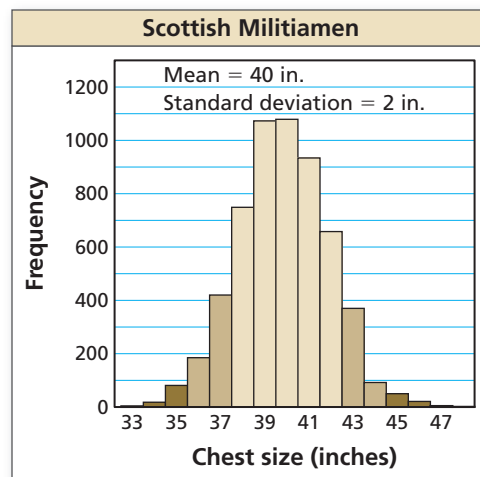


# 11.3 Shapes of Distributions

**Essential Question** How can you use a histogram to characterize the basic shape of a distribution?

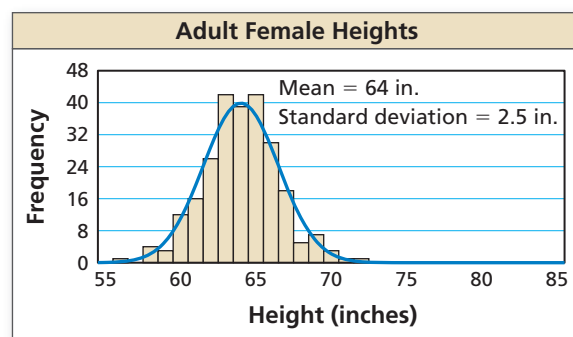
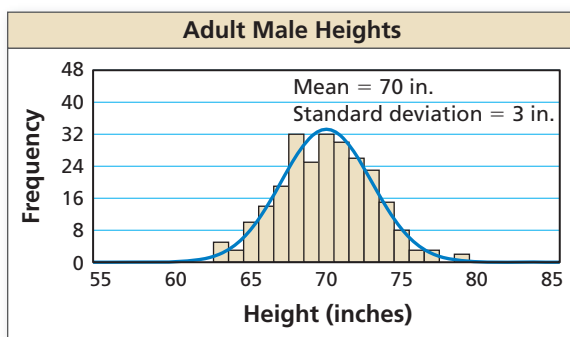
## EXPLORATION 1 Analyzing a Famous Symmetric Distribution

**Work with a partner.** A famous data set was collected in Scotland in the mid-1800s. It contains the chest sizes, measured in inches, of 5738 men in the Scottish Militia. Estimate the percent of the chest sizes that lie within (a) 1 standard deviation of the mean, (b) 2 standard deviations of the mean, and (c) 3 standard deviations of the mean. Explain your reasoning.



## EXPLORATION 2 Comparing Two Symmetric Distributions

**Work with a partner.** The graphs show the distributions of the heights of 250 adult American males and 250 adult American females.



- Which data set has a smaller standard deviation? Explain what this means in the context of the problem.
- Estimate the percent of male heights between 67 inches and 73 inches.

### ATTENDING TO PRECISION

To be proficient in math, you need to express numerical answers with a level of precision appropriate for the problem's context.

## Communicate Your Answer

- How can you use a histogram to characterize the basic shape of a distribution?
- All three distributions in Explorations 1 and 2 are roughly symmetric. The histograms are called “bell-shaped.”
  - What are the characteristics of a symmetric distribution?
  - Why is a symmetric distribution called “bell-shaped?”
  - Give two other real-life examples of symmetric distributions.

# 11.3 Lesson

## Core Vocabulary

**Previous**  
 histogram  
 frequency table

## What You Will Learn

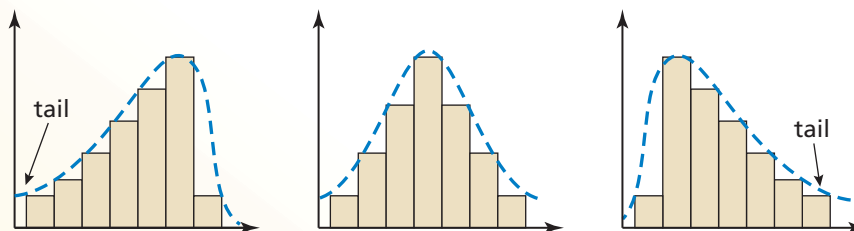
- ▶ Describe the shapes of data distributions.
- ▶ Use the shapes of data distributions to choose appropriate measures.
- ▶ Compare data distributions.

## Describing the Shapes of Data Distributions

Recall that a histogram is a bar graph that shows the frequency of data values in intervals of the same size. A histogram is another useful data display that shows the shape of a distribution.

## Core Concept

### Symmetric and Skewed Distributions



Skewed left

- The “tail” of the graph extends to the left.
- Most of the data are on the right.

Symmetric

- The data on the right of the distribution are approximately a mirror image of the data on the left of the distribution.

Skewed right

- The “tail” of the graph extends to the right.
- Most of the data are on the left.

### STUDY TIP

If all the bars of a histogram are about the same height, then the distribution is a *flat*, or *uniform*, distribution. A uniform distribution is also symmetric.

Number of tickets sold	Frequency
1–8	5
9–16	9
17–24	16
25–32	25
33–40	20
41–48	8
49–56	7

Number of pounds	Frequency
1–10	7
11–20	8
21–30	10
31–40	16
41–50	34
51–60	15

### EXAMPLE 1 Describing the Shape of a Distribution

The frequency table shows the numbers of raffle tickets sold by students in your grade. Display the data in a histogram. Describe the shape of the distribution.

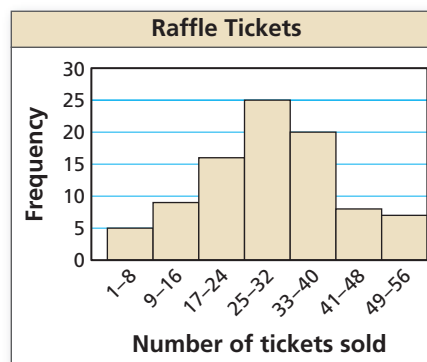
#### SOLUTION

**Step 1** Draw and label the axes.

**Step 2** Draw a bar to represent the frequency of each interval.

The data on the right of the distribution are approximately a mirror image of the data on the left of the distribution.

- ▶ So, the distribution is symmetric.



### Monitoring Progress Help in English and Spanish at [BigIdeasMath.com](http://BigIdeasMath.com)

1. The frequency table shows the numbers of pounds of aluminum cans collected by classes for a fundraiser. Display the data in a histogram. Describe the shape of the distribution.

## Choosing Appropriate Measures

Use the shape of a distribution to choose the most appropriate measure of center and measure of variation to describe the data set.

### STUDY TIP

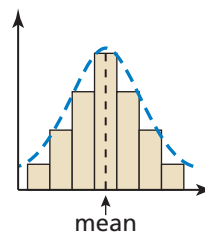
When a distribution is symmetric, the mean and median are about the same. When a distribution is skewed, the mean will be in the direction in which the distribution is skewed while the median will be less affected.

### Core Concept

#### Choosing Appropriate Measures

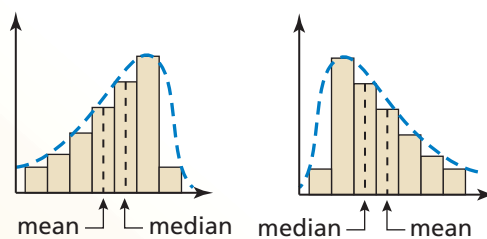
When a data distribution is symmetric,

- use the mean to describe the center and
- use the standard deviation to describe the variation.



When a data distribution is skewed,

- use the median to describe the center and
- use the five-number summary to describe the variation.



### EXAMPLE 2

#### Choosing Appropriate Measures

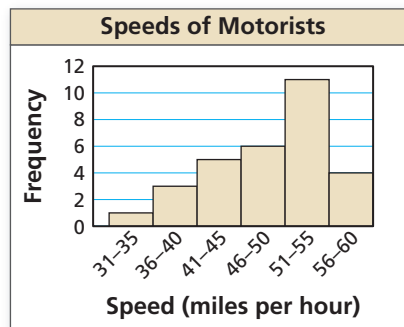
Speeds (mi/h)		
32	44	39
53	38	48
56	41	42
50	50	55
55	45	49
51	53	52
54	60	55
52	50	52
55	40	60
45	58	47

A police officer measures the speeds (in miles per hour) of 30 motorists. The results are shown in the table at the left. (a) Display the data in a histogram using six intervals beginning with 31–35. (b) Which measures of center and variation best represent the data? (c) The speed limit is 45 miles per hour. How would you interpret these results?

#### SOLUTION

- a. Make a frequency table using the described intervals. Then use the frequency table to make a histogram.

Speed (mi/h)	Frequency
31–35	1
36–40	3
41–45	5
46–50	6
51–55	11
56–60	4



- b. Because most of the data are on the right and the tail of the graph extends to the left, the distribution is skewed left. So, use the median to describe the center and the five-number summary to describe the variation.

- c. Using the frequency table and the histogram, you can see that most of the speeds are more than 45 miles per hour. So, most of the motorists were speeding.

Email Attachments Sent				
74	105	98	68	64
85	75	60	48	51
65	55	58	45	38
64	52	65	30	70
72	5	45	77	83
42	25	95	16	120

### Monitoring Progress



Help in English and Spanish at [BigIdeasMath.com](http://BigIdeasMath.com)

2. You record the numbers of email attachments sent by 30 employees of a company in 1 week. Your results are shown in the table. (a) Display the data in a histogram using six intervals beginning with 1–20. (b) Which measures of center and variation best represent the data? Explain.

## Comparing Data Distributions

### EXAMPLE 3 Comparing Data Distributions



Emoticons are graphic symbols that represent facial expressions. They are used to convey a person's mood in a text message. The double histogram shows the distributions of emoticon messages sent by a group of female students and a group of male students during 1 week. Compare the distributions using their shapes and appropriate measures of center and variation.

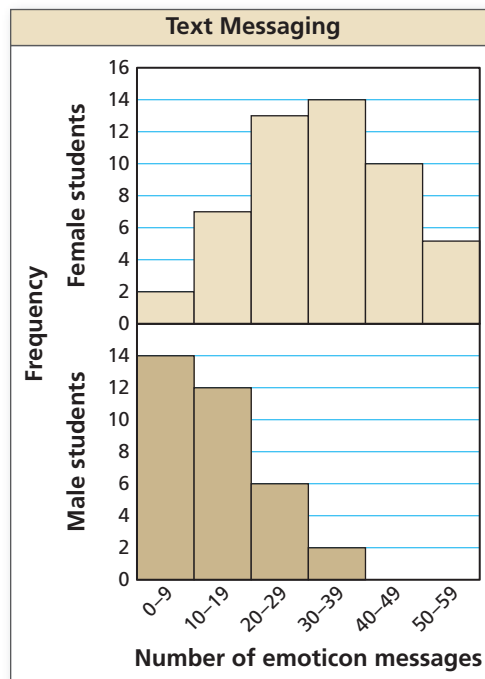
#### SOLUTION

Because the data on the right of the distribution for the female students are approximately a mirror image of the data on the left of the distribution, the distribution is symmetric. So, the mean and standard deviation best represent the distribution for female students.

Because most of the data are on the left of the distribution for the male students and the tail of the graph extends to the right, the distribution is skewed right. So, the median and five-number summary best represent the distribution for male students.

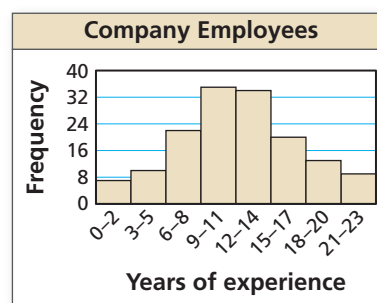
The mean of the female data set is probably in the 30–39 interval, while the median of the male data set is in the 10–19 interval. So, a typical female student is much more likely to use emoticons than a typical male student.

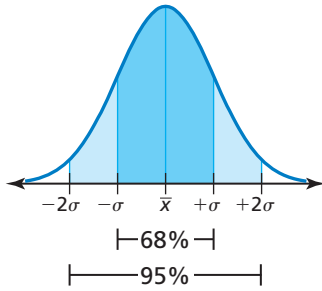
The data for the female students is more variable than the data for the male students. This means that the use of emoticons tends to differ more from one female student to the next.



### Monitoring Progress Help in English and Spanish at [BigIdeasMath.com](http://BigIdeasMath.com)

- Compare the distributions using their shapes and appropriate measures of center and variation.





Many real-life data sets have distributions that are bell-shaped and approximately symmetric about the mean. In a future course, you will study this type of distribution in detail. For now, the following rules can help you see how valuable the standard deviation can be as a measure of variation.

- About 68% of the data lie within 1 standard deviation of the mean.
- About 95% of the data lie within 2 standard deviations of the mean.
- Data values that are more than 2 standard deviations from the mean are considered unusual.

Because the data are symmetric, you can deduce that 34% of the data lie within 1 standard deviation to the left of the mean, and 34% of the data lie within 1 standard deviation to the right of the mean.

### EXAMPLE 4 Comparing Data Distributions

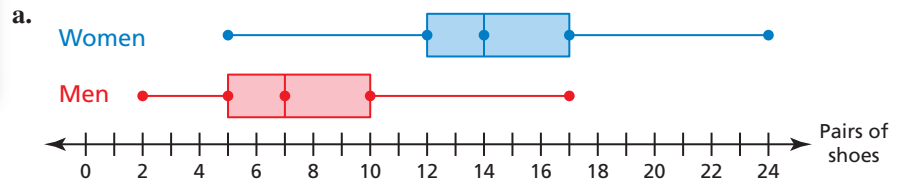


The table shows the results of a survey that asked men and women how many pairs of shoes they own.

- Make a double box-and-whisker plot that represents the data. Describe the shape of each distribution.
- Compare the number of pairs of shoes owned by men to the number of pairs of shoes owned by women.
- About how many of the women surveyed would you expect to own between 10 and 18 pairs of shoes?

	Men	Women
Survey size	35	40
Minimum	2	5
Maximum	17	24
1st Quartile	5	12
Median	7	14
3rd Quartile	10	17
Mean	8	14
Standard deviation	3	4

### SOLUTION



► The distribution for men is skewed right, and the distribution for women is symmetric.

- The centers and spreads of the two data sets are quite different from each other. The mean for women is twice the median for men, and there is more variability in the number of pairs of shoes owned by women.
- Assuming the symmetric distribution is bell-shaped, you know about 68% of the data lie within 1 standard deviation of the mean. Because the mean is 14 and the standard deviation is 4, the interval from 10 to 18 represents about 68% of the data. So, you would expect about  $0.68 \cdot 40 \approx 27$  of the women surveyed to own between 10 and 18 pairs of shoes.

### Monitoring Progress Help in English and Spanish at [BigIdeasMath.com](http://BigIdeasMath.com)

- Why is the mean greater than the median for the men?
- If 50 more women are surveyed, about how many more would you expect to own between 10 and 18 pairs of shoes?

## Vocabulary and Core Concept Check

- VOCABULARY** Describe how data are distributed in a symmetric distribution, a distribution that is skewed left, and a distribution that is skewed right.
- WRITING** How does the shape of a distribution help you decide which measures of center and variation best describe the data?

## Monitoring Progress and Modeling with Mathematics

- DESCRIBING DISTRIBUTIONS** The frequency table shows the numbers of hours that students volunteer per month. Display the data in a histogram. Describe the shape of the distribution. (See Example 1.)

Number of volunteer hours	1–2	3–4	5–6	7–8	9–10	11–12	13–14
Frequency	1	5	12	20	15	7	2

- DESCRIBING DISTRIBUTIONS** The frequency table shows the results of a survey that asked people how many hours they spend online per week. Display the data in a histogram. Describe the shape of the distribution.

Hours online	Frequency
0–3	5
4–7	7
8–11	12
12–15	14
16–19	26
20–23	45
24–27	33



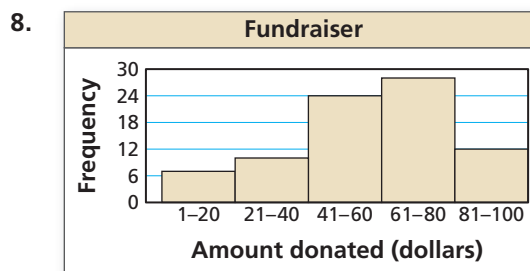
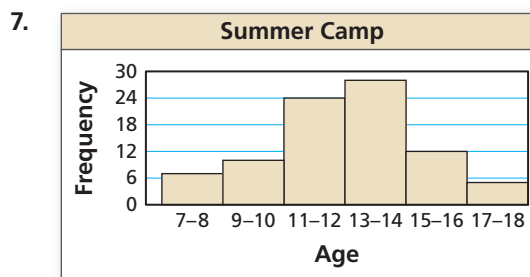
In Exercises 5 and 6, describe the shape of the distribution of the data. Explain your reasoning.

5. Stem	Leaf	6. Stem	Leaf
1	1 1 3 4 8	5	0 0 1
2	2 3 4 7 8	6	3 6 7 9
3	1 2 4 9	7	1 4 5 8 9
4	0 3 2	8	2 4 5 5 7
5	7 9	9	4 6 8 9
6	6	10	1 3 4

Key: 3|1 = 31

Key: 6|3 = 63

In Exercises 7 and 8, determine which measures of center and variation best represent the data. Explain your reasoning.



- ANALYZING DATA** The table shows the last 24 ATM withdrawals at a bank. (See Example 2.)

- Display the data in a histogram using seven intervals beginning with 26–50.
- Which measures of center and variation best represent the data? Explain.
- The bank charges a fee for any ATM withdrawal less than \$150. How would you interpret the data?

ATM Withdrawals (dollars)		
120	100	70
60	40	80
150	80	50
120	60	175
30	50	50
60	200	30
100	150	110
70	40	100



10. **ANALYZING DATA** Measuring an IQ is an inexact science. However, IQ scores have been around for years in an attempt to measure human intelligence. The table shows some of the greatest known IQ scores.

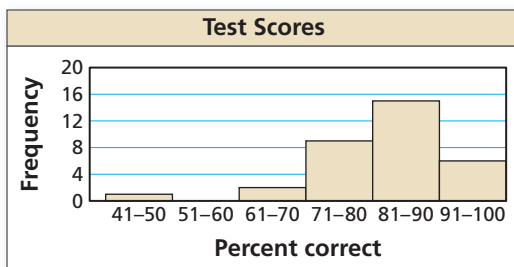
a. Display the data in a histogram using five intervals beginning with 151–166.

b. Which measures of center and variation best represent the data? Explain.

c. The distribution of IQ scores for the human population is symmetric. What happens to the shape of the distribution in part (a) as you include more and more IQ scores from the human population in the data set?

IQ Scores		
170	190	180
160	180	210
154	170	180
195	230	160
170	186	180
225	190	170

**ERROR ANALYSIS** In Exercises 11 and 12, describe and correct the error in the statements about the data displayed in the histogram.

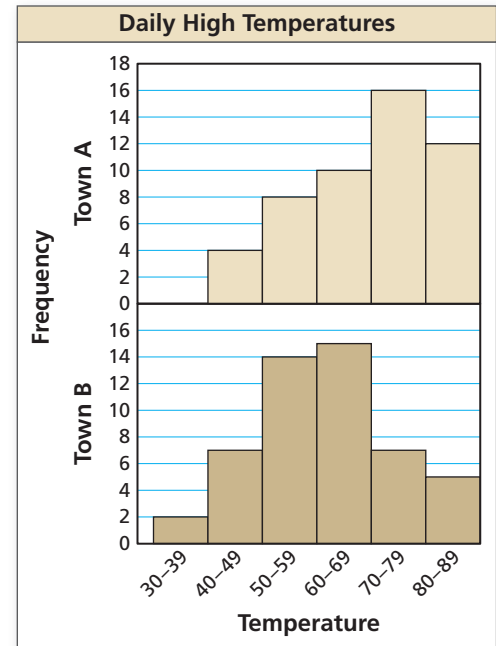


11. **X** Most of the data are on the right. So, the distribution is skewed right.

12. **X** Because the distribution is skewed, use the standard deviation to describe the variation of the data.

13. **USING TOOLS** For a large data set, would you use a stem-and-leaf plot or a histogram to show the distribution of the data? Explain.
14. **REASONING** For a symmetric distribution, why is the mean used to describe the center and the standard deviation used to describe the variation? For a skewed distribution, why is the median used to describe the center and the five-number summary used to describe the variation?

15. **COMPARING DATA SETS** The double histogram shows the distributions of daily high temperatures for two towns over a 50-day period. Compare the distributions using their shapes and appropriate measures of center and variation. (See Example 3.)



16. **COMPARING DATA SETS** The frequency tables show the numbers of entrées in certain price ranges (in dollars) at two different restaurants. Display the data in a double histogram. Compare the distributions using their shapes and appropriate measures of center and variation.

Restaurant A		Restaurant B	
Price range	Frequency	Price range	Frequency
8–10	5	8–10	0
11–13	9	11–13	2
14–16	12	14–16	5
17–19	4	17–19	7
20–22	3	20–22	8
23–25	0	23–25	6

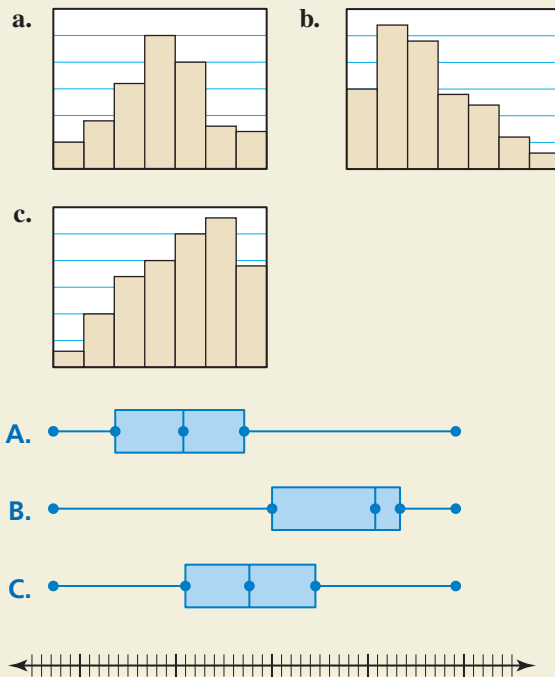
17. **OPEN-ENDED** Describe a real-life data set that has a distribution that is skewed right.
18. **OPEN-ENDED** Describe a real-life data set that has a distribution that is skewed left.

19. **COMPARING DATA SETS** The table shows the results of a survey that asked freshmen and sophomores how many songs they have downloaded on their MP3 players. (See Example 4.)

	Freshmen	Sophomores
Survey size	45	54
Minimum	250	360
Maximum	2150	2400
1st Quartile	800	780
Median	1200	2000
3rd Quartile	1600	2200
Mean	1150	1650
Standard deviation	420	480

- a. Make a double box-and-whisker plot that represents the data. Describe the shape of each distribution.
- b. Compare the number of songs downloaded by freshmen to the number of songs downloaded by sophomores.
- c. About how many of the freshmen surveyed would you expect to have between 730 and 1570 songs downloaded on their MP3 players?
- d. If you survey 100 more freshmen, about how many would you expect to have downloaded between 310 and 1990 songs on their MP3 players?
20. **COMPARING DATA SETS** You conduct the same survey as in Exercise 19 but use a different group of freshmen. The results are as follows.  
Survey size: 60; minimum: 200; maximum: 2400; 1st quartile: 640; median: 1670; 3rd quartile: 2150; mean: 1480; standard deviation: 500
- a. Compare the number of songs downloaded by this group of freshmen to the number of songs downloaded by sophomores.
- b. Why is the median greater than the mean for this group of freshmen?
21. **REASONING** A data set has a symmetric distribution. Every value in the data set is doubled. Describe the shape of the new distribution. Are the measures of center and variation affected? Explain.

22. **HOW DO YOU SEE IT?** Match the distribution with the corresponding box-and-whisker plot.

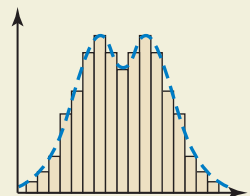


23. **REASONING** You record the following waiting times at a restaurant.

Waiting Times (minutes)									
26	38	15	8	22	42	25	20	17	18
40	35	24	31	42	29	25	0	30	13

- a. Display the data in a histogram using five intervals beginning with 0–9. Describe the shape of the distribution.
- b. Display the data in a histogram using 10 intervals beginning with 0–4. What happens when the number of intervals is increased?
- c. Which histogram best represents the data? Explain your reasoning.

24. **THOUGHT PROVOKING** The shape of a *bimodal* distribution is shown. Describe a real-life example of a bimodal distribution.



## Maintaining Mathematical Proficiency

Reviewing what you learned in previous grades and lessons

Find the domain of the function. (Section 10.1)

25.  $f(x) = \sqrt{x + 6}$

26.  $f(x) = \sqrt{2x}$

27.  $f(x) = \frac{1}{4}\sqrt{x - 7}$