1.2 TECHNOLOGY CORNER

TI-Nspire instructions in Appendix B; HP Prime instructions on the book's Web site.

2. Histograms on the calculator page 36

Section 1.2 Exercises

37. **Feeling sleepy?** Students in a college statistics class responded to a survey designed by their teacher. One of the survey questions was "How much sleep did you get last night?" Here are the data (in hours):

9	6	8	6	8	8	6	6.5	6	7	9	4	3	4
5	6	11	6	3	6	6	10	7	8	4.5	9	7	7

(a) Make a dotplot to display the data.

(b) Describe the overall pattern of the distribution and any departures from that pattern.

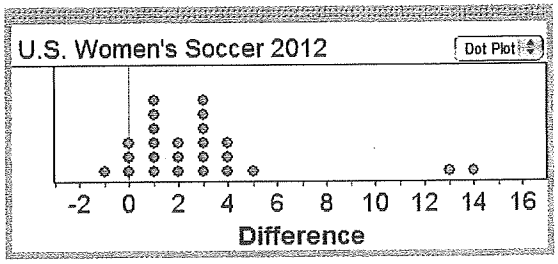
38. **Olympic gold!** The following table displays the total number of gold medals won by a sample of countries in the 2012 Summer Olympic Games in London.

Country	Gold medals	Country	Gold medals
Sri Lanka	0	Thailand	0
China	38	Kuwait	0
Vietnam	0	Bahamas	1
Great Britain	29	Kenya	2
Norway	2	Trinidad and Tobago	1
Romania	2	Greece	0
Switzerland	2	Mozambique	0
Armenia	0	Kazakhstan	7
Netherlands	6	Denmark	2
India	0	Latvia	1
Georgia	1	Czech Republic	4
Kyrgyzstan	0	Hungary	8
Costa Rica	0	Sweden	1
Brazil	3	Uruguay	0
Uzbekistan	1	United States	46

(a) Make a dotplot to display these data. Describe the overall pattern of the distribution and any departures from that pattern.

(b) Overall, 205 countries participated in the 2012 Summer Olympics, of which 54 won at least one gold medal. Do you believe that the sample of countries listed in the table is representative of this larger population? Why or why not?

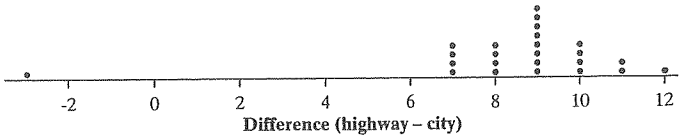
39. **U.S. women's soccer—2012** Earlier, we examined data on the number of goals scored by the U.S. women's soccer team in games played in the 12 months prior to the 2012 Olympics. The dotplot below displays the goal differential for those same games, computed as U.S. score minus opponent's score. pg 26



(a) Explain what the dot above −1 represents.

(b) What does the graph tell us about how well the team did in 2012? Be specific.

40. **Fuel efficiency** In an earlier example, we examined data on highway gas mileages of model year 2012 midsize cars. The following dotplot shows the difference (highway – city) in EPA mileage ratings for each of the 24 car models from the earlier example.

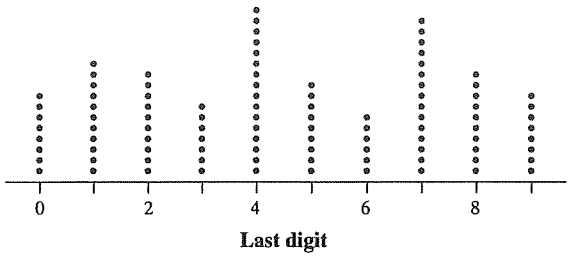


(a) Explain what the dot above 12 represents.

(b) What does the graph tell us about fuel economy in the city versus on the highway for these car models? Be specific.

41. **Dates on coins** Suppose that you and your friends emptied your pockets of coins and recorded the year marked on each coin. The distribution of dates would be skewed to the left. Explain why.

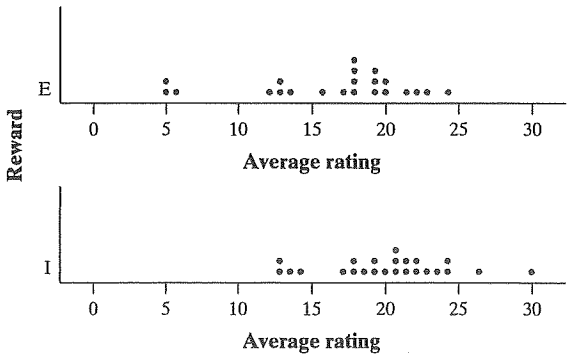
42. **Phone numbers** The dotplot below displays the last digit of 100 phone numbers chosen at random from a phone book. Describe the shape of the distribution. Does this shape make sense to you? Explain.



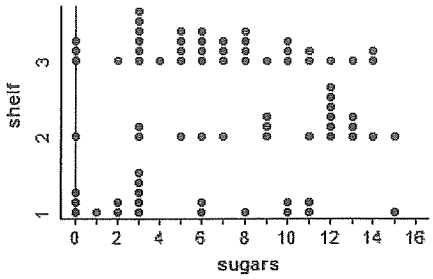
43. **Creative writing** Do external rewards—things like money, praise, fame, and grades—promote creativity? Researcher Teresa Amabile designed an experiment to find out. She recruited 47 experienced creative writers who were college students and divided them into two groups using a chance process (like drawing names from a hat). The students in one group were given a list of statements about external reasons (E) for writing, such as public recognition, making money, or pleasing their parents. Students in the other group were given a list of statements about internal reasons (I) for writing, such as expressing yourself and enjoying playing with words. Both groups were then instructed to write a poem about laughter. Each student's poem was rated separately by 12 different poets using a creativity scale.[24] These ratings were averaged to obtain an overall creativity score for each poem.

pg 30

Dotplots of the two groups' creativity scores are shown below. Compare the two distributions. What do you conclude about whether external rewards promote creativity?



44. **Healthy cereal?** Researchers collected data on 77 brands of cereal at a local supermarket.[25] For each brand, the sugar content (grams per serving) and the shelf in the store on which the cereal was located (1 = bottom, 2 = middle, 3 = top) were recorded. A dotplot of the data is shown below. Compare the three distributions. Critics claim that supermarkets tend to put sugary kids' cereals on lower shelves, where the kids can see them. Do the data from this study support this claim?



45. **Where do the young live?** Below is a stemplot of the percent of residents aged 25 to 34 in each of the 50 states. As in the stemplot for older residents (page 33), the stems are whole percents, and the leaves are tenths of a percent. This time, each stem has been split in two, with values having leaves 0 through 4 placed on one stem, and values ending in 5 through 9 placed on another stem.

```
11 | 44
11 | 66778
12 | 0134
12 | 666778888
13 | 0000001111444
13 | 7788999
14 | 0044
14 | 567
15 | 11
15 |
16 | 0
```

(a) Why did we split stems?

(b) Give an appropriate key for this stemplot.

(c) Describe the shape, center, and spread of the distribution. Are there any outliers?

46. **Watch that caffeine!** The U.S. Food and Drug Administration (USFDA) limits the amount of caffeine in a 12-ounce can of carbonated beverage to 72 milligrams. That translates to a maximum of 48 milligrams of caffeine per 8-ounce serving. Data on the caffeine content of popular soft drinks (in milligrams per 8-ounce serving) are displayed in the stemplot below.

```
1 | 556
2 | 033344
2 | 55667778888899
3 | 113
3 | 55567778
4 | 33
4 | 77
```

(a) Why did we split stems?

(b) Give an appropriate key for this graph.

(c) Describe the shape, center, and spread of the distribution. Are there any outliers?

47. **El Niño and the monsoon** It appears that El Niño, the periodic warming of the Pacific Ocean west of South America, affects the monsoon rains that are essential for agriculture in India. Here are the monsoon rains (in millimeters) for the 23 strong El Niño years between 1871 and 2004:[26]

628	669	740	651	710	736	717	698	653	604	781	784
790	811	830	858	858	896	806	790	792	957	872	

(a) To make a stemplot of these rainfall amounts, round the data to the nearest 10, so that stems are hundreds of millimeters and leaves are tens of millimeters. Make two stemplots, with and without splitting the stems. Which plot do you prefer? Why?

(b) Describe the shape, center, and spread of the distribution.

(c) The average monsoon rainfall for all years from 1871 to 2004 is about 850 millimeters. What effect does El Niño appear to have on monsoon rains?

48. **Shopping spree** A marketing consultant observed 50 consecutive shoppers at a supermarket. One variable of interest was how much each shopper spent in the store. Here are the data (in dollars), arranged in increasing order:

3.11	8.88	9.26	10.81	12.69	13.78	15.23	15.62	17.00	17.39
18.36	18.43	19.27	19.50	19.54	20.16	20.59	22.22	23.04	24.47
24.58	25.13	26.24	26.26	27.65	28.06	28.08	28.38	32.03	34.98
36.37	38.64	39.16	41.02	42.97	44.08	44.67	45.40	46.69	48.65
50.39	52.75	54.80	59.07	61.22	70.32	82.70	85.76	86.37	93.34

(a) Round each amount to the nearest dollar. Then make a stemplot using tens of dollars as the stems and dollars as the leaves.

(b) Make another stemplot of the data by splitting stems. Which of the plots shows the shape of the distribution better?

(c) Write a few sentences describing the amount of money spent by shoppers at this supermarket.

49. **Do women study more than men?** We asked the students in a large first-year college class how many minutes they studied on a typical weeknight. Here are the responses of random samples of 30 women and 30 men from the class:

Women					Men				
180	120	180	360	240	90	120	30	90	200
120	180	120	240	170	90	45	30	120	75
150	120	180	180	150	150	120	60	240	300
200	150	180	150	180	240	60	120	60	30
120	60	120	180	180	30	230	120	95	150
90	240	180	115	120	0	200	120	120	180

(a) Examine the data. Why are you not surprised that most responses are multiples of 10 minutes? Are there any responses you consider suspicious?

(b) Make a back-to-back stemplot to compare the two samples. Does it appear that women study more than men (or at least claim that they do)? Justify your answer.

50. **Basketball playoffs** Here are the numbers of points scored by teams in the California Division I-AAA high school basketball playoffs in a single day's games:[27]

71	38	52	47	55	53	76	65	77	63	65	63	68
54	64	62	87	47	64	56	78	64	58	51	91	74
71	41	67	62	106	46							

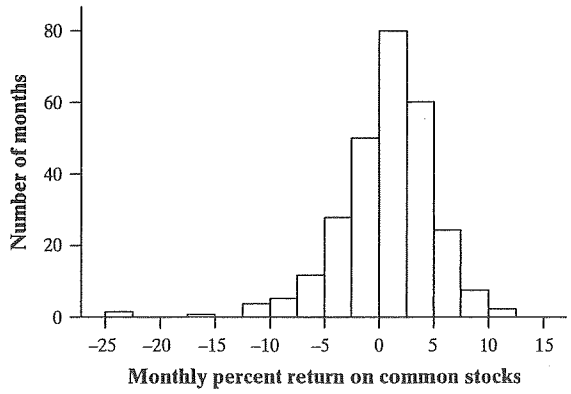
On the same day, the final scores of games in Division V-AA were

98	45	67	44	74	60	96	54	92	72	93	46
98	67	62	37	37	36	69	44	86	66	66	58

(a) Construct a back-to-back stemplot to compare the points scored by the 32 teams in the Division I-AAA playoffs and the 24 teams in the Division V-AA playoffs.

(b) Write a few sentences comparing the two distributions.

51. **Returns on common stocks** The return on a stock is the change in its market price plus any dividend payments made. Total return is usually expressed as a percent of the beginning price. The figure below shows a histogram of the distribution of the monthly returns for all common stocks listed on U.S. markets over a 273-month period.[28] The extreme low outlier represents the market crash of October 1987, when stocks lost 23% of their value in one month.



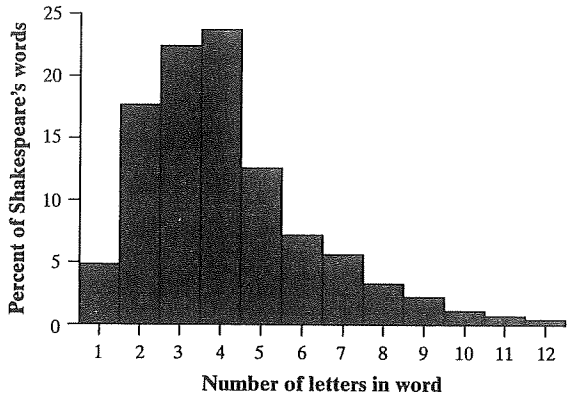
(a) Describe the overall shape of the distribution of monthly returns.

(b) What is the approximate center of this distribution?

(c) Approximately what were the smallest and largest monthly returns, leaving out the outliers?

(d) A return less than zero means that stocks lost value in that month. About what percent of all months had returns less than zero?

52. **Shakespeare** The histogram below shows the distribution of lengths of words used in Shakespeare's plays.[29] Describe the shape, center, and spread of this distribution.



53. **Traveling to work** How long do people travel each day to get to work? The following table gives the average travel times to work (in minutes) for workers in each state and the District of Columbia who are at least 16 years old and don't work at home.[30]

AL	23.6	LA	25.1	OH	22.1
AK	17.7	ME	22.3	OK	20.0
AZ	25.0	MD	30.6	OR	21.8
AR	20.7	MA	26.6	PA	25.0
CA	26.8	MI	23.4	RI	22.3
CO	23.9	MN	22.0	SC	22.9
CT	24.1	MS	24.0	SD	15.9
DE	23.6	MO	22.9	TN	23.5
FL	25.9	MT	17.6	TX	24.6
GA	27.3	NE	17.7	UT	20.8
HI	25.5	NV	24.2	VT	21.2
ID	20.1	NH	24.6	VA	26.9
IL	27.9	NJ	29.1	WA	25.2
IN	22.3	NM	20.9	WV	25.6
IA	18.2	NY	30.9	WI	20.8
KS	18.5	NC	23.4	WY	17.9
KY	22.4	ND	15.5	DC	29.2

(a) Make a histogram of the travel times using classes of width 2 minutes, starting at 14 minutes. That is, the first class is 14 to 16 minutes, the second is 16 to 18 minutes, and so on.

(b) The shape of the distribution is a bit irregular. Is it closer to symmetric or skewed? Describe the center and spread of the distribution. Are there any outliers?

54. **Carbon dioxide emissions** Burning fuels in power plants and motor vehicles emits carbon dioxide (CO_2), which contributes to global warming. The table below displays CO_2 emissions per person from countries with populations of at least 20 million.[31]

(a) Make a histogram of the data using classes of width 2, starting at 0.

(b) Describe the shape, center, and spread of the distribution. Which countries are outliers?

Carbon dioxide emissions (metric tons per person)

Country	CO_2	Country	CO_2
Algeria	2.6	Mexico	3.7
Argentina	3.6	Morocco	1.4
Australia	18.4	Myanmar	0.2
Bangladesh	0.3	Nepal	0.1
Brazil	1.8	Nigeria	0.4
Canada	17.0	Pakistan	0.8
China	3.9	Peru	1.0
Colombia	1.3	Philippines	0.9
Congo	0.2	Poland	7.8
Egypt	2.0	Romania	4.2
Ethiopia	0.1	Russia	10.8
France	6.2	Saudi Arabia	13.8
Germany	9.9	South Africa	7.0
Ghana	0.3	Spain	7.9
India	1.1	Sudan	0.3
Indonesia	1.6	Tanzania	0.1
Iran	6.0	Thailand	3.3
Iraq	2.9	Turkey	3.0
Italy	7.8	Ukraine	6.3
Japan	9.5	United Kingdom	8.8
Kenya	0.3	United States	19.6
Korea, North	3.3	Uzbekistan	4.2
Korea, South	9.3	Venezuela	5.4
Malaysia	5.5	Vietnam	1.0

55. **DRP test scores** There are many ways to measure the reading ability of children. One frequently used test is the Degree of Reading Power (DRP). In a research study on third-grade students, the DRP was administered to 44 students.[32] Their scores were:

40	26	39	14	42	18	25	43	46	27	19
47	19	26	35	34	15	44	40	38	31	46
52	25	35	35	33	29	34	41	49	28	52
47	35	48	22	33	41	51	27	14	54	45

Make a histogram to display the data. Write a paragraph describing the distribution of DRP scores.

56. **Drive time** Professor Moore, who lives a few miles outside a college town, records the time he takes to drive to the college each morning. Here are the times (in minutes) for 42 consecutive weekdays:

8.25	7.83	8.30	8.42	8.50	8.67	8.17	9.00	9.00	8.17	7.92
9.00	8.50	9.00	7.75	7.92	8.00	8.08	8.42	8.75	8.08	9.75
8.33	7.83	7.92	8.58	7.83	8.42	7.75	7.42	6.75	7.42	8.50
8.67	10.17	8.75	8.58	8.67	9.17	9.08	8.83	8.67		

Make a histogram to display the data. Write a paragraph describing the distribution of Professor Moore's drive times.

57. **The statistics of writing style** Numerical data can distinguish different types of writing and, sometimes, even individual authors. Here are data on the percent of words of 1 to 15 letters used in articles in *Popular Science* magazine:[33]

Length:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Percent:	3.6	14.8	18.7	16.0	12.5	8.2	8.1	5.9	4.4	3.6	2.1	0.9	0.6	0.4	0.2

(a) Make a histogram of this distribution. Describe its shape, center, and spread.

(b) How does the distribution of lengths of words used in *Popular Science* compare with the similar distribution for Shakespeare's plays in Exercise 52? Look in particular at short words (2, 3, and 4 letters) and very long words (more than 10 letters).

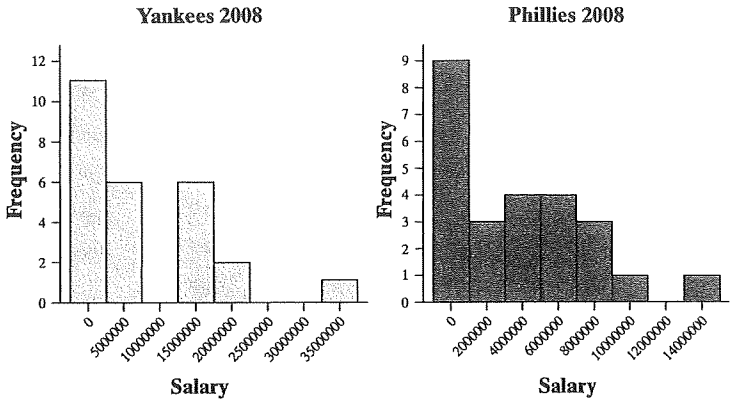
58. **Chest out, Soldier!** In 1846, a published paper provided chest measurements (in inches) of 5738 Scottish militiamen. The table below summarizes the data.[34]

Chest size	Count	Chest size	Count
33	3	41	934
34	18	42	658
35	81	43	370
36	185	44	92
37	420	45	50
38	749	46	21
39	1073	47	4
40	1079	48	1

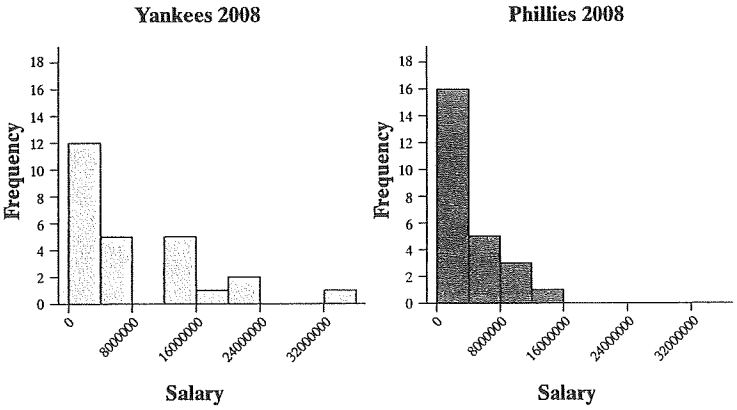
(a) Make a histogram of this distribution.

(b) Describe the shape, center, and spread of the chest measurements distribution. Why might this information be useful?

59. **Paying for championships** Does paying high salaries lead to more victories in professional sports? The New York Yankees have long been known for having Major League Baseball's highest team payroll. And over the years, the team has won many championships. This strategy didn't pay off in 2008, when the Philadelphia Phillies won the World Series. Maybe the Yankees didn't spend enough money that year. The graph below shows histograms of the salary distributions for the two teams during the 2008 season. Why can't you use this graph to effectively compare the team payrolls?



60. **Paying for championships** Refer to Exercise 59. Here is another graph of the 2008 salary distributions for the Yankees and the Phillies. Write a few sentences comparing these two distributions.



61. **Birth months** Imagine asking a random sample of 60 students from your school about their birth months. Draw a plausible graph of the distribution of birth months. Should you use a bar graph or a histogram to display the data?

62. **Die rolls** Imagine rolling a fair, six-sided die 60 times. Draw a plausible graph of the distribution of die rolls. Should you use a bar graph or a histogram to display the data?

63. **Who makes more?** A manufacturing company is reviewing the salaries of its full-time employees below the executive level at a large plant. The clerical staff is almost entirely female, while a majority of the production workers and technical staff is male. As a result, the distributions of salaries for male and female employees may be quite different. The following table gives the frequencies and relative frequencies for women and men.

	Women		Men	
Salary ($1000)	Number	%	Number	%
10–15	89	11.8	26	1.1
15–20	192	25.4	221	9.0
20–25	236	31.2	677	27.6
25–30	111	14.7	823	33.6
30–35	86	11.4	365	14.9
35–40	25	3.3	182	7.4
40–45	11	1.5	91	3.7
45–50	3	0.4	33	1.3
50–55	2	0.3	19	0.8
55–60	0	0.0	11	0.4
60–65	0	0.0	0	0.0
65–70	1	0.1	3	0.1
Total	756	100.1	2451	99.9

(a) Explain why the total for women is greater than 100%.

(b) Make histograms for these data, choosing the vertical scale that is most appropriate for comparing the two distributions.

(c) Write a few sentences comparing the salary distributions for men and women.

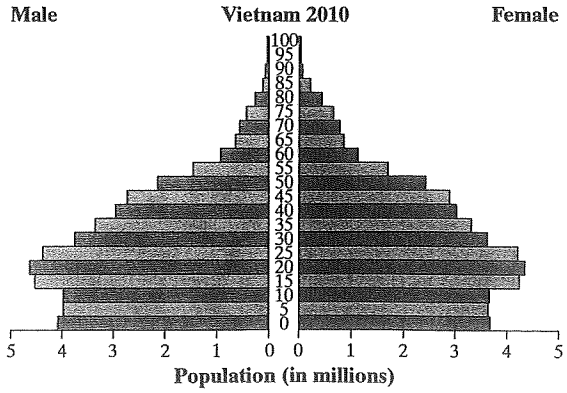
64. **Comparing AP® scores** The table below gives the distribution of grades earned by students taking the AP® Calculus AB and AP® Statistics exams in 2012.[35]

	No. of exams	Grade 5	4	3	2	1
Calculus AB	266,994	67,394	45,523	46,526	27,216	80,335
Statistics	153,859	19,267	32,521	39,355	27,684	35,032

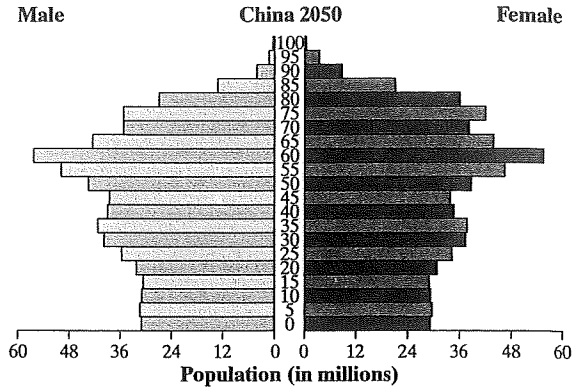
(a) Make an appropriate graphical display to compare the grade distributions for AP® Calculus AB and AP® Statistics.

(b) Write a few sentences comparing the two distributions of exam grades.

65. **Population pyramids** A population pyramid is a helpful graph for examining the distribution of a country's population. Here is a population pyramid for Vietnam in the year 2010. Describe what the graph tells you about Vietnam's population that year. Be specific.



66. **Population pyramids** Refer to Exercise 65. Here is a graph of the projected population distribution for China in the year 2050. Describe what the graph suggests about China's future population. Be specific.



67. **Student survey** A survey of a large high school class asked the following questions:

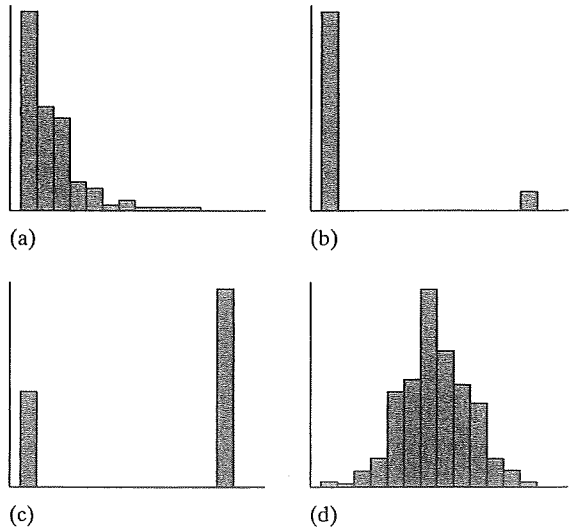
(i) Are you female or male? (In the data, male = 0, female = 1.)

(ii) Are you right-handed or left-handed? (In the data, right = 0, left = 1.)

(iii) What is your height in inches?

(iv) How many minutes do you study on a typical weeknight?

The figure below shows graphs of the student responses, in scrambled order and without scale markings. Which graph goes with each variable? Explain your reasoning.



68. **Choose a graph** What type of graph or graphs would you make in a study of each of the following issues at your school? Explain your choices.

(a) Which radio stations are most popular with students?

(b) How many hours per week do students study?

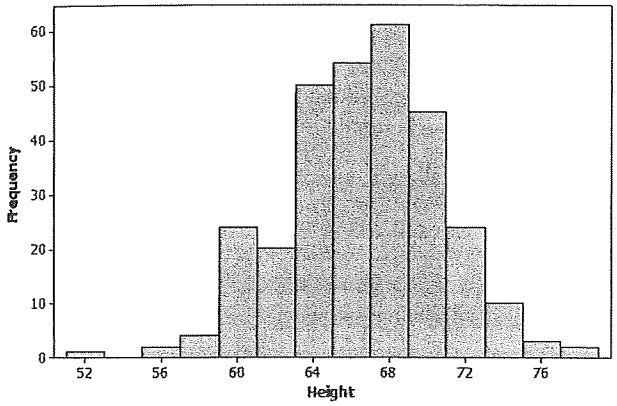
(c) How many calories do students consume per day?

Multiple choice: Select the best answer for Exercises 69 to 74.

69. Here are the amounts of money (cents) in coins carried by 10 students in a statistics class: 50, 35, 0, 97, 76, 0, 0, 87, 23, 65. To make a stemplot of these data, you would use stems

(a) 0, 1, 2, 3, 4, 5, 6, 7, 8, 9.
(b) 0, 2, 3, 5, 6, 7, 8, 9.
(c) 0, 3, 5, 6, 7.
(d) 00, 10, 20, 30, 40, 50, 60, 70, 80, 90.
(e) None of these.

70. The histogram below shows the heights of 300 randomly selected high school students. Which of the following is the best description of the shape of the distribution of heights?



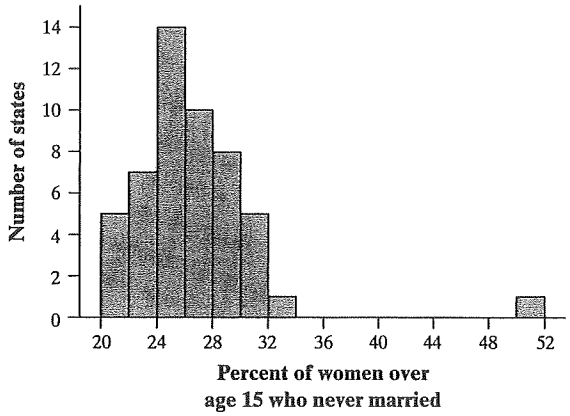
(a) Roughly symmetric and unimodal
(b) Roughly symmetric and bimodal
(c) Roughly symmetric and multimodal
(d) Skewed to the left
(e) Skewed to the right

71. You look at real estate ads for houses in Naples, Florida. There are many houses ranging from $200,000 to $500,000 in price. The few houses on the water, however, have prices up to $15 million. The distribution of house prices will be

(a) skewed to the left.
(b) roughly symmetric.
(c) skewed to the right.
(d) unimodal.
(e) too high.

72. The following histogram shows the distribution of the percents of women aged 15 and over who have never married in each of the 50 states and the District of Columbia. Which of the following statements about the histogram is correct?

(a) The center of the distribution is about 36%.
(b) There are more states with percents above 32 than there are states with percents less than 24.
(c) It would be better if the values from 34 to 50 were deleted on the horizontal axis so there wouldn't be a large gap.
(d) There was one state with a value of exactly 33%.
(e) About half of the states had percents between 24% and 28%.



73. When comparing two distributions, it would be best to use relative frequency histograms rather than frequency histograms when

(a) the distributions have different shapes.
(b) the distributions have different spreads.
(c) the distributions have different centers.
(d) the distributions have different numbers of observations.
(e) at least one of the distributions has outliers.

74. Which of the following is the best reason for choosing a stemplot rather than a histogram to display the distribution of a quantitative variable?

(a) Stemplots allow you to split stems; histograms don't.
(b) Stemplots allow you to see the values of individual observations.
(c) Stemplots are better for displaying very large sets of data.
(d) Stemplots never require rounding of values.
(e) Stemplots make it easier to determine the shape of a distribution.

75. **Baseball players** (Introduction) Here is a small part of a data set that describes Major League Baseball players as of opening day of the 2012 season:

Player	Team	Position	Age	Height	Weight	Salary
Rodriguez, Alex	Yankees	Infielder	37	6-3	225	29,000,000
Gonzalez, Adrian	Dodgers	Infielder	30	6-2	225	21,000,000
Cruz, Nelson	Rangers	Outfielder	32	6-2	240	5,000,000
Lester, Jon	Red Sox	Pitcher	28	6-4	240	7,625,000
Strasburg, Stephen	Nationals	Pitcher	24	6-4	220	3,000,000

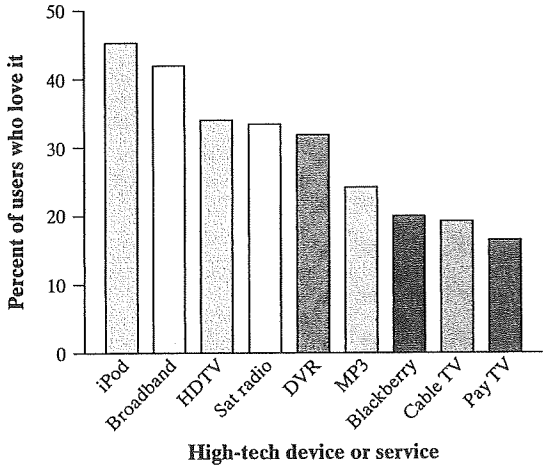
(a) What individuals does this data set describe?

(b) In addition to the player's name, how many variables does the data set contain? Which of these variables are categorical and which are quantitative?

76. **I love my iPod!** (1.1) The rating service Arbitron asked adults who used several high-tech devices and services whether they "loved" using them. Below is a graph of the percents who said they did.[36]

(a) Summarize what this graph tells you in a sentence or two.

(b) Would it be appropriate to make a pie chart of these data? Why or why not?



77. **Risks of playing soccer** (1.1) A study in Sweden looked at former elite soccer players, people who had played soccer but not at the elite level, and people of the same age who did not play soccer. Here is a two-way table that classifies these individuals by whether or not they had arthritis of the hip or knee by their mid-fifties:[37]

	Elite	Non-Elite	Did not play
Arthritis	10	9	24
No arthritis	61	206	548

(a) What percent of the people in this study were elite soccer players? What percent had arthritis?

(b) What percent of the elite soccer players had arthritis? What percent of those who had arthritis were elite soccer players?

78. **Risks of playing soccer** (1.1) Refer to Exercise 77. We suspect that the more serious soccer players have more arthritis later in life. Do the data confirm this suspicion? Give graphical and numerical evidence to support your answer.

1.3 Describing Quantitative Data with Numbers

WHAT YOU WILL LEARN **By the end of the section, you should be able to:**

- Calculate measures of center (mean, median).
- Calculate and interpret measures of spread (range, *IQR*, standard deviation).
- Choose the most appropriate measure of center and spread in a given setting.
- Identify outliers using the $1.5 \times IQR$ rule.
- Make and interpret boxplots of quantitative data.
- Use appropriate graphs and numerical summaries to compare distributions of quantitative variables.

How long do people spend traveling to work? The answer may depend on where they live. Here are the travel times in minutes for 15 workers in North Carolina, chosen at random by the Census Bureau:[38]

30 20 10 40 25 20 10 60 15 40 5 30 12 10 10