5. Instruction Time The average weekly instruction time in schools for 5 selected countries is shown. Construct a vertical bar graph and a Pareto chart for the data.

Thailand	30.5 hours
China	26.9 hours
France	24.8 hours
United States	22.2 hours
Brazil	19 hours

Source: Organization for Economic Cooperation and Development.

6. Sales of Coffee The data show the total retail sales (in billions of dollars) of coffee for 6 years. Over the years, are the sales increasing or decreasing?

Year	2001	2002	2003	2004	2005	2006
Sales	$8.3	$8.4	$9.0	$9.6	$11.1	$12.3

Source: Specialty Coffee Association of America.

7. Safety Record of U.S. Airlines The safety record of U.S. airlines for 10 years is shown. Construct a time series graph for the data.

Year	Major Accidents
1997	2
1998	0
1999	2
2000	3
2001	1
2002	1
2003	2
2004	4
2005	2
2006	2
2007	0

Source: National Transportation Safety Board.

8. Average Global Temperatures The average global temperatures for the following years are shown. Draw a time series graph and comment on the trend.

Year	2004	2005	2006	2007	2008
Temperature	57.98	58.11	57.99	58.01	57.88

Source: National Oceanic and Atmospheric Administration.

9. Carbon Dioxide Concentrations The following data for the atmospheric concentration of carbon dioxide (in ppm^2) are shown. Draw a time series graph and comment on the trend.

Year	2004	2005	2006	2007	2008
Concentration	375	377	379	381	383

Source: U.S. Department of Energy.

10. Reasons We Travel The following data are based on a survey from American Travel Survey on why people travel. Construct a pie graph for the data and analyze the results.

Purpose	Number
Personal business	146
Visit friends or relatives	330
Work-related	225
Leisure	299

Source: *USA TODAY.*

11. Characteristics of the Population 65 and Over Two characteristics of the population aged 65 and over are shown below for 2004. Illustrate each characteristic with a pie graph.

Marital status		Educational attainment	
Never married	3.9%	Less than ninth grade	13.9%
Married	57.2	Completed grades 9–12	
Widowed	30.8	but no diploma	13.0
Divorced	8.1	H.S. graduate	36.0
		Some college/ associates degree	18.4
		Bachelor's/advanced degree	18.7

Source: *New York Times Almanac.*

12. Colors of Automobiles The popular vehicle car colors are shown. Construct a pie graph for the data.

White	19%
Silver	18
Black	16
Red	13
Blue	12
Gray	12
Other	10

Source: Dupont Automotive Color Popularity Report.

13. Workers Switch Jobs In a recent survey, 3 in 10 people indicated that they are likely to leave their jobs when the economy improves. Of those surveyed, 34% indicated that they would make a career change, 29% want a new job in the same industry, 21% are going to start a business, and 16% are going to retire. Make a pie chart and a Pareto chart for the data. Which chart do you think better represents the data?

Source: National Survey Institute.

14. State which graph (Pareto chart, time series graph, or pie graph) would most appropriately represent the given situation.

a. The number of students enrolled at a local college for each year during the last 5 years.

b. The budget for the student activities department at a certain college for a specific year.
c. The means of transportation the students use to get to school.
d. The percentage of votes each of the four candidates received in the last election.
e. The record temperatures of a city for the last 30 years.
f. The frequency of each type of crime committed in a city during the year.

15. Presidents' Ages at Inauguration The age at inauguration for each U.S. President is shown. Construct a stem and leaf plot and analyze the data.

57	54	52	55	51	56	47
61	68	56	55	54	61	51
57	51	46	54	51	52	
57	49	54	42	60	69	
58	64	49	51	62	64	
57	48	51	56	43	46	
61	65	47	55	55	54	

Source: *New York Times Almanac.*

16. Calories in Salad Dressings A listing of calories per one ounce of selected salad dressings (not fat-free) is given below. Construct a stem and leaf plot for the data.

100	130	130	130	110	110	120	130	140	100
140	170	160	130	160	120	150	100	145	145
145	115	120	100	120	160	140	120	180	100
160	120	140	150	190	150	180	160		

17. Twenty Days of Plant Growth The growth (in centimeters) of two varieties of plant after 20 days is shown in this table. Construct a back-to-back stem and leaf plot for the data, and compare the distributions.

Variety 1				Variety 2			
20	12	39	38	18	45	62	59
41	43	51	52	53	25	13	57
59	55	53	59	42	55	56	38
50	58	35	38	41	36	50	62
23	32	43	53	45	55		

18. Math and Reading Achievement Scores The math and reading achievement scores from the National Assessment of Educational Progress for selected states are listed below. Construct a back-to-back stem and leaf plot with the data and compare the distributions.

Math					Reading				
52	66	69	62	61	65	76	76	66	67
63	57	59	59	55	71	70	70	66	61
55	59	74	72	73	61	69	78	76	77
68	76	73			77	77	80		

Source: *World Almanac.*

19. The sales of recorded music in 2004 by genre are listed below. Represent the data with an appropriate graph.
Answers will vary.

Rock	23.9	Jazz	2.7
Country	13.0	Classical	2.0
Rap/hip-hop	12.1	Oldies	1.4
R&B/urban	11.3	Soundtracks	1.1
Pop	10.0	New age	1.0
Religious	6.0	Other	8.9
Children's	2.8		

Source: *World Almanac.*

Extending the Concepts

20. Successful Space Launches The number of successful space launches by the United States and Japan for the years 1993–1997 is shown here. Construct a compound time series graph for the data. What comparison can be made regarding the launches?

Year	1993	1994	1995	1996	1997
United States	29	27	24	32	37
Japan	1	4	2	1	2

Source: *The World Almanac and Book of Facts.*

21. Meat Production Meat production for veal and lamb for the years 1960–2000 is shown here. (Data are in millions of pounds.) Construct a compound time series graph for the data. What comparison can be made regarding meat production?

Year	1960	1970	1980	1990	2000
Veal	1109	588	400	327	225
Lamb	769	551	318	358	234

Source: *The World Almanac and Book of Facts.*

22. Top 10 Airlines During a recent year the top 10 airlines with the most aircraft are listed. Represent these data with an appropriate graph.

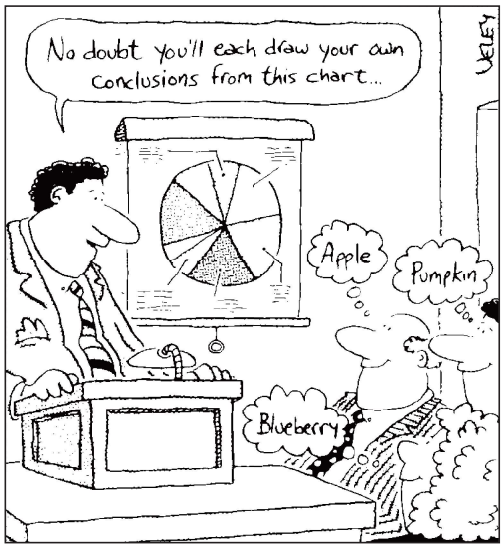
American	714	Continental	364
United	603	Southwest	327
Delta	600	British Airways	268
Northwest	424	American Eagle	245
U.S. Airways	384	Lufthansa (Ger.)	233

Source: *Top 10 of Everything.*

23. Nobel Prizes in Physiology or Medicine The top prize-winning countries for Nobel Prizes in Physiology or Medicine are listed here. Represent the data with an appropriate graph.

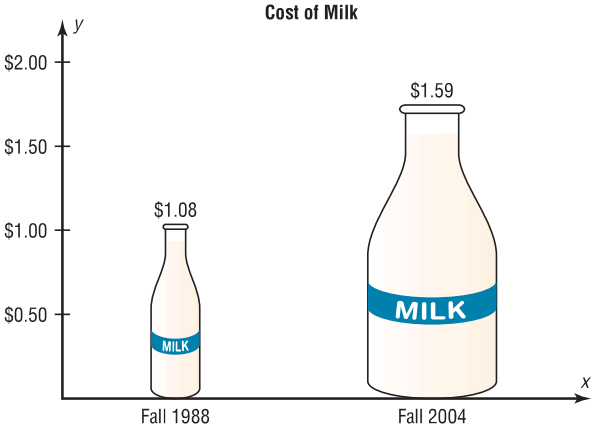
United States	80	Denmark	5
United Kingdom	24	Austria	4
Germany	16	Belgium	4
Sweden	8	Italy	3
France	7	Australia	3
Switzerland	6		

Source: *Top 10 of Everything.*

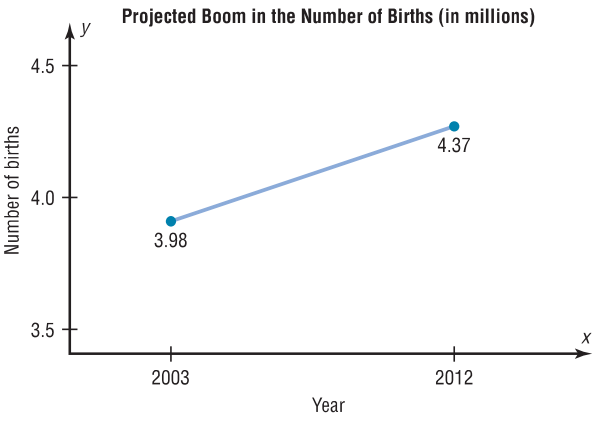


Source: Cartoon by Bradford Veley, Marquette, Michigan. Used with permission.

24. Cost of Milk The graph shows the increase in the price of a quart of milk. Why might the increase appear to be larger than it really is?



25. Boom in Number of Births The graph shows the projected boom (in millions) in the number of births. Cite several reasons why the graph might be misleading.

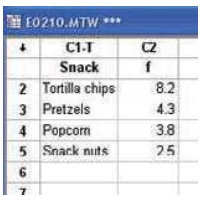


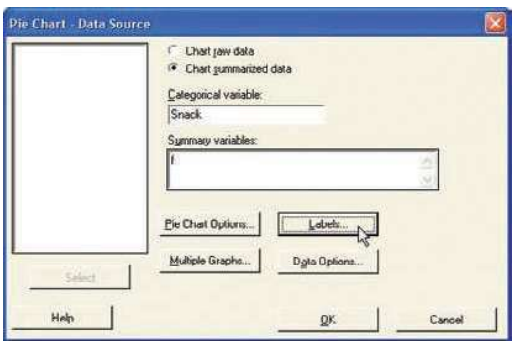
Technology *Step by Step*

MINITAB Step by Step

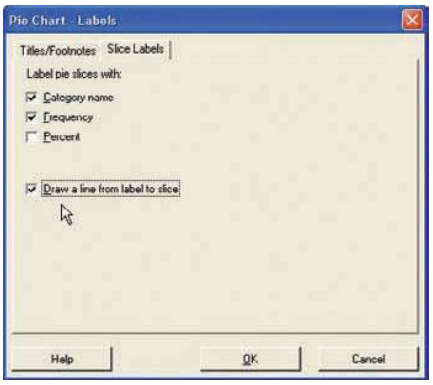
Construct a Pie Chart

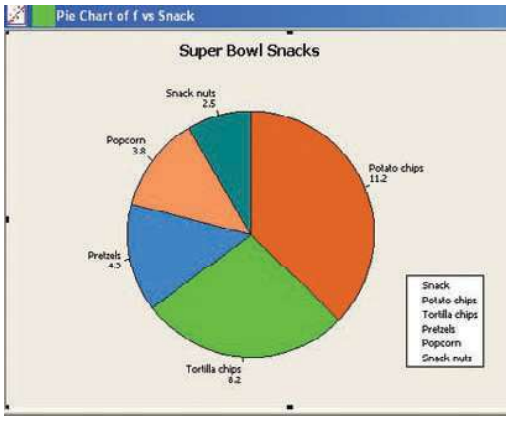
1. Enter the summary data for snack foods and frequencies from Example 2–11 into C1 and C2.





2. Name them **Snack** and **f.**
3. Select **Graph>Pie Chart.**
   a) Click the option for Chart summarized data.
   b) Press [Tab] to move to Categorical variable, then double-click C1 to select it.
   c) Press [Tab] to move to Summary variables, and select the column with the frequencies f.



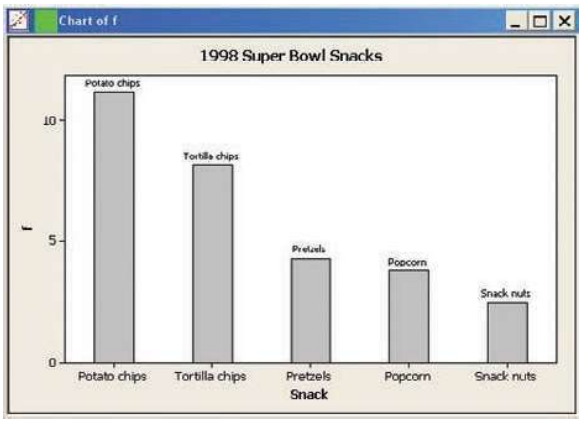


4. Click the [Labels] tab, then Titles/Footnotes.
   a) Type in the title: **Super Bowl Snacks.**
   b) Click the Slice Labels tab, then the options for Category name and Frequency.
   c) Click the option to Draw a line from label to slice.
   d) Click [OK] twice to create the chart.

Construct a Bar Chart

The procedure for constructing a bar chart is similar to that for the pie chart.

1. Select **Graph>Bar Chart.**
   a) Click on the drop-down list in Bars Represent: then select values from a table.
   b) Click on the Simple chart, then click [OK]. The dialog box will be similar to the Pie Chart Dialog Box.
2. Select the frequency column C2 f for Graph variables: and Snack for the Categorical variable.

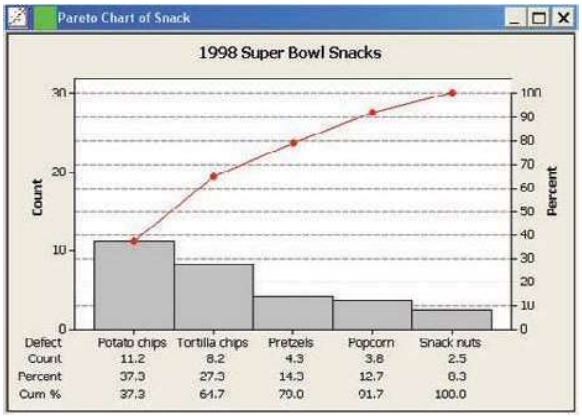


3. Click on [Labels], then type the title in the Titles/Footnote tab: **1998 Super Bowl Snacks.**
4. Click the tab for Data Labels, then click the option to Use labels from column: and select C1 Snacks.
5. Click [OK] twice.

Construct a Pareto Chart

Pareto charts are a quality control tool. They are similar to a bar chart with no gaps between the bars, and the bars are arranged by frequency.

1. Select **Stat>Quality Tools>Pareto.**
2. Click the option to Chart defects table.
3. Click in the box for the Labels in: and select Snack.
4. Click on the frequencies column C2 f.



5. Click on [Options].
   a) Check the box for Cumulative percents.
   b) Type in the title, **1998 Super Bowl Snacks.**
6. Click [OK] twice. The chart is completed.

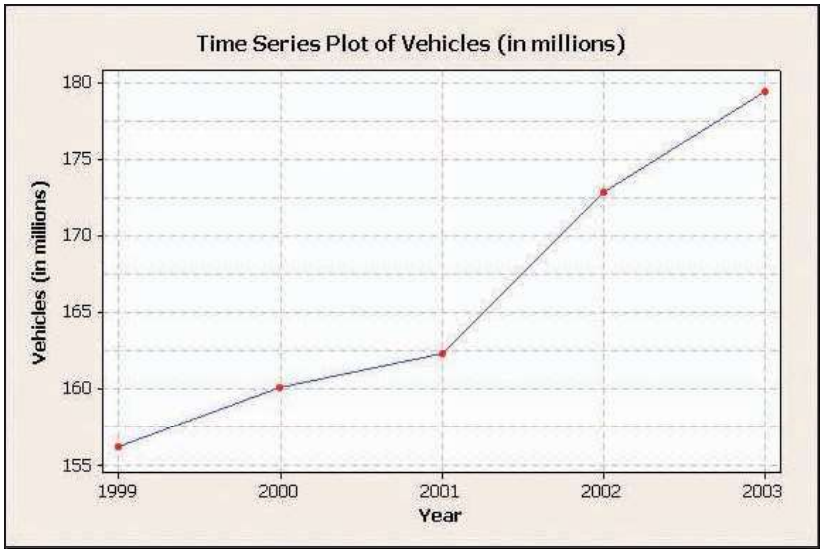
Construct a Time Series Plot

The data used are for the number of vehicles that used the Pennsylvania Turnpike.

Year	1999	2000	2001	2002	2003
Number	156.2	160.1	162.3	172.8	179.4

1. Add a blank worksheet to the project by selecting **File>New>New Worksheet.**
2. To enter the dates from 1999 to 2003 in C1, select **Calc>Make Patterned Data>Simple Set of Numbers.**
   a) Type **Year** in the text box for Store patterned data in.
   b) From first value: should be **1999.**
   c) To Last value: should be **2003.**
   d) In steps of should be **1** (for every other year). The last two boxes should be 1, the default value.
   e) Click [OK]. The sequence from 1999 to 2003 will be entered in C1 whose label will be Year.
3. Type **Vehicles (in millions)** for the label row above row 1 in C2.

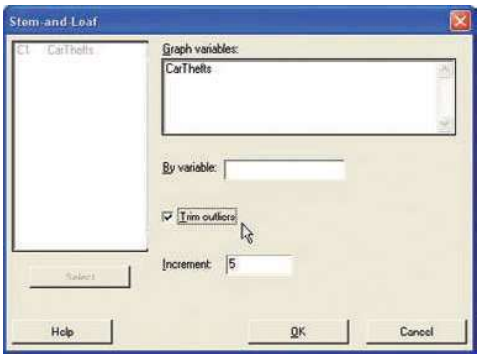
4. Type **156.2** for the first number, then press [Enter]. Never enter the commas for large numbers!
5. Continue entering the value in each row of C2.



6. To make the graph, select **Graph>Time series plot,** then Simple, and press [OK].
   a) For Series select Vehicles (in millions), then click [Time/scale].
   b) Click the Stamp option and select Year for the Stamp column.
   c) Click the Gridlines tab and select all three boxes, Y major, Y minor, and X major.
   d) Click [OK] twice. A new window will open that contains the graph.
   e) To change the title, double-click the title in the graph window. A dialog box will open, allowing you to edit the text.

Construct a Stem and Leaf Plot

1. Type in the data for Example 2–14. Label the column **CarThefts.**
2. Select **STAT>EDA>Stem-and-Leaf.** This is the same as **Graph>Stem-and-Leaf.**
3. Double-click on C1 CarThefts in the column list.
4. Click in the Increment text box, and enter the class width of **5.**
5. Click [OK]. This character graph will be displayed in the session window.



Stem-and-Leaf Display: CarThefts
Stem-and-leaf of CarThefts N = 30
Leaf Unit = 1.0

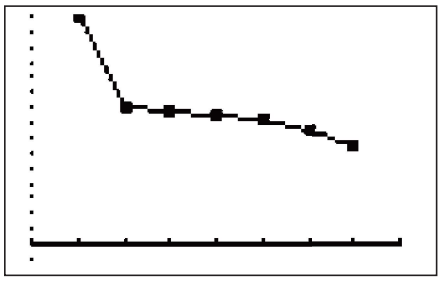
6	5	011233
13	5	5567789
15	6	23
15	6	55667899
7	7	23
5	7	55789

TI-83 Plus or TI-84 Plus Step by Step

To graph a time series, follow the procedure for a frequency polygon from Section 2–2, using the following data for the number of outdoor drive-in theaters

Year	1988	1990	1992	1994	1996	1998	2000
Number	1497	910	870	859	826	750	637

Output

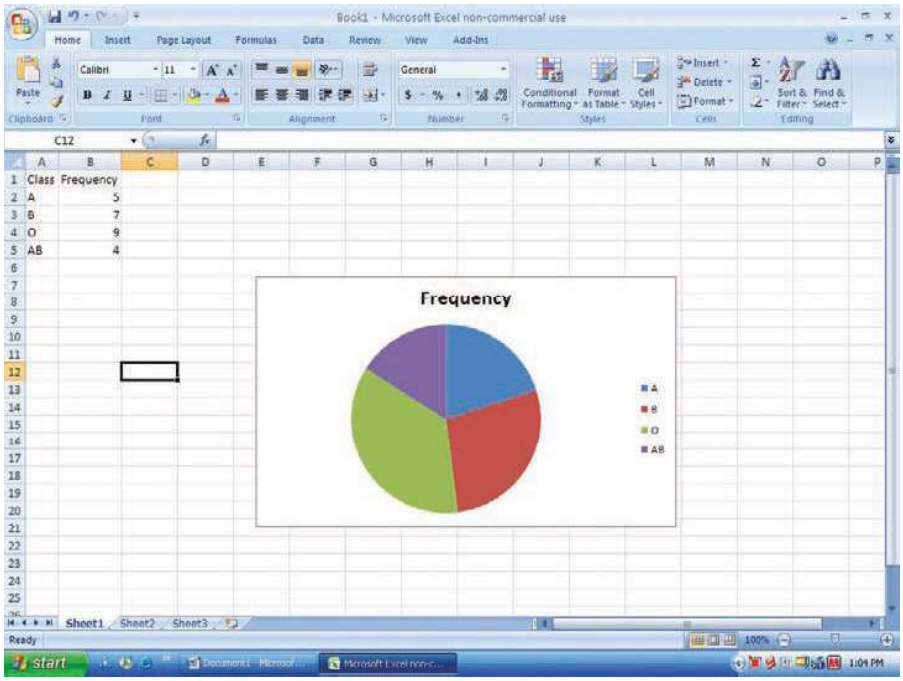


Excel Step by Step

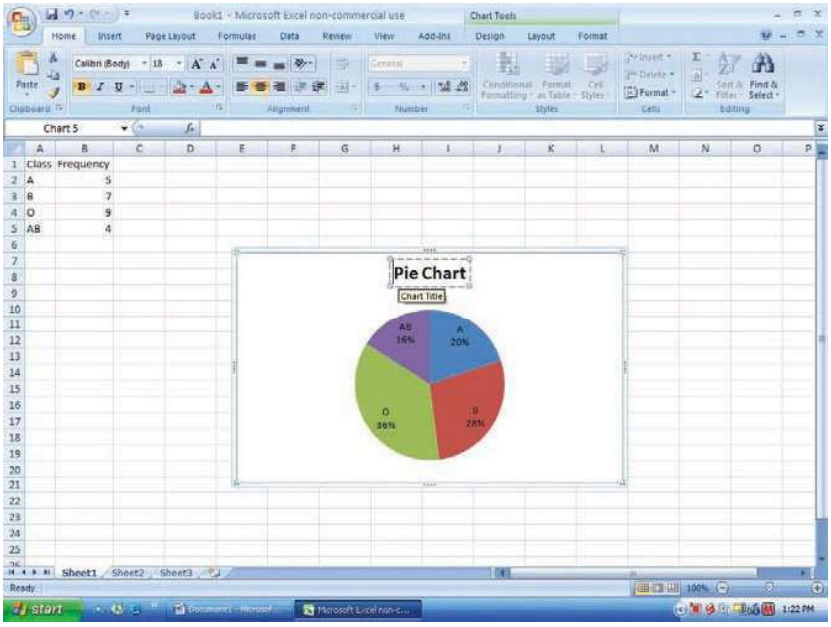
Constructing a Pie Chart

To make a pie chart:

1. Enter the blood types from Example 2–12 into column A of a new worksheet.
2. Enter the frequencies corresponding to each blood type in column B.
3. Highlight the data in columns A and B and select Insert from the toolbar, then select the Pie chart type.



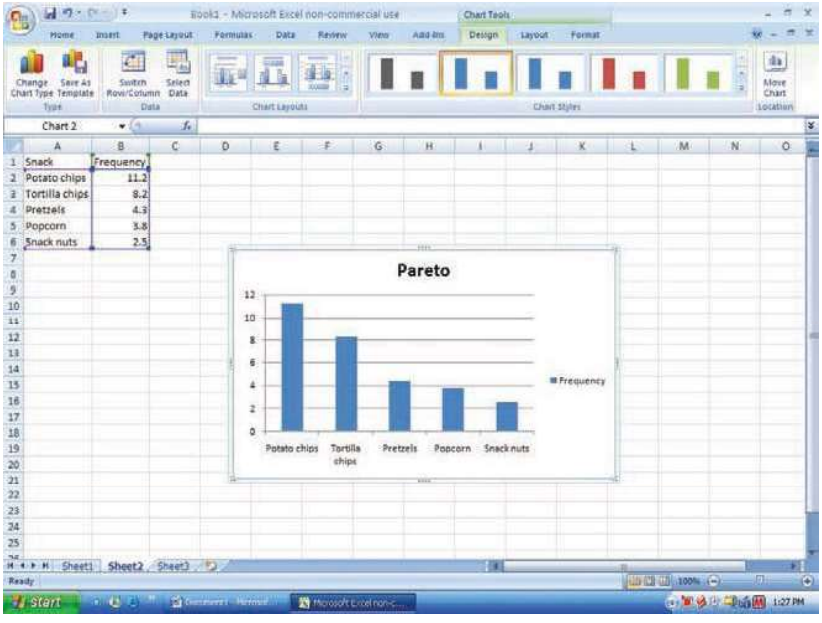
4. Click on any region of the chart. Then select Design from the Chart Tools tab on the toolbar.
5. Select Formulas from the chart Layouts tab on the toolbar.
6. To change the title of the chart, click on the current title of the chart.
7. When the text box containing the title is highlighted, click the mouse in the text box and change the title.



Constructing a Pareto Chart

To make a Pareto chart:

1. Enter the snack food categories from Example 2–11 into column A of a new worksheet.
2. Enter the corresponding frequencies in column B. The data should be entered in descending order according to frequency.
3. Highlight the data from columns A and B and select the Insert tab from the toolbar.
4. Select the Column Chart type.
5. To change the title of the chart, click on the current title of the chart.
6. When the text box containing the title is highlighted, click the mouse in the text box and change the title.



Constructing a Time Series Chart

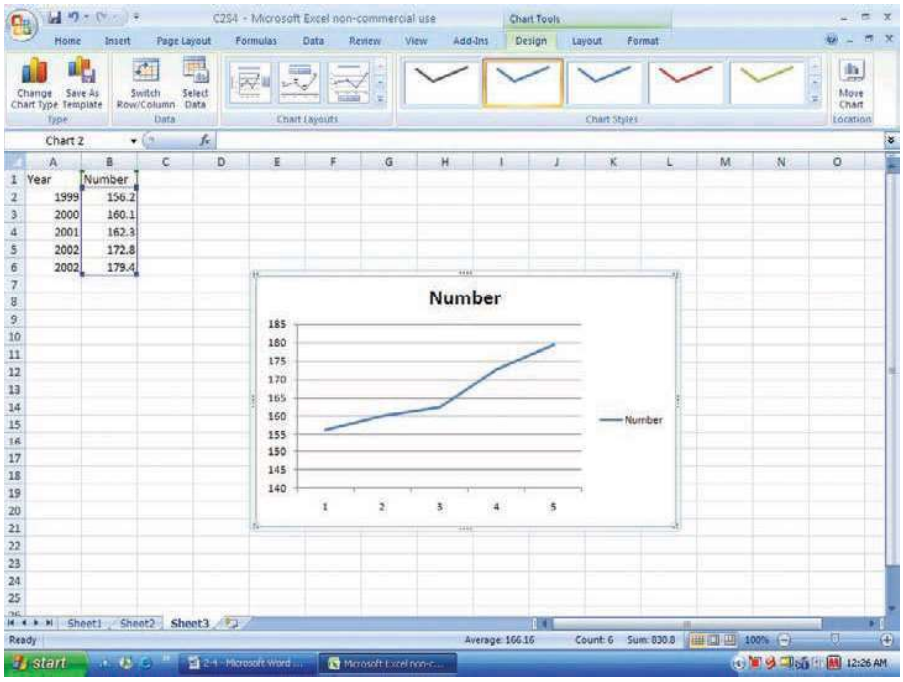
Example

Year	1999	2000	2001	2002	2003
Vehicles*	156.2	160.1	162.3	172.8	179.4

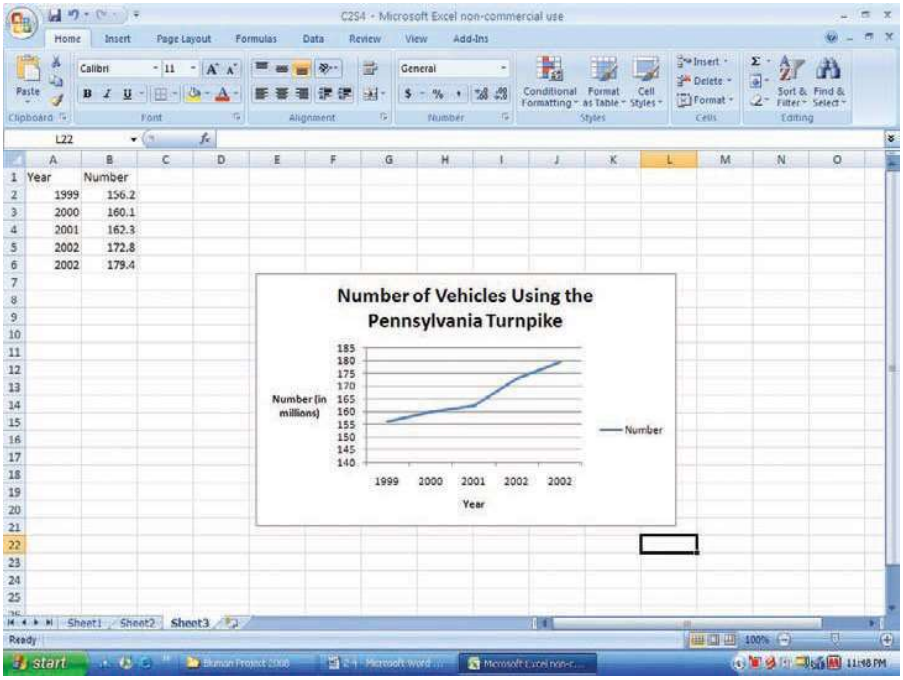
*Vehicles (in millions) that used the Pennsylvania Turnpike.
Source: *Tribune Review*.

To make a time series chart:

1. Enter the years 1999 through 2003 from the example in column A of a new worksheet.
2. Enter the corresponding frequencies in column B.
3. Highlight the data from column B and select the Insert tab from the toolbar.
4. Select the Line chart type.



5. Right-click the mouse on any region of the graph.
6. Select the Select Data option.
7. Select Edit from the Horizontal Axis Labels and highlight the years from column A, then click [OK].
8. Click [OK] on the Select Data Source box.
9. Create a title for your chart, such as Number of Vehicles Using the Pennsylvania Turnpike Between 1999 and 2003. Right-click the mouse on any region of the chart. Select the Chart Tools tab from the toolbar, then Layout.
10. Select Chart Title and highlight the current title to change the title.
11. Select Axis Titles to change the horizontal and vertical axis labels.



Summary

- When data are collected, the values are called raw data. Since very little knowledge can be obtained from raw data, they must be organized in some meaningful way. A frequency distribution using classes is the common method that is used. (2–1)
- Once a frequency distribution is constructed, graphs can be drawn to give a visual representation of the data. The most commonly used graphs in statistics are the histogram, frequency polygon, and ogive. (2–2)
- Other graphs such as the bar graph, Pareto chart, time series graph, and pie graph can also be used. Some of these graphs are frequently seen in newspapers, magazines, and various statistical reports. (2–3)
- Finally, a stem and leaf plot uses part of the data values as stems and part of the data values as leaves. This graph has the advantage of a frequency distribution and a histogram. (2–3)

Important Terms

bar graph 69
categorical frequency distribution 38
class 37
class boundaries 39
class midpoint 40
class width 39
cumulative frequency 54
cumulative frequency distribution 42
frequency 37
frequency distribution 37
frequency polygon 53
grouped frequency distribution 39
histogram 51
lower class limit 39
ogive 54
open-ended distribution 41
Pareto chart 70
pie graph 73
raw data 37
relative frequency graph 56
stem and leaf plot 80
time series graph 72
ungrouped frequency distribution 43
upper class limit 39

Important Formulas

Formula for the percentage of values in each class:

$$\% = \frac{f}{n} \cdot 100$$

where

f = **frequency of class**
n = **total number of values**

Formula for the range:

$$R = \text{highest value} - \text{lowest value}$$

Formula for the class width:

$$\text{Class width} = \text{upper boundary} - \text{lower boundary}$$

Formula for the class midpoint:

$$X_m = \frac{\text{lower boundary} + \text{upper boundary}}{2}$$

or

$$X_m = \frac{\text{lower limit} + \text{upper limit}}{2}$$

Formula for the degrees for each section of a pie graph:

$$\text{Degrees} = \frac{f}{n} \cdot 360°$$

Review Exercises

1. **How People Get Their News** The Brunswick Research Organization surveyed 50 randomly selected individuals and asked them the primary way they received the daily news. Their choices were via newspaper (N), television (T), radio (R), or Internet (I). Construct a categorical frequency distribution for the data and interpret the results. The data in this exercise will be used for Exercise 2 in this section. (2–1)

N	N	T	T	T	I	R	R	I	T
I	N	R	R	I	N	N	I	T	N
I	R	T	T	T	T	N	R	R	I
R	R	I	N	T	R	T	I	I	T
T	I	N	T	T	I	R	N	R	T

2. Construct a pie graph for the data in Exercise 1, and analyze the results. (2–3)

3. **Ball Sales** A sporting goods store kept a record of sales of five items for one randomly selected hour during a recent sale. Construct a frequency distribution for the data (B = baseballs, G = golf balls, T = tennis balls, S = soccer balls, F = footballs). (The data for this exercise will be used for Exercise 4 in this section.) (2–1)

F	B	B	B	G	T	F
G	G	F	S	G	T	
F	T	T	T	S	T	
F	S	S	G	S	B	

4. Draw a pie graph for the data in Exercise 3 showing the sales of each item, and analyze the results. (2–3)

5. **BUN Count** The blood urea nitrogen (BUN) count of 20 randomly selected patients is given here in milligrams per deciliter (mg/dl). Construct an ungrouped frequency distribution for the data. (The data for this exercise will be used for Exercise 6.) (2–1)

17	18	13	14
12	17	11	20
13	18	19	17
14	16	17	12
16	15	19	22

6. Construct a histogram, a frequency polygon, and an ogive for the data in Exercise 5 in this section, and analyze the results. (2–2)

7. The percentage (rounded to the nearest whole percent) of persons from each state completing 4 years or more of college is listed below. Organize the data into a grouped frequency distribution with 5 classes. (2–1)

Percentage of persons completing 4 years of college

23	25	24	34	22	24	27	37	33	24
26	23	38	24	24	17	28	23	30	25
30	22	33	24	28	36	24	19	25	31
34	31	27	24	29	28	21	25	26	15
26	22	27	21	25	28	24	21	25	26

Source: *New York Times Almanac.*

8. Using the data in Exercise 7, construct a histogram, a frequency polygon, and an ogive. (2–2)

9. **NFL Franchise Values** The data shown (in millions of dollars) are the values of the 30 National Football League franchises. Construct a frequency distribution for the data using 8 classes. (The data for

this exercise will be used for Exercises 10 and 12 in this section.) (2–1)

170	191	171	235	173	187	181	191
200	218	243	200	182	320	184	239
186	199	186	210	209	240	204	193
211	186	197	204	188	242		

Source: *Pittsburgh Post-Gazette.*

10. Construct a histogram, a frequency polygon, and an ogive for the data in Exercise 9 in this section, and analyze the results. (2–2)

11. Ages of the Vice Presidents at the Time of Their Death The ages of the Vice Presidents of the United States at the time of their death are listed below. Use the data to construct a frequency distribution, histogram, frequency polygon, and ogive, using relative frequencies. Use 6 classes. (2–1, 2–2)

90	83	80	73	70	51	68	79	70	71
72	74	67	54	81	66	62	63	68	57
66	96	78	55	60	66	57	71	60	85
76	98	77	88	78	81	64	66	77	70

Source: *New York Times Almanac.*

12. Construct a histogram, frequency polygon, and ogive by using relative frequencies for the data in Exercise 9 in this section. (2–2)

13. Activities While Driving A survey of 1200 drivers showed the percentage of respondents who did the following while driving. Construct a horizontal bar graph and a Pareto chart for the data. (2–3)

Drink beverage	80%
Talk on cell phone	73
Eat a meal	41
Experience road rage	23
Smoke	21

Source: Nationwide Mutual Insurance Company.

14. Air Quality The following data show the number of days the air quality for Atlanta, Georgia, was below the accepted standards. Draw a time series graph for the data. (2–3)

Year	2005	2006	2007	2008
Days	5	14	15	4

Source: U.S. Environmental Protection Agency.

15. Bank Failures The following data show the number of bank failures for recent years. Draw a time series graph and comment on the trend. (2–3)

Year	'01	'02	'03	'04	'05	'06	'07	'08	'09
Number	4	11	3	4	0	0	3	26	98

Source: Federal Deposit Insurance Corporation.

16. Public Debt The following data show the public debt in billions of dollars for recent years. Draw a time series graph for the data. (2–3)

Year	'03	'04	'05	'06	'07	'08	'09
Debt	6783.2	7379.1	7932.7	8507.0	9007.7	10,025.0	11,956.6

Source: U.S. Department of the Treasury.

17. Gold Production in Colombia The following data show the amount of gold production in thousands of troy ounces for Colombia for recent years. Draw a time series graph and comment on the trend. (2–3)

Year	'03	'04	'05	'06	'07	'08
Amount	656	701	976	1250	1270	1620

Source: U.S. Department of the Interior.

18. Spending of College Freshmen The average amounts spent by college freshmen for school items are shown. Construct a pie graph for the data. (2–3)

Electronics/computers	$728
Dorm items	344
Clothing	141
Shoes	72

Source: National Retail Federation.

19. Career Changes A survey asked if people would like to spend the rest of their careers with their present employers. The results are shown. Construct a pie graph for the data and analyze the results. (2–3)

Answer	Number of people
Yes	660
No	260
Undecided	80

20. Museum Visitors The number of visitors to the Railroad Museum during 24 randomly selected hours is shown here. Construct a stem and leaf plot for the data. (2–3)

67	62	38	73	34	43	72	35
53	55	58	63	47	42	51	62
32	29	47	62	29	38	36	41

21. Public Libraries The numbers of public libraries in operation for selected states are listed below. Organize the data with a stem and leaf plot. (2–3)

102	176	210	142	189	176	108	113	205
209	184	144	108	192	176			

Source: *World Almanac.*

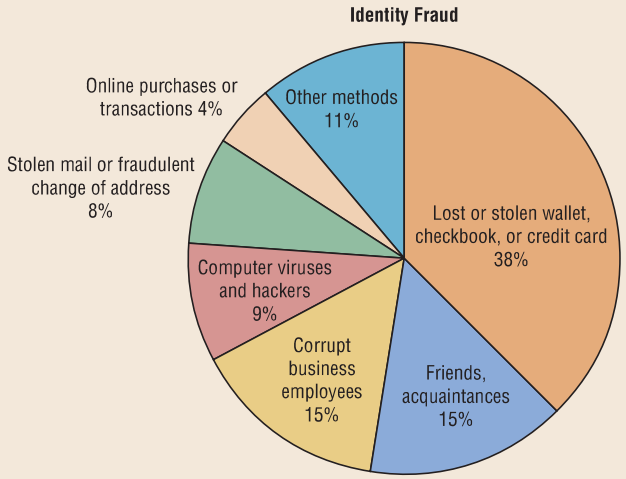
22. Job Aptitude Test A special aptitude test is given to job applicants. The data shown here represent the scores of 30 applicants. Construct a stem and leaf plot for the data and summarize the results. (2–3)

204	210	227	218	254
256	238	242	253	227
251	243	233	251	241
237	247	211	222	231
218	212	217	227	209
260	230	228	242	200

Statistics Today

How Your Identity Can Be Stolen–Revisited

Data presented in numerical form do not convey an easy-to-interpret conclusion; however, when data are presented in graphical form, readers can see the visual impact of the numbers. In the case of identity fraud, the reader can see that most of the identity frauds are due to lost or stolen wallets, checkbooks, or credit cards, and very few identity frauds are caused by online purchases or transactions.



Data Analysis

A Data Bank is found in Appendix D, or on the World Wide Web by following links from www.mhhe.com/math/stat/bluman

1. From the Data Bank located in Appendix D, choose one of the following variables: age, weight, cholesterol level, systolic pressure, IQ, or sodium level. Select at least 30 values. For these values, construct a grouped frequency distribution. Draw a histogram, frequency polygon, and ogive for the distribution. Describe briefly the shape of the distribution.
2. From the Data Bank, choose one of the following variables: educational level, smoking status, or exercise. Select at least 20 values. Construct an ungrouped frequency distribution for the data. For the distribution, draw a Pareto chart and describe briefly the nature of the chart.
3. From the Data Bank, select at least 30 subjects and construct a categorical distribution for their marital status. Draw a pie graph and describe briefly the findings.
4. Using the data from Data Set IV in Appendix D, construct a frequency distribution and draw a histogram. Describe briefly the shape of the distribution of the tallest buildings in New York City.
5. Using the data from Data Set XI in Appendix D, construct a frequency distribution and draw a frequency polygon. Describe briefly the shape of the distribution for the number of pages in statistics books.
6. Using the data from Data Set IX in Appendix D, divide the United States into four regions, as follows:

Northeast	CT ME MA NH NJ NY PA RI VT
Midwest	IL IN IA KS MI MN MS NE ND OH SD WI
South	AL AR DE DC FL GA KY LA MD NC OK SC TN TX VA WV
West	AK AZ CA CO HI ID MT NV NM OR UT WA WY

   Find the total population for each region, and draw a Pareto chart and a pie graph for the data. Analyze the results. Explain which chart might be a better representation for the data.
7. Using the data from Data Set I in Appendix D, make a stem and leaf plot for the record low temperatures in the United States. Describe the nature of the plot.

Chapter Quiz

Determine whether each statement is true or false. If the statement is false, explain why.

1. In the construction of a frequency distribution, it is a good idea to have overlapping class limits, such as 10–20, 20–30, 30–40. False
2. Histograms can be drawn by using vertical or horizontal bars. False
3. It is not important to keep the width of each class the same in a frequency distribution. False
4. Frequency distributions can aid the researcher in drawing charts and graphs. True
5. The type of graph used to represent data is determined by the type of data collected and by the researcher's purpose. True
6. In construction of a frequency polygon, the class limits are used for the *x* axis. False
7. Data collected over a period of time can be graphed by using a pie graph. False

Select the best answer.

8. What is another name for the ogive?
   a. Histogram
   b. Frequency polygon
   c. Cumulative frequency graph
   d. Pareto chart
9. What are the boundaries for 8.6–8.8?
   a. 8–9
   b. 8.5–8.9
   c. 8.55–8.85
   d. 8.65–8.75
10. What graph should be used to show the relationship between the parts and the whole?
    a. Histogram
    b. Pie graph
    c. Pareto chart
    d. Ogive
11. Except for rounding errors, relative frequencies should add up to what sum?
    a. 0
    b. 1
    c. 50
    d. 100

Complete these statements with the best answers.

12. The three types of frequency distributions are _____, _____, and _____. Categorical, ungrouped, grouped
13. In a frequency distribution, the number of classes should be between _____ and _____. 5, 20
14. Data such as blood types (A, B, AB, O) can be organized into a(n) _____ frequency distribution. Categorical
15. Data collected over a period of time can be graphed using a(n) _____ graph. Time series
16. A statistical device used in exploratory data analysis that is a combination of a frequency distribution and a histogram is called a(n) _____. Stem and leaf plot
17. On a Pareto chart, the frequencies should be represented on the _____ axis. Vertical or *y*
18. **Housing Arrangements** A questionnaire on housing arrangements showed this information obtained from 25 respondents. Construct a frequency distribution for the data (H = house, A = apartment, M = mobile home, C = condominium).

H	C	H	M	H	A	C	A	M
C	M	C	A	M	A	C	C	M
C	C	H	A	H	H	M		

19. Construct a pie graph for the data in Exercise 18.
20. **Items Purchased at a Convenience Store** When 30 randomly selected customers left a convenience store, each was asked the number of items he or she purchased. Construct an ungrouped frequency distribution for the data.

2	9	4	3	6
6	2	8	6	5
7	5	3	8	6
6	2	3	2	4
6	9	9	8	9
4	2	1	7	4

21. Construct a histogram, a frequency polygon, and an ogive for the data in Exercise 20.
22. **Murders in Selected Cities** For a recent year, the number of murders in 25 selected cities is shown. Construct a frequency distribution using 9 classes, and analyze the nature of the data in terms of shape, extreme values, etc. (The information in this exercise will be used for Exercise 23 in this section.)

248	348	74	514	597
270	71	226	41	39
366	73	241	46	34
149	68	73	63	65
109	598	278	69	27

Source: *Pittsburgh Tribune Review.*

23. Construct a histogram, frequency polygon, and ogive for the data in Exercise 22. Analyze the histogram.
24. **Recycled Trash** Construct a Pareto chart and a horizontal bar graph for the number of tons (in millions)

of trash recycled per year by Americans based on an Environmental Protection Agency study.

Type	Amount
Paper	320.0
Iron/steel	292.0
Aluminum	276.0
Yard waste	242.4
Glass	196.0
Plastics	41.6

Source: *USA TODAY.*

25. Identity Thefts The results of a survey of 84 people whose identities were stolen using various methods are shown. Draw a pie chart for the information.

Lost or stolen wallet, checkbook, or credit card	38
Retail purchases or telephone transactions	15
Stolen mail	9
Computer viruses or hackers	8
Phishing	4
Other	10
	84

Source: Javelin Strategy and Research.

26. Needless Deaths of Children *The New England Journal of Medicine* predicted the number of needless deaths due to childhood obesity. Draw a time series graph for the data.

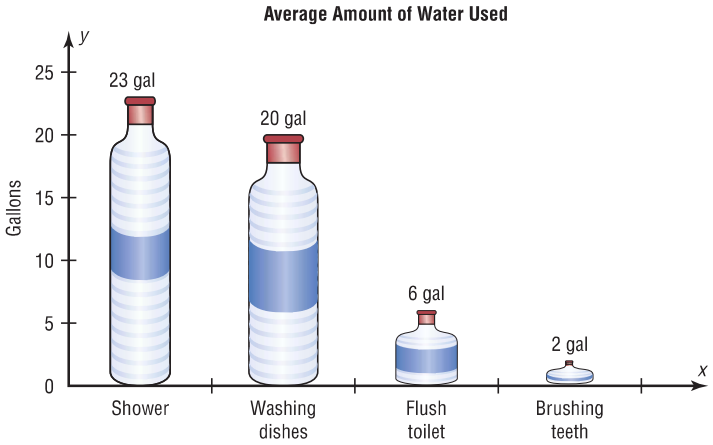
Year	2020	2025	2030	2035
Deaths	130	550	1500	3700

27. Museum Visitors The number of visitors to the Historic Museum for 25 randomly selected hours is shown. Construct a stem and leaf plot for the data.

15	53	48	19	38
86	63	98	79	38
62	89	67	39	26
28	35	54	88	76
31	47	53	41	68

Critical Thinking Challenges

1. Water Usage The graph shows the average number of gallons of water a person uses for various activities. Can you see anything misleading about the way the graph is drawn?



2. The Great Lakes Shown are various statistics about the Great Lakes. Using appropriate graphs (your choice) and summary statements, write a report analyzing the data.

	Superior	Michigan	Huron	Erie	Ontario
Length (miles)	350	307	206	241	193
Breadth (miles)	160	118	183	57	53
Depth (feet)	1,330	923	750	210	802
Volume (cubic miles)	2,900	1,180	850	116	393
Area (square miles)	31,700	22,300	23,000	9,910	7,550
Shoreline (U.S., miles)	863	1,400	580	431	300

Source: *The World Almanac and Book of Facts.*

3. Teacher Strikes In Pennsylvania there were more teacher strikes in 2004 than there were in all other states combined. Because of the disruptions, state legislators want to pass a bill outlawing teacher strikes and submitting contract disputes to binding arbitration. The graph shows the number of teacher strikes in Pennsylvania for the school years 1992 to 2004. Use the graph to answer these questions.

a. In what year did the largest number of strikes occur? How many were there?

b. In what year did the smallest number of teacher strikes occur? How many were there?

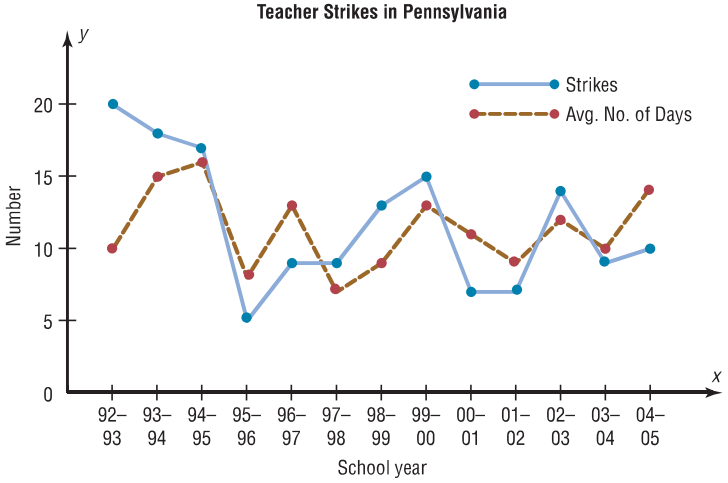
c. In what year was the average duration of the strikes the longest? What was it?

d. In what year was the average duration of the strikes the shortest? What was it?

e. In what year was the number of teacher strikes the same as the average duration of the strikes?

f. Find the difference in the number of strikes for the school years 1992–1993 and 2004–2005.

g. Do you think teacher strikes should be outlawed? Justify your conclusions.



Source: Pennsylvania School Boards Associations.

Data Projects

Where appropriate, use MINITAB, the TI-83 Plus, the TI-84 Plus, Excel, or a computer program of your choice to complete the following exercises.

1. Business and Finance Consider the 30 stocks listed as the Dow Jones Industrials. For each, find their earnings per share. Randomly select 30 stocks traded on the NASDAQ. For each, find their earnings per share. Create a frequency table with 5 categories for each data set. Sketch a histogram for each. How do the two data sets compare?

2. Sports and Leisure Use systematic sampling to create a sample of 25 National League and 25 American League baseball players from the most recently completed season. Find the number of home runs for each player. Create a frequency table with 5 categories for each data set. Sketch a histogram for each. How do the two leagues compare?

3. **Technology** Randomly select 50 songs from your music player or music organization program. Find the length (in seconds) for each song. Use these data to create a frequency table with 6 categories. Sketch a frequency polygon for the frequency table. Is the shape of the distribution of times uniform, skewed, or bell-shaped? Also note the genre of each song. Create a Pareto chart showing the frequencies of the various categories. Finally, note the year each song was released. Create a pie chart organized by decade to show the percentage of songs from various time periods.

4. **Health and Wellness** Use information from the Red Cross to create a pie chart depicting the percentages of Americans with various blood types. Also find information about blood donations and the percentage of each type donated. How do the charts compare? Why is the collection of type O blood so important?

5. **Politics and Economics** Consider the U.S. Electoral College System. For each of the 50 states, determine the number of delegates received. Create a frequency table with 8 classes. Is this distribution uniform, skewed, or bell-shaped?

6. **Your Class** Have each person in class take his or her pulse and determine the heart rate (beats in one minute). Use the data to create a frequency table with 6 classes. Then have everyone in the class do 25 jumping jacks and immediately take the pulse again after the activity. Create a frequency table for those data as well. Compare the two results. Are they similarly distributed? How does the range of scores compare?

Answers to Applying the Concepts

Section 2–1 Ages of Presidents at Inauguration

1. The data were obtained from the population of all Presidents at the time this text was written.
2. The oldest inauguration age was 69 years old.
3. The youngest inauguration age was 42 years old.
4. Answers will vary. One possible answer is

Age at inauguration	Frequency
42–45	2
46–49	7
50–53	8
54–57	16
58–61	5
62–65	4
66–69	2

5. Answers will vary. For the frequency distribution given in Question 4, there is a peak for the 54–57 bin.
6. Answers will vary. This frequency distribution shows no outliers. However, if we had split our frequency into 14 bins instead of 7, then the ages 42, 43, 68, and 69 might appear as outliers.
7. Answers will vary. The data appear to be unimodal and fairly symmetric, centering on 55 years of age.

Section 2–2 Selling Real Estate

1. A histogram of the data gives price ranges and the counts of homes in each price range. We can also talk about how the data are distributed by looking at a histogram.
2. A frequency polygon shows increases or decreases in the number of home prices around values.
3. A cumulative frequency polygon shows the number of homes sold at or below a given price.
4. The house that sold for $321,550 is an extreme value in this data set.
5. Answers will vary. One possible answer is that the histogram displays the outlier well since there is a gap in the prices of the homes sold.
6. The distribution of the data is skewed to the right.

Section 2–3 Leading Cause of Death

1. The variables in the graph are the year, cause of death, and rate of death per 100,000 men.
2. The cause of death is qualitative, while the year and death rates are quantitative.
3. Year is a discrete variable, and death rate is continuous. Since cause of death is qualitative, it is neither discrete nor continuous.
4. A line graph was used to display the data.
5. No, a Pareto chart could not be used to display the data, since we can only have one quantitative variable and one categorical variable in a Pareto chart.
6. We cannot use a pie chart for the same reasons as given for the Pareto chart.
7. A Pareto chart is typically used to show a categorical variable listed from the highest-frequency category to the category with the lowest frequency.
8. A time series chart is used to see trends in the data. It can also be used for forecasting and predicting.

This page intentionally left blank

CHAPTER 3

Data Description

Objectives

After completing this chapter, you should be able to

1. Summarize data, using measures of central tendency, such as the mean, median, mode, and midrange.
2. Describe data, using measures of variation, such as the range, variance, and standard deviation.
3. Identify the position of a data value in a data set, using various measures of position, such as percentiles, deciles, and quartiles.
4. Use the techniques of exploratory data analysis, including boxplots and five-number summaries, to discover various aspects of data.

Outline