Exercises

Section 3.1

1. **Details** Here are histograms for four manufactured sets of numbers. The histograms look rather different, but all four sets 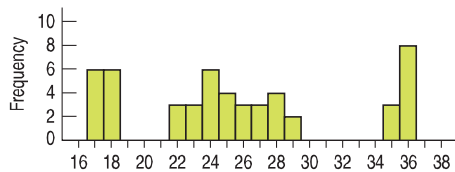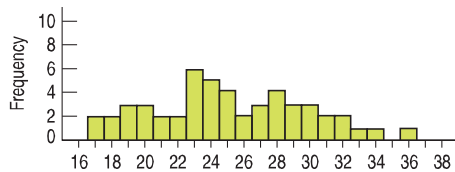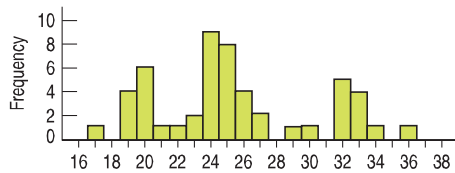have the same 5-number summary, so the boxplots for all four sets are identical to the one shown. Using these plots as examples, explain some limitations in using boxplots to describe the shape of a distribution.
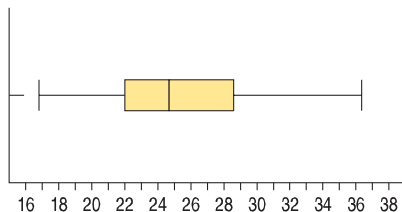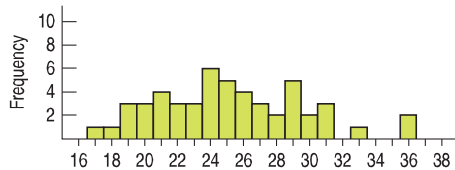
2. **Opposites** In a way, boxplots are the opposite of histograms. A histogram divides the number line into equal intervals and displays the number of data values in each interval. A boxplot divides the *data* into equal parts and displays the portion of the number line each part covers. These two plots display the number of incarcerated prisoners in each state as of June 2006.
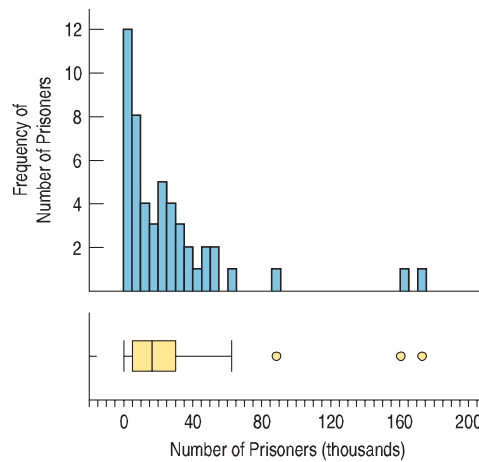
a) Explain how you could tell, by looking at a boxplot, where the tallest bars on the histogram would be located.
b) Explain how both the boxplot and the histogram can indicate a skewed distribution.
c) Identify one feature of the distribution that the histogram shows but a boxplot does not.
d) Identify one feature of the distribution that the boxplot shows but the histogram does not.

Section 3.2

3. **Outliers** The 5-number summary for the run times in minutes of the 150 highest grossing movies of 2010 looks like this:

Min	QI	Med	Q3	Max
43	98	104.5	116	160

Are there any outliers in these data? How can you tell?

4. **Adoptions** The following table shows the mean and median ages in years of children adopted in different states in 2010. In all 50 states, the mean age is higher than the median. What might that imply about the shape of the distribution of ages of adopted children in each state?

State	Mean Age	Median Age
Alabama	7.3	6.5
Alaska	7.1	6.6
Arizona	6.4	5.1
Arkansas	6.0	4.5
California	5.9	4.6
Colorado	5.9	4.7

Section 3.3

5. Adoptions II Here is a histogram showing the total number of adoptions in each of the 50 states and the District of Columbia.
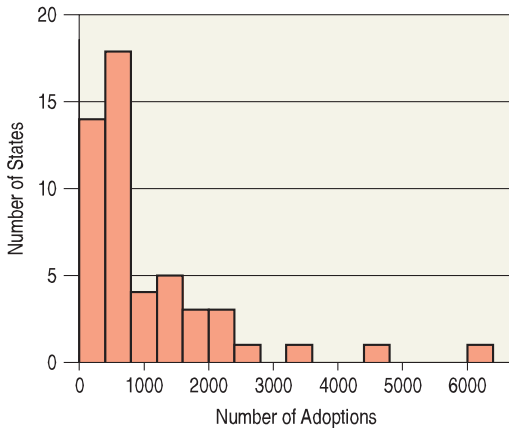
Would you expect the mean number of adoptions or the median number of adoptions to be higher? Why?

6. Test score centers After entering the test scores from her Statistics class of 25 students, the instructor calculated the mean and the median of the scores. Upon checking, she discovered that she had entered the top score as 46, but it should have been 56. When she corrects this score, how will the mean and median be affected?

Section 3.4

7. Test score spreads The professor in question 6 had also calculated the IQR and standard deviation of the test scores before realizing her mistake. (She had entered the top score as 46 but it was actually 56.) What effect will correcting the error have on the IQR and the standard deviation?

8. Fuel economy The boxplot shows the fuel economy ratings for 67 model year 2012 subcompact cars. Some summary statistics are also provided. The extreme outlier is the **Mitsubishi i-MiEV**, an electric car whose electricity usage is *equivalent* to 112 miles per gallon.
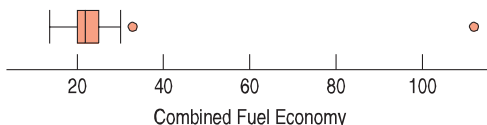

Mean	sd	Min	Q1	Med	Q3	Max	n
23.76	11.87	14	20	22	25	112	67

If that electric car is removed from the data set, how will the standard deviation be affected? The IQR?

Chapter Exercises

**9. Histogram** Find a histogram that shows the distribution of a variable in a newspaper, a magazine, or the Internet.
  a) Does the article identify the W's?
  b) Discuss whether the display is appropriate.
  c) Discuss what the display reveals about the variable and its distribution.
  d) Does the article accurately describe and interpret the data? Explain.

10. Not a histogram Find a graph other than a histogram that shows the distribution of a quantitative variable in a newspaper, a magazine, or the Internet.
  a) Does the article identify the W's?
  b) Discuss whether the display is appropriate for the data.
  c) Discuss what the display reveals about the variable and its distribution.
  d) Does the article accurately describe and interpret the data? Explain.

11. In the news Find an article in a newspaper, a magazine, or the Internet that discusses an "average."
  a) Does the article discuss the W's for the data?
  b) What are the units of the variable?
  c) Is the average used the median or the mean? How can you tell?
  d) Is the choice of median or mean appropriate for the situation? Explain.

12. In the news II Find an article in a newspaper, a magazine, or the Internet that discusses a measure of spread.
  a) Does the article discuss the W's for the data?
  b) What are the units of the variable?
  c) Does the article use the range, IQR, or standard deviation?
  d) Is the choice of measure of spread appropriate for the situation? Explain.

13. Thinking about shape Would you expect distributions of these variables to be uniform, unimodal, or bimodal? Symmetric or skewed? Explain why.
  a) The number of speeding tickets each student in the senior class of a college has ever had.
  b) Players' scores (number of strokes) at the U.S. Open golf tournament in a given year.
  c) Weights of female babies born in a particular hospital over the course of a year.
  d) The length of the average hair on the heads of students in a large class.

14. More shapes Would you expect distributions of these variables to be uniform, unimodal, or bimodal? Symmetric or skewed? Explain why.
  a) Ages of people at a Little League game.
  b) Number of siblings of people in your class.
  c) Pulse rates of college-age males.
  d) Number of times each face of a die shows in 100 tosses.

T **15. Sugar in cereals** The histogram displays the sugar content (as a percent of weight) of 49 brands of breakfast cereals.
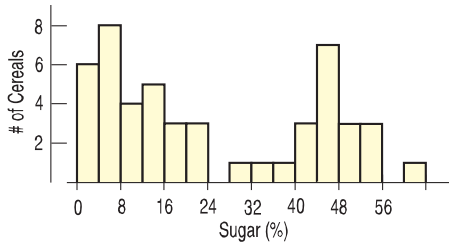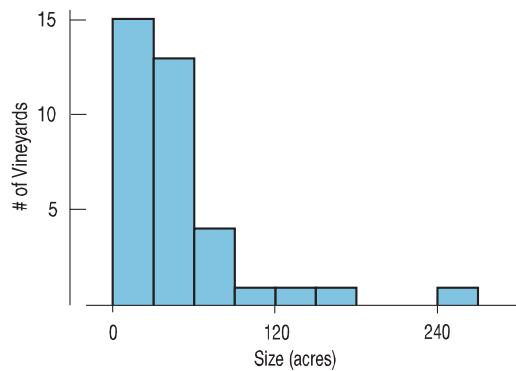

a) Describe this distribution.
b) What do you think might account for this shape?

T **16. Singers** The display shows the heights of some of the singers in a chorus, collected so that the singers could be positioned on stage with shorter ones in front and taller ones in back.
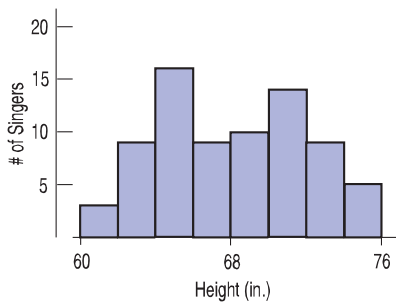
a) Describe the distribution.
b) Can you account for the features you see here?

T **17. Vineyards** The histogram shows the sizes (in acres) of 36 vineyards in the Finger Lakes region of New York.

a) Approximately what percentage of these vineyards are under 60 acres?
b) Write a brief description of this distribution (shape, center, spread, unusual features).

T **18. Run times** One of the authors collected the times (in minutes) it took him to run 4 miles on various courses during a 10-year period. Here is a histogram of the times.
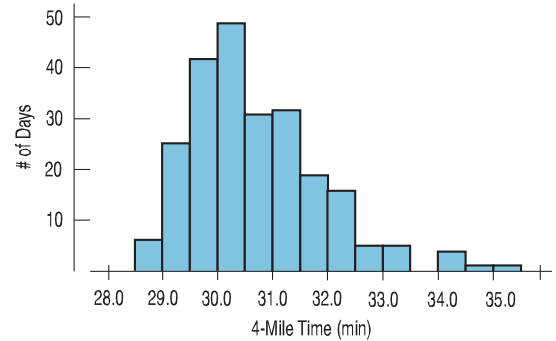
Describe the distribution and summarize the important features. What is it about running that might account for the shape you see?

T **19. Heart attack stays** The histogram shows the lengths of hospital stays (in days) for all the female patients admitted to hospitals in New York during one year with a primary diagnosis of acute myocardial infarction (heart attack).
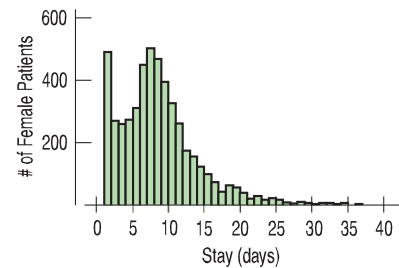
a) From the histogram, would you expect the mean or median to be larger? Explain.
b) Write a few sentences describing this distribution (shape, center, spread, unusual features).
c) Which summary statistics would you choose to summarize the center and spread in these data? Why?

T **20. E-mails** A university teacher saved every e-mail received from students in a large Introductory Statistics class during an entire term. He then counted, for each student who had sent him at least one e-mail, how many e-mails each student had sent.
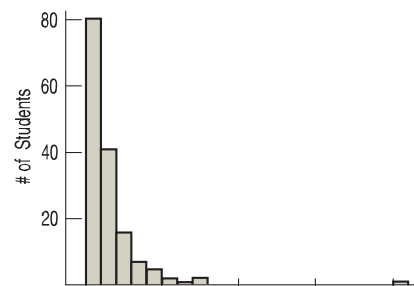
a) From the histogram, would you expect the mean or the median to be larger? Explain.

b) Write a few sentences describing this distribution (shape, center, spread, unusual features).

c) Which summary statistics would you choose to summarize the center and spread in these data? Why?

21. Super Bowl points 2010 How many points do football teams score in the Super Bowl? Here are the total numbers of points scored by both teams in each of the first 45 Super Bowl games:

45, 47, 23, 30, 29, 27, 21, 31, 22, 38, 46, 37, 66, 50, 37, 47, 44, 47, 54, 56, 59, 52, 36, 65, 39, 61, 69, 43, 75, 44, 56, 55, 53, 39, 41, 37, 69, 61, 45, 31, 46, 31, 50, 48, 56

a) Find the median.

b) Find the quartiles.

c) Write a description based on the 5-number summary.

22. Super Bowl wins 2010 In the Super Bowl, by how many points does the winning team outscore the losers? Here are the winning margins for the first 45 Super Bowl games:

25, 19, 9, 16, 3, 21, 7, 17, 10, 4, 18, 17, 4, 12, 17, 5, 10, 29, 22, 36, 19, 32, 4, 45, 1, 13, 35, 17, 23, 10, 14, 7, 15, 7, 27, 3, 27, 3, 3, 11, 12, 3, 4, 14, 6

a) Find the median.

b) Find the quartiles.

c) Write a description based on the 5-number summary.

23. Summaries Here are costs of nine compact refrigerators rated very good or excellent by *Consumer Reports* on their website www.consumerreports.org.

$ 150 150 160 180 150 140 120 130 120

Find these statistics by *hand* (no calculator!):

a) mean

b) median and quartiles

c) range and IQR

24. Tornadoes 2011 Here are the annual numbers of deaths from tornadoes in the United States from 1998 through 2011 (Source: NOAA):

130 94 40 40 555 54 35 38 67 81 125 21 45 544

Find these statistics by *hand* (no calculator!):

a) mean

b) median and quartiles

c) range and IQR

25. Mistake A clerk entering salary data into a company spreadsheet accidentally put an extra "0" in the boss's salary, listing it as $2,000,000 instead of $200,000. Explain how this error will affect these summary statistics for the company payroll:

a) measures of center: median and mean.

b) measures of spread: range, IQR, and standard deviation.

26. Sick days During contract negotiations, a company seeks to change the number of sick days employees may take, saying that the annual "average" is 7 days of absence per employee. The union negotiators counter that the

"average" employee misses only 3 days of work each year. Explain how both sides might be correct, identifying the measure of center you think each side is using and why the difference might exist.

27. Standard deviation I For each lettered part, a through c, examine the two given sets of numbers. Without doing any calculations, decide which set has the larger standard deviation and explain why. Then check by finding the standard deviations *by hand*.

	Set 1	Set 2
a)	3, 5, 6, 7, 9	2, 4, 6, 8, 10
b)	10, 14, 15, 16, 20	10, 11, 15, 19, 20
c)	2, 6, 6, 9, 11, 14	82, 86, 86, 89, 91, 94

28. Standard deviation II For each lettered part, a through c, examine the two given sets of numbers. Without doing any calculations, decide which set has the larger standard deviation and explain why. Then check by finding the standard deviations *by hand*.

	Set 1	Set 2
a)	4, 7, 7, 7, 10	4, 6, 7, 8, 10
b)	100, 140, 150, 160, 200	10, 50, 60, 70, 110
c)	10, 16, 18, 20, 22, 28	48, 56, 58, 60, 62, 70

T **29. Pizza prices** The histogram shows the distribution of the prices of plain pizza slices (in $) for 156 weeks in Dallas, TX.
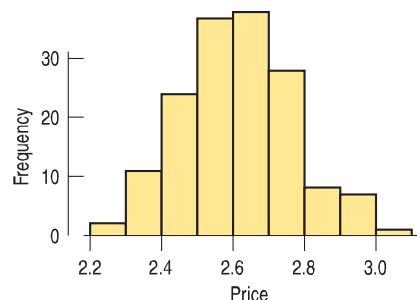
Which summary statistics would you choose to summarize the center and spread in these data? Why?

T **30. Neck size** The histogram shows the neck sizes (in inches) of 250 men recruited for a health study in Utah.
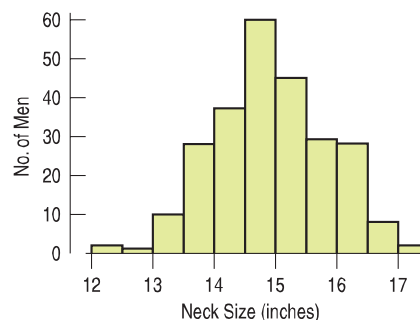
Which summary statistics would you choose to summarize the center and spread in these data? Why?

31. Pizza prices again Look again at the histogram of the pizza prices in Exercise 29.

a) Is the mean closer to $2.40, $2.60, or $2.80? Why?

b) Is the standard deviation closer to $0.15, $0.50, or $1.00? Explain.

32. Neck sizes again Look again at the histogram of men's neck sizes in Exercise 30.

a) Is the mean closer to 14, 15, or 16 inches? Why?

b) Is the standard deviation closer to 1 inch, 3 inches, or 5 inches? Explain.

33. Movie lengths 2010 The histogram shows the running times in minutes of the 150 top-grossing feature films released in 2010.
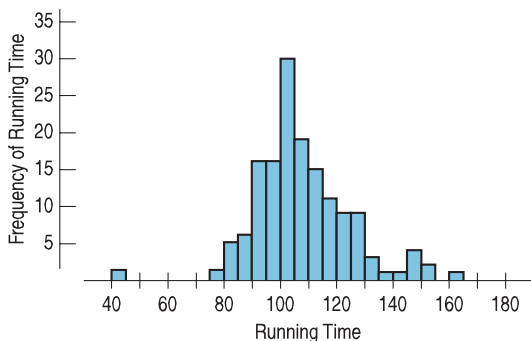

a) You plan to see a movie this weekend. Based on these movies, how long do you expect a typical movie to run?

b) Would you be surprised to find that your movie ran for $2\frac{1}{2}$ hours (150 minutes)?

c) Which would you expect to be higher: the mean or the median run time for all movies? Why?

34. Golf drives 2011 The display shows the average drive distance (in yards) for 186 professional golfers on the men's PGA tour in 2011.
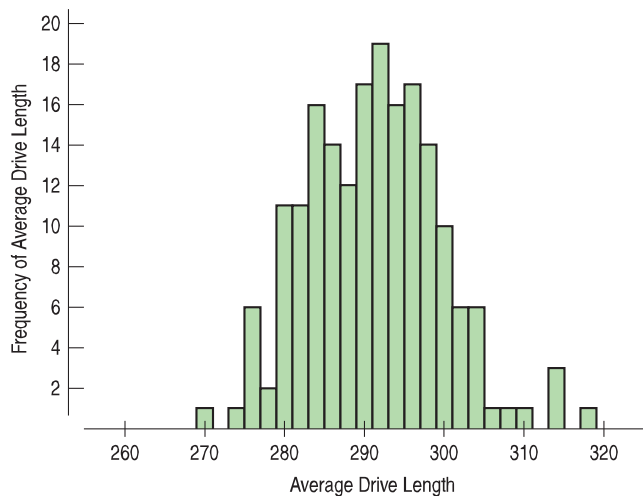
a) Describe this distribution.

b) Approximately what proportion of professional male golfers drive, on average, 280 yards or less?

c) Estimate the mean by examining the histogram.

d) Do you expect the mean to be smaller than, approximately equal to, or larger than the median? Why?

35. Movie lengths II 2010 Exercise 33 looked at the running times of movies released in 2010. The standard deviation of these running times is 16.6 minutes, and the quartiles are $Q_1 = 98$ minutes and $Q_3 = 116$ minutes.

a) Write a sentence or two describing the spread in running times based on
  i) the quartiles.
  ii) the standard deviation.

b) Do you have any concerns about using either of these descriptions of spread? Explain.

36. Golf drives II 2011 Exercise 34 looked at distances PGA golfers can hit the ball. The standard deviation of these average drive distances is 8.4 yards, and the quartiles are $Q_1 = 285.2$ yards and $Q_3 = 297.5$ yards.

a) Write a sentence or two describing the spread in distances based on
  i) the quartiles.
  ii) the standard deviation.

b) Do you have any concerns about using either of these descriptions of spread? Explain.

37. Movie earnings The histogram shows total gross earnings (in millions of dollars) of major release movies in 2011.
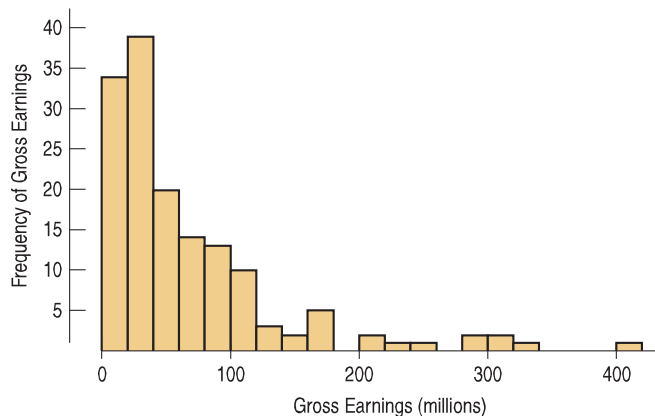
An industry publication reports that the average movie makes 41.7 million, but a watchdog group concerned with rising ticket prices says that the average earnings are 66.9 million. What statistic do you think each group is using? Explain.

38. Cold weather A meteorologist preparing a talk about global warming compiled a list of weekly low temperatures (in degrees Fahrenheit) he observed at his southern Florida home last year. The coldest temperature for any week was

value of 2°. Assuming that he correctly listed all the other temperatures, explain how this error will affect these summary statistics:

a) measures of center: mean and median.
b) measures of spread: range, IQR, and standard deviation.

39. Payroll A small warehouse employs a supervisor at $1200 a week, an inventory manager at $700 a week, six stock boys at $400 a week, and four drivers at $500 a week.

a) Find the mean and median wage.
b) How many employees earn more than the mean wage?
c) Which measure of center best describes a typical wage at this company: the mean or the median?
d) Which measure of spread would best describe the payroll: the range, the IQR, or the standard deviation? Why?

T 40. Singers The frequency table shows the heights (in inches) of 130 members of a choir.

Height	Count	Height	Count
60	2	69	5
61	6	70	11
62	9	71	8
63	7	72	9
64	5	73	4
65	20	74	2
66	18	75	4
67	7	76	1
68	12		

a) Find the median and IQR.
b) Find the mean and standard deviation.
c) Display these data with a histogram.
d) Write a few sentences describing the distribution.

41. Gasoline 2011 In October 2011, 16 gas stations in eastern Wisconsin posted these prices for a gallon of regular gasoline:

|			
3.43	3.46	3.43	3.59
3.65	3.63	3.62	3.65
3.66	3.31	3.35	3.42
3.41	3.46	3.47	3.48

a) Make a stem-and-leaf display of these gas prices. Use split stems; for example, use two 3.3 stems—one for prices between $3.30 and $3.34 and the other for prices from $3.35 to $3.39.
b) Describe the shape, center, and spread of this distribution.
c) What unusual feature do you see?

42. The Great One During his 20 seasons in the NHL, Wayne Gretzky scored 50% more points than anyone who ever played professional hockey. He accomplished this amazing feat while playing in 280 fewer games than Gordie Howe, the previous record holder. Here are the number of games Gretzky played during each season:

79, 80, 80, 80, 74, 80, 80, 79, 64, 78, 73, 78, 74, 45, 81, 48, 80, 82, 82, 70

a) Create a stem-and-leaf display for these data, using split stems.
b) Describe the shape of the distribution.
c) Describe the center and spread of this distribution.
d) What unusual feature do you see? What might explain this?

43. States The stem-and-leaf display shows populations of the 50 states, in millions of people, according to the 2010 census.

```
0 1111111111122223333334444
0 555556666667789
1 000233
1 99
2 
2 5
3 
3 7
(2|5 means 25)
```
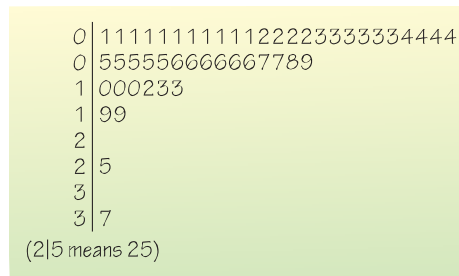
a) What measures of center and spread are most appropriate?
b) Without doing any calculations, which must be larger: the median or the mean? Explain how you know.
c) From the stem-and-leaf display, find the median and the interquartile range.
d) Write a few sentences describing this distribution.

44. Wayne Gretzky In Exercise 42, you examined the number of games played by hockey great Wayne Gretzky during his 20-year career in the NHL.

a) Would you use the median or the mean to describe the center of this distribution? Why?
b) Find the median.
c) Without actually finding the mean, would you expect it to be higher or lower than the median? Explain.

45. A-Rod 2010 Alex Rodriguez (known to fans as A-Rod) was the youngest player ever to hit 500 home runs. Here is a stem-and-leaf display of the number of home runs hit by A-Rod during the 1994–2010 seasons. (www.baseball-reference.com/players/r/rodrial01.shtml.) Describe the distribution, mentioning its shape and any unusual features.

```
0 05
1 
2 3
3 005566
4 12278
5 247
(5 2 means 52)
```
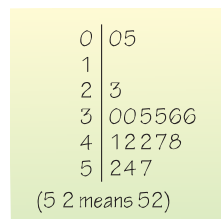
46. Bird species 2010 The Cornell Lab of Ornithology holds an annual Christmas Bird Count (www.birdsource.org), in which bird watchers at various locations around the country see how many different species of birds they can spot. Here are some of the counts reported from sites in Texas during the 2010 event:

```
150  216  177  150
166  156  159  160
164  169
175  150  178  183
199  154  164  203
158  231
```

a) Create a stem-and-leaf display of these data.
b) Write a brief description of the distribution. Be sure to discuss the overall shape as well as any unusual features.

T 47. Hurricanes 2010 The following data give the number of hurricanes classified as major hurricanes in the Atlantic Ocean each year from 1944 through 2010, as reported by NOAA (www.nhc.noaa.gov):
3, 3, 1, 2, 4, 3, 8, 5, 3, 4, 2, 6, 2, 2, 5, 2, 2, 7, 1, 2, 6, 1, 3, 1, 0, 5, 2, 1, 0, 1, 2, 3, 2, 1, 2, 2, 2, 3, 1, 1, 1, 3, 0, 1, 3, 2, 1, 2, 1, 1, 0, 5, 6, 1, 3, 5, 3, 4, 2, 3, 6, 7, 2, 2, 5, 2, 5

a) Create a dotplot of these data.
b) Describe the distribution.

T 48. Horsepower Create a stem-and-leaf display for these horsepowers of autos reviewed by *Consumer Reports* one year, and describe the distribution:

```
155  103  130   80   65
142  125  129   71   69
125  115  138   68   78
150  133  135   90   97
 68  105   88  115  110
 95   85  109  115   71
 97  110   65   90
 75  120   80   70
```

49. A-Rod again 2010 Students were asked to make a histogram of the number of home runs hit by Alex Rodriguez from 1994 to 2010 (see Exercise 45). One student submitted the following display (he noted that A Rod had hit no home runs in 1994):
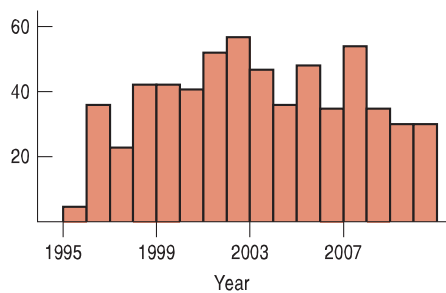
a) Comment on this graph.
b) Create your own histogram of the data.

50. Return of the birds 2010 Students were given the assignment to make a histogram of the data on bird counts reported in Exercise 46. One student submitted the following display:
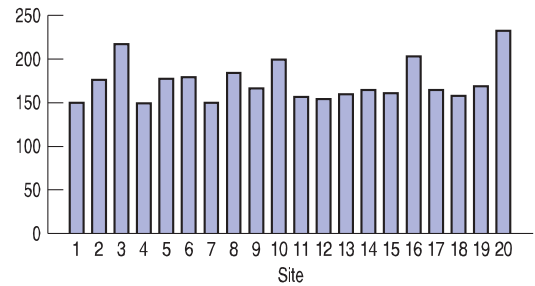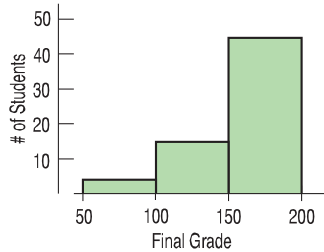

a) Comment on this graph.
b) Create your own histogram of the data.

51. Acid rain Two researchers measured the pH (a scale on which a value of 7 is neutral and values below 7 are acidic) of water collected from rain and snow over a 6-month period in Allegheny County, Pennsylvania. Describe their data with a graph and a few sentences:
4.57 5.62 4.12 5.29 4.64 4.31 4.30 4.39 4.45 5.67
4.39 4.52 4.26 4.26 4.40 5.78 4.73 4.56 5.08 4.41
4.12 5.51 4.82 4.63 4.29 4.60

T 52. Marijuana 2007 In 2007 the Council of Europe published a report entitled *The European School Survey Project on Alcohol and Other Drugs* (www.espad.org). Among other issues, the survey investigated the percentages of 16-year-olds who had used marijuana. Shown here are the results for 34 European countries. Create an appropriate graph of these data, and describe the distribution.

Country	Percentage	Country	Percentage
Armenia	3	Italy	23
Austria	17	Latvia	18
Belgium	24	Lithuania	18
Bulgaria	22	Malta	13
Croatia	18	Monaco	28
Cyprus	5	Netherlands	28
Czech Republic	45	Norway	6
Estonia	26	Poland	16
Faroe Islands	6	Portugal	13
Finland	8	Romania	4
France	31	Russia	19
Germany	20	Slovak Republic	32
Greece	6	Slovenia	22
Hungary	13	Sweden	7
Iceland	9	Switzerland	33
Ireland	20	Ukraine	14
Isle of Man	34	United Kingdom	29

53. Final grades A professor (of something other than Statistics!) distributed the following histogram to show the distribution of grades on his 200-point final exam. Comment on the display.

54. Final grades revisited After receiving many complaints about his final-grade histogram from students currently taking a Statistics course, the professor from Exercise 53 distributed the following revised histogram:
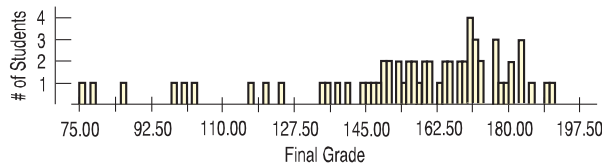
a) Comment on this display.
b) Describe the distribution of grades.
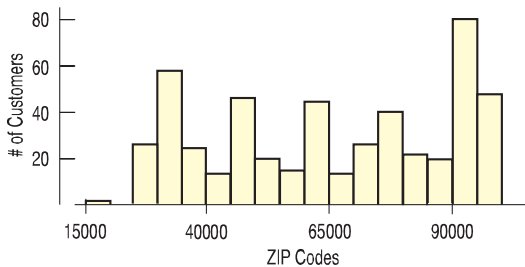
55. ZIP codes Holes-R-Us, an Internet company that sells piercing jewelry, keeps transaction records on its sales. At a recent sales meeting, one of the staff presented a histogram of the ZIP codes of the last 500 customers, so that the staff might understand where sales are coming from. Comment on the usefulness and appropriateness of the display.

56. ZIP codes revisited Here are some summary statistics to go with the histogram of the ZIP codes of 500 customers from the Holes-R-Us Internet Jewelry Salon that we saw in Exercise 55:

Count	500
Mean	64,970.0
StdDev	23,523.0
Median	64,871
IQR	44,183
Q1	46,050
Q3	90,233

What can these statistics tell you about the company's sales?

T 57. Math scores 2009 The National Center for Education Statistics (nces.ed.gov/nationsreportcard/) reported 2005 average mathematics achievement scores for eighth graders in all 50 states:

State	Score	State	Score
Alabama	269	Montana	292
Alaska	283	Nebraska	284
Arizona	277	Nevada	274
Arkansas	276	New Hampshire	292
California	270	New Jersey	293
Colorado	287	New Mexico	270
Connecticut	289	New York	283
Delaware	284	North Carolina	284
Florida	279	North Dakota	293
Georgia	278	Ohio	286
Hawaii	274	Oklahoma	276
Idaho	287	Oregon	285
Illinois	282	Pennsylvania	288
Indiana	287	Rhode Island	278
Iowa	284	South Carolina	280
Kansas	289	South Dakota	291
Kentucky	279	Tennessee	275
Louisiana	272	Texas	287
Maine	286	Utah	284
Maryland	288	Vermont	293
Massachusetts	299	Virginia	286
Michigan	278	Washington	289
Minnesota	294	West Virginia	270
Mississippi	265	Wisconsin	288
Missouri	286	Wyoming	286

a) Find the median, the IQR, the mean, and the standard deviation of these state averages.
b) Which summary statistics would you report for these data? Why?
c) Write a brief summary of the performance of eighth graders nationwide.

58. Boomtowns 2011 In 2011, the website NewGeography .com listed its ranking of the best cities for job growth in the United States. Here is the magazine's top 20 larger cites, along with their weighted job rating index.

1	Austin-Round Rock-San Marcos, TX	89.1
2	New Orleans-Metairie-Kenner, LA	88.9
3	Houston-Sugar Land-Baytown, TX	82.1
4	San Antonio-New Braunfels, TX	80.7
5	Dallas-Plano-Irving, TX Metropolitan Division	80.6
6	Washington-Arlington-Alexandria, DC-VA-MD-WV Metropolitan Division	74.1
7	Northern Virginia, VA	73.8
8	Nashville-Davidson–Murfreesboro–Franklin, TN	73.1
9	New York City, NY	72.6
10	Philadelphia City, PA	71.9
11	Pittsburgh, PA	71.7
12	Bethesda-Rockville-Frederick, MD Metropolitan Division	70.7
13	Boston-Cambridge-Quincy, MA NECTA Division	70.4
14	Raleigh-Cary, NC	67.8
15	Fort Worth-Arlington, TX Metropolitan Division	67.8
16	Rochester, NY	66.7
17	Nassau-Suffolk, NY Metropolitan Division	65.5
18	Buffalo-Niagara Falls, NY	64.7
19	Columbus, OH	64.3
20	Salt Lake City, UT	64.2

a) Make a suitable display of the weighted growth indices.

b) Summarize the typical growth index among these cities with a median and mean. Why do they differ?

c) Given what you know about the distribution, which of the measures in part b does the better job of summarizing the growth indices? Why?

d) Summarize the spread of the growth index distribution with a standard deviation and with an IQR.

e) Given what you know about the distribution, which of the measures in part d does the better job of summarizing the growth rates? Why?

f) Suppose we subtract from each of the preceding growth rates the average U.S. large-city growth index of 49.23%, so that we can look at how much these indices exceed the U.S. rate. How would this change the values of the summary statistics you calculated above? (*Hint:* You need not recompute any of the summary statistics from scratch.)

g) If we were to omit Austin-Round Rock-San Marcos, TX from the data, how would you expect the mean, median, standard deviation, and IQR to change? Explain your expectations for each.

h) Write a brief report about all of these growth indices.

59. Population growth The following data show the percentage change in population for the 50 states and the District of Columbia from 2000 to 2009. Using appropriate graphical displays and summary statistics, write a report on the percentage change in population by state.

Alabama	5.9
Alaska	11.4
Arizona	28.6
Arkansas	8.1
California	9.1
Colorado	16.8
Connecticut	3.3
Delaware	13.0
District of Columbia	4.8
Florida	16.0
Georgia	20.1
Hawaii	6.9
Idaho	19.5
Illinois	4.0
Indiana	5.6
Iowa	2.8
Kansas	4.8
Kentucky	6.7
Louisiana	0.5
Maine	3.4
Maryland	7.6
Massachusetts	3.9
Michigan	0.3
Minnesota	7.0
Mississippi	3.8
Missouri	7.0
Montana	8.1
Nebraska	5.0
Nevada	32.3
New Hampshire	7.2
New Jersey	3.5
New Mexico	10.5
New York	3.0
North Carolina	16.6
North Dakota	0.7
Ohio	1.7
Oklahoma	6.9
Oregon	11.8
Pennsylvania	2.6
Rhode Island	0.5
South Carolina	13.7
South Dakota	7.6
Tennessee	10.7
Texas	18.8
Utah	24.7
Vermont	2.1
Virginia	11.4
Washington	13.1
West Virginia	0.6
Wisconsin	5.4
Wyoming	10.2

Source: www.census.gov/compendia/statab/rankings.html

T 60. Prisons 2006 A report from the U.S. Department of Justice (www.ojp.usdoj.gov/bjs/) reported the percent changes in federal prison populations in 21 northeastern and midwestern states during 2006. Using appropriate graphical displays and summary statistics, write a report on the changes in prison populations.

State	Percent Change	State	Percent Change
Connecticut	1.6	Iowa	0.9
Maine	−7.7	Kansas	−1.2
Massachusetts	4.8	Michigan	3.4
New Hampshire	2.5	Minnesota	6.4
New Jersey	1.1	Missouri	−2.8
New York	0.0	Nebraska	4.5
Pennsylvania	3.7	North Dakota	6.1
Rhode Island	7.8	Ohio	5.6
Vermont	8.3	South Dakota	5.3
Illinois	1.7	Wisconsin	−2.0
Indiana	3.8		

✓ Just Checking ANSWERS

(Thoughts will vary.)

1. Slightly skewed to the right. Center around 3 miles? Few over 10 miles.

2. Bimodal. Center between 1 and 2 hours? Many people watch no football; others watch most of one or more games. Probably only a few values over 5 hours.

3. Strongly skewed to the right, with almost everyone at $0; a few small prizes, with the winner an outlier.

4. Fairly symmetric, somewhat uniform, perhaps slightly skewed to the right. Center in the 40s? Few ages below 25 or above 70.

5. Uniform, symmetric. Center near 5. Roughly equal counts for each digit 0–9.

6. Incomes are probably skewed to the right and not symmetric, making the median the more appropriate measure of center. The mean will be influenced by the high end of family incomes and not reflect the "typical" family income as well as the median would. It will give the impression that the typical income is higher than it is.

7. An IQR of 30 mpg would mean that only 50% of the cars get gas mileages in an interval 30 mpg wide. Fuel economy doesn't vary that much. 3 mpg is reasonable. It seems plausible that 50% of the cars will be within about 3 mpg of each other. An IQR of 0.3 mpg would mean that the gas mileage of half the cars varies little from the estimate. It's unlikely that cars, drivers, and driving conditions are that consistent.

8. We'd prefer a standard deviation of 2 months. Making a consistent product is important for quality. Customers want to be able to count on the MP3 player lasting somewhere close to 5 years, and a standard deviation of 2 years would mean that life spans were highly variable.