Positive
Malaysia	Member	15,600	-39,420,064	A-	Stable
Mexico	Member	15,100	1,288,112	BBB	Stable
Peru	Member	10,000	-7,888,993	BBB	Stable
Philippines	Member	4,100	15,667,209	BB+	Stable
Poland	Member	20,100	19,552,976	A-	Stable
Portugal	Member	23,200	21,060,508	BBB-	Negative
South Korea	Member	31,700	-37,509,141	A+	Stable
Romania	Member	12,300	13,323,709	BBB-	Stable
Russia	Observer	16,700	-151,400,000	BBB	Positive
Rwanda	Member	1,300	939,222	B	Stable
Serbia	Observer	10,700	8,275,693	BB-	Stable
Seychelles	Observer	24,700	666,026	B	Stable
Singapore	Member	59,900	-27,110,421	AAA	Stable
Slovakia	Member	23,400	-2,110,626	A+	Stable
Slovenia	Member	29,100	2,310,617	AA-	Negative

**WEB file****Nations**

Data sets such as Nations are available on the website for this text.

South Africa	Member	11,000	3,321,801	BBB+	Stable
Sweden	Member	40,600	-10,903,251	AAA	Stable
Switzerland	Member	43,400	-27,197,873	AAA	Stable
Thailand	Member	9,700	2,049,669	BBB	Stable
Turkey	Member	14,600	71,612,947	BB+	Positive
UK	Member	35,900	162,316,831	AAA	Negative
Uruguay	Member	15,400	2,662,628	BB	Positive
USA	Member	48,100	784,438,559	AAA	Stable
Zambia	Member	1,600	-1,805,198	B+	Stable

Member, 5,400, 2,673,359, BB-, and Stable. The second observation contains the following measurements: Member, 40,800, -33,304,157, AAA, Stable, and so on. A data set with 60 elements contains 60 observations.

## Scales of Measurement

Data collection requires one of the following scales of measurement: nominal, ordinal, interval, or ratio. The scale of measurement determines the amount of information contained in the data and indicates the most appropriate data summarization and statistical analyses.

When the data for a variable consist of labels or names used to identify an attribute of the element, the scale of measurement is considered a **nominal scale**. For example, referring to the data in Table 1.1, the scale of measurement for the WTO Status variable is nominal because the data “member” and “observer” are labels used to identify the status category for the nation. In cases where the scale of measurement is nominal, a numerical code as well as a nonnumerical label may be used. For example, to facilitate data collection and to prepare the data for entry into a computer database, we might use a numerical code for the WTO Status variable by letting 1 denote a member nation in the World Trade Organization and 2 denote an observer nation. The scale of measurement is nominal even though the data appear as numerical values.

The scale of measurement for a variable is considered an **ordinal scale** if the data exhibit the properties of nominal data and in addition, the order or rank of the data is meaningful. For example, referring to the data in Table 1.1, the scale of measurement for the Fitch Rating is ordinal because the rating labels which range from AAA to F can be rank ordered from best credit rating AAA to poorest credit rating F. The rating letters provide the labels similar to nominal data, but in addition, the data can also be ranked or ordered based on the credit rating, which makes the measurement scale ordinal. Ordinal data can also be recorded by a numerical code, for example, your class rank in school.

The scale of measurement for a variable is an **interval scale** if the data have all the properties of ordinal data and the interval between values is expressed in terms of a fixed unit of measure. Interval data are always numeric. College admission SAT scores are an example of interval-scaled data. For example, three students with SAT math scores of 620, 550, and 470 can be ranked or ordered in terms of best performance to poorest performance in math. In addition, the differences between the scores are meaningful. For instance, student 1 scored  $620 - 550 = 70$  points more than student 2, while student 2 scored  $550 - 470 = 80$  points more than student 3.

The scale of measurement for a variable is a **ratio scale** if the data have all the properties of interval data and the ratio of two values is meaningful. Variables such as distance, height, weight, and time use the ratio scale of measurement. This scale requires that a zero value be included to indicate that nothing exists for the variable at the zero point. For

example, consider the cost of an automobile. A zero value for the cost would indicate that the automobile has no cost and is free. In addition, if we compare the cost of \$30,000 for one automobile to the cost of \$15,000 for a second automobile, the ratio property shows that the first automobile is  $\$30,000/\$15,000 = 2$  times, or twice, the cost of the second automobile.

## Categorical and Quantitative Data

Data can be classified as either categorical or quantitative. Data that can be grouped by specific categories are referred to as **categorical data**. Categorical data use either the nominal or ordinal scale of measurement. Data that use numeric values to indicate how much or how many are referred to as **quantitative data**. Quantitative data are obtained using either the interval or ratio scale of measurement.

*The statistical method appropriate for summarizing data depends upon whether the data are categorical or quantitative.*

A **categorical variable** is a variable with categorical data, and a **quantitative variable** is a variable with quantitative data. The statistical analysis appropriate for a particular variable depends upon whether the variable is categorical or quantitative. If the variable is categorical, the statistical analysis is limited. We can summarize categorical data by counting the number of observations in each category or by computing the proportion of the observations in each category. However, even when the categorical data are identified by a numerical code, arithmetic operations such as addition, subtraction, multiplication, and division do not provide meaningful results. Section 2.1 discusses ways for summarizing categorical data.

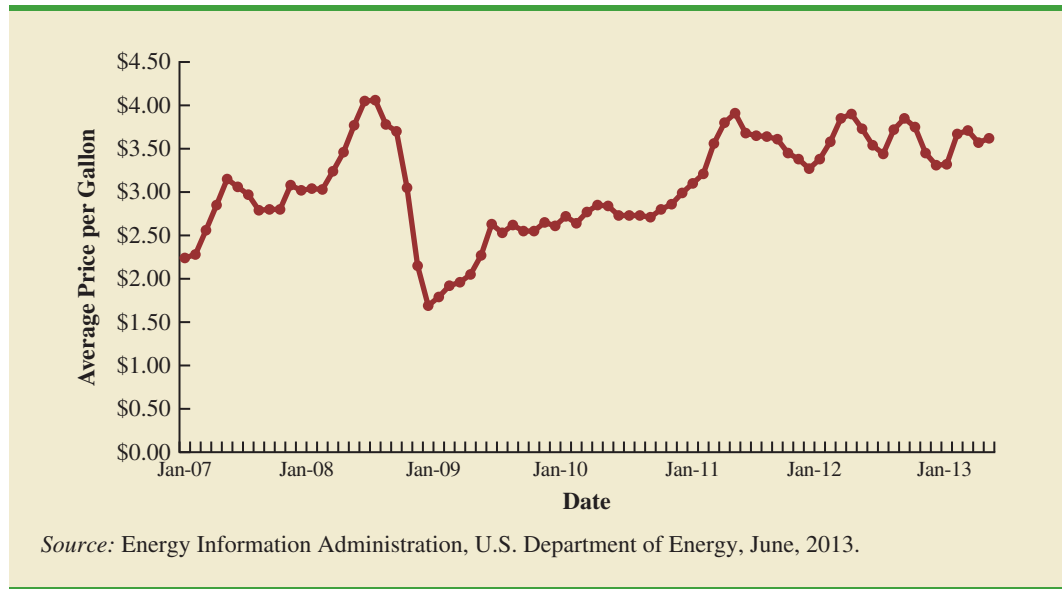
Arithmetic operations provide meaningful results for quantitative variables. For example, quantitative data may be added and then divided by the number of observations to compute the average value. This average is usually meaningful and easily interpreted. In general, more alternatives for statistical analysis are possible when data are quantitative. Section 2.2 and Chapter 3 provide ways of summarizing quantitative data.

## Cross-Sectional and Time Series Data

For purposes of statistical analysis, distinguishing between cross-sectional data and time series data is important. **Cross-sectional data** are data collected at the same or approximately the same point in time. The data in Table 1.1 are cross-sectional because they describe the five variables for the 60 World Trade Organization nations at the same point in time. **Time series data** are data collected over several time periods. For example, the time series in Figure 1.1 shows the U.S. average price per gallon of conventional regular gasoline between 2007 and 2013. Note that gasoline prices peaked in the summer of 2008 and then dropped sharply in the fall of 2008. Between January, 2009 and May, 2011, the average price per gallon continued to climb steadily. Since then prices have shown more fluctuation, reaching an average price per gallon of \$3.62 in May, 2013.

Graphs of time series data are frequently found in business and economic publications. Such graphs help analysts understand what happened in the past, identify any trends over time, and project future values for the time series. The graphs of time series data can take on a variety of forms, as shown in Figure 1.2. With a little study, these graphs are usually easy to understand and interpret. For example, Panel (A) in Figure 1.2 is a graph that shows the Dow Jones Industrial Average Index from 2002 to 2013. In April 2002, the popular stock market index was near 10,000. Over the next five years the index rose to slightly over 14,000 in October 2007. However, notice the sharp decline in the time series after the high in 2007. By March 2009, poor economic conditions had caused the Dow Jones Industrial Average Index to return to the 7000 level. This was a scary and discouraging period for investors. However, by late 2009, the index was showing a recovery by reaching 10,000. The index has climbed steadily since then and was above 15,000 in early 2013.

**FIGURE 1.1** U.S. AVERAGE PRICE PER GALLON FOR CONVENTIONAL REGULAR GASOLINE



© Cengage Learning

The graph in Panel (B) shows the net income of McDonald's Inc. from 2005 to 2012. The declining economic conditions in 2008 and 2009 were actually beneficial to McDonald's as the company's net income rose to all-time highs. The growth in McDonald's net income showed that the company was thriving during the economic downturn as people were cutting back on the more expensive sit-down restaurants and seeking less-expensive alternatives offered by McDonald's. McDonald's net income continued to new all-time highs in 2010 and 2011, but decreased slightly in 2012.

Panel (C) shows the time series for the occupancy rate of hotels in South Florida over a one-year period. The highest occupancy rates, 95% and 98%, occur during the months of February and March when the climate of South Florida is attractive to tourists. In fact, January to April of each year is typically the high-occupancy season for South Florida hotels. On the other hand, note the low occupancy rates during the months of August to October, with the lowest occupancy rate of 50% occurring in September. High temperatures and the hurricane season are the primary reasons for the drop in hotel occupancy during this period.

## NOTES AND COMMENTS

1. An observation is the set of measurements obtained for each element in a data set. Hence, the number of observations is always the same as the number of elements. The number of measurements obtained for each element equals the number of variables. Hence, the total number of data items can be determined by multiplying the number of observations by the number of variables.
2. Quantitative data may be discrete or continuous. Quantitative data that measure how many (e.g., number of calls received in 5 minutes) are discrete. Quantitative data that measure how much (e.g., weight or time) are continuous because no separation occurs between the possible data values.