EXPLORING DATA AND STATISTICS

7.7

Statistics and Statistical Graphs

GOAL 1 MEASURES OF CENTRAL TENDENCY AND DISPERSION

In this lesson you will use the following two data sets. They show the free-throw percentages for the players in the Women's National Basketball Association (WNBA) for 1998. INTERNET **DATA UPDATE** of WNBA data at www.mcdougallittell.com

Women's National Basketball Association Free-Throw Percentages	
Eastern Conference	**Western Conference**
46, 47, 48, 50, 50, 50, 52, 53, 55, 57, 57, 58, 60, 61, 61, 62, 63, 63, 63, 63, 63, 63, 64, 64, 67, 67, 67, 69, 71, 72, 72, 72, 72, 73, 75, 75, 75, 75, 75, 76, 77, 78, 79, 79, 80, 81, 81, 82, 82, 83, 83, 85, 89, 91, 92, 100	36, 50, 50, 56, 56, 57, 57, 58, 61, 61, 62, 62, 63, 63, 64, 64, 65, 66, 66, 66, 66, 67, 69, 69, 70, 70, 70, 70, 71, 71, 71, 71, 72, 72, 73, 73, 74, 74, 74, 74, 75, 76, 76, 76, 77, 77, 78, 80, 81, 83, 83, 83, 85, 85, 87, 100, 100, 100

Statistics are numerical values used to summarize and compare sets of data. The following **measures of central tendency** are three commonly used statistics.

1. The **mean**, or *average*, of n numbers is the sum of the numbers divided by n. The mean is denoted by $\bar{x}$, which is read as "x-bar." For the data $x_1, x_2, \ldots, x_n$, the mean is $\bar{x} = \dfrac{x_1 + x_2 + \cdots + x_n}{n}$.
2. The **median** of n numbers is the middle number when the numbers are written in order. (If n is even, the median is the mean of the two middle numbers.)
3. The **mode** of n numbers is the number or numbers that occur most frequently. There may be one mode, no mode, or more than one mode.

EXAMPLE 1 *Finding Measures of Central Tendency*

Find the mean, median, and mode of the two data sets listed above.

SOLUTION

EASTERN CONFERENCE: Mean: $\bar{x} = \dfrac{46 + 47 + \cdots + 100}{57} = \dfrac{3931}{57} \approx 69.0$

Median: 69 Mode: 63

WESTERN CONFERENCE: Mean: $\bar{x} = \dfrac{36 + 50 + \cdots + 100}{58} = \dfrac{4106}{58} \approx 70.8$

Median: 71 Modes: 66, 70, 71, 74

All three measures of central tendency for the Western Conference are greater than those for the Eastern Conference. So, the Western Conference has better free-throw percentages overall.

What you should learn

GOAL 1 Use measures of central tendency and measures of dispersion to describe data sets.

GOAL 2 Use box-and-whisker plots and histograms to represent data graphically, as applied in **Exs. 40–42**.

Why you should learn it

▼ To use statistics and statistical graphs to analyze **real-life** data sets, such as the free-throw percentages for the players in the WNBA in **Examples 1–6**.

REAL LIFE

Measures of central tendency tell you what the *center* of the data is. Other commonly used statistics are called **measures of dispersion**. They tell you how *spread out* the data are. One simple measure of dispersion is the **range**, which is the difference between the greatest and least data values.

EXAMPLE 2 Finding Ranges of Data Sets

The ranges of the free-throw percentages in the two data sets on the previous page are:

EASTERN CONFERENCE: Range $= 100 - 46 = 54$

WESTERN CONFERENCE: Range $= 100 - 36 = 64$

Because the Western Conference's range of free-throw percentages is greater, its free-throw percentages are more spread out.

. . . . . .

Another measure of dispersion is **standard deviation**, which describes the typical difference (or *deviation*) between the mean and a data value.

STANDARD DEVIATION OF A SET OF DATA

The standard deviation σ (read as "sigma") of $x_1, x_2, \ldots, x_n$ is:

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

EXAMPLE 3 Finding Standard Deviations of Data Sets

STUDENT HELP

INTERNET HOMEWORK HELP

Visit our Web site www.mcdougallittell.com for extra examples.

The standard deviations of the free-throw percentages in the two data sets on the previous page are:

EASTERN CONFERENCE:

$$\sigma \approx \sqrt{\frac{(46 - 69.0)^2 + (47 - 69.0)^2 + \cdots + (100 - 69.0)^2}{57}}$$

$$\approx \sqrt{\frac{8660}{57}}$$

$$\approx \sqrt{152}$$

$$\approx 12.3$$

WESTERN CONFERENCE:

$$\sigma \approx \sqrt{\frac{(36 - 70.8)^2 + (50 - 70.8)^2 + \cdots + (100 - 70.8)^2}{58}}$$

$$\approx \sqrt{\frac{7910}{58}}$$

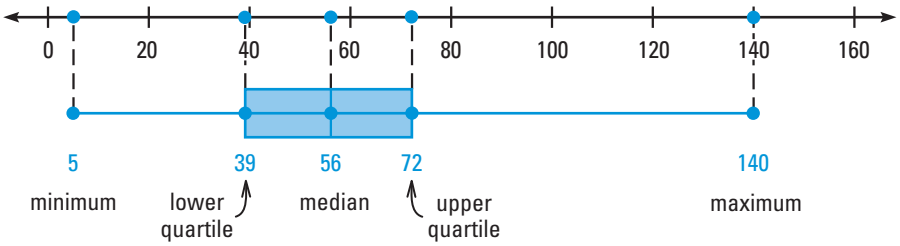
$$\approx \sqrt{136}$$

$$\approx 11.7$$

Because the Eastern Conference's standard deviation is greater, its free-throw percentages are more spread out *about the mean*.

GOAL 2 USING STATISTICAL GRAPHS

Although statistics are useful in describing a data set, sometimes a graph of the data can be more informative. One type of statistical graph is a **box-and-whisker plot**. The "box" encloses the middle half of the data set and the "whiskers" extend to the minimum and maximum data values.



The median divides the data set into two halves. The **lower quartile** is the median of the lower half, and the **upper quartile** is the median of the upper half. You can use the following steps to draw a box-and-whisker plot.

1. Order the data from least to greatest.
2. Find the minimum and maximum values.
3. Find the median.
4. Find the lower and upper quartiles.
5. Plot these five numbers below a number line.
6. Draw the box, the whiskers, and a line segment through the median.

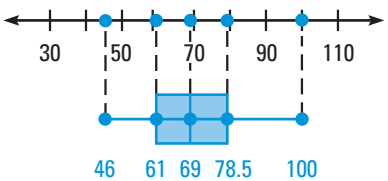
EXAMPLE 4 *Drawing Box-and-Whisker Plots*

Draw a box-and-whisker plot of each data set on page 445.

SOLUTION

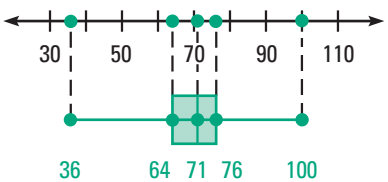
EASTERN CONFERENCE

The minimum is 46 and the maximum is 100. The median is 69. The lower quartile is 61 and the upper quartile is 78.5.



WESTERN CONFERENCE

The minimum is 36 and the maximum is 100. The median is 71. The lower quartile is 64 and the upper quartile is 76.



FOCUS ON PEOPLE

REAL LIFE **SANDY BRONDELLO** was ranked first in the WNBA in 1998 for free-throw percentage (among players who attempted at least 10 free throws). As a player for the Detroit Shock, she made 96 out of 104 free throws for a free-throw percentage of 92.

Like the computations in Examples 2 and 3, the box-and-whisker plots show you that the Western Conference's free-throw percentages are more spread out overall (comparing the entire plots) and that the Eastern Conference's free-throw percentages are more spread out about the mean (comparing the boxes).

Another way to display numerical data is with a special type of bar graph called a **histogram**. In a histogram data are grouped into intervals of *equal* width. The number of data values in each interval is the **frequency** of the interval. To draw a histogram, begin by making a **frequency distribution**, which shows the frequency of each interval.

EXAMPLE 5 *Making Frequency Distributions*

Make a frequency distribution of each data set on page 445. Use seven intervals beginning with the interval 31–40.

SOLUTION Begin by writing the seven intervals. Then tally the data values by interval. Finally, count the tally marks to get the frequencies.

Eastern Conference			Western Conference		
Interval	Tally	Frequency	Interval	Tally	Frequency
31–40		0	31–40	|	1
41–50	𝍸 |	6	41–50	||	2
51–60	𝍸 ||	7	51–60	𝍸	5
61–70	𝍸 𝍸 𝍸 |	16	61–70	𝍸 𝍸 𝍸 𝍸	20
71–80	𝍸 𝍸 𝍸 ||	17	71–80	𝍸 𝍸 𝍸 𝍸	20
81–90	𝍸 |||	8	81–90	𝍸 ||	7
91–100	|||	3	91–100	|||	3

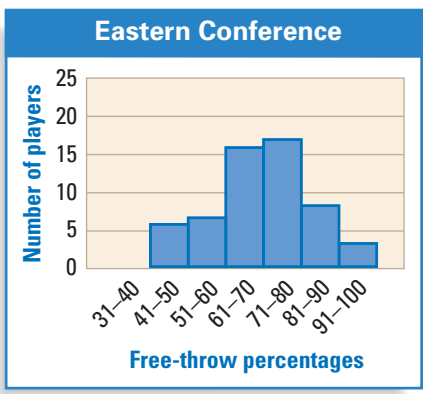
EXAMPLE 6 *Drawing Histograms*

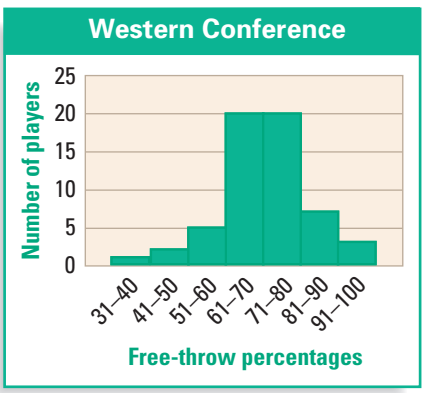
Draw a histogram of each data set on page 445.

SOLUTION Use the frequency distributions in Example 5. Draw a horizontal axis, divide it into seven equal sections, and label the sections with the intervals. Then draw a vertical axis for measuring the frequencies. Finally, draw bars of appropriate heights to represent the frequencies of the intervals.

STUDENT HELP

Skills Review
For help with statistical graphs, see p. 934.



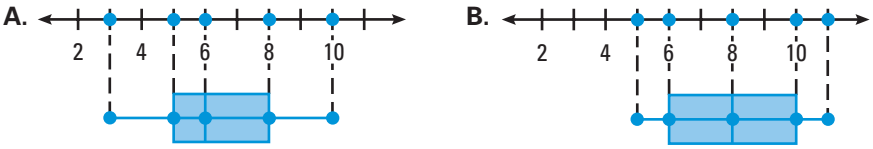


GUIDED PRACTICE

Vocabulary Check ✓

Concept Check ✓

1. Define mean, median, mode, and range of a set of n numbers.
2. Give an example of two sets of four numbers, each with a mean of 5. Choose the numbers so that one set has a range of 3 and the other set has a range of 7. Which has the greater standard deviation?
3. The following box-and-whisker plots represent two sets of data. Which data set has the greater range? Explain.



Skill Check ✓

HISTORY SCORES **In Exercises 4–9, use the following data set of scores received by students on a history exam.**

68, 72, 76, 81, 84, 86, 86, 86, 89, 91, 95, 99

4. Find the mean, median, and mode of the data.
5. Find the range of the data.
6. Find the standard deviation of the data.
7. Draw a box-and-whisker plot of the data.
8. Make a frequency distribution of the data. Use five intervals beginning with 61–68.
9. Draw a histogram of the data.

PRACTICE AND APPLICATIONS

STUDENT HELP

Extra Practice to help you master skills is on p. 950.

STUDENT HELP

HOMEWORK HELP

Example 1: Exs. 10–15, 33, 36, 39
Examples 2, 3: Exs. 16–21, 34, 37
Example 4: Exs. 22–27, 40, 43
Examples 5, 6: Exs. 28–32, 41, 44

MEASURES OF CENTRAL TENDENCY **Find the mean, median, and mode of the data set.**

10. 6, 7, 9, 9, 9, 10
11. 43, 46, 47, 47, 51, 54, 59
12. 88, 83, 91, 82, 78, 81, 91, 91
13. 220, 250, 210, 290, 310, 230, 230
14. 2.9, 2.1, 2.6, 2.9, 3.0, 2.5, 3.4
15. 0, 0, 0.1, 0.2, 0.3, 0.5, 0.5, 0.6, 1.0

MEASURES OF DISPERSION **Find the range and standard deviation of the data set.**

16. 10, 20, 30, 40, 50, 60
17. 1, 1, 3, 4, 5, 9
18. 10, 12, 7, 11, 20, 7, 6, 8, 9
19. 1202, 1229, 1012, 1014, 1120, 1429
20. 3.1, 2.7, 6.0, 5.6, 2.3, 2.0, 1.3
21. 19.4, 16.3, 12.7, 24.8, 19.2, 15.4

BOX-AND-WHISKER PLOTS **Draw a box-and-whisker plot of the data set.**

22. 1, 2, 2, 4, 5, 7
23. 19, 19, 19, 89, 93, 95
24. 47, 88, 89, 61, 70, 71, 79
25. 40, 100, 20, 40, 100, 70, 90
26. 1.7, 8.5, 1.2, 3.8, 8.5, 5.2, 6.9
27. 61.2, 23.0, 72.7, 74.3, 19.1, 6.6, 28.4

FOCUS ON CAREERS

CHEMICAL ENGINEER

Chemical engineers work in a variety of fields such as chemical manufacturing, electronics, and photographic equipment. Some test methods of production, as discussed in Exs. 33–35.

CAREER LINK
www.mcdougallittell.com

HISTOGRAMS **Use the given intervals to make a frequency distribution of the data set. Then draw a histogram of the data set.**

28. Use five intervals beginning with 1–2.
1, 1, 1, 2, 2, 2, 3, 3, 4, 4, 4, 4, 5, 6, 6, 7, 7, 7, 7, 8, 8, 8, 8, 9, 9

29. Use five intervals beginning with 10–19.
10, 12, 14, 15, 15, 23, 26, 26, 27, 28, 37, 37, 37, 37, 38, 39, 39, 40, 48, 58

30. Use ten intervals beginning with 0–9.
66, 9, 43, 28, 5, 7, 90, 9, 78, 6, 69, 43, 28, 55, 13, 64, 45, 10, 54, 96

31. Use fifteen intervals beginning with 0.0–0.4.
2.2, 5.6, 2.2, 4.4, 2.2, 6.6, 2.4, 5.8, 2.4, 4.8, 2.4, 6.1, 5.4, 5.4

32. Use eight intervals beginning with 15.5–20.4.
22.2, 22.2, 22.6, 24.3, 24.3, 24.8, 54.5, 54.4, 54.1, 52.3, 52.3, 52.9

SCRATCH PROTECTION **In Exercises 33–35, use the following information.**
Computer chips have a silicon dioxide coating that gives them scratch protection. Phosphorus is contained in this coating and is a key to its effectiveness. A company that produces silicon dioxide has two machines that vary from run to run on the phosphorus content. The phosphorus contents for four runs on each machine are shown below.

Machine 1	
Run number	Percent of coating that is phosphorous (by weight)
1	2.63
2	2.72
3	2.47
4	2.55

Machine 2	
Run number	Percent of coating that is phosphorous (by weight)
1	1.09
2	3.06
3	4.10
4	2.12

33. Find the mean, median, and mode of each data set.

34. Find the range and standard deviation of each data set.

35. **CRITICAL THINKING** Based on the data, which machine is more consistent? Explain your reasoning.

REAL ESTATE **In Exercises 36–38, suppose a real estate agent has sold seven homes priced at $104,900, $119,900, $134,900, $142,000, $179,900, $199,900, and $750,000.**

36. Find the mean, median, and mode of the selling prices.

37. Find the range and standard deviation of the selling prices.

38. **CRITICAL THINKING** How does the selling price $750,000 affect the mean? How does this explain why the most commonly used measure of central tendency for housing prices is the median rather than the mean or the mode?

39. **POLITICAL SURVEY** A random survey of 1000 people asked their opinions on a controversial issue. The opinions are categorized as 0 = a negative opinion, 1 = a positive opinion, and 2 = a neutral opinion. The results of the survey are 212 category 0 responses, 627 category 1 responses, and 161 category 2 responses. Which measure of central tendency is most appropriate to represent the data? Explain.

HISTORY CONNECTION **In Exercises 40–42, use the tables below which give the ages of the Presidents and Vice Presidents of the United States.**

Ages of the first 42 Presidents of the United States when they first took office
42, 43, 46, 46, 47, 48, 49, 49, 50, 51, 51, 51, 51, 51, 52, 52, 54, 54, 54, 54, 55, 55, 55, 55, 56, 56, 56, 57, 57, 57, 57, 58, 60, 61, 61, 61, 62, 64, 64, 65, 68, 69

Ages of the first 47 Vice Presidents of the United States when they first took office
36, 40, 41, 42, 42, 42, 44, 45, 45, 46, 48, 49, 49, 50, 50, 51, 51, 51, 52, 52, 52, 52, 52, 53, 53, 53, 53, 56, 56, 56, 57, 57, 58, 59, 60, 60, 61, 64, 64, 65, 65, 66, 66, 68, 69, 69, 71

▶ Source: *Facts About the Presidents*

40. Draw a box-and-whisker plot of each data set.

41. Make a frequency distribution of each data set using five intervals beginning with 30–39. Then draw a histogram of each data set.

42. VISUAL THINKING What is one conclusion you can draw about the ages of the presidents and vice presidents based on your graphs?

STATISTICS CONNECTION **In Exercises 43–45, use the tables below which give the margins of victory for each championship game in the AFC and in the NFC for the 1966–1998 seasons.**

AFC Championship margins of victory
24, 33, 4, 10, 10, 21, 4, 17, 11, 6, 17, 3, 29, 14, 7, 20, 14, 16, 17, 17, 3, 5, 11, 16, 48, 3, 19, 17, 4, 4, 14, 3, 13

NFC Championship margins of victory
7, 4, 34, 20, 7, 11, 23, 17, 4, 30, 11, 17, 28, 9, 13, 1, 14, 3, 23, 24, 17, 7, 25, 27, 2, 31, 10, 17, 10, 11, 17, 13, 3

43. Draw a box-and-whisker plot of each data set.

44. Make a frequency distribution of each data set using five intervals beginning with 1–10. Then draw a histogram of each data set.

45. VISUAL THINKING What is one conclusion you *cannot* draw about the margins of victory in the AFC and NFC championship games based on your graphs?

46. RESEARCH Research two sets of data that you can compare. Find the measures of central tendency and measures of dispersion of each data set. Then draw a box-and-whisker plot and a histogram of each data set. What do you observe?

Test Preparation

47. MULTIPLE CHOICE What is the mean of 2, 2, 6, 7, 9, 10?

(A) 2 (B) 6 (C) 6.5 (D) 7 (E) 7.2

48. MULTIPLE CHOICE What is the median of 0.5, 0.6, 0.7, 1.2, 1.5, 1.5?

(A) 0.7 (B) 0.95 (C) 1 (D) 1.2 (E) 1.5

49. MULTIPLE CHOICE What is (are) the mode(s) of 12, 13, 13, 15, 16, 16?

(A) 13 (B) 14 (C) 13, 16 (D) 17 (E) none

★ Challenge

50. ALTERNATE FORMULA The formula for standard deviation can also be written as follows:

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

For $n = 3$, show that this formula is equivalent to the formula given on page 446. (*Hint:* You will need to show that $x_1 + x_2 + x_3 = 3\bar{x}$.)

EXTRA CHALLENGE
www.mcdougallittell.com

MIXED REVIEW

EVALUATING EXPRESSIONS **Evaluate the expression for the given value of *x*.** (Review 1.2 for 8.1)

51. $x^5 - 8$ when $x = 2$

52. $3x^3 + 7$ when $x = \frac{3}{7}$

53. $(7x)^3 + 17$ when $x = -1$

54. $\frac{4x}{x^4 - 1}$ when $x = 0.5$

PROPERTIES OF EXPONENTS **Evaluate the expression. Tell which properties of exponents you used.** (Review 6.1 for 8.1)

55. $3^3 \cdot 3^4$

56. $(4^{-3})^2$

57. $(-2)(-2)^{-3}$

58. $(5^{-2})^{-2}$

59. $10^{-2} \cdot 10^0$

60. $7^0 \cdot 7^2 \cdot 7^{-2}$

GRAPHING POLYNOMIALS **Graph the polynomial function.** (Review 6.2)

61. $f(x) = 3x^5$

62. $f(x) = x^3 - 4$

63. $f(x) = -x^4 + 3x^2 - 5$

64. $f(x) = -x^3 + 2x$

65. $f(x) = x^4 - 6x^2 - 9$

66. $f(x) = x^5 - 2x + 4$

QUIZ 3

Self Test for Lessons 7.5–7.7

Graph the function. Then state the domain and range. (Lesson 7.5)

1. $y = \sqrt{x + 8}$

2. $y = (x + 7)^{1/2} - 2$

3. $y = 3\sqrt[3]{x - 6}$

Solve the equation. Check for extraneous solutions. (Lesson 7.6)

4. $\sqrt[4]{2x} = 5$

5. $\sqrt{3x + 7} = x - 1$

6. $\sqrt[3]{2x} - 2\sqrt[3]{x} = 0$

Find the mean, median, mode, range, and standard deviation of the data set. (Lesson 7.7)

7. 0, 1, 2, 2, 5, 6, 6, 6, 7, 9

8. 15, 32, 18, 21, 26, 12, 43

9. **GEOMETRY CONNECTION** The radius r of a sphere is related to the volume V of the sphere by the formula $r = 0.620\sqrt[3]{V}$. Graph the formula. Then estimate the volume of a sphere with a radius of 10 units. (Lesson 7.5)

10. **ASTRONOMY** Kepler's third law of planetary motion states that a planet's orbital period P (in days) is related to its orbit's semi-major axis a (in millions of kilometers) by the formula $P = 0.199a^{3/2}$. The orbital period of Mars is about 1.88 *years*. What is its semi-major axis? (Lesson 7.6)

HISTORY CONNECTION **In Exercises 11 and 12, use the following data set of years when each state was admitted to statehood.** (Lesson 7.7)

1819, 1959, 1912, 1836, 1850, 1876, 1788, 1787, 1845, 1788,
1959, 1890, 1818, 1816, 1846, 1861, 1792, 1812, 1820, 1788,
1788, 1837, 1858, 1817, 1821, 1889, 1867, 1864, 1788, 1787,
1912, 1788, 1789, 1889, 1803, 1907, 1859, 1787, 1790, 1788,
1889, 1796, 1845, 1896, 1791, 1788, 1889, 1863, 1848, 1890

11. Draw a box-and-whisker plot of the data set.

12. Make a frequency distribution of the data set using five intervals beginning with 1750–1799. Then draw a histogram of the data set.