P.6 Representing Data Graphically

Line Plots

Statistics is the branch of mathematics that studies techniques for collecting, organizing, and interpreting data. In this section, you will study several ways to organize data. The first is a **line plot,** which uses a portion of a real number line to order numbers. Line plots are especially useful for ordering small sets of numbers (about 50 or less) by hand.

Many statistical measures can be obtained from a line plot. Two such measures are the *frequency* and *range* of the data. The **frequency** measures the number of times a value occurs in a data set. The **range** is the difference between the greatest and smallest data values. For example, consider the data values

20, 21, 21, 25, 32.

The frequency of 21 in the data set is 2 because 21 occurs twice. The range is 12 because the difference between the greatest and smallest data values is $32 - 20 = 12$.

Example 1 Constructing a Line Plot

Use a line plot to organize the following test scores. Which score occurs with the greatest frequency? What is the range of scores?

93, 70, 76, 67, 86, 93, 82, 78, 83, 86, 64, 78, 76, 66, 83
83, 96, 74, 69, 76, 64, 74, 79, 76, 88, 76, 81, 82, 74, 70

Solution

Begin by scanning the data to find the smallest and largest numbers. For the data, the smallest number is 64 and the largest is 96. Next, draw a portion of a real number line that includes the interval $[64, 96]$. To create the line plot, start with the first number, 93, and enter an $\times$ above 93 on the number line. Continue recording $\times$'s for each number in the list until you obtain the line plot shown in Figure P.29. From the line plot, you can see that 76 occurs with the greatest frequency. Because the range is the difference between the greatest and smallest data values, the range of scores is $96 - 64 = 32$.

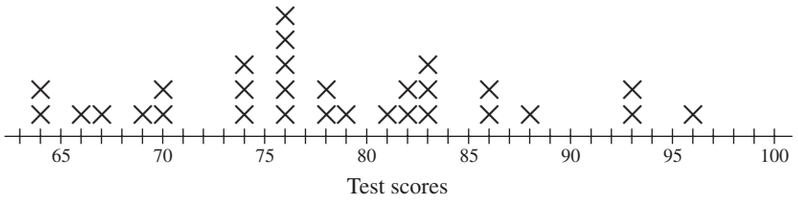


Figure P.29

CHECKPOINT Now try Exercise 1.

What you should learn

- Use line plots to order and analyze data.
- Use histograms to represent frequency distributions.
- Use bar graphs to represent and analyze data.
- Use line graphs to represent and analyze data.

Why you should learn it

Double bar graphs allow you to compare visually two sets of data over time. For example, in Exercises 9 and 10 on page 65, you are asked to estimate the difference in tuition between public and private institutions of higher education.

Cindy Charles/PhotoEdit

Note that methods for representing data graphically also include the scatter plot, already mentioned in Section P.5.

Histograms and Frequency Distributions

When you want to organize large sets of data, it is useful to group the data into intervals and plot the frequency of the data in each interval. A **frequency distribution** can be used to construct a **histogram.** A histogram uses a portion of a real number line as its horizontal axis. The bars of a histogram are not separated by spaces.

Example 2 Constructing a Histogram

The table at the right shows the percent of the resident population of each state and the District of Columbia that was at least 65 years old in 2004. Construct a frequency distribution and a histogram for the data. (Source: U.S. Census Bureau)

AK	6.4	MT	13.7
AL	13.2	NC	12.1
AR	13.8	ND	14.7
AZ	12.7	NE	13.3
CA	10.7	NH	12.1
CO	9.8	NJ	12.9
CT	13.5	NM	12.1
DC	12.1	NV	11.2
DE	13.1	NY	13.0
FL	16.8	OH	13.3
GA	9.6	OK	13.2
HI	13.6	OR	12.8
IA	14.7	PA	15.3
ID	11.4	RI	13.9
IL	12.0	SC	12.4
IN	12.4	SD	14.2
KS	13.0	TN	12.5
KY	12.5	TX	9.9
LA	11.7	UT	8.7
MA	13.3	VA	11.4
MD	11.4	VT	13.0
ME	14.4	WA	11.3
MI	12.3	WI	13.0
MN	12.1	WV	15.3
MO	13.3	WY	12.1
MS	12.2		

Solution

To begin constructing a frequency distribution, you must first decide on the number of intervals. There are several ways to group the data. However, because the smallest number is 6.4 and the largest is 16.8, it seems that six intervals would be appropriate. The first would be the interval $[6, 8)$, the second would be $[8, 10)$, and so on. By tallying the data into the six intervals, you obtain the frequency distribution shown below. You can construct the histogram by drawing a vertical axis to represent the number of states and a horizontal axis to represent the percent of the population 65 and older. Then, for each interval, draw a vertical bar whose height is the total tally, as shown in Figure P.30.

Interval	*Tally*
$[6, 8)$	|
$[8, 10)$	||||
$[10, 12)$	𝍸 ||
$[12, 14)$	𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 ||
$[14, 16)$	𝍸 |
$[16, 18)$	|

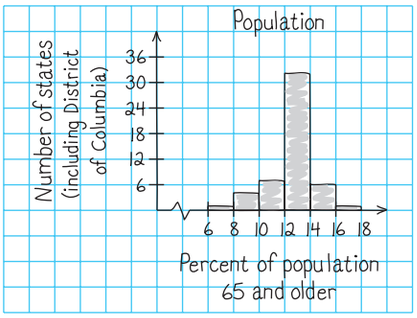


Figure P.30

CHECKPOINT Now try Exercise 5.

Example 3 Constructing a Histogram

A company has 48 sales representatives who sold the following numbers of units during the first quarter of 2008. Construct a frequency distribution for the data.

107	162	184	170	177	102	145	141
105	193	167	149	195	127	193	191
150	153	164	167	171	163	141	129
109	171	150	138	100	164	147	153
171	163	118	142	107	144	100	132
153	107	124	162	192	134	187	177

Interval	*Tally*
100–109	𝍸 |||
110–119	|
120–129	|||
130–139	|||
140–149	𝍸 ||
150–159	𝍸
160–169	𝍸 |||
170–179	𝍸 |
180–189	||
190–199	𝍸

Solution

To begin constructing a frequency distribution, you must first decide on the number of intervals. There are several ways to group the data. However, because the smallest number is 100 and the largest is 195, it seems that 10 intervals would be appropriate. The first interval would be 100–109, the second would be 110–119, and so on. By tallying the data into the 10 intervals, you obtain the distribution shown at the right above. A histogram for the distribution is shown in Figure P.31.

CHECKPOINT Now try Exercise 6.

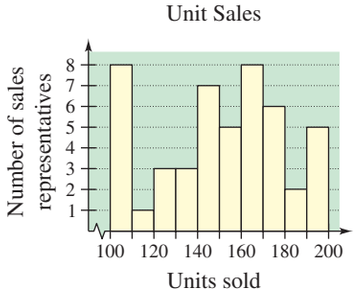


Figure P.31

Bar Graphs

A **bar graph** is similar to a histogram, except that the bars can be either horizontal or vertical and the labels of the bars are not necessarily numbers. Another difference between a bar graph and a histogram is that the bars in a bar graph are usually separated by spaces.

Example 4 Constructing a Bar Graph

The data below show the monthly normal precipitation (in inches) in Houston, Texas. Construct a bar graph for the data. What can you conclude? (Source: National Climatic Data Center)

January	3.7	*February*	3.0	*March*	3.4
April	3.6	*May*	5.2	*June*	5.4
July	3.2	*August*	3.8	*September*	4.3
October	4.5	*November*	4.2	*December*	3.7

Solution

To create a bar graph, begin by drawing a vertical axis to represent the precipitation and a horizontal axis to represent the month. The bar graph is shown in Figure P.32. From the graph, you can see that Houston receives a fairly consistent amount of rain throughout the year—the driest month tends to be February and the wettest month tends to be June.

CHECKPOINT Now try Exercise 7.

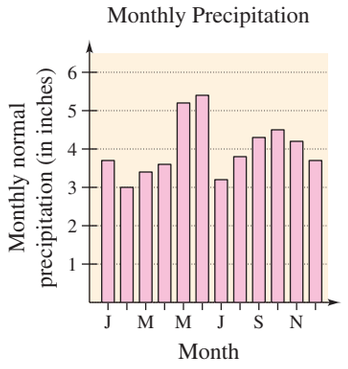


Figure P.32

Example 5 Constructing a Double Bar Graph

The table shows the percents of associate degrees awarded to males and females for selected fields of study in the United States in 2003. Construct a double bar graph for the data. (Source: U.S. National Center for Education Statistics)

Field of Study	% Female	% Male
Agriculture and Natural Resources	36.4	63.6
Biological Sciences/ Life Sciences	70.4	29.6
Business and Management	66.8	33.2
Education	80.5	19.5
Engineering	16.5	83.5
Law and Legal Studies	89.6	10.4
Liberal/General Sciences	63.1	36.9
Mathematics	36.5	63.5
Physical Sciences	44.7	55.3
Social Sciences	65.3	34.7

Solution

For the data, a horizontal bar graph seems to be appropriate. This makes it easier to label and read the bars. Such a graph is shown in Figure P.33.

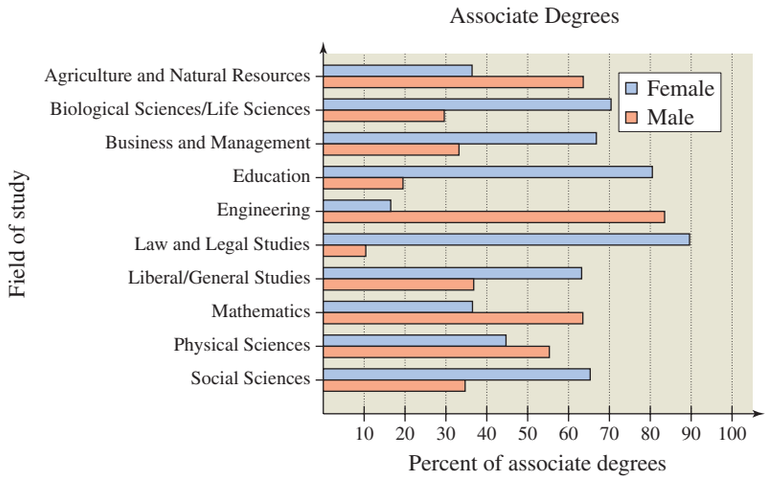


Figure P.33

CHECKPOINT Now try Exercise 11.

Line Graphs

A **line graph** is similar to a standard coordinate graph. Line graphs are usually used to show trends over periods of time.

Example 6 Constructing a Line Graph

The table at the right shows the number of immigrants (in thousands) entering the United States for each decade from 1901 to 2000. Construct a line graph for the data. What can you conclude? (Source: U.S. Immigration and Naturalization Service)

Decade	Number
1901–1910	8795
1911–1920	5736
1921–1930	4107
1931–1940	528
1941–1950	1035
1951–1960	2515
1961–1970	3322
1971–1980	4493
1981–1990	7338
1991–2000	9095

Solution

Begin by drawing a vertical axis to represent the number of immigrants in thousands. Then label the horizontal axis with decades and plot the points shown in the table. Finally, connect the points with line segments, as shown in Figure P.34. From the line graph, you can see that the number of immigrants hit a low point during the depression of the 1930s. Since then the number has steadily increased.

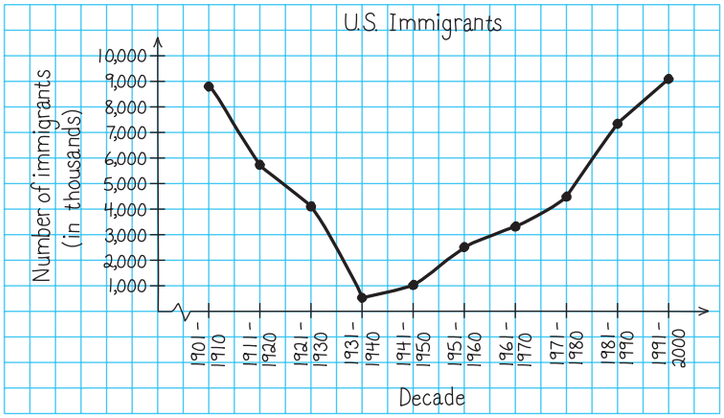


Figure P.34

CHECKPOINT Now try Exercise 17.

TECHNOLOGY TIP You can use a graphing utility to create different types of graphs, such as line graphs. For instance, the table at the right shows the numbers N (in thousands) of women on active duty in the United States military for selected years. To use a graphing utility to create a line graph of the data, first enter the data into the graphing utility's *list editor*, as shown in Figure P.35. Then use the *statistical plotting* feature to set up the line graph, as shown in Figure P.36. Finally, display the line graph (use a viewing window in which $1970 \le x \le 2010$ and $0 \le y \le 250$), as shown in Figure P.37. (Source: U.S. Department of Defense)

Year	Number
1975	97
1980	171
1985	212
1990	227
1995	196
2000	203
2005	203

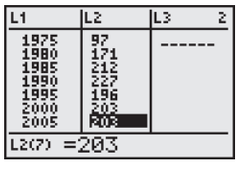


Figure P.35

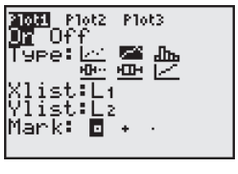


Figure P.36

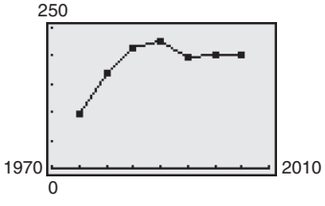


Figure P.37

P.6 Exercises

See www.CalcChat.com for worked-out solutions to odd-numbered exercises.

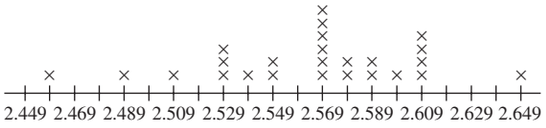
Vocabulary Check

Fill in the blanks.

1. _____ is the branch of mathematics that studies techniques for collecting, organizing, and interpreting data.
2. _____ are useful for ordering small sets of numbers by hand.
3. A _____ uses a portion of a real number line as its horizontal axis, and the bars are not separated by spaces.
4. You use a _____ to construct a histogram.
5. The bars in a _____ can be either vertical or horizontal.
6. _____ show trends over periods of time.

---

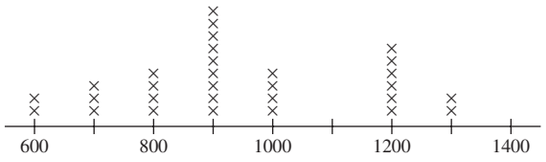
1. ***Consumer Awareness*** The line plot shows a sample of prices of unleaded regular gasoline in 25 different cities.



(a) What price occurred with the greatest frequency?

(b) What is the range of prices?

2. ***Agriculture*** The line plot shows the weights (to the nearest hundred pounds) of 30 head of cattle sold by a rancher.



(a) What weight occurred with the greatest frequency?

(b) What is the range of weights?

Quiz and Exam Scores **In Exercises 3 and 4, use the following scores from an algebra class of 30 students. The scores are for one 25-point quiz and one 100-point exam.**

Quiz **20, 15, 14, 20, 16, 19, 10, 21, 24, 15, 15, 14, 15, 21, 19, 15, 20, 18, 18, 22, 18, 16, 18, 19, 21, 19, 16, 20, 14, 12**

Exam **77, 100, 77, 70, 83, 89, 87, 85, 81, 84, 81, 78, 89, 78, 88, 85, 90, 92, 75, 81, 85, 100, 98, 81, 78, 75, 85, 89, 82, 75**

3. Construct a line plot for the quiz. Which score(s) occurred with the greatest frequency?
4. Construct a line plot for the exam. Which score(s) occurred with the greatest frequency?

5. ***Agriculture*** The list shows the numbers of farms (in thousands) in the 50 states in 2004. Use a frequency distribution and a histogram to organize the data. (Source: U.S. Department of Agriculture)

AK 1	AL 44	AR 48	AZ 10
CA 77	CO 31	CT 4	DE 2
FL 43	GA 49	HI 6	IA 90
ID 25	IL 73	IN 59	KS 65
KY 85	LA 27	MA 6	MD 12
ME 7	MI 53	MN 80	MO 106
MS 42	MT 28	NC 52	ND 30
NE 48	NH 3	NJ 10	NM 18
NV 3	NY 36	OH 77	OK 84
OR 40	PA 58	RI 1	SC 24
SD 32	TN 85	TX 229	UT 15
VA 48	VT 6	WA 35	WI 77
WV 21	WY 9		

6. ***Schools*** The list shows the numbers of public high school graduates (in thousands) in the 50 states and the District of Columbia in 2004. Use a frequency distribution and a histogram to organize the data. (Source: U.S. National Center for Education Statistics)

AK 7.1	AL 37.6	AR 26.9	AZ 57.0
CA 342.6	CO 42.9	CT 34.4	DC 3.2
DE 6.8	FL 129.0	GA 69.7	HI 10.3
IA 33.8	ID 15.5	IL 121.3	IN 57.6
KS 30.0	KY 36.2	LA 36.2	MA 57.9
MD 53.0	ME 13.4	MI 106.3	MN 59.8
MO 57.0	MS 23.6	MT 10.5	NC 71.4
ND 7.8	NE 20.0	NH 13.3	NJ 88.3
NM 18.1	NV 16.2	NY 150.9	OH 116.3
OK 36.7	OR 32.5	PA 121.6	RI 9.3
SC 32.1	SD 9.1	TN 43.6	TX 236.7
UT 29.9	VA 71.7	VT 7.0	WA 60.4
WI 62.3	WV 17.1	WY 5.7	

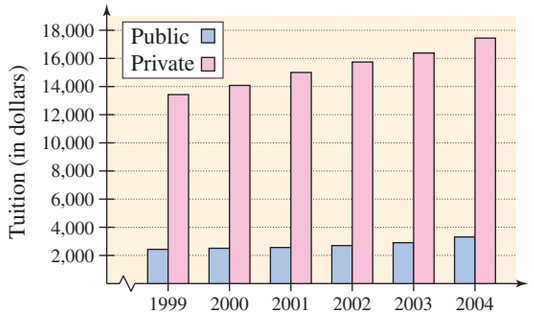
7. ***Business*** The table shows the numbers of Wal-Mart stores from 1995 to 2006. Construct a bar graph for the data. Write a brief statement regarding the number of Wal-Mart stores over time. (Source: Value Line)

Year	Number of stores
1995	2943
1996	3054
1997	3406
1998	3599
1999	3985
2000	4189
2001	4414
2002	4688
2003	4906
2004	5289
2005	5650
2006	6050

8. ***Business*** The table shows the revenues (in billions of dollars) for Costco Wholesale from 1995 to 2006. Construct a bar graph for the data. Write a brief statement regarding the revenue of Costco Wholesale stores over time. (Source: Value Line)

Year	Revenue (in billions of dollars)
1995	18.247
1996	19.566
1997	21.874
1998	24.270
1999	27.456
2000	32.164
2001	34.797
2002	38.762
2003	42.546
2004	48.107
2005	52.935
2006	58.600

Tuition **In Exercises 9 and 10, the double bar graph shows the mean tuitions (in dollars) charged by public and private institutions of higher education in the United States from 1999 to 2004.** (Source: U.S. National Center for Education Statistics)



9. Approximate the difference in tuition charges for public and private schools for each year.

10. Approximate the increase or decrease in tuition charges for each type of institution from year to year.

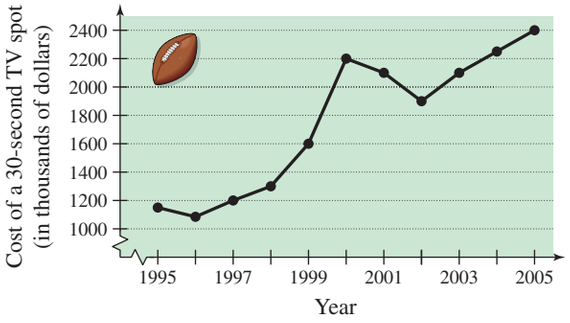
11. ***College Enrollment*** The table shows the total college enrollments (in thousands) for women and men in the United States from 1997 to 2003. Construct a double bar graph for the data. (Source: U.S. National Center for Education Statistics)

Year	Women (in thousands)	Men (in thousands)
1997	8106.3	6396.0
1998	8137.7	6369.3
1999	8300.6	6490.6
2000	8590.5	6721.8
2001	8967.2	6960.8
2002	9410.0	7202.0
2003	9652.0	7259.0

12. ***Population*** The table shows the populations (in millions) in the coastal regions of the United States in 1970 and 2003. Construct a double bar graph for the data. (Source: U.S. Census Bureau)

Region	1970 population (in millions)	2003 population (in millions)
Atlantic	52.1	67.1
Gulf of Mexico	10.0	18.9
Great Lakes	26.0	27.5
Pacific	22.8	39.4

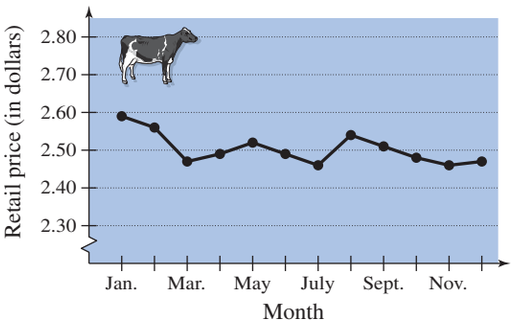
Advertising **In Exercises 13 and 14, use the line graph, which shows the costs of a 30-second television spot (in thousands of dollars) during the Super Bowl from 1995 to 2005.** (Source: The Associated Press)



13. Approximate the percent increase in the cost of a 30-second spot from Super Bowl XXX in 1996 to Super Bowl XXXIX in 2005.

14. Estimate the increase or decrease in the cost of a 30-second spot from (a) Super Bowl XXIX in 1995 to Super Bowl XXXIII in 1999, and (b) Super Bowl XXXIV in 2000 to Super Bowl XXXIX in 2005.

Retail Price **In Exercises 15 and 16, use the line graph, which shows the average retail price (in dollars) of one pound of 100% ground beef in the United States for each month in 2004.** (Source: U.S. Bureau of Labor Statistics)



15. What is the highest price of one pound of 100% ground beef shown in the graph? When did this price occur?

16. What was the difference between the highest price and the lowest price of one pound of 100% ground beef in 2004?

17. ***Labor*** The table shows the total numbers of women in the work force (in thousands) in the United States from 1995 to 2004. Construct a line graph for the data. Write a brief statement describing what the graph reveals. (Source: U.S. Bureau of Labor Statistics)

Year	Women in the work force (in thousands)
1995	60,944
1996	61,857
1997	63,036
1998	63,714
1999	64,855
2000	66,303
2001	66,848
2002	67,363
2003	68,272
2004	68,421

18. ***SAT Scores*** The table shows the average Scholastic Aptitude Test (SAT) Math Exam scores for college-bound seniors in the United States for selected years from 1970 to 2005. Construct a line graph for the data. Write a brief statement describing what the graph reveals. (Source: The College Entrance Examination Board)

Year	SAT scores
1970	512
1975	498
1980	492
1985	500
1990	501
1995	506
2000	514
2005	520

19. ***Hourly Earnings*** The table on page 67 shows the average hourly earnings (in dollars) of production workers in the United States from 1994 to 2005. Use a graphing utility to construct a line graph for the data. (Source: U.S. Bureau of Labor Statistics)

Year	Hourly earnings (in dollars)
1994	11.19
1995	11.47
1996	11.84
1997	12.27
1998	12.77
1999	13.25
2000	13.73
2001	14.27
2002	14.73
2003	15.19
2004	15.48
2005	15.90

Table for 19

20. ***Internet Access*** The list shows the percent of households in each of the 50 states and the District of Columbia with Internet access in 2003. Use a graphing utility to organize the data in the graph of your choice. Explain your choice of graph. (Source: U.S. Department of Commerce)

AK 67.6	AL 45.7	AR 42.4	AZ 55.2
CA 59.6	CO 63.0	CT 62.9	DC 53.2
DE 56.8	FL 55.6	GA 53.5	HI 55.0
IA 57.1	ID 56.4	IL 51.1	IN 51.0
KS 54.3	KY 49.6	LA 44.1	MA 58.1
MD 59.2	ME 57.9	MI 52.0	MN 61.6
MO 53.0	MS 38.9	MT 50.4	NC 51.1
ND 53.2	NE 55.4	NH 65.2	NJ 60.5
NM 44.5	NV 55.2	NY 53.3	OH 52.5
OK 48.4	OR 61.0	PA 54.7	RI 55.7
SC 45.6	SD 53.6	TN 48.9	TX 51.8
UT 62.6	VA 60.3	VT 58.1	WA 62.3
WI 57.4	WV 47.6	WY 57.7	

Cellular Phones **In Exercises 21 and 22, use the table, which shows the average monthly cellular telephone bills (in dollars) in the United States from 1999 to 2004.** (Source: Telecommunications & Internet Association)

Year	Average monthly bill (in dollars)
1999	41.24
2000	45.27
2001	47.37
2002	48.40
2003	49.91
2004	50.64

21. Organize the data in an appropriate display. Explain your choice of graph.

22. The average monthly bills in 1990 and 1995 were $80.90 and $51.00, respectively. How would you explain the trend(s) in the data?

23. ***High School Athletes*** The table shows the numbers of participants (in thousands) in high school athletic programs in the United States from 1995 to 2004. Organize the data in an appropriate display. Explain your choice of graph. (Source: National Federation of State High School Associations)

Year	Female athletes (in thousands)	Male athletes (in thousands)
1995	2240	3536
1996	2368	3634
1997	2474	3706
1998	2570	3763
1999	2653	3832
2000	2676	3862
2001	2784	3921
2002	2807	3961
2003	2856	3989
2004	2865	4038

Synthesis

24. ***Writing*** Describe the differences between a bar graph and a histogram.

25. ***Think About It*** How can you decide which type of graph to use when you are organizing data?

26. ***Graphical Interpretation*** The graphs shown below represent the same data points. Which of the two graphs is misleading, and why? Discuss other ways in which graphs can be misleading. Try to find another example of a misleading graph in a newspaper or magazine. Why is it misleading? Why would it be beneficial for someone to use a misleading graph?