each graph, the population would be all households of Lee County for the "Household income" graph, all residents of Lee County for the "Age" graph, and all registered adults for the "Political party" graph. The variables needed to complete these graphs are income, age, and political party affiliation. Income and age are continuous variables, whereas party affiliation is a nominal variable. The circle graph data were most likely collected by categories and then the percentages calculated. Counts were used to create the Political party bar graph.

Statistics⊜Now™

FYI Watch a supplement video example at **http://1pass.thomson.com** or on your CD for more on basic terms.

SECTION 1.2 EXERCISES

Statistics⊜Now™

Skillbuilder Applet Exercises must be worked using an accompanying applet found on your Student's Suite CD-ROM or at the StatisticsNow website at **http://1pass.thomson.com.**

1.5 *Statistics* is defined on page 4 as "the science of collecting, describing, and interpreting data." Using your own words, write a sentence describing each of the three statistical activities. Retain your work for Exercise 1.87.

1.6 Determine which of the following statements is descriptive in nature and which is inferential. Refer to "Even in kindergarten, social skills trump" in Applied Example 1.1 (pp. 4–5).

a. Of all U.S. kindergarten teachers, 32% say that "Knowing the alphabet" is an essential skill.

b. Of the 800 U.S. kindergarten teachers polled, 32% say that "Knowing the alphabet" is an essential skill.
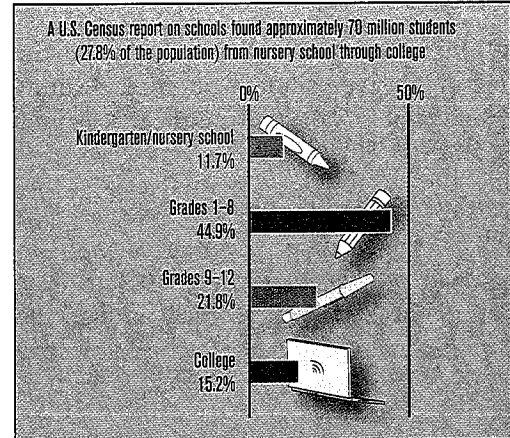
1.7 Determine which of the following statements is descriptive in nature and which is inferential. Refer to "Pampering people produces profit" in Applied Example 1.2 (p. 5).

a. Of the surveyed spas in 2004, 72% were categorized as day spas.

b. Of all visits to spas in 2003, 23% were visits by men.

1.8 Refer to the graphic "America's students by grade level."

The latest Census report on schools found about 70 million students (27.8% of the population) from nursery school through college.

AMERICA'S STUDENTS BY GRADE LEVEL

A U.S. Census report on schools found approximately 70 million students (27.8% of the population) from nursery school through college

Data from USA Today, 5/9/2000.

a. What is the population?

b. What information was obtained from each person?

c. Using the information given, estimate the number of students in college.

d. Using the information given, estimate the size of the entire U.S. population.

1.9 International Communications Research (ICR) conducted the 2004 National Spring Cleaning Survey for The Soap and Detergent Association. ICR questioned 1000 American male and female heads of household regarding their house cleaning attitudes. The survey has a margin of error of plus or minus 5%.

a. What is the population?

b. How many people were polled?

c What information was obtained from each person?

1.26 The admissions office wants to estimate the cost of textbooks for students at our college. Let the variable x be the total cost of all textbooks purchased by a student this semester. The plan is to randomly identify 100 students and obtain their total textbook costs. The average cost for the 100 students will be used to estimate the average cost for all students.

a. Describe the parameter the admissions office wishes to estimate.

b. Describe the population.

c. Describe the variable involved.

d. Describe the sample.

e. Describe the statistic and how you would use the 100 data values collected to calculate the statistic.

1.27 A quality-control technician selects assembled parts from an assembly line and records the following information concerning each part:

A: defective or nondefective

B: the employee number of the individual who assembled the part

C: the weight of the part

a. What is the population?

b. Is the population finite or infinite?

c. What is the sample?

d. Classify the three variables as either attribute or numerical.

1.28 Select 10 students currently enrolled at your college and collect data for these three variables:

X: number of courses enrolled in

Y: total cost of textbooks and supplies for courses

*Z:* method of payment used for textbooks and supplies

a. What is the population?

b. Is the population finite or infinite?

c. What is the sample?

d. Classify the three variables as nominal, ordinal, discrete, or continuous.

1.29 A study was conducted by Aventis Pharmaceuticals Inc. to measure the adverse side effects of Allegra, a drug used for the treatment of seasonal allergies. A sample of 679 allergy sufferers in the United States was given 60 mg of the drug twice a day. The patients were to report whether they experienced relief from their allergies as well as any adverse side effects (viral infection, nausea, drowsiness, etc.).

Source: Good Housekeeping, February 2005, p. 120.

a. What is the population being studied?

b. What is the sample?

c. What are the characteristics of interest about each element in the population?

d. Are the data being collected qualitative or quantitative?

1.30 Skillbuilder Applet Exercise simulates taking a sample of size 10 from a population of 100 college students. Take a sample of size 10.
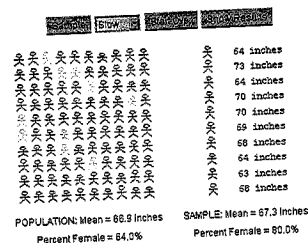
a. What is the population?

b. Is the population finite or infinite?

c. Name two parameters and give their values?

d. What is the sample?

e. Name the two corresponding statistics and give their values?

f. Take another sample of size 10. Which of the preceding items remain fixed and which changed?

1.31 Identify each of the following as an example of (1) attribute (qualitative) or (2) numerical (quantitative) variables:

a. the breaking strength of a given type of string

b. the hair color of children auditioning for the musical *Annie*

c. the number of stop signs in towns of less than 500 people

d. whether or not a faucet is defective

e. the number of questions answered correctly on a standardized test

f. the length of time required to answer a telephone call at a certain real estate office

1.32 Identify each of the following as examples of (1) nominal, (2) ordinal, (3) discrete, or (4) continuous variables:
a. A poll of registered voters as to which candidate they support
b. The length of time required for a wound to heal when a new medicine is being used
c. The number of televisions within a household
d. The distance first-year college women can kick a football
e. The number of pages per job coming off a computer printer

f. The kind of tree used as a Christmas tree

1.33 Suppose a 12-year-old asked you to explain the difference between a sample and a population.
a. What information should your answer include?
b. What reasons would you give for why one would take a sample instead of surveying every member of the population?

1.34 Suppose a 12-year-old asked you to explain the difference between a statistic and a parameter.
a. What information should your answer include?
b. What reasons would you give for why one would report the value of a statistic instead of a parameter?

1.3 Measurability and Variability

Within a set of measurement data, we always expect variation. If little or no variation is found, we would guess that the measuring device is not calibrated with a small enough unit. For example, we take a carton of a favorite candy bar and weigh each bar individually. We observe that each of the 24 candy bars weighs $7/8$ ounce, to the nearest $1/8$ ounce. Does this mean that the bars are all identical in weight? Not really! Suppose we were to weigh them on an analytical balance that weighs to the nearest ten-thousandth of an ounce. Now the 24 weights will most likely show **variability.**

It does not matter what the response variable is; there will most likely be variability in the data if the tool of measurement is precise enough. One of the primary objectives of statistical analysis is measuring variability. For example, in the study of quality control, measuring variability is absolutely essential. Controlling (or reducing) the variability in a manufacturing process is a field all its own—namely, statistical process control.

SECTION 1.3 EXERCISES

Statistics⬨Now™

Skillbuilder Applet Exercises must be worked using an accompanying applet found on your Student's Suite CD-ROM or at the StatisticsNow website at **http://1pass.thomson.com.**

1.35 Suppose we measure the weights (in pounds) of the individuals in each of the following groups:
Group 1: cheerleaders for National Football League teams
Group 2: players for National Football League teams

✓ Understand how convenience and volunteer samples result in biased samples.　　　　pp. 18–19, Ex. 1.43

✓ Understand the differences among and be able to identify experiments, observational studies, and judgment samples.　　　　pp. 19–21

✓ Understand and be able to describe the single-stage sampling methods of "simple random sample" and "systematic sampling."　　　　pp. 21–24

✓ Understand and be able to describe the multistage sampling methods of "stratified sampling" and "cluster sampling."　　　　pp. 24–25

✓ Understand and be able to explain the difference between probability and statistics.　　　　p. 27, Ex. 1.61

✓ Understand that variability is inherent in everything and in the sampling process.　　　　p. 17, Ex. 1.36

Chapter Exercises

Statistics ⊘ Now™
Go to the StatisticsNow website **http://1pass.thomson.com** to
- Assess your understanding of this chapter
- Check your readiness for an exam by taking the Pre-Test quiz and exploring the resources in the Personalized Learning Plan

1.67 We want to describe the so-called typical student at your college. Describe a variable that measures some characteristic of a student and results in:

a. Attribute data

b. Numerical data

1.68 A candidate for a political office claims that he will win the election. A poll is conducted, and 35 of 150 voters indicate that they will vote for the candidate, 100 voters indicate that they will vote for his opponent, and 15 voters are undecided.

a. What is the population parameter of interest?

b. What is the value of the sample statistic that might be used to estimate the population parameter?

c. Would you tend to believe the candidate based on the results of the poll?

1.69 A researcher studying consumer buying habits asks every 20th person entering Publix Su-

permarket how many times per week he or she goes grocery shopping. She then records the answer as T.

a. Is $T = 3$ an example of a sample, a variable, a statistic, a parameter, or a data value?

Suppose the researcher questions 427 shoppers during the survey.

b. Give an example of a question that can be answered using the tools of descriptive statistics.

c. Give an example of a question that can be answered using the tools of inferential statistics.

1.70 A researcher studying the attitudes of parents of preschool children interviews a random sample of 50 mothers, each having one preschool child. He asks each mother, "How many times did you compliment your child yesterday?" He records the answer as C.

a. Is $C = 4$ an example of a data value, a statistic, a parameter, a variable, or a sample?

b. Give an example of a question that can be answered using the tools of descriptive statistics.

c. Give an example of a question that can be answered using the tools of inferential statistics.

1.71 Harris Interactive conducted an online poll of U.S. adults during December 2004 for the *Wall Street Journal Online's Health Industry Edition*.

Study the graph and answer the questions.

a. Does it seem reasonable that the Population line and the Drivers line run basically parallel to each other and that the Population line is above the Drivers line? Explain what it means for the lines to be parallel. What would it mean if they were not parallel?

b. The Drivers and Motor Vehicles lines cross. What does this mean? When did the lines intersect, and what does the point of intersection represent?

1.76 The 700 Club: Barry Bonds of the San Francisco Giants is on pace to become baseball's home-run king this season or next. Last year, he joined Hank Aaron and Babe Ruth as the only major-league players to have hit more than 700 home runs in their careers. Here is a look at how they amassed their totals.

a. Describe and compare the overall appearance of the three graphs. Include thoughts about such things as length of career, when the most home-runs per year were hit and their relationship to the aging process, and any others you think of.

b. Does it appear that one of them was more consistent with annual homerun production?

c. From the evidence presented here, who do you think should be called the "Homerun King"?

d. Was Barry Bonds' 73 homeruns in one season a fluke?

e. If you were a team owner and most interested in homerun production over the next several years, which one would you want on your team? Let's say you were signing him at age 21. At age 35.

1.77 In the autumn of 2003, the National Safe Kids Campaign conducted a study of helmet use among children ages 5 to 14 who participate in wheeled sports. Data were collected from various sites across the United States that were designated as places where children often engage in wheeled sports. Activity, apparent gender, and estimated age were recorded for each rider, along with information on helmet use. It was found that, overall, 41% of children were wearing a helmet while participating in a wheeled sport.

a. Was this study an experiment or an observational study?

b. Identify the parameter of interest.

c. Identify the statistic and give its value.

d. Classify the four variables as numerical or attribute.

1.78 *USA Today,* in a December 2004 article titled "There's no place like work for the holidays," presented the results of a study of 600 full-time U.S. workers done by Penn Schoen & Berland Associates. Results revealed that 33% of the respondents were taking no time off during the holidays. Of those surveyed, 28% reported taking 1 to 2 days off during the holidays.

a. What is the population?

b. What is the sample?

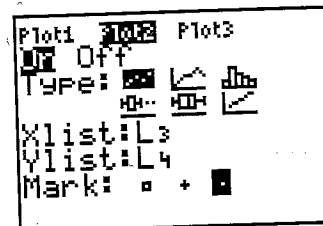c. Is this a judgment sample or a probability sample?

Exercise 1.76
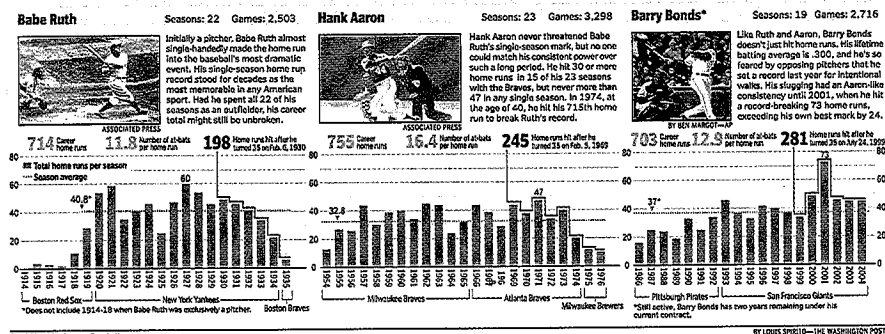
Source: *The Washington Post*

Choose: 2nd > STAT PLOT > 2:PLOT2

```
Plot1 PLOT2 Plot3
On Off
Type: ▨ ⬈ ⬛
      ⬛⬛ ⬛⬛ ⬈
Xlist:L3
Ylist:L4
Mark: ▫ + ■
```

Choose: Window
Enter: at most lowest value for both, at least highest value for
           both, 0 or increment, −2, at least highest counting number,1,1
Choose: Graph > Trace > > > > (gives data values)

SECTION 2.2 EXERCISES

Statistics⬧Now™

Datasets can be found on your Student's Suite CD-ROM or at the StatisticsNow website at **http://1pass.thomson.com.**

2.3 As a statistical graph the circle graph has limitations. Examine the circle graph in Figure 2.1 and the bar graph in Figure 2.2.

a. What information do they both demonstrate?

b. What information is shown in the circle graph that cannot be shown in the bar graph?

c. "Generally speaking, the bar graph is a better choice for use than the circle graph." Justify this statement.

2.4 How Americans prefer to eat an apple was reported in *USA Today*, November 11, 2004: Bite into it—47%, Cut it into slices—39%, Peel—11%, Don't Know—3%.

a. Construct a circle graph showing how Americans prefer to eat an apple.

b. Construct a bar graph showing how Americans prefer to eat an apple.

c. In your opinion, does the circle graph in part a or the bar graph in part b result in a better representation of the information? Explain.

2.5 American Payroll Association got a big response to this question about company dress code: "The current dress code at my company is . . ."

Final results: a. A little too relaxed—27%
                b. A little too formal—15%
                c. Just right—58%

Most people mentioned the importance of "comfort" in their explanations. The vast majority of respondents were very happy with their company's dress code or policy.

a. Construct a circle graph depicting this information. Label completely.

b. Construct a bar graph depicting this same information. Label completely.

c. Compare the previous two graphs, describing what you see in each one now that the graphs have been drawn and completely labeled. Do you get the same impression about these people's feelings from both graphs? Does one emphasize anything the other one does not?

2.6 [EX02-006] The American Community Survey is limited to the household population and excludes the population living in institutions, college dormitories, and other group quarters. Montana's 2003 household make-up is listed here.

	Household Population
Householder	374,879
Spouse	197,379
Child	243,609
Other relatives	27,583
Nonrelatives	49,047
Total	892,497

Source: U.S. Census Bureau

force-
Trans-
ber of
1es by

Number

2031
454
1106
1159
322

ent of
1/

:his in-

:nd the
:cting if
.e over-
ow the
1tes the

assem-
n.
port?

100
80
60
40
20
0

Percent

hers
6
1.0
).0

value of
ns.

1 line the
What two
pecial at-
? Explain.

2.16 [EX02-016] The world's cocoa production in 2004–2005 is predicted to decline following a record 3396 MT crop in 2003–2004. Most of the reduction is expected to occur in West Africa, with Asian and Latin American production remaining relatively unchanged. West Africa will continue to be the leader in world cocoa production in 2005.

World Cocoa Production (1000 tonnes)

Country	2003–2004	2004–2005 (forecast)
Brazil	163	164
Cameroons	150	150
Ecuador	100	93
Ghana	605	530
Indonesia	420	415
Ivory Coast	1500	1275
Malaysia	25	26
Nigeria	165	170
Other Africa	38	39
Other America	165	170
Other Asia	65	67

Source: World Cocoa Foundation, http://www.chocolateandcocoa.org/stats/supply/default.asp

a. By how much is the total production expected to decrease? What percent decrease is that?

b. Construct a Pareto diagram depicting the 2003–2004 production.

c. Construct a Pareto diagram depicting the 2004–2005 production forecast.

d. African countries are expected to supply what percentage of the world's cocoa for 2004–2005? The Americas? Asia?

2.17 [EX02-017] The number of points scored during each game by a high school basketball team last season was as follows: 56, 54, 61, 71, 46, 61, 55, 68, 60, 66, 54, 61, 52, 36, 64, 51. Construct a dotplot of these data.

2.18 [EX02-018] The table here lists the median house selling prices (in $1000s) for 20 U.S. cities as listed on Realtor.com.

120	120	120	119	117	117	116	116	115	115
114	114	112	112	112	111	109	107	105	105

Source: http://www.realtor.org/Research.nsf/pages/EconHousingData

a. Construct a dotplot of these data.

b. Describe the distribution displayed by the dotplot found in part a.

2.19 [EX02-019] HoopsHype.com regularly posts the latest on the NBA. Following are the heights (in inches) of the basketball players who were the first round picks by the professional teams on June 24, 2004:

82	82	74	79	75	79	80	83	78	79
83	85	71	81	81	78	80	78	79	72
89	81	80	74	76	79	78	75	84	

Source: http://www.hoopshype.com/draft.htm

a. Construct a dotplot of the heights of these players.

b. Use the dotplot to uncover the shortest and the tallest players.

c. What is the most common height, and how many players share that height?

d. What feature of the dotplot illustrates the most common height?

2.20 [EX02-020] As baseball players, Babe Ruth and Hank Aaron were well known for their ability to hit home runs. Mark McGwire and Sammy Sosa became well known for their ability to hit home runs during the "great home run chase" of 1998. Bobby Bonds gained his fame in 2001. Listed here are the number of home runs each player hit in each major-league season in which he played.

Exercise 2.20

Ruth	4	3	2	11	29	54	59	35	41	46	25	47	60	54	46	49	46	41	34	22	6	
Aaron	13	27	26	44	30	39	40	34	45	44	24	32	44	39	29	44	38	47	34	40	20	12	10
McGwire	3	49	32	33	39	22	42	9	9	39	52	58	70	65	32	29						
Sosa	4	15	10	8	33	25	36	40	36	66	63	50	64	49	40	35						
Bonds	16	25	24	19	33	25	34	46	37	33	42	40	37	34	49	73	46	45	45			

a. Construct a dotplot of the data for Ruth and Aaron, using the same axis.

b. Using the dotplots found in part a, make a case for each of the following statements with regard to past players: "Aaron is the homerun king!" "Ruth is the homerun king!"

c. Construct a dotplot of the data for McGwire, Sosa, and Bonds using the same axis.

d. Using the dotplots found in part c, make a case for the statements "McGwire is the homerun king!" "Bonds is the homerun king!" and "Sosa is not currently the homerun king!" with regard to the present players. In what way do the dotplots support each statement?

FYI If you use your computer or calculator, use the commands on page 49.

2.21 [EX02-021] Delco Products, a division of General Motors, produces commutators designed to be 18.810 mm in overall length. (A commutator is a device used in the electrical system of an automobile.) The following sample of 35 commutator lengths was taken while monitoring the manufacturing process:

18.802	18.810	18.780	18.757	18.824	18.827	18.825
18.809	18.794	18.787	18.844	18.824	18.829	18.817
18.785	18.747	18.802	18.826	18.810	18.802	18.780
18.830	18.874	18.836	18.758	18.813	18.844	18.861
18.824	18.835	18.794	18.853	18.823	18.863	18.808

Source: With permission of Delco Products Division, GMC

Use a computer to construct a dotplot of these data values.

2.22 A computer was used to construct the dotplot at the bottom of the page.

a. How many data values are shown?

b. List the values of the five smallest data.

c. What is the value of the largest data item?

d. What value occurred the greatest number of times? How many times did it occur?

2.23 [EX02-017] Construct a stem-and-leaf display of the number of points scored during each basketball game last season:

56	54	61	71	46	61	55	68
60	66	54	61	52	36	64	51

2.24 [EX02-024] Forbes.com posted the 5-year (2000 to 2004) total returns, in percents, for 17 banking industry companies.

| Name	Return %	Name	Return %
Astoria Financial	23.9	Popular	15.3
Banknorth Group	18.6	State Street	5.0
Bank of America	13.2	Synovus Finl	8.5
BB&T	8.1	UnionBanCal	10.4
Compass Bancshares	16.7	Wachovia	10.5
Golden West Finl	29.1	Wells Fargo	8.8
M&T Bank	19.1	Westcorp	25.9
National City	12.6	Zions Bancorp	2.9
North Fork Bancorp	19.8		

Source: http://www.forbes.com/lists/results.jhtml

a. Construct a stem-and-leaf display of the data.

b. Based on the stem-and-leaf display, describe the distribution of percentages of profitability.

2.25 [EX02-025] The amounts shown here are the fees charged by Quik Delivery for the 40 small packages it delivered last Thursday afternoon:

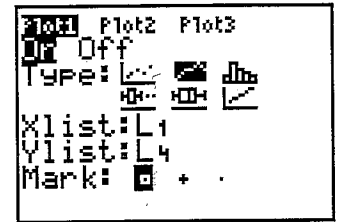|									
4.03	3.56	3.10	6.04	5.62	3.16	2.93	3.82	4.30	3.86
4.57	3.59	4.57	6.16	2.88	5.03	5.46	3.87	6.81	4.91
3.62	3.62	3.80	3.70	4.15	2.07	3.77	5.77	7.86	4.63
4.81	2.86	5.02	5.24	4.02	5.44	4.65	3.89	4.00	2.99

a. Construct a stem-and-leaf display.

b. Based on the stem-and-leaf display, describe the distribution of the data.

Figure for Exercise 2.22
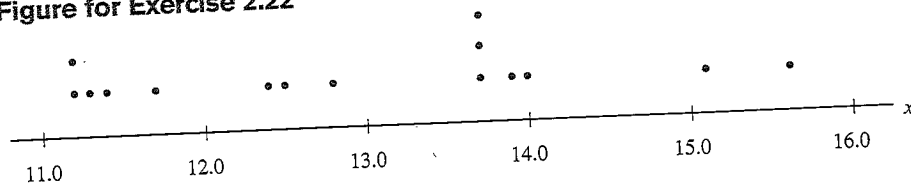
Choose: **2nd > STAT PLOT > 1:Plot**

Choose: **Zoom > 9:ZoomStat > Trace > > >**

Adjust window if needed for better readability.

SECTION 2.3 EXERCISES

Statistics⬡Now™

Skillbuilder Applet Exercises must be worked using an accompanying applet found on your Student's Suite CD-ROM or at the StatisticsNow website at **http://1pass.thomson.com.**

Datasets can be found on your Student's Suite CD-ROM or at the StatisticsNow website at **http://1pass.thomson.com.**

2.29 a. Form an ungrouped frequency distribution of the following data:

1, 2, 1, 0, 4, 2, 1, 1, 0, 1, 2, 4

Referring to the preceding distribution:

b. Explain what $f = 5$ represents.

c. What is the sum of the frequency column?

d. What does this sum represent?

2.30 Bar graphs and histograms are not the same thing. Explain their similarities and differences.

2.31 [EX02-031] The players on the Rochester Raging Rhinos professional soccer team scored 38 goals during the 2002 season.

Player	1	2	3	4	5	6	7	8	9	10	11	12	13
Goals	2	8	1	2	2	6	2	1	5	2	3	2	2

Source: http://www.rhinossoccer.com/team.asp

a. If you want to show the number of goals scored by each player, would it be more appropriate to display this information on a bar graph or a histogram? Explain.

b. Construct the appropriate graph for part a.

c. If you wanted to show (emphasize) the distribution of scoring by the team, would it be

more appropriate to display this information on a bar graph or a histogram? Explain.

d. Construct the appropriate graph for part c.

2.32 [EX02-032] The California Department of Education gives an annual report on the Advanced Placement (AP) test results for each year. In the 2003–2004 school year, Modoc County had students with the following scores:

2	3	2	1	2	3	3	2	2	3	5	3	2	2
2	4	5	1	2	2	4	3	2	2	5	4	1	4
2	4	5	2	3	2	3	2	3	3	5	1	5	5

Source: http://data1.cde.ca.gov/dataquest/

a. Construct an ungrouped frequency distribution for the test scores.

b. Construct a frequency histogram of this distribution.

c. Prepare a relative frequency distribution for these same data.

d. If AP scores of at least 3 are often required for college transferability, what percentage of Modoc AP scores will receive college credit?

(Retain these solutions to use in Exercise 2.52 on p. 72.)

2.33 [EX02-033] The U.S. Women's Olympic Soccer team had a great year in 2004. One way to describe the players on that team is by their individual heights.

Height (inches)

68	67	65	66	65	67	64	69	69
65	64	71	66	67	68	66	65	71

Source: http://www.SoccerTimes.com

(left margin fragments)

ages into C2
ary of the first
boundary).

column B (in-

cells)

-

ages 61–62.

mn A and the
dditional class
ncy equal to 0

Next
remove all

1.)

1s) > Next

ish

e data).

L2 (include an
ero); then con-

a. Construct an ungrouped frequency distribution for the heights.

b. Construct a frequency histogram of this distribution.

c. Prepare a relative frequency distribution for these same data.

d. What percentage of the team is at least 5 feet, 6 inches tall?

2.34 [EX02-034] The U.S. Census Bureau posted the following 2003 Report on America's Families and Living Arrangements for all races.

No. in Household	Percentage	No. in Household	Percentage
1	26.4%	5	6.3%
2	33.3%	6	2.3%
3	16.1%	7+	1.2%
4	14.3%		

Source: http://www.census.gov/population/www/socdemo/hh-fam/cps2003.html

a. Draw a relative frequency histogram for the number of people per household.

b. What shape distribution does the histogram suggest?

c. Based on the graph, what do you know about the households in the United States?

2.35 [EX02-035] The 2003 American Community Survey universe is limited to the household population and excludes the population living in institutions, college dormitories, and other group quarters. The accompanying table lists the number of rooms in each of the 8,658,290 housing units in Texas.

Rooms	Housing Units	Rooms	Housing Units
1 room	124,486	6 rooms	1,649,479
2 rooms	349,496	7 rooms	913,138
3 rooms	1,007,873	8 rooms	520,248
4 rooms	1,548,984	9+ rooms	485,506
5 rooms	2,059,080		

Source: U.S. Census Bureau, American Community Survey Office

a. Draw a relative frequency histogram for the number of rooms per household.

b. What shape distribution does the histogram suggest?

c. Based on the graph, what do you know about the number of rooms per household in Texas?

2.36 [EX02-036] Here are the ages of 50 dancers who responded to a call to audition for a musical comedy:

|									
21	19	22	19	18	20	23	19	19	20
19	20	21	22	21	20	22	20	21	20
21	19	21	21	19	19	20	19	19	19
20	20	19	21	21	22	19	19	21	19
18	21	19	18	22	21	24	20	24	17

a. Prepare an ungrouped frequency distribution of these ages.

b. Prepare an ungrouped relative frequency distribution of the same data.

c. Prepare a relative frequency histogram of these data.

d. Prepare a cumulative relative frequency distribution of the same data.

e. Prepare an ogive of these data.

2.37 [EX02-037] The opening-round scores for the Ladies' Professional Golf Association tournament at Locust Hill Country Club were posted as follows:

69	73	72	74	77	80	75	74	72	83	68	73	75	78
76	74	73	68	71	72	75	79	74	75	74	74	68	79
75	76	75	77	74	74	75	75	72	73	73	72	72	71
71	70	82	77	76	73	72	72	72	75	75	74	74	74
76	76	74	73	74	73	72	72	74	71	72	73	72	72
74	74	67	69	71	70	72	74	76	75	75	74	73	74
74	78	77	81	73	73	74	68	71	74	78	70	68	71
72	72	75	74	76	77	74	74	73	73	70	68	69	71
77	78	68	72	73	78	77	79	79	77	75	75	74	73
73	72	71	68	70	71	78	78	76	74	75	72	72	72
75	74	76	77	78	78								

a. Form an ungrouped frequency distribution of these scores.

b. Draw a histogram of the first-round golf scores. Use the frequency distribution from part a.

2.38 Figuring *where* lightning strikes will occur is a near impossible task. *When* lightning strikes occur, however, has become more predictable based on research. For a small area in Colorado, data were collected and the results are displayed in the histogram that follows.
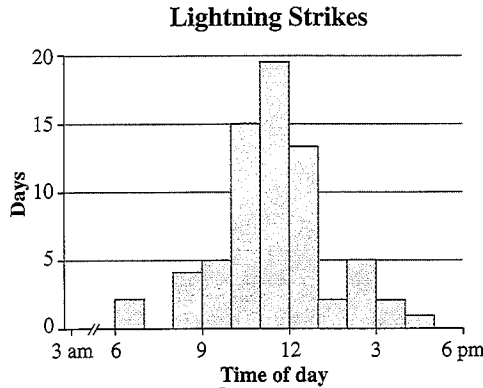
Lightning Strikes

Based on the histogram:

a. Data were collected for what variable?

b. What does each bar (interval) represent?

c. What conclusion can be reached about the "when" of lightning strikes in this small area of Colorado?

d. What characteristics of the graph support the conclusion?

2.39 [EX02-039] A survey of 100 resort club managers on their annual salaries resulted in the following frequency distribution:

Annual Salary ($1000s)	15–25	25–35	35–45	45–55	55–65
No. of Managers	12	37	26	19	6

a. The data value "35" belongs to which class?

b. Explain the meaning of "35–45."

c. Explain what "class width" is, give its value, and describe three ways that it can be determined.

d. Draw a frequency histogram of the annual salaries for resort club managers. Label class boundaries.

(Retain these solutions to use in Exercise 2.51 on p. 71.)

2.40 Skillbuilder Applet Exercise demonstrates the procedure of transforming a stem-and-leaf display into a histogram.
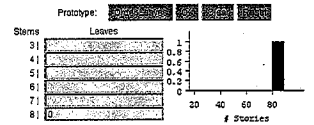
Type the leaves for the number of stories into the stem-and-leaf display. Click OK to see the corresponding histogram. Comment on the similarities and differences.

2.41[EX02-041] The KSW computer science aptitude test was given to 50 students. The following frequency distribution resulted from their scores:

KSW Test Score	0–4	4–8	8–12	12–16	16–20	20–24	24–28
Frequency	4	8	8	20	6	3	1

a. What are the class boundaries for the class with the largest frequency?

b. Give all the class midpoints associated with this frequency distribution.

c. What is the class width?

d. Give the relative frequencies for the classes.

e. Draw a relative frequency histogram of the test scores.

2.42 [EX02-042] The USA Snapshot titled "Nuns an aging order" reports that the median age of the 94,022 Roman Catholic nuns in the United States is 65 years and the percentages of U.S. nuns by age group are as follows:

Under 50	51–70	Over 70	Refused to give age
16%	42%	37%	5%

This information is based on a survey of 1049 Roman Catholic nuns. Suppose the survey had resulted in the following frequency distribution (52 ages unknown):

Age	20–30	30–40	40–50	50–60	60–70	70–80	80–90
Frequency	34	58	76	187	254	241	147

a. Draw and completely label a frequency histogram.

b. Draw and completely label a relative frequency histogram of the same distribution.

c. Carefully examine the two histograms in parts a and b and explain why one of them might be

easier to understand. (Retain these solutions to use in Exercise 2.166, p. 123.)

FYI Use the computer or calculator commands on pages 61–63 to construct a histogram of a frequency distribution.

2.43 [EX02-043] The speeds of 55 cars were measured by a radar device on a city street:

27	23	22	38	43	24	35	26	28	18	20
25	23	22	52	31	30	41	45	29	27	43
29	28	27	25	29	28	24	37	28	29	18
26	33	25	27	25	34	32	36	22	32	33
21	23	24	18	48	23	16	38	26	21	23

a. Classify these data into a grouped frequency distribution by using class boundaries 12–18, 18–24, ... , 48–54.

b. Find the class width.

c. For the class 24–30, find the class midpoint, the lower class boundary, and the upper class boundary.

d. Construct a frequency histogram of these data.

FYI Use the computer or calculator commands on pages 61–63 to construct a histogram for a given set of data.

2.44 [EX02-044] The hemoglobin A_{1c} test, a blood test given to diabetics patients during their periodic checkups, indicates the level of control of blood sugar during the past 2 to 3 months. The following data values were obtained for 40 different diabetics patients at a university clinic:

6.5	5.0	5.6	7.6	4.8	8.0	7.5	7.9	8.0	9.2
6.4	6.0	5.6	6.0	5.7	9.2	8.1	8.0	6.5	6.6
5.0	8.0	6.5	6.1	6.4	6.6	7.2	5.9	4.0	5.7
7.9	6.0	5.6	6.0	6.2	7.7	6.7	7.7	8.2	9.0

a. Classify these A_{1c} values into a grouped frequency distribution using the classes 3.7–4.7, 4.7–5.7, and so on.

b. What are the class midpoints for these classes?

c. Construct a frequency histogram of these data.

2.45 [EX02-045] All of the third-graders at Roth Elementary School were given a physical-fitness strength test. The following data resulted:

12	22	6	9	2	9	5	9	3	5	16	1	22
18	6	12	21	23	9	10	24	21	17	11	18	19
17	5	14	16	19	19	18	3	4	21	16	20	15
14	17	4	5	22	12	15	18	20	8	10	13	20
6	9	2	17	15	9	4	15	14	19	3	24	

a. Construct a dotplot.

b. Prepare a grouped frequency distribution using classes 1–4, 4–7, and so on, and draw a histogram of the distribution. (Retain the solution for use in answering Exercise 2.75, p. 81.)

c. Prepare a grouped frequency distribution using classes 0–3, 3–6, 6–9, and so on, and draw a histogram of the distribution.

d. Prepare a grouped frequency distribution using class boundaries −2.5, 2.5, 7.5, 12.5, and so on, and draw a histogram of the distribution.

e. Prepare a grouped frequency distribution using classes of your choice, and draw a histogram of the distribution.

f. Describe the shape of the histogram found in parts b–e separately. Relate the distribution seen in the histogram to the distribution seen in the dotplot.

g. Discuss how the number of classes used and the choice of class boundaries used affect the appearance of the resulting histogram.

2.46 [EX02-046] People have marveled for years about the continuing eruptions of the geyser Old Faithful in Yellowstone National Park. The times of duration, in minutes, for a sample of 50 eruptions of Old Faithful are listed here.

4.00	3.75	2.25	1.67	4.25	3.92
4.53	1.85	4.63	2.00	1.80	4.00
4.33	3.77	3.67	3.68	1.88	1.97
4.00	4.50	4.43	3.87	3.43	4.13
4.13	2.33	4.08	4.35	2.03	4.57
4.62	4.25	1.82	4.65	4.50	4.10
4.28	4.25	1.68	3.43	4.63	2.50
4.58	4.00	4.60	4.05	4.70	3.20
4.60	4.73				

Source: http://www.stat.sc.edu/~west/javahtml/Histogram.html

: Roth
fitness

1	22
18	19
20	15
13	20
24	

n using
r a his-
olution
;1.)

n using
draw a

on using
and so
bution.

on using
ogram of

found in
tribution
:ion seen

used and
affect the
1.

. for years
eyser Old
te times of
eruptions

;	3.92
)	4.00
3	1.97
3	4.13
3	4.57
0	4.10
3	2.50
0	3.20

il

a. Draw a dotplot displaying the eruption-length data.

b. Draw a histogram of the eruption-length data using class boundaries 1.6–2.0–2.4–· ·–4.8.

c. Draw another histogram of the data using different class boundaries and widths.

d. Repeat part c.

e. Repeat parts a and b using the larger set of 107 eruptions found on the CD. **[DS02-046]**

f. Which graph, in your opinion, does the best job of displaying the distribution? Why?

g. Write a short paragraph describing the distribution.

2.47 [EX02-047] The Office of Coal, Nuclear, Electric and Alternate Fuels reported the following data as the costs (in cents) of the average revenue per kilowatt-hour for sectors in Arkansas:

6.61	7.61	6.99	7.48	5.10	7.56	6.65	5.93	7.92
5.52	7.47	6.79	8.27	7.50	7.44	6.36	5.20	5.48
7.69	8.74	5.75	6.94	7.70	6.67	4.59	5.96	7.26
5.38	8.88	7.49	6.89	7.25	6.89	6.41	5.86	8.04

a. Prepare a grouped frequency distribution for the average revenue per kilowatt-hour using class boundaries 4, 5, 6, 7, 8, 9.

b. Find the class width.

c. List the class midpoints.

d. Construct a relative frequency histogram of these data.

2.48 [EX02-048] Education has long been considered the ticket for upward mobility in the United States. In today's information age, a college education has become the minimum level of educational attainment to enter an increasingly competitive market for jobs with more than subsistence wages. A report from the SUNY Downstate Medical Center included a study of the suburban areas that surround U.S. cities. One variable that was reported was the percentage of suburban residents 25 years of age and older who attended at least some college:

49.3	75.2	64.7	66.1	51.8

••• Remainder of data on Student's Suite CD-ROM

Source: SUNY Downstate Medical Center, 2004

a. Prepare a grouped frequency distribution for the percent of suburban populations age 25 and older with any college attendance using class midpoints 25, 30, 35, . . . 75.

b. List the class boundaries.

c. Construct a relative frequency histogram of these data.

2.49 Can you think of variables whose distribution might yield the following different shapes? (See Figure 2.13 on p. 63 if necessary.)

a. A symmetrical, or normal, shape

b. A uniform shape

c. A skewed-to-the-right shape

d. A skewed-to-the-left shape

e. A bimodal shape

2.50 Skillbuilder Applet Exercise demonstrates the effect that the number of classes or bins has on the shape of a histogram.
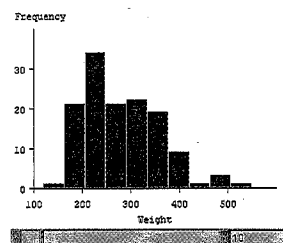
a. What shape distribution does using one class or bin produce?

b. What shape distribution does using two classes or bins produce?

c. What shape distribution does using 10 or 20 bins produce?

2.51 A survey of 100 resort club managers on their annual salaries resulted in the following frequency distribution. (See Exercise 2.39 on p. 69.)

Annual Salary ($1000s)	15-25	25-35	35-45	45-55	55-65
No. of Managers	12	37	26	19	6

a. Prepare a cumulative frequency distribution for the annual salaries.

b. Prepare a cumulative relative frequency distribution for the annual salaries.

c. Construct an ogive for the cumulative relative frequency distribution found in part b.

FYI Use the computer or calculator commands on page 66 to construct an ogive for a given set of data

2.52 [EX02-032] a. Prepare a cumulative relative frequency distribution for the variable "AP score" in Exercise 2.32.

b. Construct an ogive of the distribution.

2.53 [EX02-041] a. Prepare a cumulative relative frequency distribution for the variable "KSW test score" in Exercise 2.41.

b. Construct an ogive of the distribution.

2.54 Undergraduates who use loans to pay for college average $16,500 in debt when they graduate. The relative frequency distribution of their monthly debt after graduation is shown here:

Monthly debt, $	Less than 100	100–149	150–199	200–249	250–299	300 or more
Relative Frequency	0.17	0.17	0.17	0.19	0.1	0.2

Source: USA Today Snapshot, December 23, 2004

a. Prepare a cumulative relative frequency distribution for the monthly debt.

b. Construct an ogive for the cumulative relative frequency distribution found in part a.

2.55 [EX02-055] The Quality of Life in the Nation's 100 Largest Cities and Their Suburbs: New and Continuing Challenges for Improving Health and Well-Being, June 2004, reports on the percent of poor population living in a high-poverty neighborhood in 82 U.S. cities:

|				
29.8	21.4	32.0	5.9	27.8

••• Remainder of data on Student's Suite CD-ROM

Source: SUNY Downstate Medical Center

a. Prepare a grouped frequency distribution of the percent data using class midpoints of 0, 5, 10, . . . 45.

b. Prepare a grouped relative frequency distribution of these data.

c. Draw a relative frequency histogram of these data.

d. Prepare a cumulative relative frequency distribution of the same data.

e. Draw an ogive of these data.

2.56 The levels of various compounds resulted in the distribution graphs that follow. They all seem to be quite symmetrical about their centers, but they differ in their spreads.

a. For which histogram, A, B, C, or D, would you anticipate the numerical measure of spread to be the largest? the smallest?

b. Which two of the four histograms would you anticipate about the same difference between their smallest values and their largest values?

Figures for Exercise 2.56

Histogram A

Histogram B

Histogram C

Histogram D
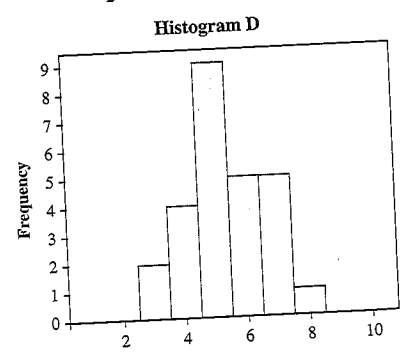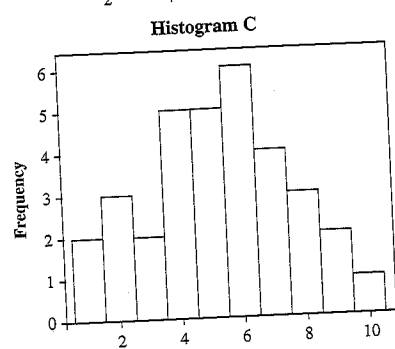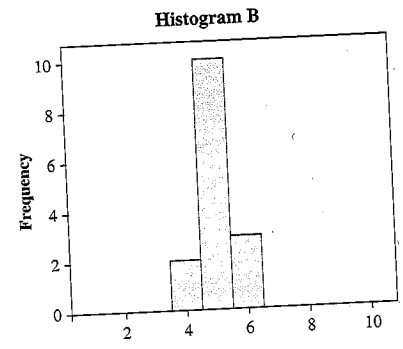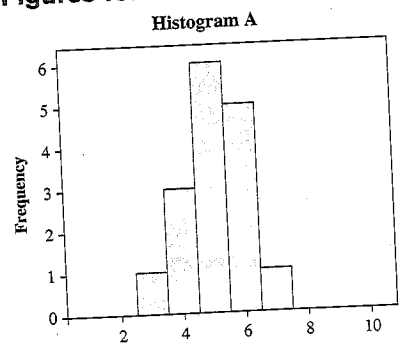
b. Move the red dot to the far left. What happens to the mean? What happens to the median?

c. Which measure of central tendency, the mean or the median, gives a better sense of the center when a maverick (or outlier) is present in the data?

2.65 The number of cars per apartment owned by a sample of tenants in a large complex is 1, 2, 1, 2, 2, 2, 1, 2, 3, 2. What is the mode?

2.66 Each year, approximately 160 colleges compete in the National Concrete Canoe Competition. Each team must design a seaworthy canoe from a substance not known for its capacity to float. The canoes must weigh between 100 and 350 pounds. Find the midrange.
Source: Reader's Digest, March 2005

2.67 a. Find the mean, median, mode, and midrange for the sample data 9, 6, 7, 9, 10, 8.

b. Verify and discuss the relationship between the answers in part a, as shown in Figure 2.20 on page 77.

2.68 Consider the sample 2, 4, 7, 8, 9. Find the following:

a. mean, $\bar{x}$ b. median, $\tilde{x}$

c. mode   d. midrange

2.69 Consider the sample 6, 8, 7, 5, 3, 7. Find the following:

a. mean, $\bar{x}$ b. median, $\tilde{x}$

c. mode   d. midrange

2.70 Fifteen randomly selected college students were asked to state the number of hours they slept the previous night. The resulting data values are 5, 6, 6, 8, 7, 7, 9, 5, 4, 8, 11, 6, 7, 8, 7. Find the following:

a. mean, $\bar{x}$ b. median, $\tilde{x}$

c. mode   d. midrange

2.71 [EX02-071] A random sample of 10 of the 2005 Nextel Cup NASCAR drivers produced the following ages: 33, 48, 41, 29, 40, 48, 44, 42, 49, 28.

a. Find the mean age for the 10 NASCAR drivers of the 2005 Nextel Cup.

b. Find the median age for the 10 NASCAR drivers of the 2005 Nextel Cup.

c. Find the midrange of age for the 10 NASCAR drivers of the 2005 Nextel Cup.

d. Find the mode, if one exists, for age for the 10 NASCAR drivers of the 2005 Nextel Cup.

2.72 [EX02-072] A constant objective in the manufacture of contact lenses is to improve those features that affect lens power and visual acuity. One such feature involves the tooling from which lenses are ultimately manufactured. The results of initial process development runs were examined for critical feature X. The resulting data are listed here:

0.026	0.027	0.024	0.023	0.034	0.035	0.035	0.033	0.034
0.033	0.032	0.038	0.041	0.041	0.021	0.022	0.027	0.032
0.023	0.023	0.024	0.017	0.023	0.019	0.027		

Source: Bausch & Lomb (variable not named and data coded at B&L's request)

a. Draw both a dotplot and a histogram of the critical feature X data.

b. Find the mean for critical feature X.

c. Find the median for critical feature X.

d. Find the midrange for critical feature X.

e. Find the mode, if one exists, for critical feature X.

f. What feature of the distribution, as shown by the graphs found in part a, seems unusual? Where do the answers found in parts b, c, and d fall relative to the distribution? Explain.

g. Identify at least one possible cause for this seemingly unusual situation.

2.73 [EX02-073] One measure of airline performance is overall flight on-time rates. For January 2005, the on-time arrival rates of domestic

flights at the 31 largest U.S. airports were as follows.

ATL	69.09	BWI	74.01	BOS	62.14

••• Remainder of data on Student's Suite CD-ROM

Source: U.S. Department of Transportation, Bureau of Transportation Statistics

a. Find the mean on-time arrival rate for January 2005.

b. Find the median on-time arrival rate for January 2005.

c. Construct a stem-and-leaf display of the data.

d. Describe the relationship between the mean and the median and what properties of the data cause the mean to be lower than the median.

(Retain these solutions to use in Exercise 2.99 on p. 92.)

2.74 The "average" is a commonly reported statistic. This single bit of information can be very informative or very misleading, with the mean and median being the two most commonly reported.

a. The mean is a useful measure, but it can be misleading. Describe a circumstance when the mean is very useful as the average and a circumstance when the mean is very misleading as the average.

b. The median is a useful measure, but it can be misleading. Describe a circumstance when the median is very useful as the average and a circumstance when the median is very misleading as the average.

2.75 [EX02-075] All the third-graders at Roth Elementary School were given a physical fitness strength test. These data resulted:

12	22	6	9	2	9	5	9	3	5	16	1	22
18	6	12	21	23	9	10	24	21	17	11	18	19
17	5	14	16	19	19	18	3	4	21	16	20	15
14	17	4	5	22	12	15	18	20	8	10	13	20
6	9	2	17	15	9	4	15	14	19	3	24	

a. Construct a dotplot.

b. Find the mode.

c. Prepare a grouped frequency distribution using classes 1–4, 4–7, and so on, and draw a histogram of the distribution.

d. Describe the distribution; specifically, is the distribution bimodal (about what values)?

e. Compare your answers in parts a and c, and comment on the relationship between the mode and the modal values in these data.

f. Could the discrepancy found in the comparison in part e occur when using an ungrouped frequency distribution? Explain.

g. Explain why, in general, the mode of a set of data does not necessarily give us the same information as the modal values do.

2.76 [EX02-076] Consumers are frequently cautioned against eating too much food that is high in calories, fat, and sodium for numerous health and fitness reasons. *Nutrition Action HealthLetter* published a list of popular low-fat brands of hot dogs commonly labeled "fat-free," "reduced fat," "low-fat," "light," and so on, together with their calories, fat content, and sodium. All quantities are for one hot dog:

Hot Dog Brand	Calories	Fat (g)	Sodium (mg)
Ball Park Fat Free Beef Franks	50	0	460
Butterball Fat Free Franks	40	0	490

••• Remainder of data on Student's Suite CD-ROM

Source: Nutrition Action HealthLetter, "On the Links," July/August 1998, pp. 12–13

a. Find the mean, median, mode, and midrange of the calories, fat, and sodium contents of all the frankfurters listed. Use a table to summarize your results.

b. Construct a dotplot of the fat contents. Locate the mean, median, mode, and midrange on the plot.

c. In the summer of 2005, the winner of Nathan's Famous Fourth of July Hot Dog Eating Contest consumed 49 hot dogs in 12 minutes. If he had been served the median hot dog, how many calories, grams of fat, and milligrams of sodium did he consume in the single sitting? If the recommended daily allowance for sodium intake is 2400 mg, did he likely exceed it? Explain.

2.77 [EX02-077] The number of runs scored by major league baseball (MLB) teams is likely influenced by whether the game is played at home or at the opponents' ballpark. In an attempt to measure

j. Compare the results found in part i to those found in parts d and f.

k. How accurate were your predications to part h? Explain.

2.80 You are responsible for planning the parking needed for a new 256-unit apartment complex, and you're told to base the needs on the statistic "average number of vehicles per household is 1.9."

a. Which average (mean, median, mode, midrange) will be helpful to you? Explain.

b. Explain why 1.9 cannot be the median, the mode, or the midrange for the variable "number of vehicles."

c. If the owner wants parking that will accommodate 90% of all the tenants who own vehicles, how many spaces must you plan for?

2.81 In what states do the residents pay the most taxes? the least? Perhaps it depends on the variable used to measure amount of taxes paid. In 2004 the Tax Policy Center reported the following statistics about the average annual 2002 taxes and percent of personal income paid per person by state.

	Taxes per Capita	Rank	% Personal Income	Rank
Hawaii	$2748	1	9.6	1
South Dakota	$1283	50	4.8	47
New Hampshire	$1478	45	4.4	50

Sources: Federation of Tax Administrators (2004) and U.S. Bureau of the Census and Bureau of Economic Analysis, http://taxpolicycenter.org/TaxFacts/TFDB/TFTemplate.cfm?Docid=309&Topic2id=90

a. Compare and contrast the variables "taxes per capita" and "percent of personal income." How do you account for the differences in ranks for South Dakota and New Hampshire?

b. Based on this information, using the highest and lowest per state amount of taxes paid per person, what was the "average" amount paid per person?

c. Based on this information, using the highest and lowest percent of income per state paid per person, what was the "average" percent paid per person?

d. Explain why your answers to parts b and c are the only average value you can determine from the given information. What is the name of this average?

2.82 Your instructor and your class have made a deal on the exam just taken and currently being graded. If the class attains a mean score of 74 or better, there will be no homework on the coming weekend. If the class mean is 72 or below, then not only will there be homework as usual but all of the class members will have to show up on Saturday and do 2 hours of general cleanup around the school grounds as a community service project. There are 15 students in your class. Your instructor has graded the first 14 exams, and their mean score is 73.5. Your exam is the only one left to grade.

a. What score must you get in order for the class to win the deal?

b. What score must you get in order that the class will not have to do the community service work?

2.83 Starting with the data values 70 and 100, add three data values to the sample so that the sample has the following: (Justify your answer in each case.)

a. Mean of 100   b. Median of 70

c. Mode of 87   d. Midrange of 70

e. Mean of 100 and a median of 70

f. Mean of 100 and a mode of 87

g. Mean of 100 and a midrange of 70

h. Mean of 100, a median of 70, and a mode of 87

2.84 Skillbuilder Applet Exercise matches means with corresponding histograms. After several practice rounds using "New Plots," explain your method of matching.