

Notes and Data Sources

Chapter 1

1. [visualcapitalist.com/video-ico-explosion-one-
-animated-timeline](http://visualcapitalist.com/video-ico-explosion-one-animated-timeline)
2. census.gov
3. [creditcards.com/credit-card-news/change-favorite-
-credit-card.php](http://creditcards.com/credit-card-news/change-favorite-credit-card.php)
4. This example is used in a template for creating Pareto charts in Excel. You can download the template from [office.microsoft.com/en-us/templates/
/cost-analysis-with-pareto-chart-TC006082757.aspx](http://office.microsoft.com/en-us/templates/cost-analysis-with-pareto-chart-TC006082757.aspx)
5. Pareto charts are named for the Italian economist Vilfredo Pareto (1848–1923). Pareto was one of the first to analyze economic problems with mathematical tools. The Pareto Principle (sometimes called the 80/20 rule) takes various forms, such as “80% of the work is done by 20% of the people.” Pareto charts are a graphical version of the principle. The chart identifies the few important categories (the 20%) that account for most of the responses (the 80%). Of course, in any given setting, the actual percents will vary.
6. Estimates from Statistics Canada; statcan.gc.ca
7. Federal Reserve Bank of St. Louis; fred.stlouisfed.org/series/WTB6MS
8. Our eyes do respond to area, but not quite linearly. It appears that we perceive the ratio of two bars to be about the 0.7 power of the ratio of their actual areas. See William S. Cleveland, *The Elements of Graphing Data*, Wadsworth, 1985, pp. 278–284.
9. Haipeng Shen, “Nonparametric regression for problems involving lognormal distributions,” PhD thesis, University of Pennsylvania, 2003. Thanks to Haipeng Shen and Larry Brown for sharing the data.
10. See note 7.
11. Table 1 in U.S. Environmental Protection Agency, *Advancing Sustainable Materials Management: 2014 Fact Sheet*, November 2016.
12. April 2017 to March 2018 data from marketshare.hitslink.com
13. See, for example, facebook.com/milliondollarlisting
14. July 2017 data from [thenextweb.com/contributors/
/2017/07/13/india-overtakes-usa-become-facebooks-
-top-country](http://thenextweb.com/contributors/2017/07/13/india-overtakes-usa-become-facebooks-top-country)
15. Downloaded from the U.S. Census Bureau website, May 2018; census.gov
16. Data for 2017 based on the Canadian Census; statcan.gc.ca
17. Rankings for 2017 from the World Bank website; doingbusiness.org/rankings
18. U.S. Census Bureau website, May 2018; census.gov
19. World Bank website; wdi.worldbank.org/table/4.1
20. World Bank website; wdi.worldbank.org/table/4.2
21. forbes.com/best-countries-for-business/list, May 2018.
22. *Forbes* website; forbes.com/powerful-brands
23. See, for example, the bibliographic entry for Gosset in the School of Mathematics and Statistics of the University of St. Andrews, Scotland, MacTutor History of Mathematics archive; history.mcs.st-andrews.ac.uk/Biographies/Gosset.html
24. These and other data that were collected and used by Gosset can be found in the Guinness Archives in Dublin; guinness-storehouse.com/en/Archive.aspx
25. Downloaded from beer100.com/calories_in_beer.htm, on May 8, 2018.
26. census.gov/programs-surveys/asm.html
27. Information about the National Assessment of Educational Progress program can be found at nationsreportcard.gov
28. Some software calls these graphs *Normal probability plots*. There is a technical distinction between the two types of graphs, but the terms are often used loosely.
29. The idea that all distributions are normal in the middle is attributed to Charlie Winsor. See John W. Tukey, “A survey of sampling from contaminated distributions,” in Ingram Olkin, Sudhish G. Ghurye, Wassily Hoeffding, William G. Madow, and Henry B. Mann, eds., *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, Stanford University Press, 1960 pp. 448–485.
30. Matthias R. Mehl et al. “Are women really more talkative than men?” *Science*, 317, no. 5834 (2007), p. 82.
31. [careerbuilder.com/jobs/keyword/business-
-administration](http://careerbuilder.com/jobs/keyword/business-administration)
32. [online.wsj.com/articles/the-world-rankings-
-of-flopping-1403660175](http://online.wsj.com/articles/the-world-rankings-of-flopping-1403660175)

33. World Bank website; data.worldbank.org/indicator/CM.MKT.LDOM.NO
34. Color popularity retrieved from the Kelley Blue Book, kbb.com/car-advice/articles/best-color-to-buy/, May 20, 2018.

Chapter 2

1. Data for 2014 from usgovernmentspending.com/compare_state_education_spend
2. Data for 2018 from nrcan.gc.ca/energy/efficiency/transportation/21002
3. A sophisticated treatment of improvements and additions to scatterplots is described by William S. Cleveland and Robert McGill, "The many faces of a scatterplot," *Journal of the American Statistical Association*, 79 (1984), pp. 807–822.
4. data.worldbank.org/indicator/CM.MKT.LDOM.NO
5. beer100.com
6. www12.statcan.ca
7. spectrumtechniques.com/isotope_generator.htm
8. These data were collected under the supervision of Zach Grigsby, Science Express Coordinator, College of Science, Purdue University.
9. A careful study of this phenomenon is reported by William S. Cleveland, Persi Diaconis, and Robert McGill, "Variables on scatterplots look more highly correlated when the scales are increased," *Science*, 216 (1982), pp. 1138–1141.
10. *The Inclusive Development Index 2018 Summary and Data Highlights*, World Economic Forum, 2009; weforum.org
11. From a presentation by Charles Knauf, Monroe County (New York) Environmental Health Laboratory.
12. Frank J. Anscombe, "Graphs in statistical analysis," *American Statistician*, 27 (1973), pp. 17–21.
13. target.com/site/en/corporate
14. See, for example, ibm.com/analytics/hadoop/big-data-analytics and sas.com/en_us/insights/big-data/what-is-big-data.html
15. See, for example, niehs.nih.gov/health/topics/agents/emf
16. C. M. Ryan, Christine A. Northrup-Clewes, Barbara Knox, and David I. Thurnham, "The effect of in-store music on consumer choice of wine," *Proceedings of the Nutrition Society*, 57 (1998), p. 1069A.
17. *Education Indicators: An International Perspective*, Institute of Education Studies, National Center for Education Statistics; nces.ed.gov/surveys/international
18. pewinternet.org/2018/05/31/teens-social-media-technology-2018
19. Based on *The Ethics of American Youth—2012*, available from the Josephson Institute; charactercounts.org/programs/reportcard
20. M.-Y. Chen et al., "Adequate sleep among adolescents is positively associated with health status and health-related behaviors," *BMC Public Health*, 6, no. 59 (2006); biomedicalcentral.com/1471-2458/6/59
21. See the U.S. Bureau of Census website at census.gov for these and similar data.
22. From the 2013–14 edition of the Purdue University Data Digest; purdue.edu/datadigest
23. From the *2012 Statistical Abstract of the United States*; census.gov/compendia/statab/cats/population.html
24. See note 3.
25. *OECD StatExtracts*, Organization for Economic Cooperation Development, stats.oecd.org/wbos
26. Information about this procedure was provided by Samuel Flanigan of *U.S. News & World Report*. See usnews.com/usnews/rankguide/rghome.htm for a description of the variables used to construct the ranks and for the most recent ranks.
27. Based on data provided by Professor Michael Hunt and graduate student James Bateman of the Purdue University Department of Forestry and Natural Resources.
28. Reported in *The New York Times*, July 20, 1989, from an article appearing that day in the *New England Journal of Medicine*.
29. Condensed from David R. Appleton, Joyce M. French, and Mark P. J. Vanderpump, "Ignoring a covariate: An example of Simpson's paradox," *American Statistician*, 50 (1996), pp. 340–341.
30. Lien-Ti Bei, *Consumers' Purchase Behavior toward Recycled Products: An Acquisition-Transaction Utility Theory Perspective*, MS thesis, Purdue University, 1993.

Chapter 3

1. www.bls.gov, May 21, 2018.
2. See, for example, mathsreports.wordpress.com/overall-narrative/mathematics-is-important
3. nationsreportcard.gov
4. See the NORC web pages at norc.uchicago.edu
5. statista.com
6. project-redcap.org
7. "Did you know," *Consumer Reports*, February 2013, p. 10.
8. Bruce Barrett et al., "Echinacea for treating the common cold," *Annals of Internal Medicine*, 153 (2010), pp. 769–777.

9. Based on a study conducted by Tammy Younts and directed by Professor Deb Bennett of the Purdue University Department of Educational Studies. For more information about Reading Recovery, see readingrecovery.org
 10. Based on a study conducted by Rajendra Chaini under the direction of Professor Bill Hoover of the Purdue University Department of Forestry and Natural Resources.
 11. See the Harvard Business Review Blog Network entry, blogs.hbr.org/2013/04/the-hidden-biases-in-big-data
 12. sm.rutgers.edu/pubs/Grinberg-SMPatterns-ICWSM2013.pdf
 13. census.gov/programs-surveys/asm.html
 14. *Billboard's* Hot Ringtones list, billboard.com, May 23, 2018.
 15. *Billboard's* The Hot 100, billboard.com, May 23, 2018.
 16. From the online version of the Bureau of Labor Statistics, *Handbook of Methods*, bls.gov. The details of the design are more complicated than the text describes.
 17. The nonresponse rate for the CPS can be found at census.gov/programs-surveys/cps/technical-documentation/methodology/non-response-rates.html
 18. The Pew Research Center for People and the Press designs careful surveys and is an excellent source of information about nonresponse. See pewresearch.org/2017/05/15/what-low-response-rates-mean-for-telephone-surveys
 19. 50states.com/areacodes
 20. For a full description of the STAR program and its follow-up studies, see www.heros-inc.org/star.htm
 21. Simplified from Arno J. Rethans, John L. Swasy, and Lawrence J. Marks, "Effects of television commercial repetition, receiver knowledge, and commercial length: A test of the two-factor model," *Journal of Marketing Research* 23 (February 1986), pp. 50–61.
 22. Based on an experiment performed by Jake Gandolph under the direction of Professor Lisa Mauer in the Purdue University Department of Food Science.
 23. Based on an experiment performed by Evan Whalen under the direction of Professor Patrick Connolly in the Purdue University Department of Computer Graphics Technology.
 24. Based on a study conducted by Brent Ladd, a water quality specialist with the Purdue University Department of Agricultural and Biological Engineering.
 25. Based on a study conducted by Sandra Simonis under the direction of Professor Jon Harbor from the Purdue University Earth and Atmospheric Sciences Department.
 26. John C. Bailar III, "The real threats to the integrity of science," *Chronicle of Higher Education* (April 21, 1995), pp. B1–B2.
 27. See the details on the website of the Office for Human Research Protections of the Department of Health and Human Services, hhs.gov/ohrp
 28. The difficulties of interpreting guidelines for informed consent and for the work of institutional review boards in medical research are a main theme of Beverly Woodward, "Challenges to human subject protections in U.S. medical research," *Journal of the American Medical Association* 282 (1999), pp. 1947–1952. The references in this paper point to other discussions. Updated regulations and guidelines appear on the OHRP website (see note 2).
 29. Quotation from the *Report of the Tuskegee Syphilis Study Legacy Committee*, May 20, 1996. A detailed history is provided by James H. Jones, *Bad Blood: The Tuskegee Syphilis Experiment*, Free Press, 1993.
 30. Dr. Hennekens's words are from an interview in the Annenberg/Corporation for Public Broadcasting video series *Against All Odds: Inside Statistics*.
 31. ftc.gov/opa/2009/04/kellogg.shtm
 32. findarticles.com/p/articles/mi_m0CYD/is_8_40/ai_n13675065
 33. R. Dennis Middlemist, Eric Knowles, and Charles Matter, "Personal space invasions in the lavatory: Suggestive evidence for arousal," *Journal of Personality and Social Psychology* 33 (1976), pp. 541–546.
 34. For a review of domestic violence experiments, see C. D. Maxwell et al., *The Effects of Arrest on Intimate Partner Violence: New Evidence from the Spouse Assault Replication Program*, U.S. Department of Justice, NCH188199, 2001, ojp.usdoj.gov/nij/pubs-sum/188199.htm
 35. See the Federal Trade Commission website, ftc.gov, for more information about online behavioral advertising.
- ### Chapter 4
1. "The AMA Gold Report 2017 Top 50 Market Research Firms," ama.org
 2. Closing price data are available from several sources, including finance.yahoo.com
 3. Global Automotive 2017 Color Popularity Report, axaltacs.com/corporate/en_US.html
 4. The full 2016 Canadian Medical Association report, 16th Annual National Report Card on Health Care, cma.ca

5. Association of Certified Fraud Examiners, 2016 Report to the Nations on Occupational Fraud and Abuse, acfe.com
6. U.S. Energy Information Administration, Annual Energy Outlook 2018, eia.gov
7. Employment drug testing statistics are reported at pre-employment.com
8. “Liars index shows almost one in four résumés falsified education: Lyin’ king chosen,” pre-employment.com
9. 2017 Canadians and Digital Health Survey, conducted by Ipsos, ipsos.com
10. The Gallup Organization, *Confidence in Institutions, 2017*, news.gallup.com
11. Based on 2016 census data from the website of Statistics Canada, www12.statcan.gc.ca
12. 2017 Housing Survey, Toronto Community Housing, torontohousing.ca
13. Internet usage statistics, internetworldstats.com
14. Canadian transportation statistics from Statistics Canada, statcan.gc.ca. U.S. transportation statistics from U.S. Bureau of Transportation Statistics, bts.gov
15. See Note 3.
16. www.lendingclub.com/info/download-data.action
17. 2017 Ofcom study, “Adults media use and attitudes,” ofcom.org.uk
18. “NFL took in \$13 billion in revenue last season—see how it stacks up against other pro sports leagues,” July 2, 2016, moneywatch.com
19. Estimated probabilities from the National Collegiate Athletic Association (NCAA), see ncaa.org
20. See Note 19.
21. www.hsbaseballweb.com/probability.htm
22. IRS Tax Statistics, see www.irs.gov/statistics
23. Pew Research Center, “Automation in everyday life,” October 4, 2017, pewinternet.org
24. 2018 Nielsen survey, “A marketers guide to podcasting,” nielsen.com
25. U.S. Census Bureau, census.gov

Chapter 5

1. We use \bar{x} both for the random variable, which takes different values in repeated sampling, and for the numerical value of the random variable in a particular sample. Similarly, s stands both for a random variable and for a specific value. This notation is mathematically imprecise but statistically convenient.
2. Based on a study by Layth Alwan of data obtained from a Midwest blood center.

3. The statement that the mean is at the center of symmetric density curves is typically the case. However, there are usual symmetric density curves (e.g., Cauchy) for which the mean does not exist.
4. The mean of a continuous random variable X with density function $f(x)$ is defined by the following integral:

$$\mu_X = \int xf(x)dx$$

This integral is a kind of weighted average, analogous to the discrete-case mean:

$$\mu_X = \sum xP(X = x)$$

The variance of a continuous random variable X is the average squared deviation of the values of X from their mean and is defined by the following integral:

$$\sigma_X^2 = \int (x - \mu)^2 f(x)dx$$

5. See Amos Tversky and Daniel Kahneman, “Belief in the law of small numbers,” *Psychological Bulletin* 76, no. 2 (1971), pp. 105–110; and other works by these authors for a full account of our misperception of randomness.
6. Probabilities involving runs can be quite difficult to compute. That the probability of a run of three or more heads in 10 independent tosses of a fair coin is $(1/2) + (1/128) = 0.508$ can be found by clever counting, as can the other results given in the text. A general treatment using advanced methods appears in Section XIII.7 of William Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1, 3rd ed., Wiley, 1968.
7. Thomas Gilovich et al., “The hot hand in basketball: On the misperception of random sequences,” *Cognitive Psychology* 17, no. 3 (1985), pp. 295–314. A later series of articles that debate the independence question includes Amos Tversky and Thomas Gilovich, “The cold facts about the ‘hot hand’ in basketball,” *Chance* 2, no. 1 (1989), pp. 16–21; Patrick Larkey et al., “It’s OK to believe in the ‘hot hand,’” *Chance* 2, no. 4 (1989), pp. 22–30; and Amos Tversky and Thomas Gilovich, “The ‘hot hand’: Statistical reality or cognitive illusion?” *Chance* 2, no. 4 (1989), pp. 31–34.
8. As an example, the Charles Schwab online website (schwab.com) provides mean returns and standard deviations of returns for all its managed mutual funds under Investment Help.
9. Closing price data are available from several sources, including finance.yahoo.com
10. See Note 9.
11. Kevin Hammermeister et al., “Outcomes 15 years after valve replacement with a mechanical

vs. a prosthetic valve: Final report of the Veterans Administration randomized trial," *Journal of American College of Cardiology* 36, no. 4 (2000), pp. 1152–1158.

12. about.usps.com/what-we-are-doing/service-performance/welcome.htm
13. www.football-data.co.uk/englandm.php
14. Data provided by Professor Maria Goranova of the University of Wisconsin–Milwaukee.
15. Bank of America. "Trends in consumer mobility report 2017," newsroom.bankofamerica.com
16. Statistics from smartasset.com/mortgage/best-cities-for-public-transportation
17. Cody Cook et al., "The gender earnings gap in the gig economy: Evidence from over a million ride-share drivers," Stanford Graduate School of Business Working Paper (2018).
18. Brent Bowen and Dean Headley, "Airline Quality Rating 2017," commons.erau.edu
19. City of Chicago performance metrics found at chicago.gov/city/en/narr/foia/key_performance_indicators0.html
20. From 2017 EDUCAUSE report, "Student and faculty technology research studies," library.educause.edu
10. Nielson, "Mobile acts as a nerve center of the Canadian shopping experience," nielsen.com
11. Accenture. "Time to navigate the super myway," accenture.com
12. EDUCAUSE, "Student and faculty technology research studies," 2017, library.educause.edu
13. National Center for Education Statistics, "The condition of education 2018," 2018, nces.ed.gov
14. "Blurred vision, burning eyes: This is a Lasik success?" *The New York Times*, June 11, 2018.
15. academicintegrity.org/statistics
16. A summary of Larry Wright's study can be found at www.nytimes.com/2009/03/04/sports/basketball/04freethrow.html
17. Common Sense Media, "Plugged-in parents of tweens and teens," 2017, commonsensemedia.org
18. Incentive Research Foundation, "Gift card study," 2018, theirf.org/research
19. American Public Transportation Association, "Who rides public transportation," 2017, apta.com
20. Deloitte, "Global mobile consumer survey," 2017, deloitte.com

Chapter 6

1. nrf.com/resources/consumer-research-and-data
2. Results from "2017 U.S. workplace bullying survey," workplacebullying.org
3. Andrew J. K. Phillips et al., "Irregular sleep/wake patterns are associated with poorer academic performance and delayed circadian and sleep/wake timing," *Scientific Reports* 7 (2017), www.nature.com/srep
4. Tim Althoff et al., "Large-scale physical activity data reveal worldwide activity inequality," *Nature* 547 (2017), pp. 1152–1158.
5. Citibike system data website, citibikenyc.com/system-data
6. The exact distribution of the average of n exponential random variables with rate parameter λ is a gamma distribution with shape parameter n and rate parameter $n\lambda$.
7. USC Center for the Digital Future, "Surveying the digital future," 2017, digitalcenter.org
8. Steve Mazie, "Do you have too many Facebook friends?" *Big Think*, 2017, bigthink.com/experts/stevenmazie
9. From the grade distribution database of the Indiana University Office of the Registrar, gradedistribution.registrar.indiana.edu

Chapter 7

1. Stock volumes and prices (including those of Verizon Communications) can be found at finance.yahoo.com
2. The 2015–2016 statistics for California were obtained from the California Department of Education website, www.cde.ca.gov
3. Pamela Foohey et al., "No money down bankruptcy," *Southern California Law Review* 90 (2017), pp. 1055–1110. Standard deviations in the study provided by Robert Lawless, University of Illinois College of Law.
4. Average starting salary taken from the Class of 2018 Salary Survey by the National Association of Colleges and Employers, naceweb.org
5. See thekaraokechannel.com
6. Average starting salaries for different business majors for students from the University of Texas at Austin are found at www.mcombs.utexas.edu/Career-Services/Statistics
7. The vehicle is a 2006 Toyota Highlander Hybrid owned by the first author.
8. Data obtained from the Philippine Statistics Authority, psa.gov.ph
9. Information reported in "State of American well-being: 2017 well-being rankings," at wellbeingindex.sharecare.com

10. Iqbal T. Hawaldar et al., "Pricing and performance of IPOs: Evidence from Indian stock market," *Cogent Economics & Finance* 6, no. 1, (2018), pp. 1–20.
11. See February 2015 New York State Attorney General's press release regarding sending cease and desist letters to GNC, Target, Walgreens, and Walmart regarding contaminants contained in store brand supplements at ag.ny.gov
12. From a study conducted by the Division of Financial Aid Purdue University.
13. Based on a 2017 report of the Office of Institutional Research and Assessment of Peru State College, "Campus climate, safety, and sexual assault survey: Students," found at www.peru.edu/oira
14. Ronald A. Fisher, "The arrangement of field experiments," *Journal of the Ministry of Agriculture of Great Britain* 33 (1926), p. 504, quoted in Leonard J. Savage, "On rereading R. A. Fisher," *Annals of Statistics* 4 (1976), p. 471. Fisher's work is described in a biography by his daughter: Joan Fisher Box, *R. A. Fisher: The Life of a Scientist*, Wiley, 1978.
15. David J. Benjamin et al., "Redefine statistical significance," *Nature* 2 (2018), pp. 6–10.
16. See *Scientific American* blog, "5 sigma: What's that?" July 17, 2012, at blogs.scientificamerican.com/observations/five-sigmawhats-that
17. Warren Leary, "Cell phones: Questions but no answers," *New York Times*, October 26, 1999.
18. "Pop-Tarts before a Hurricane," *Country Living*, August, 25, 2017, at countryliving.com/food-drinks
19. Data from city.milwaukee.gov/Property-Information.htm
20. Game log statistics on NBA players can be found at stats.nba.com
21. 2018 Census Bureau Training Presentation titled "Using ACS estimates and margins of error," found at census.gov
22. From "Health insurance coverage in the United States: 2016," found at census.gov
23. From the CMO Survey "Top Line Results," February 2018, found at cmosurvey.org
24. From "2017 survey of prosthodontists," found at prosthodontics.org
25. Barchi Gillai et al., "The relationship between responsible supply chain practices and performance," *Insights from the Stanford Initiative for the Study of Supply Chain Responsibility (SISSCR)*, November 2013.
26. Data provided by Mugdha Gore and Joseph Thomas, Purdue University School of Pharmacy.
27. Annual Kaiser Family Foundation (Kaiser) and the Health Research & Educational Trust (HRET) health survey reports found at kff.org/health-costs

Chapter 8

1. Information regarding their test can be found at www.consumerreports.org/smartphones/smartphones-with-the-best-battery-life
2. Data are based on C. Don Wiggins, "The legal perils of 'underdiversification': A case study," *Personal Financial Planning* 1, no. 6 (1999), pp. 16–18.
3. Data provided by Bill Berezowitz and James Malloy of GE Healthcare.
4. Shashi K. Pankaj et al., "Effects of cold plasma on food quality: A review," *Foods* 7, no. 1 (2018), Article 4, doi.org/10.3390/foods7010004
5. These recommendations are based on extensive computer work. See, for example, Harry O. Posten, "The robustness of the one-sample *t*-test over the Pearson system," *Journal of Statistical Computation and Simulation* 9 (1979), pp. 133–149 and E. S. Pearson and N. W. Please, "Relation between the shape of population distribution and the robustness of four simple test statistics," *Biometrika* 62 (1975), pp. 223–241.
6. The standard reference here is Bradley Efron and Robert J. Tibshirani, *An Introduction to the Bootstrap*, Chapman Hall, 1993. A less technical overview is in Bradley Efron and Robert J. Tibshirani, "Statistical data analysis in the computer age," *Science* 253 (1991), pp. 390–395.
7. "Insolvency statistics in Canada 2017: Annual report," at www.ic.gc.ca/eic/site/bsf-osb.nsf/eng/br03879.html
8. This announcement can be found at www.epa.gov/fueleconomy/basic-information-fuel-economy-labeling
9. Niels van de Ven et al., "The return trip effect: Why the return trip often seems to take less time," *Psychonomic Bulletin and Review* 18, no. 5 (2011), pp. 827–832.
10. 2017 Midyear Consumer Expenditure tables released in September 2018 and found at www.bls.gov/cex/#tables
11. Data from Ray Weaver and Shane Frederick, "A reference price theory of the endowment effect," *Journal of Marketing Research* 49 (October 2012), pp. 696–707.
12. Constantine Dzhabarov et al., "Sell in May and go away: The evidence in the international equity index futures markets," *Quantitative Finance* 18, no. 2 (2017), pp. 171–181.

13. Christine L. Porath and Amir Erez, "Overlooked but not untouched: How rudeness reduces onlookers' performance on routine and creative tasks," *Organizational Behavior and Human Decision Processes* 109 (2009), pp. 29–44.

14. Data provided by Timothy Sturm.

15. The Satterthwaite degrees of freedom are given by

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)^2}{\frac{1}{n_1 - 1} \left(\frac{s_1^2}{n_1} \right)^2 + \frac{1}{n_2 - 1} \left(\frac{s_2^2}{n_2} \right)^2}$$

This t distribution approximation is quite accurate when both sample sizes n_1 and n_2 are 5 or larger.

16. Detailed information about the conservative t procedures can be found in Paul Leaverton and John J. Birch, "Small sample power curves for the two sample location problem," *Technometrics* 11 (1969), pp. 299–307; Henry Scheffé, "Practical solutions of the Behrens-Fisher problem," *Journal of the American Statistical Association* 65 (1970), pp. 1501–1508; and D. J. Best and J. C.W. Rayner, "Welch's approximate solution for the Behrens-Fisher problem," *Technometrics* 29 (1987), pp. 205–210.
17. Koert van Ittersum et al., "Smart shopping carts: How real-time feedback influences spending," *Journal of Marketing* 77 (November 2013), pp. 21–36.
18. Extensive simulation studies are reported in Harry O. Posten, "The robustness of the two-sample t test over the Pearson system," *Journal of Statistical Computation and Simulation* 6 (1978), pp. 295–311; Harry O. Posten, H. Yeh, and D. B. Owen, "Robustness of the two-sample t -test under violations of the homogeneity assumption," *Communications in Statistics* 11 (1982), pp. 109–126; and Harry O. Posten, "Robustness of the two-sample t -test under violations of the homogeneity assumption, part II," *Journal of Statistical Computation and Simulation* 8 (1992), pp. 2169–2184.
19. Data were obtained from insideairbnb.com/get-the-data.html
20. Based on information made available December 2018, "Wheat data: Yearbook tables: Wheat: Average price received by farmers, United States," at www.ers.usda.gov/data-products/wheat-data.aspx#.U7LgyihCz_c
21. Based on Mary H. Keener, "Predicting the financial failure of retail companies in the United States," *Journal of Business & Economic Research* 11, no. 8 (2013), pp. 373–380.
22. Marcos Economides et al., "Improvements in stress, affect, and irritability following brief use of

a mindfulness-based smartphone app: A randomized controlled trial," *Mindfulness* 9, no. 5 (2018), pp. 1584–1593.

23. Karel Kleisner et al., "Trustworthy-looking face meets brown eyes," *PLoS ONE* 8, no. 1 (2013), e53285, doi:10.1371/journal.pone.0053285
24. Cynthia E. Cryfer et al., "Misery is not miserly: Sad and self-focused individuals spend more," *Psychological Science* 19 (2008), pp. 525–530.
25. Elizabeth F Beach and Valerie Nie, "Noise levels in fitness classes are still too high: Evidence from 1997–1998 and 2009–2011," *Archives of Environmental & Occupational Health* 69, no. 4 (2014), pp. 223–230.
26. The 2017 study can be found at www.qsrmagazine.com/reports/2017-drive-thru-performance-study
27. Berit Bakke et al., "Cumulative exposure to dust and gases as determinants of lung function decline in tunnel construction workers," *Occupational Environmental Medicine* 61 (2004), pp. 262–269.
28. Y. Charles Zhang and Norbert Schwarz, "How and why 1 year differs from 365 days: A conversational logic analysis of inferences from the granularity of quantitative expressions," *Journal of Consumer Research* 39 (August 2012), pp. S212–S223.
29. Based on A. H. Ismail and R. John Young, "The effect of chronic exercise on the personality of middle-aged men," *Journal of Human Ergology* 2 (1973), pp. 47–57.
30. Jacob Cohen, *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.) Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.
31. 2018 press release from *The Student Monitor*, at www.studentmonitor.com
32. The average starting salary taken from the April 2018 salary survey by the National Association of Colleges and Employers (NACE).
33. This city's restaurant inspection data can be found at www.jsonline.com/watchdog/dataondemand
34. Based on 2017 information from the USDA Feed Grains Database, at www.ers.usda.gov
35. Peter Glick et al., "Evaluations of sexy women in low- and high-status jobs," *Psychology of Women Quarterly* 29 (2005), pp. 389–395.
36. Morgan K. Ward and Darren W. Dahl, "Should the devil sell Prada? Retail rejection increases aspiring consumers' desire for the brand," *Journal of Consumer Research* 41, no. 3 (2014), pp. 590–609.
37. Ajay Ghei, "An empirical analysis of psychological androgyny in the personality profile of the successful hotel manager," MS thesis, Purdue University, 1992.
38. Data from the "wine" database in the archive of machine learning data bases at the University of

California, Irvine, <ftp.ics.uci.edu/pub/machine-learning-databases>

39. Sommer Christie et al., "Individual alpha peak frequency in ice hockey shooting performance," *Frontiers in Psychology* 8 (May 2017), Article 762.
40. Yvan R. Germain, "The dyeing of ramie with fiber reactive dyes using the cold pad-batch method," MS thesis, Purdue University, 1988.
41. This exercise is based on events that are real. The data and details have been altered to protect the privacy of the individuals involved.
42. Data provided by Joseph A. Wipf, Department of Foreign Languages and Literatures, Purdue University.
43. Based on Gavriel Salvendy, "Selection of industrial operators: The one-hole test," *International Journal of Production Research* 13 (1973), pp. 303–321.

Chapter 9

1. Based on "Research shows promise for new source of rubber for tires," August 11, 2017, www.truckinginfo.com
2. Based on Amit Bhattacharjee et al., "Tip of the hat, wag of the finger: How moral decoupling enables consumers to admire and admonish," *Journal of Consumer Research* 39, no. 6 (2013), pp. 1167–1184.
3. This rule is intended to provide a general guideline for deciding when serious errors may result by applying ANOVA procedures. When the sample sizes in each group are very small, this rule may be a little too conservative. For unequal sample sizes, particular difficulties can arise when a relatively small sample size is associated with a population having a relatively large standard deviation.
4. Ethan Pancer et al., "Isolated environmental cues and product efficacy penalties: The color green and eco-labels," *Journal of Business Ethics* 143 (2017), pp. 159–177.
5. Discussion on this and other tests can be found in M.H. Kutner et al., *Applied Linear Models*, 5th ed., McGraw-Hill/Irwin, 2005.
6. Katariina Mäenpää et al., "Consumer perceptions of Internet banking in Finland: The moderating role of familiarity," *Journal of Retailing and Consumer Services* 15 (2008), pp. 266–276.
7. Garry Kuan et al., "Effects of relaxing and arousing music during imagery training on dart-throwing performance, physiological arousal indices, and competitive state anxiety," *Frontiers in Psychology* 9 (February 2018), Article 14.
8. Ryan Hamilton et al., "We'll be honest, this won't be the best article you've ever read: The use of

dispreferred markers in word-of-mouth communication," *Journal of Consumer Research* 41 (2014), pp. 197–212.

9. This example is based on data from a study conducted by Jim Baumann and Leah Jones of the Purdue University School of Education.
10. Brian E. Saelens et al., "Relation between higher physical activity and public transit use," *American Journal of Public Health* 104, no. 5 (2014), pp. 854–859.
11. Lucas Nesselhauf et al., "Information and involvement: The influence on the acceptance of innovative wine packaging," *International Journal of Wine Business Research* 29, no. 3 (2018), pp. 285–298.
12. Kendall J. Eskine, "Wholesome foods and wholesome morals? Organic foods reduce prosocial behavior and harshen moral judgments," *Social Psychological and Personality Science*, 2012, doi: 10.1177/1948550612447114
13. Robert Rouwenhorst and Liang Zhao, "Zipped commercials, zapped memory? Not necessarily," *Business and Management Research* 6, no. 3 (2017), pp. 85–93.
14. Jeffrey T. Kullgren et al., "Individual- versus group-based financial incentives for weight loss," *Annals of Internal Medicine* 158, no. 7 (2013), pp. 505–514.
15. Peter R. Bartel et al., "Attention and working memory in resident anaesthetists after night duty: Group and individual effects," *Occupational and Environmental Medicine* 61 (2004), pp. 167–170.
16. Based on Manuhar U. Kalwani and Chi K. Yim, "Consumer price and promotion expectations: An experimental study," *Journal of Marketing Research* 29 (1992), pp. 90–100.

Chapter 10

1. See pewresearch.org/fact-tank/2018/01/23/americans-see-both-good-and-bad-in-trends-that-are-changing-the-workplace/ft_18-01-17_laborforcetrends_1/. For more information about the Pew Research Center American Trends Panel, see pewsocialtrends.org/2017/10/18/methodology-12
2. For more information about the survey, see aba.com/Products/Surveys/Pages/2017-Bank-Insurance.aspx
3. See Alan Agresti and Brent Coull, "Approximate is better than 'exact' for interval estimation of binomial proportions," *American Statistician* 52 (1998), pp. 119–126. A detailed theoretical study is Lawrence D. Brown, Tony Cai, and Anirban DasGupta, "Confidence intervals for a binomial

- proportion and asymptotic expansions,” *Annals of Statistics* 30 (2002), pp. 160–201.
4. This example is adapted from a survey directed by Professor Joseph N. Uhl of the Department of Agricultural Economics, Purdue University. The survey was sponsored by the Indiana Christmas Tree Growers Association.
 5. See southerncross.co.nz/about-the-group/media-releases/2013.aspx
 6. See “The myth of social media” at wsj.com/public/resources/documents/sac_report_11_socialmedia_061114.pdf
 7. See nielsen.com/us/en/insights/news/2015/what-moves-todays-teenage-canadian-music-fan.html
 8. Heather Tait, *Aboriginal Peoples Survey, 2006: Inuit Health and Social Conditions* (2008), Social and Aboriginal Statistics Division, Statistics Canada. Available from statcan.gc.ca/pub
 9. See news.teamxbox.com/xbox/18254
 10. Oliver Meixner et al., “The use of social media within the Austrian supply chain for food and beverages,” *Proceedings in System Dynamics and Innovations in Food Networks* (2013), pp. 1–13. See centmapress.ilb.uni-bonn.de/ojs/index.php/proceedings/index
 11. See Alan Agresti and Brian Caffo, “Simple and effective confidence intervals for proportions and differences of proportions result from adding two successes and two failures,” *American Statistician* 45 (2000), pp. 280–288. The Wilson interval is a bit conservative (true coverage probability is higher than the confidence level) when p_1 and p_2 are equal and close to 0 or 1, but the traditional interval is much less accurate and has the fatal flaw that the true coverage probability is less than the confidence level.
 12. Nicolas Gueguen and Celine Jacob, “Clothing color and tipping: Gentlemen patrons give more tips to waitresses with red clothes,” *Journal of Hospitality & Tourism Research* 38, no. 2 (2014), pp. 275–280.
 13. See Steven W. Lagakos, Barbara J. Wessen, and Marvin Zelen, “An analysis of contaminated well water and health effects in Woburn, Massachusetts,” *Journal of the American Statistical Association* 81 (1986), pp. 583–596, and the following discussion. This case is the basis for the movie *A Civil Action*.
 14. See, for example, gartner.com/it-glossary/internet-of-things
 15. From pewinternet.org/2014/05/14/internet-of-things, posted May 14, 2014.
 16. Reported in Stephanie Goldberg, “Benefits integration picks up steam: Compliance drives interest in combining workplace absence programs,” *Business Insurance* 48, no. 17 (2014), p. 0013. Also, see “2014 Aon Newitt Health Care Survey” at aon.com
 17. Jiao Xu et al., “News media channels: Complements or substitutes? Evidence from mobile phone usage,” *Journal of Marketing* 78 (2014), pp. 97–112. The methodology used in the study has been simplified for our purposes.
 18. From Rick B. van Baaren, “The parrot effect: How to increase tip size,” *Cornell Hotel and Restaurant Administration Quarterly* 46 (2005), pp. 79–84.
 19. Some details are given in David H. Kaye and Mikel Aickin (eds.), *Statistical Methods in Discrimination Litigation*, Marcel Dekker, 1986.
 20. The report, dated May 18, 2012, is available from pewinternet.org/Reports/2012/Future-of-Gamification/Overview.aspx
 21. From the Pew Research Center’s Project for Excellence in Journalism, *The State of the News Media 2012*, available from stateofthemediamedia.org/?src=prc-headline
 22. Data are from the NOAA Satellite and Information Service at ncdc.noaa.gov/special-reports/groundhog-day.php

Chapter 11

1. Marek Matejun, “The role of flexibility in building the competitiveness of small and medium enterprises,” *Management* 18, no. 1 (2014), pp. 154–168.
2. When the expected cell counts are small, it is best to use a test based on the exact distribution rather than the chi-square approximation, particularly for 2×2 tables. Many statistical software systems offer an “exact” test as well as the chi-square test for 2×2 tables.
3. The full report of the study appeared in George H. Beaton et al., “Effectiveness of vitamin A supplementation in the control of young child morbidity and mortality in developing countries,” United Nations ACC/SCN State-of-the-Art Series, Nutrition Policy Discussion Paper no. 13 (1993).
4. Nicolas Gueguen and Celine Jacob, “Clothing color and tipping: Gentlemen patrons give more tips to waitresses with red clothes,” *Journal of Hospitality & Tourism Research* 38, no. 2 (2014), pp. 275–280.
5. Based on Shan Feng et al., “Does classical music relieve math anxiety? Role of tempo on price computational avoidance,” *Psychology & Marketing* 31, no. 7 (2014), pp. 489–499.
6. From pewresearch.org/fact-tank/2018/03/16/how-millennials-compare-with-their-grandparents
7. The analysis could also be performed by using a two-way table to compare the ages of the selected

and not-selected customers. Because the selected customers are a relatively small percent of the total number of customers, the results will be approximately the same.

8. See the M&M Mars website at us.mms.com/us/about/products for this and other information.
9. See pewinternet.org/about.asp
10. Based on pewinternet.org/fact-sheet/internet-broadband
11. See nhcaa.org
12. These data are a composite based on several actual audits of this type.
13. From Robert J. M. Dawson, "The 'unusual episode' data revisited," *Journal of Statistics Education* 3, no. 3 (1995). Electronic journal available at the American Statistical Association website, www.amstat.org

Chapter 12

1. In practice, x may also be a random quantity. Inferences can then be interpreted as *conditional* on a given value of x .
2. Mirjam Van Praag et al., "The higher returns to formal education for entrepreneurs versus employees," *Small Business Economics* 40 (2013), pp. 375–396.
3. National Science Foundation, National Center for Science and Engineering Statistics, *Higher Education Research and Development: Fiscal Year 2017*. Detailed tables released in November 2018 and available at www.nsf.gov/statistics/srvyherd
4. As the text notes, the residuals are not independent observations. They also have somewhat different standard deviations. For practical purposes of examining a regression model, we can nonetheless interpret the normal quantile plot as if the residuals were data from a single distribution.
5. Inflation is measured by the December-to-December change in the Consumer Price Index. These data were found at www.bls.gov/cpi/. Interest rates for the six-month secondary market Treasury bill were obtained at fred.stlouisfed.org/series/TB6MS#0
6. See the essay "Regression toward the mean," in Stephen M. Stigler, *Statistics on the Table*, Harvard University Press, 1999. The quotation from Milton Friedman appears in this essay.
7. In fact, the Excel regression output does not report the sign of the correlation r . The scatterplot in Figure 10.3 shows that r is positive. To get the correlation with the correct sign in Excel, you must use the `=CORREL()` function.
8. Selling prices and assessment values are available at php.jconline.com/propertysales/propertysales.php
9. Pradit Withisuphakorn and Pornsit Jiraporn, "CEO age and CEO gender: Are female CEOs older than their male counterparts?" *Finance Research Letters* 22 (2017), pp. 129–135.
10. Tuition and fees were obtained from www.chronicle.com/interactives/tuition-and-fees
11. Michal Plotnicki and Adam Szyszka, "IPO market timing: The evidence of the disposition effect among corporate managers," *Global Finance Journal* 25, no.1 (2014), pp. 48–55.
12. Mike Mondello and Joel Maxcy, "The impact of salary dispersion and performance bonuses in NFL organizations," *Management Decision*, 47, no.1 (2009), pp. 110–123. These data were collected from www.cbssports.com/nfl/playerrankings/regularseason/ and content usatoday.com/sports/football/nfl/salaries
13. Ziming Xuan et al., "Tax policy, adult binge drinking, and youth alcohol consumption in the United States," *Alcoholism: Clinical and Experimental Research* 37, no. 10 (2013), pp. 1713–1719.
14. Data on net new cash flow of long-term mutual funds obtained from "Chapter 3: U.S. Mutual Funds," *2018 Investment Company Fact Book*, Investment Company Institute, www.icifactbook.org
15. Data were provided by the Ames City Assessor, Ames, IA.
16. Data on fuel consumption ratings made available by the Government of Canada, data.gc.ca/data/en/dataset/98f1a129-f628-4ce4-b24d-6f16bf24dd64
17. Data obtained from the *The Inclusive Development Index 2018* report, World Economic Forum, www.weforum.org/reports
18. Roya. Rahimi, "Organizational culture and customer relationship management: A simple linear regression analysis," *Journal of Hospitality Marketing & Management* 26, no. 4 (2017), pp. 443–449.
19. This annual report can be found at www.kiplinger.com
20. Data available at www.ncdc.noaa.gov
21. Woo Gon Kim and Hong-Bumm Kim, "Measuring customer-based restaurant brand equity," *Cornell Hotel and Restaurant Administration Quarterly* 45, no. 2 (2004), pp. 115–131.
22. Stefan Groschl, "Persons with disabilities: A source of nontraditional labor for Canada's hotel industry," *Cornell Hotel and Restaurant Administration Quarterly* 46, no. 2 (2005), pp. 258–274.

23. Table of values available at www.ers.usda.gov/Data/AgProductivity
24. Data obtained from the CNBC report titled “Companies with closer CEO pay ratios may generate higher profit per worker,” posted on May 13, 2018, www.cnbc.com/2018/05/13/closer-ceo-pay-ratios-may-generate-higher-profit-per-worker.html

Chapter 13

1. Based on the Space Management Model for Purdue University implemented by Keith Murray, Director of Space Management and Academic Scheduling.
2. Data obtained from the *The Inclusive Development Index 2018* report, World Economic Forum, www.weforum.org/reports
3. U.S. Federal Deposit Insurance Corp., *Statistics on Banking*, issued annually. Information for current year can be found online at www5.fdic.gov/sdi/main.asp?formname=compare
4. Elaine M. Wong et al., “A face only an investor could love: CEOs’ facial structure predicts their firms’ financial performance,” *Psychological Science* 22, no. 12 (2011), pp. 1478–1483.
5. Alain Cohn et al., “Do professional norms in the banking industry favor risk-taking?” *Review of Financial Studies* 30, no. 11 (2017), pp. 3801–3823.
6. Available at *ConsumerReports.org*. Latest summary posted November 2018.
7. These data were obtained from “The QSR 50,” an annual report provided by *QSR* magazine, qsrmagazine.com/reports
8. Data provided by the owners of Duck Worth Wearing, Ames, Iowa.
9. From “Largest Indianapolis-area architectural firms,” *Indianapolis Business Journal*, June 15, 2014.
10. The data were obtained from the Internet Movie Database (IMDb), www.imdb.com, on August 20, 2018.
11. The KISS principle refers to the empirical principle “Keep it simple, stupid.” In regression, this refers to keeping the models simple and avoiding unnecessary complexity.
12. From Maxime Vachon et al., “Associations between systemic quality of life and burnout among French Canadian workers,” *Social Indicators Research* 142, no. 3 (2018), pp. 1193–1210.
13. From Alex Traugutt et al., “Salary disparities between male and female head coaches: An investigation of the NCAA power five conferences,” *Journal of Sport* 6, no. 1 (2018), pp. 40–58.
14. From Michael E. Staten et al., “Information costs and the organization of credit markets: A theory of indirect lending,” *Economic Inquiry* 28 (1990), pp. 508–529.
15. Summary information taken from “FINAL REPORT: Canada Small Business Financing Program (CSBFP) Awareness and Satisfaction Study,” prepared for Industry Canada by R.A. Malatest & Associates Ltd., July 2013.
16. From Susan Stites-Doe and James J. Cordeiro, “An empirical assessment of the determinants of bank branch manager compensation,” *Journal of Applied Business Research* 15 (1999), pp. 55–66.
17. The data were collected from www.realtor.com on October 8, 2001.
18. Karen Misquitta et al., “The relationship between brain atrophy and cognitive-behavioural symptoms in retired Canadian football players with multiple concussions,” *NeuroImage: Clinical* 19 (2018), pp. 551–558.
19. Bill Merrilees and Tino Fenech, “From catalog to Web: B2B multi-channel marketing strategy,” *Industrial Marketing Management* 36 (2007), pp. 44–49.
20. Based on Manuhar U. Kalwani and Chi K. Yim, “Consumer price and promotion expectations: An experimental study,” *Journal of Marketing Research* 29 (1992), pp. 90–100.
21. Tung-Shan Liao and John Rice, “Innovation investments, market engagement and financial performance: A study among Australian manufacturing SMEs,” *Research Policy* 39, no. 1 (2010), pp. 117–125.
22. Katharine Kelley et al., “Estimating consumer spending on tickets, merchandise, and food and beverage: A case study of a NHL team,” *Journal of Sport Management* 28 (2014), pp. 253–265.
23. Yield data can be obtained at www.nass.usda.gov/Quick_Stats
24. A description of this case, as well as other examples of the use of statistics in legal settings, is given in Michael O. Finkelstein, *Quantitative Methods in Law*, Free Press, 1978.

Chapter 14

1. Stock prices (including those for Adidas) can be found at finance.yahoo.com
2. Amazon quarterly net sales data were extracted from quarterly reports found by following the link “Investor Relations” at amazon.com
3. The differences in correlation values are due to the fact that the ACF computes the correlations by using the same sample mean for the y_t variable

- and the lag variable y_{t-k} , namely, the sample of the whole series. The y_t variable has n observations, while the y_{t-k} variable has $(n - k)$ observations. The standard correlation formula would treat these variables differently and thus use two different sample means.
4. NBA game results are available at espn.com/nba
 5. See Note 1.
 6. Data available from the Economic Research website of the Federal Reserve Bank of St. Louis, research.stlouisfed.org
 7. See Note 1.
 8. Table available from W. A. Fuller, *Introduction to Statistical Time Series*, 2nd ed., Wiley, 1995.
 9. A comprehensive treatment of unit root tests can be found in J. D. Hamilton, *Introduction to Statistical Time Series*, Princeton University Press, 1995.
 10. Data from the University of Michigan Surveys of Consumers website, sca.isr.umich.edu
 11. A variety of historical data on gold can be found at the World Gold Council website, gold.org
 12. See Note 6.
 13. Data obtained from the National Oceanic and Atmospheric Administration Great Lakes Environmental Research Laboratory, glerl.noaa.gov
 14. See Note 6.
 15. Data obtained from railroadpm.org
 16. See Note 6.
 17. Data obtained from the City of Chicago data portal, data.cityofchicago.org
 18. Data gathered from institutehorsupplymanagement.org
 19. The exponential smoothing model is the forecasting equation for an ARIMA(0, 1, 1) model. There are many excellent books on ARIMA modeling, including the authoritative reference book of G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*, 4th ed., Wiley, 2008.
 20. Data obtained from the Bureau of Transportation Statistics, bts.gov
 21. Data obtained from the Australian Bureau of Statistics, abs.gov.au
 22. Data obtained from the City of Philadelphia data portal, data.phila.gov
 23. See Note 6.
 24. Data provided by David Robinson.
 25. Data from the National Bureau of Statistics of China website, data.stats.gov.cn/english
 26. See Note 25.
 27. Data extracted from the International Telecommunication Union (United Nations specialized agency), itu.int
 28. Data extracted using the data tools found at the U.S. Census Bureau, www.census.gov
 29. For more details on the issues associated with log transformation in regression, see D. Miller, "Reducing transformation bias in curve fitting," *The American Statistician* 38 (1984), pp. 124–126.
 30. See Note 1.
 31. See Note 1.
 32. Data extracted from quarterly reports found by following the link "Investor Relations" at att.com
 33. Data found at the statistics portal (statista.com); original source is the hotel data tracking company STR, str.com
 34. See Note 6.
 35. Data obtained from the baseball statistics website www.baseball-reference.com
 36. Data obtained from the City of Los Angeles data portal, data.lacity.org
 37. See Note 20.
 38. Data obtained from the National Oceanic and Atmospheric Administration climate site, climate.gov
 39. Data obtained from the NFL statistics website www.pro-football-reference.com
 40. See Note 38.
 41. See Note 28.
 42. Data obtained from the Midwestern Regional Climate Center, mrcc.illinois.edu
 43. Densities of populations over time of most countries in the world can be found at the World Bank website, worldbank.org
 44. Data obtained from OPEC website, opec.org
 45. See Note 22.

Chapter 15

1. As of 2019, the American Society for Quality (ASQ) had honored 25 individuals by conferring on them the status of Honorary Member. A detailed summary of the background and contributions of the individuals noted here along with other pioneers can be found at an ASQ website, asq.org/about-asq/honorary-members
2. The cause-and-effect diagram was prepared by S. K. Bhat of the General Motors Technical Center as part of a course assignment at Purdue University.
3. Control charts were invented in the 1920s by Walter Shewhart at the Bell Telephone Laboratories. Shewhart's classic book, *Economic Control of Quality of Manufactured Product* (Van Nostrand, 1931), organized the application of statistics to improving quality.

4. In his classic book, *Out of the Crisis* (MIT Center for Advanced Engineering Study, 1986), W. Edwards Deming demonstrates the effects of counterproductive adjustment to an in-control process by means of a physical experiment based on dropping marbles through a funnel onto a tabletop. Participants in the experiment learn that the least scatter on the tabletop is obtained by not moving the funnel, that is, by means of “no action.”
5. In statistics, the term “efficient” relates to the variance of the sampling distribution of the estimator. The estimator with the smallest variation is referred to as an *efficient* estimator.
6. Simulated data based on information appearing in Arvind Salvekar, “Application of six sigma to DRG 209,” found at the Smarter Solutions website, smarterolutions.com
7. S&P closing prices can be found at finance.yahoo.com
8. Game log statistics on NBA players can be found at stats.nba.com
9. Data provided by Linda McCabe, Purdue University.
10. The exact formula for c_4 is given by

$$c_4 = \sqrt{\frac{2}{n-1}} \frac{\left(\frac{k}{2}-1\right)!}{\left(\frac{k-1}{2}-1\right)!}$$

where k is the number of observations. If the argument of the factorial is a non-integer it is computed as follows

$$\left(\frac{k}{2}\right)! = \left(\frac{k}{2}\right)! \left(\frac{k}{2}-1\right)! \left(\frac{k}{2}-2\right)! \cdots \left(\frac{1}{2}\right) \sqrt{\pi}$$

11. Data provided by Colleen O’Brien, Team Leader Quality Resource and Privacy and Safety Officer, Bellin Health.
12. Data on aviation accidents can be found at the Federal Aviation Administration (FAA) Data & Research page, faa.gov/data_research

Chapter 16

1. Rebecca A. Ferrer et al., “Revisiting the effects of anger on risk-taking: Empirical and meta-analytic evidence for differences between males and females,” *Journal of Behavioral Decision Making* 30 (2017), pp. 516–526.
2. This example is based on Laura Herrewijn and Karolien Poels, “Recall and recognition of in-game advertising: The role of game control,” *Frontiers in Psychology* 4 (2014), pp. 1–14.
3. Based on a student project by Stefannie Garcia, Stephanie Morgan, Jeremy Sanders, Taylor Hooper, and Natalie Rowe titled “The effect

of scarcity on consumer purchase intentions,” University of New Orleans, 2014.

4. This example is based on Iana A. Castro et al., “The influence of disorganized shelf displays and limited product quantity on consumer purchase,” *Journal of Marketing* 77 (2013), pp. 118–133.
5. We present the two-way ANOVA model and analysis for the general case in which the sample sizes may be unequal. If the sample sizes vary a great deal, serious complications can arise. There is no longer a single standard ANOVA analysis. Most computer packages offer several options for the computation of the ANOVA table when group counts are unequal. When the counts are approximately equal, all methods give essentially the same results.
6. U.S. Census Bureau, *American Community Survey*, 2017 American Community Survey 1-Year Estimates.
7. This example is based on Shibin Sheng and Yue Pan, “Bundling as a new product introduction strategy: The role of brand image and bundle features,” *Journal of Retailing and Consumer Services* 16 (2009), pp. 367–376.
8. Example 16.10 is based on a study described in Todd Green and John Peloza, “Finding the right shade of green: The effect of advertising appeal type on environmentally friendly consumption,” *Journal of Advertising* 43, no. 2 (2014), pp. 128–141.
9. Based on Manuhar U. Kalwani and Chi K. Yim, “Consumer price and promotion expectations: An experimental study,” *Journal of Marketing Research* 29 (1992), pp. 90–100.
10. Henrik Hagtvedt and S. Adam Brasel, “Color saturation increases perceived product size,” *Journal of Consumer Research* 44, no. 2 (2017), pp. 396–413.
11. Jane Kolodinsky et al., “Sex and cultural differences in the acceptance of functional foods: A comparison of American, Canadian, and French college students,” *Journal of American College Health* 57 (2008), pp. 143–149.
12. Tomas Brodin et al., “Dilute concentrations of a psychiatric drug alter behavior of fish from natural populations,” *Science* 339 (2013), pp. 814–815.
13. Vincent P. Magnini and Kiran Karande, “The influences of transaction history and thank you statements in service recovery,” *International Journal of Hospitality Management* 28 (2009), pp. 540–546.
14. Willemijn M. van Dolen, Ko de Ruyter, and Sandra Streukens. “The effect of humor in electronic service encounters,” *Journal of Economic Psychology* 29 (2008), pp. 160–179.

15. Tamar Kugler et al., "Trust between individuals and groups: Groups are less trusting than individuals but just as trustworthy," *Journal of Economic Psychology* 28 (2007), pp. 646–657.
16. S. Leroy, "Why is it so hard to do my work? The challenge of attention residue when switching between work tasks," *Organizational Behavior and Human Decision Processes* 109 (2009), pp. 168–181.
17. Margaret C. Campbell and Kevin Lane Keller, "Brand familiarity and advertising repetition effects," *Journal of Consumer Research* 30 (2003), pp. 292–304.
18. Based on a problem from Renée A. Jones and Regina P. Becker, Department of Statistics, Purdue University.
19. Debora V. Thompson and Prashant Malaviya, "Consumer-generated ads: Does awareness of advertising co-creation help or hurt persuasion?" *Journal of Marketing* 77 (2013), pp. 33–47.
20. Lijia Lin et al., "Animated agents and learning: Does the type of verbal feedback they provide matter?" *Computers and Education*, 2013, doi: 10.1016/j.compedu.2013.04.017
8. Data obtained from data.worldbank.org/indicator
9. Data loosely based on Alexander Redlein and Michael Zobl, "ERP systems within facility management," *Advanced Research in Scientific Areas Proceedings*, December 2013, pp. 153–155.
10. Data provided by Warren Page, New York City Technical College, from a study done by John Hudesman.
11. Data from "Results report on the vitamin C pilot program," prepared by SUSTAIN (Sharing United States Technology to Aid in the Improvement of Nutrition) for the U.S. Agency for International Development.
12. Go to www.futurity.org/fried-food-taste-without-all-the-fat for more information.
13. Simplified from the EESEE story "Stepping Up Your Heart Rate."
14. Data provided by Diana Schellenberg, Purdue University School of Health Sciences.
15. Data provided by Sam Phillips, Purdue University.
16. Data from Olga Goncalves, "Efficiency and productivity of French ski resorts," *Tourism Management* 36 (2013), pp. 650–657.
17. Data provided by Helen Park. See H. Park et al., "Fortifying bread with each of three antioxidants," *Cereal Chemistry* 74 (1997), pp. 202–206.
18. Data provided by Jo Welch, Purdue University Department of Foods and Nutrition.
19. See Note 7.
20. The QSR magazine study can be found at www.qsrmagazine.com/content/drive-thru-performance-study-customer-service
21. *Consumer Reports*, June 1986, pp. 366–367.
22. Data from Parviz Asheghian, "The managerial efficiencies of Indian firms as compared to American firms," *International Journal of Economics and Management Sciences* 6 (2012), pp. 45–55.
23. Based on A. A. Adish et al., "Effect of consumption of food cooked in iron pots on iron status and growth of young children: A randomised trial," *The Lancet* 353 (1999), pp. 712–716.

Chapter 17

1. A complete listing of all hotels by ratings can be found at tripadvisor.com
2. MLB season statistics can be found at baseball-reference.com
3. This test was invented by Frank Wilcoxon (1892–1965) in 1945. Wilcoxon was a chemist who met statistical problems in his work at the research laboratories of American Cyanamid Company.
4. For purists, here is the precise definition: X_1 is *stochastically larger* than X_2 if

$$P(X_1 > a) \geq P(X_2 > a)$$

for all a , with strict inequality for at least one a . The Wilcoxon rank sum test is effective against this alternative in the sense that the power of the test approaches 1 (i.e., the test becomes more certain to reject the null hypothesis) as the number of observations increases.

5. Data from Huey Chern Boo, "Consumers' perceptions and concerns about safety and healthfulness of food served at fairs and festivals," MS thesis, Purdue University, 1997.
6. Discussion forum count taken from the University of Wisconsin–Milwaukee MBA course Predictive Analytics for Managers.
7. From Sapna Aneja, "Biodeterioration of textile fibers in soil," MS thesis, Purdue University, 1994.

Chapter 18

1. Logistic regression models for the general case where there are more than two possible values for the response variable have been developed. These are considerably more complicated and are beyond the scope of our present study. For more information on logistic regression, see A. Agresti, *An Introduction to Categorical Data Analysis*, 3rd ed., Wiley, 2019; and R. X. Sturdivant, S. Lemeshow, and D. W. Hosmer Jr., *Applied Logistic Regression*, 3rd ed., Wiley, 2013.

2. Nicolas Guéguen and Céline Jacob, "Clothing color and tipping: Gentlemen patrons give more tips to waitresses with red clothes," *Journal of Hospitality & Tourism Research* 38, no. 2 (2014), pp. 275–280.
 3. Michael Lynn and Shou Wang, "The indirect effects of tipping policies on patronage intentions through perceived expensiveness, fairness, and quality," *Journal of Economic Psychology* 39 (2013), pp. 62–71.
 4. The Millennials & Investing Survey was conducted online January 16–25, 2018. "Battle of the sexes: How millennials' financial attitudes, habits differ by gender" was released by PNV Investments on June 13, 2018.
 5. Data taken from Oliver Rowntree, "The Mobile Gender Gap Report 2018," Connect Women GSMA.
 6. The press release for this survey can be found at the Best Western website, www.bestwestern.com/about-us/press-media/press-release-details.asp?NewsID=910
 7. From Karin Weber and Weley S. Roehl, "Profiling people searching for and purchasing travel products on the World Wide Web," *Journal of Travel Research* 37 (1999), pp. 291–298.
 8. This example is taken from a classical text written by a contemporary of R. A. Fisher. (Fisher developed many of the fundamental ideas of statistical inference that we use today.) The reference is D. J. Finney, *Probit Analysis*, Cambridge University Press, 1947. Although not included in the analysis, it is important to note that the experiment included a control group that received no insecticide. No aphids died in this group. Also, although we have chosen to call the response "killed," in Finney's text, the category is described as "apparently dead, moribund, or so badly affected as to be unable to walk more than a few steps." This is an example of the need to make careful judgments when defining variables to be used in a statistical analysis.
- Nevertheless, an insect that is "unable to walk more than a few steps" is unlikely to eat very much of a chrysanthemum plant.
9. Based on results from "Understanding how U.S. online shoppers are reshaping the retail experience," 2012, PricewaterhouseCoopers LLP.
 10. Getachew N. Tolesa et al., "Logistic regression analysis of marketability of tomato fruit harvested at different maturity stages and subjected to disinfection, storage condition and storage period treatments," *Biological Agriculture & Horticulture* 34, no. 1 (2018), pp. 40–52.
 11. Laura G. Brown et al., "Food safety practices linked with proper refrigerator temperatures in retail delis," *Foodborne Pathogens and Disease* 15, no. 5 (2018), pp. 300–307.
 12. Based on Leigh J. Maynard and Malvern Mupandawana, "Tipping behavior in Canadian restaurants," *International Journal of Hospitality Management* 28 (2009), pp. 597–603.
 13. Anthony A. Noce and Larry McKeown, "A new benchmark for Internet use: A logistic modeling of factors influencing Internet use in Canada, 2005," *Government Information Quarterly* 25 (2008), pp. 462–476.
 14. Results from Rayna Brown and Neal Sarma, "CEO overconfidence, CEO dominance and corporate acquisitions," *Journal of Economics and Business* 59 (2007), pp. 358–379.
 15. Based on Greg Clinch, "Employee compensation and firms' research and development activity," *Journal of Accounting Research* 29 (1991), pp. 59–78.
 16. Tom Reichert, "The prevalence of sexual imagery in ads targeted to young adults," *Journal of Consumer Affairs* 37 (2003), pp. 403–412.
 17. Based on information in "NCAA 2003 national study of collegiate sports wagering and associated health risks," which can be found at the NCAA website, www.ncaa.org