

1. p: responses of all us 13-17 old
 contained 1055 13-17 old
 8. REVIEW EXERCISES 19

8. REVIEW EXERCISES

- 1.1 ✓ A newspaper headline reads,
U.S. TEENS TRUST, FEAR THEIR PEERS
 and the article explains that a telephone poll was conducted of 1055 persons 13 to 17 years old. Identify a statistical population and the sample.
- 1.2 Consider the population of all students at your college. You want to learn about total monthly entertainment expenses for a student.
 (a) Specify the population unit.
 (b) Specify the variable of interest.
 (c) Specify the statistical population.
- 1.3 ✓ Consider the population of persons living in Chicago. You want to learn about the proportion which are illegal aliens.
 (a) Specify the population unit. — person
 (b) Specify the variable of interest. illegal (Y)
 (c) Specify the statistical population. all living in Chicago
- 1.4 A student is asked to estimate the mean height of all male students on campus. She decides to use the heights of members of the basketball team because they are conveniently printed in the game program.
 (a) Identify the statistical population and the sample.
 (b) Comment on the selection of the sample.
 (c) How should a sample of males be selected?
- 1.5 ✓ Psychologists³ asked 46 golfers, after they played a round, to estimate the diameter of the hole on the green by visually selecting one of nine holes cut in a board.
 (a) Specify the population unit. any golfer
 (b) Specify the statistical population and sample. all possible estimates of hole diameter size
- 1.6 A phone survey in 2008⁴ of 1010 adults included a response to the number of leisure hours per week. Identify the population unit, statistical population, and sample.
- 1.7 ✓ It is often easy to put off doing an unpleasant task. At a Web site,⁵ persons can take a test and receive a score that determines if they have a serious problem with procrastination. Should the scores from people who take this test on-line be considered a random sample? Explain your reasoning.
- 1.8 A magazine that features the latest electronics and computer software for homes enclosed a short questionnaire on a postcard. Readers were asked to answer questions concerning their use and ownership of various software and hardware products, and to then send the card to the publisher. A summary of the results appeared in a later issue of the magazine that used the data to make statements such as 40% of readers have purchased program X. Identify a population and sample and comment on the representativeness of the sample. Are readers who have not purchased any new products mentioned in the questionnaire as likely to respond as those who have purchased?
- 1.9 Each year a local weekly newspaper gives out "Best of the City" awards in categories such as restaurant, deli, pastry shop, and so on. Readers are asked to fill in their favorites on a form enclosed in this free weekly paper and then send it to the publisher. The establishment receiving the most votes is declared the winner in its category. Identify the population and sample and comment on the representativeness of the sample.
- 1.10 Which of the following are anecdotal and which are based on sample?
 (a) Out of 200 students questioned, 40 admitted they lied regularly.
 (b) Bobbie says the produce at Market W is the freshest in the city.
 (c) Out of 50 persons interviewed at a shopping mall, 18 had made a purchase that day.
- 1.11 Which of the following are anecdotal and which are based on a sample?
 (a) Tom says he gets the best prices on electronics at the www.bestelc.com Internet site.

³J. Witt et al. "Putting to a bigger hole: Golf performance relates to perceived size," *Psychonomic Bulletin and Review* 15(3) (2008), pp. 581-586.

⁴Harris Interactive telephone survey (October 16-19, 2008).

⁵http://psychologytoday.psychtests.com/tests/procrastination_access.html

In the stem-and-leaf display, the column of first digits to the left of the vertical line is viewed as the stem, and the second digits as the leaves. Viewed sideways, it looks like a histogram with a cell width equal to 10. However, it is more informative than a histogram because the actual data points are retained. In fact, every observation can be recovered exactly from this stem-and-leaf display.

A stem-and-leaf display retains all the information in the leading digits of the data. When the leaf unit = .01, $3.5 | 0 \ 2 \ 3 \ 7 \ 8$ presents the data 3.50, 3.52, 3.53, 3.57, and 3.58. Leaves may also be two-digit at times. When the first leaf digit = .01, $.4 | 07 \ 13 \ 82 \ 90$ presents the data 0.407, 0.413, 0.482, and 0.490.

Further variants of the stem-and-leaf display are described in Exercises 2.25 and 2.26. This versatile display is one of the most applicable techniques of exploratory data analysis.

When the sample size is small or moderate, no information is lost with the stem-and-leaf diagram because you can see every data point. The major disadvantage is that, when the sample size is large, diagrams with hundreds of numbers in a row cannot be constructed in a legible manner.

Exercises

- 2.1 Cities must find better ways to dispose of solid waste. According to the Environmental Protection Agency, the composition of the 254 million tons of solid municipal waste created in 2007 was

Paper and paperboard	32.7%
Yard waste	12.8%
Food waste	12.5%
Plastics	12.1%
Metals	8.2%
Other materials	

- Determine the percentage of other materials in the solid waste. This category includes glass, wood, rubber, and so on.
- Create a Pareto chart.
- What percentage of the total solid waste is paper or paperboard? What percentage is from the top two categories? What percentage is from the top five categories?

- 2.2 Recorded here are the blood types of 40 persons who have volunteered to donate blood at a plasma center. Summarize the data in a frequency table. Include calculations of the relative frequencies.

O O A B A O A A A O
B O B O O A O O A A
A A AB A B A A O O A
O O A A A O A O O AB

2.3

A student at the University of Wisconsin surveyed 40 students in her dorm concerning their participation in extracurricular activities during the past week. The data on number of activities are

1 5 0 1 4 3 0 2 1 6 1 1 0 0
2 0 0 3 1 2 1 2 2 2 2 2 1 0
2 2 3 4 2 7 2 2 3 3 1 1

Present these data in a frequency table and in a relative frequency bar chart.

2.4

The number of automobile accidents reported per month helps to identify intersections that require improvement. Beginning January 2004 and ending November 2008, the number of crashes per month reported at an intersection near a university campus in Madison, Wisconsin, are

1 3 3 3 2 2 3 1 2 4 1 4
1 3 1 1 1 0 1 2 2 5 5 2
5 5 4 3 3 6 1 2 3 2 4 3
4 4 3 5 3 3 3 5 1 5 5 3
4 2 2 0 0 1 4 1 0 2 0

Present these data in a frequency table and in a relative frequency bar chart.

- 2.5 The following table shows how workers in one department get to work.

Mode of Transportation	Frequency
Drive alone	25
Car pool	3
Ride bus	7
Other	5

- (a) Calculate the relative frequency of each mode of transportation.
(b) Construct a pie chart.

- 2.6 Of the \$207 million raised by a major university's fund drive, \$117 million came from individuals and bequests, \$24 million from industry and business, and \$66 million from foundations and associations. Present this information in the form of a pie chart.

- 2.7 Data from one campus dorm on the number of burglaries are collected each week of the semester. These data are to be grouped into the classes 0-1, 2-3, 3-5, 6 or more. Both endpoints included. Explain where a difficulty might arise.

- 2.8 Data from one campus dorm, on the number of complaints about the dorm food are collected each week of the semester. These weekly counts are to be grouped into the classes 0-1, 2-3, 4-5, 7 or more. Both endpoints are included. Explain where a difficulty might arise.

- 2.9 A sample of persons will each be asked to give the number of their close friends. The responses are to be grouped into the following classes: 0, 1-3, 3-5, 6 or more. Left endpoint is included. Explain where difficulties might arise.

- 2.10 The weights of the players on the university football team (to the nearest pound) are to be grouped into the following classes: 160-175, 175-190, 190-205, 205-220, 220-235, 235 or more. The left endpoint is included but not the right endpoint. Explain where difficulties might arise.

- 2.11 On flights from San Francisco to Chicago, the number of empty seats are to be grouped into the following classes: 0-4, 5-9, 10-14, 15-19, more than 19.

Is it possible to determine from this frequency distribution the exact number of flights on which there were:

- (a) Fewer than 10 empty seats?
(b) More than 14 empty seats?
(c) At least 5 empty seats?
(d) Exactly 9 empty seats?
(e) Between 5 and 15 empty seats inclusively?

- 2.12 A major West Coast power company surveyed 50 customers who were asked to respond to the statement, "People should rely mainly on themselves to solve problems caused by power outages" with one of the following responses.

1. Definitely agree.
2. Somewhat agree.
3. Somewhat disagree.
4. Definitely disagree.

The responses are as follows:

4 2 1 3 3 2 4 2 1 1 2 2 2 2 1 3 4
1 4 4 1 3 2 4 1 4 3 3 1 1 1 2 1 1
4 4 4 4 4 1 2 2 2 4 4 4 1 3 4 2

Construct a frequency table.

- 2.13 A sample of 50 departing airline passengers at the main check-in counter produced the following number of bags checked through to final destinations.

0 1 2 2 1 2 1 2 3 0 1 0
1 1 0 1 3 0 1 2 1 1 1 2
1 2 2 1 2 0 0 2 2 1 1 1
1 1 1 1 2 0 1 3 0 1 2 1
1 3

- (a) Make a relative frequency line diagram.
(b) Comment on the pattern.
(c) What proportion of passengers who check in at the main counter fail to check any bags?

- 2.14 A person with asthma took measurements by blowing into a peak-flow meter on seven consecutive days.

429 425 471 422 432 444 454

Display the data in a dot diagram.

- 2.15 Before microwave ovens are sold, the manufacturer must check to ensure that the radiation coming through the door is below a specified safe limit. The amounts of radiation leakage (mW/cm^2) with the door closed from 25 ovens are as follows (courtesy of John Cryer):

15 9 18 10 5 12 8
5 8 10 7 2 1
5 3 5 15 10 15
9 8 18 1 2 11

Display the data in a dot diagram.

- 2.16 A campus area merchant recorded the number of bad checks received per month, for five months

4 5 4 7 6

Display the data in a dot diagram.

- 2.17 The city of Madison regularly checks the water quality at swimming beaches located on area lakes. The concentration of fecal coliforms, in number of colony forming units (CFU) per 100 ml of water, was measured on fifteen days during the summer at one beach.

180 1600 90 140 50 260 400 90
380 110 10 60 20 340 80

- (a) Make a dot diagram.
(b) Comment on the pattern and any unusual features.
(c) The city closes any swimming beach if a count is over 1350. What proportion of days, among the fifteen, was this beach closed?

2.18 Tornadoes kill many people every year in the United States. The yearly number of lives lost during the 59 years 1950 through 2008 are summarized in the following table.

Number of Deaths	Frequency
24 or less	2
25–49	20
50–74	18
75–99	7
100–149	6
150–199	2
200–249	1
250 or more	3
Total	59

- (a) Calculate the relative frequency for the intervals $[0, 25)$, $[25, 50)$ and so on where the right-hand endpoint is excluded. Take the last interval to be $[250, 550)$.
(b) Plot the relative frequency histogram. (*Hint:* Since the intervals have unequal widths, make the height of each rectangle equal to the relative frequency divided by the width of the interval.)
(c) What proportion of the years had 49 or fewer deaths due to tornadoes?
(d) Comment on the shape of the distribution.
- 2.19 A zoologist collected wild lizards in the Southwestern United States. Thirty lizards from the genus *Phrynosoma* were placed on a treadmill and their speed measured. The recorded speed (meters/second) is the fastest time to run a half meter. (Courtesy of K. Bonine.)

1.28 1.36 1.24 2.47 1.94 2.52 2.67 1.29
1.56 2.66 2.17 1.57 2.10 2.54 1.63 2.11
2.57 1.72 0.76 1.02 1.78 0.50 1.49 1.57
1.04 1.92 1.55 1.78 1.70 1.20

- (a) Construct a frequency distribution using the class intervals 0.45–0.90, 0.90–1.35, and so on, with the endpoint convention that the left endpoint is included and the right endpoint is excluded. Calculate the relative frequencies.
- (b) Make a histogram.

- 2.20 The United States Geological Survey maintains data on large earthquakes including those of magnitude greater than 6.0 in California. Through 2008, the ordered magnitudes of the 55 quakes are

6.1 6.1 6.1 6.1 6.1 6.2 6.2 6.2 6.2 6.3 6.3
 6.3 6.4 6.4 6.4 6.4 6.4 6.4 6.4 6.5 6.5 6.5
 6.5 6.5 6.6 6.6 6.6 6.7 6.7 6.7 6.8 6.8 6.8
 6.8 6.8 6.9 6.9 7.0 7.0 7.0 7.1 7.1 7.1 7.2
 7.2 7.2 7.2 7.3 7.3 7.3 7.3 7.4 7.8 7.8 7.9

Construct a histogram using equal-length intervals starting with $[6.0, 6.3]$ where the right-hand endpoint is included but not the left-hand endpoint.

- 2.21 Referring to Exercise 2.20, construct a density histogram using the intervals $[6.0, 6.3]$, $[6.3, 6.6]$, $[6.6, 6.9]$, $[6.9, 7.2]$, and $[7.2, 7.9]$.
- 2.22 The following data represent the scores of 40 students on a college qualification test (courtesy of R. W. Johnson).

162 171 138 145 144 126 145 162 174 178
 167 98 161 152 182 136 165 137 133 143
 184 166 115 115 95 190 119 144 176 135
 194 147 160 158 178 162 131 106 157 154

Make a stem-and-leaf display.

- 2.23 A federal government study of the oil reserves in Elk Hills, CA, included a study of the amount of iron present in the oil.

Amount of Iron (percent ash)

20	18	25	26	17
14	20	14	18	15
22	15	17	25	22
12	52	27	24	41
34	20	17	20	19
20	16	20	15	34
22	29	29	34	27
13	6	24	47	32
12	17	36	35	41
36	32	46	30	51

Make a stem-and-leaf display.

The following is a stem-and-leaf display with two-digit leaves. (The leading leaf digit = 10.0.)

1	
2	46 68 93
3	19 44 71 82 97
4	05 26 43 90
5	04 68
6	13

List the corresponding measurements.

- 2.25 If there are too many leaves on some stems in a stem-and-leaf display, we might double the number of stems. The leaves 0–4 could hang on one stem and 5–9 on the repeated stem. For the observations

193 198 200 202 203 203 205 205 206 207
 207 208 212 213 214 217 219 220 222 226 237

we would get the double-stem display

19	3
19	8
20	0233
20	556778
21	234
21	79
22	02
22	6
23	
23	7

$$100(.25) =$$

Table 8 The Lengths of Long-Distance Phone Calls in Minutes

1.6	1.7	1.8	1.8	1.9	2.1	2.5	3.0	3.0	4.4
4.5	4.5	5.9	7.1	7.4	7.5	7.7	8.6	9.3	9.5
12.7	15.3	15.5	15.9	15.9	16.1	16.5	17.3	17.5	19.0
19.4	22.5	23.5	24.0	31.7	32.8	43.5	53.3		

SOLUTION To determine the first quartile, we take $p = .25$ and calculate the product $38 \times .25 = 9.5$. Because 9.5 is not an integer, we take the next largest integer, 10. In Table 8, we see that the 10th ordered observation is 4.4 so the first quartile is $Q_1 = 4.4$ minutes.

We confirm that this observation has 10 values *at or below* it and 29 values *at or above* so that it does satisfy the conceptual definition of the first quartile.

For the median, we take $p = .5$ and calculate $38 \times .5 = 19$. Because this is an integer, we average the 19th and 20th smallest observations to obtain the median, $(9.3 + 9.5)/2 = 9.4$ minutes.

Next, to determine the third quartile, we take $p = .75$ and calculate $38 \times .75 = 28.5$. The next largest integer is 29, so the 29th ordered observation is the third quartile $Q_3 = 17.5$ minutes. More simply, we could mimic the calculation of the first quartile but now count down 10 observations starting with the largest value.

For the 90th percentile, we determine $38 \times .90 = 34.2$, which we increase to 35. The 90th percentile is 31.7 minutes. Only 10% of calls last 31.7 minutes or longer.

Exercises

2.28 Calculate the mean and median for each of the following data sets.

(a) 3 7 4 11 5

(b) 3 1 7 3 1

2.29 Calculate the mean and median for each of the following data sets.

(a) 2 5 1 4 3

(b) 26 30 38 32 26 31

(c) -1 2 0 1 4 -1 2

2.30 The height that bread rises may be one indicator of how light it will be. As a first step, before

modifying her existing recipe, a student cook measured the raise height (cm) on eight occasions:

6.3 6.9 5.7 5.4 5.6 5.5 6.6 6.5

Find the mean and median of the raised heights.

2.31 With reference to the water quality in Exercise 2.17:

(a) Find the sample mean.

(b) Does the sample mean or the median give a better indication of the water quality of a "typical" day? Why?

- 2.32 The monthly income in dollars for seven sales persons at a car dealership are

2450 2275 2425 4700 2650 2350 2475

- Calculate the mean and median salary.
- Which of the two is preferable as a measure of center and why?

- 2.33 Records show that in Las Vegas, NV, the normal daily maximum temperature ($^{\circ}\text{F}$) for each month starting in January is

56 62 68 77 87 99 105 102 95 82 66 57

Verify that the mean of these figures is 79.67. Comment on the claim that the daily maximum temperature in Las Vegas averages a pleasant 79.67.

- 2.34 A major wine producer reported sales (in hundreds of cases) for two-week periods during one summer:

85 82 77 83 80 77 94

Obtain the sample mean and median.

- 2.35 With reference to the radiation leakage data given in Exercise 2.15:

- Calculate the sample mean.
- Which gives a better indication of the amount of radiation leakage, the sample mean or the median?

- 2.36 Recent crime reports on the number of aggravated assaults at each of the 27 largest universities reporting for the year are summarized in the computer output

Descriptive Statistics: AggAslt

Variable	N	Mean	Median	StDev
AggAslt	27	10.30	10.00	7.61

Locate two measures of center tendency, or location, and interpret the values.

- 2.37 The weights (oz) of nineteen babies born in Madison, Wisconsin, are summarized in the computer output

Descriptive Statistics: Weight

Variable	N	Mean	Median	StDev
Weight	19	118.05	117.00	15.47

Locate two measures of center tendency, or location, and interpret the values.

- 2.38 With reference to the extracurricular activities data in Exercise 2.3, obtain the

- sample mean.
- sample median.
- Comment on the effect of a large observation.

- 2.39 With reference to the number of returns in Example 3, obtain the sample (a) mean and (b) median.

- 2.40 Old Faithful, the most famous geyser in Yellowstone Park, had the following durations (measured in seconds) in six consecutive eruptions:

240 248 113 268 117 253

- Find the sample median.
- Find the sample mean.

- 2.41 Loss of calcium is a serious problem for older women. To investigate the amount of loss, a researcher measured the initial amount of bone mineral content in the radius bone of the dominant hand of elderly women and then the amount remaining after one year. The differences, representing the loss of bone mineral content, are given in the following table (courtesy of E. Smith).

8	7	13	3	6
4	8	6	3	4
0	1	11	7	1
8	6	12	13	10
9	11	3	2	9
7	1	16	3	2
10	15	2	5	8
17	8	2	5	5

- Find the sample mean.
- Does the sample mean or the median give a better indication of the amount of mineral loss?

- 2.42 Physical education researchers interested in the development of the overarm throw measured the horizontal velocity of a thrown ball

at the time of release. The results for first-grade children (in feet/second) (courtesy of L. Halverson and M. Robertson) are

Males

54.2	39.6	52.3	48.4	35.9	30.4	25.2	45.4	48.9	48.9
45.8	44.0	52.5	48.3	59.9	51.7	38.6	39.1	49.9	38.3

Females

30.3	43.0	25.7	26.7	27.3	31.9	53.7	32.9	19.4	23.7
23.3	23.3	37.8	39.5	33.5	30.4	28.5			

- Find the sample median for males.
- Find the sample median for females.
- Find the sample median for the combined set of males and females.

2.43 On opening day one season, 10 major league baseball games were played and they lasted the following numbers of minutes.

167 211 187 176 170 158 198 218 145 232

Find the sample median.

2.44 If you were to use the data on the length of major league baseball games in Exercise 2.43 to estimate the total amount of digital memory needed to film another 10 major league baseball games, which is the more meaningful description, the sample mean or the sample median? Explain.

2.45 The following measurements of the diameters (in feet) of Indian mounds in southern Wisconsin were gathered by examining reports in the *Wisconsin Archeologist* (courtesy of J. Williams).

22 24 24 30 22 20 28 30 24 34 36 15 37

- Create a dot diagram.
- Calculate the mean and median and then mark these on the dot diagram.
- Calculate the quartiles.

2.46 With reference to Exercise 2.3, calculate the quartiles.

2.47 Refer to the data of college qualification test scores given in Exercise 2.22.

- Find the median.
- Find Q_1 and Q_3 .

2.48 A large mail-order firm employs numerous persons to take phone orders. Computers on which orders are entered also automatically collect data on phone activity. One variable useful for planning staffing levels is the number of calls per shift handled by each employee. From the data collected on 25 workers, calls per shift were (courtesy of Land's End)

118	118	57	92	127	109	96	68	73
69	106	91	93	94	102	105	100	104
80	50	96	82	72	108	73		

Calculate the sample mean.

2.49 With reference to Exercise 2.48, calculate the quartiles.

2.50 The speedy lizard data, from Exercise 2.19, are

1.28	1.36	1.24	2.47	1.94	2.52	2.67	1.29
1.56	2.66	2.17	1.57	2.10	2.54	1.63	2.11
2.57	1.72	0.76	1.02	1.78	0.50	1.49	1.57
1.04	1.92	1.55	1.78	1.70	1.20		

- Find the sample median, first quartile, and third quartile.
- Find the sample 90th percentile.

2.51 With reference to the water quality data in Exercise 2.17:

- Find the sample median, first quartile, and third quartile.
- Find the sample 90th percentile.

2.52 Some properties of the mean and median.

- If a fixed number c is added to all measurements in a data set, then the mean of the new measurements is $c +$ (the original mean).
- If all measurements in a data set are multiplied by a fixed number d , then the mean of the new measurements is $d \times$ (the original mean).

Figure 13 displays the amount of reflected light in the near-infrared band as recorded by satellite when flying over forest areas and urban areas, respectively. Because high readings tend to correspond to forest and low readings to urban areas, the readings have proven useful in classifying unknown areas.

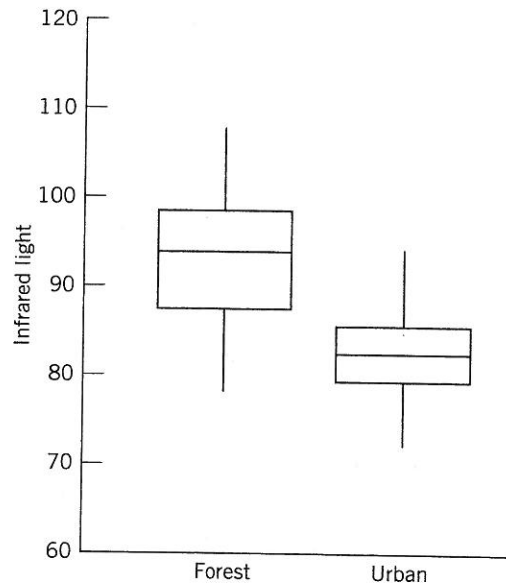


Figure 13 Boxplots of near-infrared light reflected from forest and urban areas.

Example 17 Comparing Boxplots for the Reflected Light Data

Refer to the boxplots for the amount of reflected near-infrared light in Figure 13.

- Do forests or urban areas produce the largest median reflected light?
- Which has the largest interquartile range, IQR?

- SOLUTION**
- It is clear that forests have the largest median. Its median is over 10 units higher than that of urban areas.
 - The height of the box is the IQR. The IQR of the forest data is over twice that of the IQR for urban areas.

Exercises

2.57 For the data set

7 2 3

- Calculate the deviations $(x - \bar{x})$ and check to see that they add up to 0.

- Calculate the sample variance and the standard deviation.

2.58 Repeat (a) and (b) of Exercise 2.57 for the data set

4 9 2

- 2.59 For the data set 8 6 14 4:
- Calculate the deviations $(x - \bar{x})$ and check to see that they add up to 0.
 - Calculate the variance and the standard deviation.

- 2.60 Repeat (a) and (b) of Exercise 2.59 for the data set

2.5 1.7 2.1 1.5 1.7

- 2.61 For the data of Exercise 2.57, calculate s^2 by using the alternative formula.

- 2.62 For the data of Exercise 2.59, calculate s^2 by using the alternative formula.

- 2.63 For each data set, calculate s^2 .

- 1 4 3 2 2
- 2 1 -1 -3 0 -2
- 9 8 8 9 8 8 9

- 2.64 The monthly rents for 7 one-bedroom apartments located in one area of the city, are

625 740 805 670 705 740 870

- Give two possible factors that may contribute to variation in the monthly rents.

Calculate

- The sample variance.
- The sample standard deviation.

- 2.65 Find the standard deviation of the measurements of diameters given in Exercise 2.45.

- 2.66 A campus area merchant recorded the number of bad checks received per month, for five months

4 5 4 7 6

Calculate:

- The sample variance.
- The sample standard deviation.

- 2.67 The city of Madison regularly checks the quality of water at swimming beaches located on area lakes. Fifteen times the concentration of fecal coliforms, in number of colony forming units (CFU) per 100 ml of water, was measured during the summer at one beach.

180 1600 90 140 50 260 400 90
380 110 10 60 20 340 80

- Calculate the sample variance.
- Calculate the sample standard deviation.
- One day, the water quality was bad—the reading was 1600 CFU—and the beach was closed. Drop this value and calculate the sample standard deviation for the days where the water quality was suitable for swimming. Comment on the change.

- 2.68 With reference to the radiation leakage data given in Exercise 2.15, calculate:

- The sample variance.
- The sample standard deviation.

- 2.69 With reference to the data on the length of 10 major league baseball games in Exercise 2.43:

- Find the sample mean.
- Find the sample variance.
- Find the sample standard deviation.

- 2.70 With reference to checked bags in Exercise 2.13,

- Find the sample mean.
- Find the sample standard deviation.

- 2.71 A sample of seven compact discs at the music store stated the performance times as lasting the following numbers of minutes for Beethoven's Ninth Symphony.

66.9 66.2 71.0 68.6 65.4 68.4 71.9

- Find the sample median.
- Find the sample mean.
- Find the sample standard deviation.

- 2.72 Recent crime reports on the number of aggravated assaults at each of the 27 largest universities reporting for the year are summarized in the computer output.

Descriptive Statistics: AggAslt

Variable	N	Mean	Median	StDev
AggAslt	27	10.30	10.00	7.61

Variable	Minimum	Maximum	Q1	Q3
AggAslt	0.00	29.00	5.00	14.00

- Locate a measure of variation and also calculate the sample variance.

- (b) Calculate the interquartile range and interpret this value.
- (c) Give a value for a standard deviation that would correspond to greater variation in the numbers of aggravated assaults.
- 2.73 The weights (oz) of nineteen babies born in Madison, Wisconsin, are summarized in the computer output.

Descriptive Statistics: Weight

Variable	N	Mean	Median	StDev
Weight	19	118.05	117.00	15.47

Variable	Minimum	Maximum	Q1	Q3
Weight	89.00	144.00	106.00	131.00

- (a) Locate a measure of variation and also calculate the sample variance.
- (b) Calculate the interquartile range and interpret this value.
- (c) Give a value for a standard deviation that would correspond to smaller variation in the weights.
- 2.74 *Some properties of the standard deviation.*
1. If a fixed number c is added to all measurements in a data set, the deviations $(x - \bar{x})$ remain unchanged (see Exercise 2.52). Consequently, s^2 and s remain unchanged.
 2. If all measurements in a data set are multiplied by a fixed number d , the deviations $(x - \bar{x})$ get multiplied by d . Consequently, s^2 gets multiplied by d^2 , and s by $|d|$. (Note: The standard deviation is never negative.)

Verify these properties for the data set

5 9 9 8 10 7

taking $c = 4$ in property (1) and $d = 2$ in (2).

- 2.75 For the data set of Exercise 2.22, calculate the interquartile range.
- 2.76 For the extracurricular data of Exercise 2.3, calculate the interquartile range.
- 2.77 Should you be surprised if the range is larger than twice the interquartile range? Explain.

2.78 Calculations with the test scores data of Exercise 2.22 give $\bar{x} = 150.125$ and $s = 24.677$.

- (a) Find the proportion of the observations in the intervals $\bar{x} \pm 2s$ and $\bar{x} \pm 3s$.
- (b) Compare your findings in part (a) with those suggested by the empirical guidelines for bell-shaped distributions.
- 2.79 Refer to the data on bone mineral content in Exercise 2.41.

- (a) Calculate $\bar{x}$ and s .
- (b) Find the proportion of the observations that are in the intervals $\bar{x} \pm s$, $\bar{x} \pm 2s$, and $\bar{x} \pm 3s$.
- (c) Compare the results of part (b) with the empirical guidelines.
- 2.80 Refer to the data on lizards in Exercise 2.19.
- (a) Calculate $\bar{x}$ and s .
- (b) Find the proportion of the observations that are in the intervals $\bar{x} \pm s$, $\bar{x} \pm 2s$, and $\bar{x} \pm 3s$.
- (c) Compare the results of part (b) with the empirical guidelines.

2.81 Refer to the data on number of returns in Example 3.

- (a) Calculate $\bar{x}$ and s .
- (b) Find the proportions of the observations that are in the intervals $\bar{x} \pm s$, $\bar{x} \pm 2s$, and $\bar{x} \pm 3s$.
- (c) Compare the results of part (b) with the empirical guidelines.

2.82 Sample z score. The z scale (or standard scale) measures the position of a data point relative to the mean and in units of the standard deviation. Specifically,

$$z \text{ value of a measurement} = \frac{\text{Measurement} - \bar{x}}{s}$$

When two measurements originate from different sources, converting them to the z scale helps to draw a sensible interpretation of their relative magnitudes. For instance, suppose a student scored 65 in a math course and 72 in a history course. These (raw) scores tell little about the student's performance. If the class averages and standard deviations were $\bar{x} = 60$, $s = 20$ in math and $\bar{x} = 78$, $s = 10$ in history, this student's

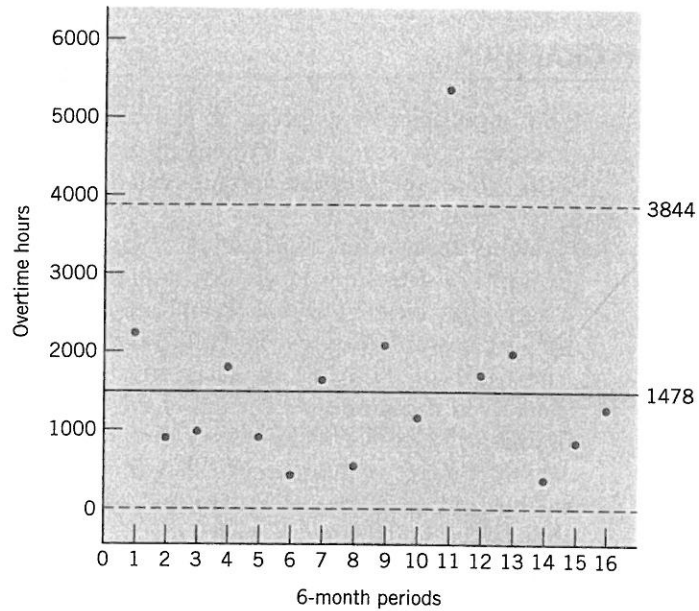


Figure 16 Control chart for extraordinary event overtime hours.

Exercises

- 2.92 Make a time plot of the phone call data in Exercise 2.48 and comment on the statistical control.

- 2.93 A city department has introduced a quality improvement program and has allowed employees to get credit for overtime hours when attending meetings of their quality groups. The total number of overtime meeting hours for each of the 26 pay periods in one year by row were

30	215	162	97	194	163	60	41	100
43	96	69	80	42	162	75	95	65
57	131	54	114	64	114	38	140	

Make a time plot of the overtime meeting hours data.

- 2.94 Make a control chart for the data referred to in Exercise 2.92 and comment.

- 2.95 Make a control chart for the data in Exercise 2.93 and comment.

$$\bar{x} \pm 2s$$

- 2.96 The exchange rate between the United States and Canada can be stated as the number of Canadian dollars that can be purchased with \$1. The official values for the year are

Year	1992	1993	1994	1995	1996	1997
Exchange rate	1.21	1.29	1.37	1.37	1.36	1.38
	1998	1999	2000	2001	2002	2003
	1.48	1.49	1.48	1.55	1.57	1.40
	2004	2005	2006	2007	2008	
	1.30	1.21	1.13	1.07	1.07	

Is this exchange rate in statistical control? Make a time plot and comment.

- 2.97 Make a control chart of the data in Exercise 2.96 and comment.

- 2.98 Make a control chart for the extraordinary event overtime data in Example 18 after removing the outlier identified in that example. You need to recalculate the mean and standard deviation.