

at the Mount Washington Observatory are presented for each month. These are likely to be measures of central tendency for each month over time. Explain why these “normal” temperatures might be calculated as means or medians. What would be the reasoning for using one statistic over the other?

- 4.28 Back in Exercises 4.13 and 4.14 we saw how the mean and median changed when an outlier was included in the computations. If you were reporting the typical salary at a company, how might the mean and median give different impressions to potential applicants?
- 4.29 The “normal” weather data from the Mount Washington Observatory are broken down by months. Why might you not want to average across all months in a year? How else could you summarize the year?
- 4.30 There appears to be an outlier in the data for peak wind gust recorded on top of Mount Washington (see data in Exercise 4.20). Where do you see an outlier and how would excluding this data point affect the different calculations of central tendency?
- 4.31 Here are winning percentages for 11 players for their best four-year pitching performances: 0.755, 0.721, 0.708, 0.773, 0.782, 0.747, 0.477, 0.817, 0.617, 0.650, 0.651.

- What is the mean of these scores?
 - What is the median of these scores?
 - Compare the mean and median. Does the difference between them suggest that the data are skewed very much?
- 4.32 Briefly describe a real-life situation in which the median is preferable to the mean. Give hypothetical numbers for mean and median in your explanation. Be original! (Don't use home prices or another example from the chapter.)
- 4.33 Find an advertisement for a weight-loss product either online or in the print media—the more unbelievable the claims, the better!
- What does the ad promise that this product will do for the consumer?
 - What data does it offer for its promised benefits? Does it offer any descriptive statistics or merely testimonials? If it offers descriptive statistics, what are the limitations of what they report?
 - If you were considering this product, what measures of central tendency would you most like to see? Explain your answer, noting why not all measures of central tendency would be helpful.
 - If a friend with no statistical background were considering this product, what would you tell him or her?

- 4.34 When you see an ad on TV for a body-shaping product (e.g., an abdominal muscle machine), often a

person with a wonderful success story is featured in the ad. The statement “individual results may vary” hints at what kind of data the advertisement may be presenting.

- What kind of data is being presented in these ads?
 - What statistics could be presented to help inform the public about how much “individual results might vary”?
- 4.35 The National Survey of Student Engagement asked students how often they asked questions in class or participated in classroom discussions. The options were “never,” “sometimes,” “often,” and “very often.” Here are the percentages, reported in 2005, of students who responded “very often” for the 31 institutions classified as liberal arts colleges that allowed their 2004 data to become public through the *U.S. News & World Report* Web site.

58	45	53	45	65	41	50	46	54
59	52	60	59	62	54	52	53	54
83	60	32	62	50	50	43	32	53
60	52	55	53					

- What is the range of these data?
 - The top college is Marlboro College in Vermont, and the two tied for lowest are Randolph-Macon Women's College in Virginia and Texas A&M University in Galveston. What research questions do these data suggest to you? State at least one research question generated by these data.
- 4.36 Here again are the data from the National Survey of Student Engagement for a sample of 19 national universities, as reported in 2005. These are the percentages of students who said they were assigned between 5 and 10 20-page papers.

0	5	3	3	1	10	2
2	3	1	2	4	2	1
1	1	4	3	5		

- Calculate the mean of these data using the symbols and formula.
 - Calculate the variance of these data using the symbols and formula, but also using columns to show all calculations.
 - Calculate the standard deviation using the symbols and formula.
 - In your own words, describe what the mean and standard deviation of these data tell us about these scores.
- 4.37 For each of the following situations, state whether the mean would be a statistic or a parameter. Explain your answer.

- According to 1991 Canadian census data, the mean income (from employment only) of French-speaking Canadians living in Ontario was \$29,527, higher than the general population mean of \$28,838.
- In the 2004–2005 National Basketball Association season, the 30 teams won a mean 41.00 games.
- The General Social Survey (GSS) includes a vocabulary test in which participants are given a series of words and asked to choose the appropriate synonym from a multiple-choice list of five words (e.g., *beast* with the choices *afraid*, *words*, *large*, *animal*, and *separate*). The mean vocabulary test score was 5.98.
- The National Survey of Student Engagement (NSSE) asked students at participating institutions how often they discussed ideas or readings with their professors outside of class. Among the 19 national universities that made their data public, the mean percentage of students who responded “very often” was 8%.

4.38 Consider the many possible distributions of grades on a quiz in a statistics class; imagine that the grades could range from 0 to 100. For each of the following situations, give a hypothetical mean and median (that is, make up a mean and a median that might occur with a distribution that has this shape). Explain your answer.

- Normal distribution
- Positively skewed distribution
- Negatively skewed distribution

4.39 For each of the following distributions, state whether it's more likely to be unimodal or bimodal. Explain your answer.

- Age of patients in a hospital maternity ward
- Depression scores on a Beck Depression Inventory
- GRE scores of applicants to sociology graduate programs
- The cost of an AIDS drug that is sold in developed countries in Europe, as well as in developing countries in Africa

4.40 Here are the numbers of wins for the 30 National Basketball Association teams in the 2004–2005 season.

45 43 42 33 33 54 47 44 42 30
 59 45 36 18 13 52 49 44 27 26
 62 50 37 34 34 59 58 51 45 18

- Determine the mean, median, and mode of these data. Use symbols and the formula when showing your calculation of the mean.
- Using software, calculate the range and standard deviation of these data.
- Write a one- to two-paragraph summary describing the distribution of these data. Mention center,

variability, and shape. Be sure to discuss the number of modes (i.e., unimodal, bimodal, multimodal), any possible outliers, and the presence and direction of any skew.

4.41 The U.S. Census Bureau collects and analyzes data on numerous aspects of American life by state, including the percentage of people with high school degrees, bachelor's degrees, and advanced degrees. If you wanted to calculate the “average” percentage of people with advanced degrees across all states, would you report a mean, median, or mode? Explain your answer clearly.

4.42 According to a 2007 article on the Economist.com Web site, Americans are the international leaders in TV viewing, averaging 8 hours and 11 minutes a day. Below are approximate, daily average viewing times for 12 countries based on this source:

United States—8.2 hours
 Turkey—5 hours
 Italy—4.05 hours
 Japan—3.75 hours
 Spain—3.6 hours
 Portugal—3.5 hours
 Australia—3.2 hours
 South Korea—3.16 hours
 Canada—3.1 hours
 Britain—3 hours
 Denmark—3 hours
 Finland—2.8 hours

- Compute the mean and median across these 12 data points.
- How are these statistics affected by including or excluding the United States?

4.43 Refer to the data from Exercise 4.42.

- How do you think these daily “averages” were calculated—using means or medians?
- Do you think TV viewing habits might vary by other personal or demographic characteristics? Could these represent confounds?
- How might you collect samples to more specifically describe TV viewing habits as a function of other personal characteristics?

4.44 When the typical height or typical weight of children is plotted to create growth charts, do you think it would be appropriate to use the mean for these data? There are often outliers for height, but why might we not have to be concerned with their effect on these data?

4.45 Guinness World Records relies on what kind of data for its amazing claims?