

Main Ideas

- Use measures of central tendency to represent a set of data.
- Find measures of variation for a set of data.

New Vocabulary

univariate data
 measure of central tendency
 measure of variation
 dispersion
 variance
 standard deviation

Study Tip**Look Back**

To review **outliers**, see Lesson 2-5.

GET READY for the Lesson

On Mr. Dent's most recent Algebra 2 test, his students earned the following scores.

72	70	77	76	90	68	81	86	34	94
71	84	89	67	19	85	75	66	80	94

When his students ask how they did on the test, which measure of central tendency should Mr. Dent use to describe the scores?

Measures of Central Tendency Data with one variable, such as the test scores, are called **univariate data**. Sometimes it is convenient to have one number that describes a set of data. This number is called a **measure of central tendency**, because it represents the center or middle of the data. The most commonly used measures of central tendency are the *mean*, *median*, and *mode*.

When deciding which measure of central tendency to use to represent a set of data, look closely at the data itself.

CONCEPT SUMMARY*Measures of Tendency*

Use	When ...
mean	the data are spread out, and you want an average of the values
median	the data contain outliers
mode	the data are tightly clustered around one or two values

EXAMPLE Choose a Measure of Central Tendency

- 1 SWEEPSTAKES** A sweepstakes offers a first prize of \$10,000, two second prizes of \$100, and one hundred third prizes of \$10. Which measure of central tendency best represents the available prizes?

Since 100 of the 103 prizes are \$10, the mode (\$10) best represents the available prizes. Notice that in this case the median is the same as the mode.

CHECK Your Progress

1. Which measure of central tendency would the organizers of the sweepstakes be most likely to use in their advertising?

Reading Math

Symbols The symbol σ is the lower case Greek letter *sigma*. \bar{x} is read *x bar*.

Measures of Variation Measures of variation or dispersion measure how spread out or scattered a set of data is. The simplest measure of variation to calculate is the *range*, the difference between the greatest and the least values in a set of data. Variance and standard deviation are measures of variation that indicate how much the data values differ from the mean.

To find the **variance** σ^2 of a set of data, follow these steps.

1. Find the mean, \bar{x} .
2. Find the difference between each value in the set of data and the mean.
3. Square each difference.
4. Find the mean of the squares.

The **standard deviation** σ is the square root of the variance.

KEY CONCEPT

Standard Deviation

If a set of data consists of the n values x_1, x_2, \dots, x_n and has mean \bar{x} , then the standard deviation σ is given by the following formula.

$$\sigma = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}}$$

EXAMPLE Standard Deviation

- 2 STATES** The table shows the populations in millions of 11 eastern states as of the 2000 Census. Find the variance and standard deviation of the data to the nearest tenth.

State	Population	State	Population	State	Population
NY	19.0	MD	5.3	RI	1.0
PA	12.3	CT	3.4	DE	0.8
NJ	8.4	ME	1.3	VT	0.6
MA	6.3	NH	1.2	—	—

Source: U.S. Census Bureau

- Step 1** Find the mean. Add the data and divide by the number of items.

$$\bar{x} = \frac{19.0 + 12.3 + 8.4 + 6.3 + 5.3 + 3.4 + 1.3 + 1.2 + 1.0 + 0.8 + 0.6}{11}$$

$$\approx 5.418 \quad \text{The mean is about 5.4 million people.}$$

- Step 2** Find the variance.

$$\sigma^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n} \quad \text{Variance formula}$$

$$= \frac{(19.0 - 5.4)^2 + (12.3 - 5.4)^2 + \dots + (8.0 - 5.4)^2 + (0.6 - 5.4)^2}{11}$$

$$= \frac{344.4}{11} \quad \text{Simplify.}$$

$$\approx 31.309 \quad \text{The variance is about 31.3.}$$

Step 3 Find the standard deviation.

$$\sigma^2 \approx 31.3$$

Take the square root of each side.

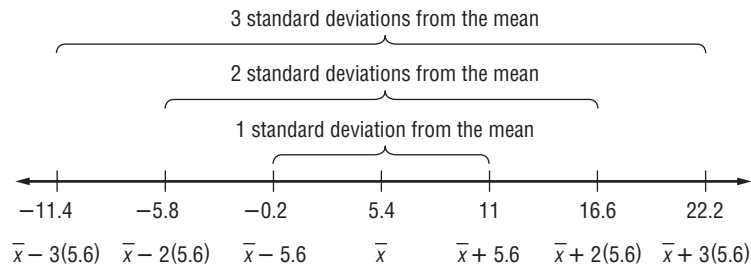
$$\sigma \approx 5.594640292 \quad \text{The standard deviation is about 5.6 million people.}$$

CHECK Your Progress

2. The leading number of home runs in Major League Baseball for the 1994–2004 seasons were 43, 50, 52, 56, 70, 65, 50, 73, 57, 47, and 48. Find the variance and standard deviation of the data to the nearest tenth.

 **Personal Tutor at** algebra2.com

Most of the members of a set of data are within 1 standard deviation of the mean. The data in Example 2 can be broken down as shown below.



Looking at the original data, you can see that most of the states' populations were between 2.4 million and 20.2 million. That is, the majority of members of the data set were within 1 standard deviation of the mean.

You can use a TI-83/84 Plus graphing calculator to find statistics for the data in Example 2.

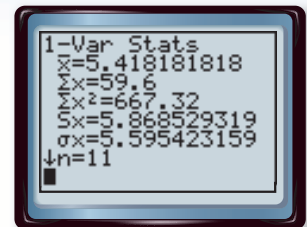
GRAPHING CALCULATOR LAB

One-Variable Statistics

The TI-83/84 Plus can compute a set of one-variable statistics from a list of data. These statistics include the mean, variance, and standard deviation. Enter the data into L1.

KEYSTROKES: **STAT** **ENTER** 19.0 **ENTER** 12.3 **ENTER** ...

Then use **STAT** **▶** 1 **ENTER** to show the statistics. The mean \bar{x} is about 5.4, the sum of the values $\sum x$ is 59.6, the standard deviation σx is about 5.6, and there are $n = 11$ data items. If you scroll down, you will see the least value ($\min X = .6$), the three quartiles (1, 3.4, and 8.4), and the greatest value ($\max X = 19$).



THINK AND DISCUSS

1. Find the variance of the data set.
2. Enter the data set in list L1 but without the outlier 19.0. What are the new mean, median, and standard deviation?
3. Did the mean or median change less when the outlier was deleted?

Example 1
(pp. 717–718)

EDUCATION For Exercises 1 and 2, use the following information.

The table below shows the amounts of money spent on education per student in a recent year in two regions of the United States.

Pacific States		Southwest Central States	
State	Expenditures per Student (\$)	State	Expenditures per Student (\$)
Alaska	9564	Texas	6771
California	7405	Arkansas	6276
Washington	7039	Louisiana	6567
Oregon	7642	Oklahoma	6229

Source: *The World Almanac*

- Find the mean for each region.
- For which region is the mean more representative of the data? Explain.

Example 2
(pp. 718–719)

Find the variance and standard deviation of each set of data to the nearest tenth.

- {48, 36, 40, 29, 45, 51, 38, 47, 39, 37}
- {321, 322, 323, 324, 325, 326, 327, 328, 329, 330}
- {43, 56, 78, 81, 47, 42, 34, 22, 78, 98, 38, 46, 54, 67, 58, 92, 55}

Exercises

HOMEWORK HELP	
For Exercises	See Examples
6–13, 24–30	2
14–23	1

Find the variance and standard deviation of each set of data to the nearest tenth.

- {400, 300, 325, 275, 425, 375, 350}
- {5, 4, 5, 5, 5, 5, 6, 6, 6, 6, 7, 7, 7, 7, 8, 9}
- {2.4, 5.6, 1.9, 7.1, 4.3, 2.7, 4.6, 1.8, 2.4}
- {4.3, 6.4, 2.9, 3.1, 8.7, 2.8, 3.6, 1.9, 7.2}
- {234, 345, 123, 368, 279, 876, 456, 235, 333, 444}
- {13, 14, 15, 16, 17, 18, 19, 20, 21, 23, 67, 56, 34, 99, 44, 55}

12.	Stem	Leaf	13.	Stem	Leaf
	4	4 5 6 7 7		5	7 7 7 8 9
	5	3 5 6 7 8 9		6	3 4 5 5 6 7
	6	7 7 8 9 9 9 4 5 = 45		7	2 3 4 5 6 6 3 = 63

BASKETBALL For Exercises 14 and 15, use the following information.

The table below shows the rebounding totals for the members of the 2005 Charlotte Sting.

162	145	179	37	44	53	70	65	47	35	71	5	5
-----	-----	-----	----	----	----	----	----	----	----	----	---	---

Source: WNBA

- Find the mean, median, and mode of the data to the nearest tenth.
- Which measure of central tendency best represents the data? Explain your answer.

ADVERTISING For Exercises 16–18, use the following information.

An electronics store placed an ad in the newspaper showing five flat-screen TVs for sale. The ad says, “Our flat-screen TVs average \$695.” The prices of the flat-screen TVs are \$1200, \$999, \$1499, \$895, \$695, \$1100, \$1300, and \$695.

16. Find the mean, median, and mode of the prices.
17. Which measure is the store using in its ad? Why did they choose it?
18. As a consumer, which measure would you want to see advertised? Explain.

EDUCATION For Exercises 19 and 20, use the following information.

The Millersburg school board is negotiating a pay raise with the teacher’s union. Three of the administrators have salaries of \$90,000 each. However, a majority of the teachers have salaries of about \$45,000 per year.

19. You are a member of the school board and would like to show that the current salaries are reasonable. Would you quote the mean, median, or mode as the “average” salary to justify your claim? Explain.
20. You are the head of the teacher’s union and maintain that a pay raise is in order. Which of the mean, median, or mode would you quote to justify your claim? Explain your reasoning.

**Real-World Link**

While the Mall of America does not have the most gross leasable area, it is the largest fully enclosed retail and entertainment complex in the United States.

Source: Mall of America

SHOPPING MALLS For Exercises 21–23, use the following information.

The table lists the areas of some large shopping malls in the United States.

Mall	Gross Leasable Area (ft ²)
1 Del Amo Fashion Center, Torrance, CA	3,000,000
2 South Coast Plaza/Crystal Court, Costa Mesa, CA	2,918,236
3 Mall of America, Bloomington, MN	2,472,500
4 Lakewood Center Mall, Lakewood, CA	2,390,000
5 Roosevelt Field Mall, Garden City, NY	2,300,000
6 Gurnee Mills, Gurnee, IL	2,200,000
7 The Galleria, Houston, TX	2,100,000
8 Randall Park Mall, North Randall, OH	2,097,416
9 Oakbrook Shopping Center, Oak Brook, IL	2,006,688
10 Sawgrass Mills, Sunrise, FL	2,000,000
10 The Woodlands Mall, The Woodlands, TX	2,000,000
10 Woodfield, Schaumburg, IL	2,000,000

Source: Blackburn Marketing Service

21. Find the mean, median, and mode of the gross leasable areas.
22. You are a realtor who is trying to lease mall space in different areas of the country to a large retailer. Which measure would you talk about if the customer felt that the malls were too large for his store? Explain.
23. Which measure would you talk about if the customer had a large inventory? Explain.

SCHOOL For Exercises 24–26, use the frequency table at the right that shows the scores on a multiple-choice test.

24. Find the variance and standard deviation of the scores.
25. What percent of the scores are within one standard deviation of the mean?
26. What percent of the scores are within two standard deviations of the mean?

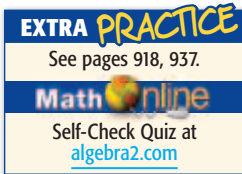
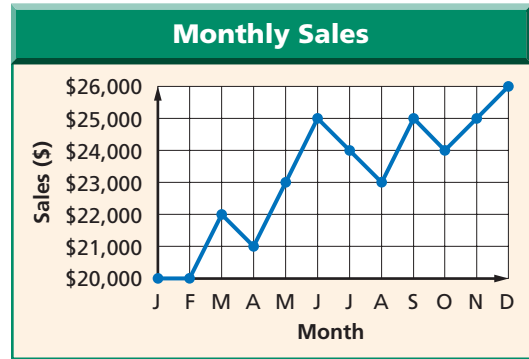
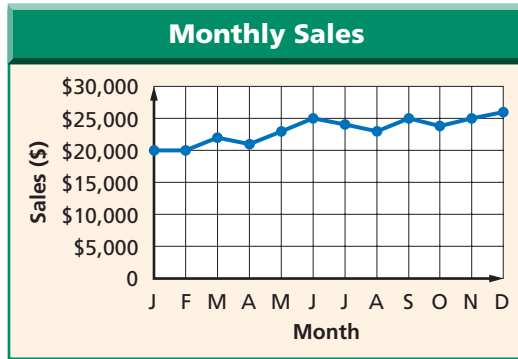
Score	Frequency
90	3
85	2
80	3
75	7
70	6
65	4

FOOTBALL For Exercises 27–30, use the weights in pounds of the starting offensive linemen of the football teams from three high schools.

Jackson	Washington	King
170, 165, 140, 188, 195	144, 177, 215, 225, 197	166, 175, 196, 206, 219

27. Find the standard deviation of the weights for Jackson High.
28. Find the standard deviation of the weights for Washington High.
29. Find the standard deviation of the weights for King High.
30. Which team had the most variation in weights? How do you think this variation will impact their play?

For Exercises 31–33, consider the two graphs below.



31. Explain why the graphs made from the same data look different.
32. Describe a situation where the first graph might be used.
33. Describe a situation where the second graph might be used.

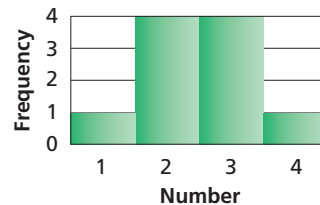
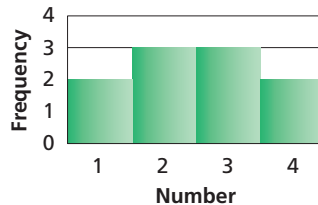
H.O.T. Problems

34. **OPEN ENDED** Give a sample set of data with a variance and standard deviation of 0.
35. **REASONING** Find a counterexample for the following statement.
The standard deviation of a set of data is always less than the variance.

CHALLENGE For Exercises 36 and 37, consider the two sets of data.

$$A = \{1, 2, 2, 2, 2, 3, 3, 3, 3, 4\}, B = \{1, 1, 2, 2, 2, 3, 3, 3, 4, 4\}$$

36. Find the mean, median, variance, and standard deviation of each set of data.
37. Explain how you can tell which histogram below goes with each data set without counting the frequencies in the sets.



38. **Which One Doesn't Belong?** Identify the term that does not belong with the other three. Explain your reasoning.

mode

variance

mean

median

39. **Writing in Math** Use the information on page 717 to explain what statistics a teacher should tell the class after a test. Include the mean, median, and mode of the given data set and which measure of central tendency you think best represents the test scores and why. How will the measures of central tendency be affected if Mr. Dent adds 5 points to each score?

STANDARDIZED TEST PRACTICE

40. **ACT/SAT** What is the mean of the numbers represented by $x + 1$, $3x - 2$, and $2x - 5$?
- A $2x - 2$
B $\frac{6x - 7}{3}$
C $\frac{x + 1}{3}$
D $x + 4$
41. **REVIEW** A school has two backup generators having probabilities of 0.9 and 0.95, respectively, of operating in case of power outage. Find the probability that at least one backup generator operates during a power outage.
- F 0.855
G 0.89
H 0.95
J 0.995

Spiral Review

Determine whether the events are *mutually exclusive* or *inclusive*. Then find the probability. (Lesson 12-5)

42. A card is drawn from a standard deck of cards. What is the probability that it is a 5 or a spade?
43. A jar of change contains 5 quarters, 8 dimes, 10 nickels, and 19 pennies. If a coin is pulled from the jar at random, what is the probability that it is a nickel or a dime?

Two cards are drawn from a standard deck of cards. Find each probability. (Lesson 12-4)

44. $P(\text{ace, then king})$ if replacement occurs
45. $P(\text{ace, then king})$ if no replacement occurs
46. $P(\text{heart, then club})$ if no replacement occurs
47. $P(\text{heart, then club})$ if replacement occurs
48. **BUSINESS** The Energy Booster Company keeps their stock of Health Aid liquid in a tank that is a rectangular prism. Its sides measure $x - 1$ centimeters, $x + 3$ centimeters, and $x - 2$ centimeters. Suppose they would like to bottle their Health Aid in $x - 3$ containers of the same size. How much liquid in cubic centimeters will remain unbottled? (Lesson 6-6)

GET READY for the Next Lesson

PREREQUISITE SKILL Find each percent.

- | | | |
|----------------|----------------|----------------|
| 49. 68% of 200 | 50. 68% of 500 | 51. 95% of 400 |
| 52. 95% of 500 | 53. 99% of 400 | 54. 99% of 500 |