88	88	110	88	80	69	102	78	70	55
79	85	80	100	60	90	77	55	75	55
54	60	75	64	105	56	71	70	65	72

Source: *New York Times Almanac.*

11. GRE Scores at Top-Ranked Engineering Schools The average quantitative GRE scores for the top 30 graduate schools of engineering are listed. Construct a grouped frequency distribution and a cumulative frequency distribution with 5 classes.

767	770	761	760	771	768	776	771	756	770
763	760	747	766	754	771	771	778	766	762
780	750	746	764	769	759	757	753	758	746

Source: *U.S. News & World Report, Best Graduate Schools.*

12. Airline Passengers The number of passengers (in thousands) for the leading U.S. passenger airlines in 2004 is indicated below. Use the data to construct a grouped frequency distribution and a cumulative frequency distribution with a reasonable number of classes, and comment on the shape of the distribution.

91,570	86,755	81,066	70,786	55,373	42,400
40,551	21,119	16,280	14,869	13,659	13,417
13,170	12,632	11,731	10,420	10,024	9,122
7,041	6,954	6,406	6,362	5,930	5,585
5,427					

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

13. Ages of Declaration of Independence Signers The ages of the signers of the Declaration of Independence are shown. (Age is approximate since only the birth year appeared in the source, and one has been omitted since his birth year is unknown.) Construct a grouped frequency distribution and a cumulative frequency distribution for the data using 7 classes. (The data in this exercise will be used in Exercise 23 in Section 3–1.)

41	54	47	40	39	35	50	37	49	42	70	32
44	52	39	50	40	30	34	69	39	45	33	42
44	63	60	27	42	34	50	42	52	38	36	45
35	43	48	46	31	27	55	63	46	33	60	62
35	46	45	34	53	50	50					

Source: *The Universal Almanac.*

14. Unclaimed Expired Prizes The number of unclaimed expired prizes (in millions of dollars) for lottery tickets bought in a sample of states as shown. Construct a frequency distribution for the data using 5 classes. (The data in this exercise will be used for Exercise 22 in Section 3–1.)

28.5	51.7	19	5
2	1.2	14	14.6
0.8	11.6	3.5	30.1
1.7	1.3	13	14

15. Presidential Vetoes The number of total vetoes exercised by the past 20 Presidents is listed below. Use the data to construct a grouped frequency distribution and a cumulative frequency distribution with 5 classes. What is challenging about this set of data?

44	39	37	21	31	170	44	635	30	78
42	6	250	43	10	82	50	181	66	37

16. Salaries of College Coaches The data are the salaries (in hundred thousands of dollars) of a sample of 30 colleges and university coaches in the United States. Construct a frequency distribution for the data using 8 classes. (The data in this exercise will be used for Exercise 11 in Section 2–2.)

164	225	225	140	188
210	238	146	201	544
550	188	415	261	164
478	684	330	307	435
857	183	381	275	578
450	385	297	390	515

17. NFL Payrolls The data show the NFL team payrolls (in millions of dollars) for a specific year. Construct a frequency distribution for the payroll using 7 classes. (The data in this exercise will be used in Exercise 17 in Section 3–2.)

99	105	106	102
102	93	109	106
77	91	103	118
97	100	107	103
94	109	100	98
84	92	98	110
94	104	98	123
102	99	100	107

Source: *NFL.*

18. State Gasoline Tax The state gas tax in cents per gallon for 25 states is given below. Construct a grouped frequency distribution and a cumulative frequency distribution with 5 classes.

7.5	16	23.5	17	22
21.5	19	20	27.1	20
22	20.7	17	28	20
23	18.5	25.3	24	31
14.5	25.9	18	30	31.5

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

Extending the Concepts

19. JFK Assassination A researcher conducted a survey asking people if they believed more than one person was involved in the assassination of John F. Kennedy. The results were as follows: 73% said yes, 19% said no, and 9% had no opinion. Is there anything suspicious about the results?

Technology *Step by Step*

MINITAB Step by Step

Make a Categorical Frequency Table (Qualitative or Discrete Data)

1. Type in all the blood types from Example 2–1 down C1 of the worksheet.

   A B B AB O O O B AB B B B O A O A O O O AB AB A O B A

2. Click above row 1 and name the column **BloodType.**
3. Select **Stat>Tables>Tally Individual Values.**

   The cursor should be blinking in the Variables dialog box. If not, click inside the dialog box.
4. Double-click C1 in the Variables list.
5. Check the boxes for the statistics: Counts, Percents, and Cumulative percents.
6. Click [OK]. The results will be displayed in the Session Window as shown.

Tally for Discrete Variables: BloodType

BloodType	Count	Percent	CumPct
A	5	20.00	20.00
AB	4	16.00	36.00
B	7	28.00	64.00
O	9	36.00	100.00
N=	25		

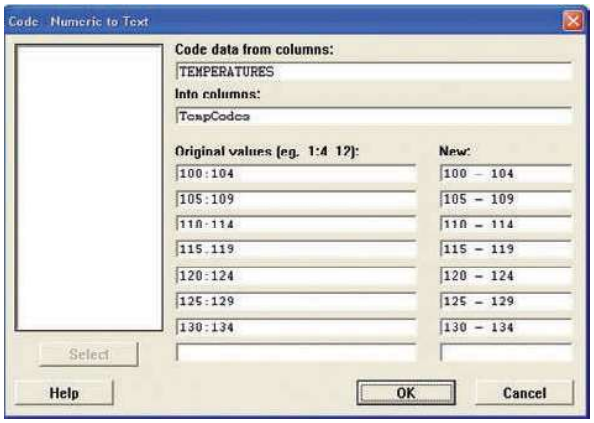
Make a Grouped Frequency Distribution (Quantitative Variable)

1. Select **File>New>New Worksheet.** A new worksheet will be added to the project.
2. Type the data used in Example 2–2 into C1. Name the column **TEMPERATURES.**
3. Use the instructions in the textbook to determine the class limits.

   In the next step you will create a new column of data, converting the numeric variable to text categories that can be tallied.
4. Select **Data>Code>Numeric to Text.**
   a) The cursor should be blinking in Code data from columns. If not, click inside the box, then double-click C1 Temperatures in the list. Only quantitative variables will be shown in this list.
   b) Click in the Into columns: then type the name of the new column, **TempCodes.**
   c) Press [Tab] to move to the next dialog box.
   d) Type in the first interval **100:104.**

      Use a colon to indicate the interval from 100 to 104 with no spaces before or after the colon.
   e) Press [Tab] to move to the New: column, and type the text category **100–104.**
   f) Continue to tab to each dialog box, typing the interval and then the category until the last category has been entered.

The dialog box should look like the one shown.



5. Click [OK]. In the worksheet, a new column of data will be created in the first empty column, C2. This new variable will contain the category for each value in C1. The column C2-T contains alphanumeric data.
6. Click **Stat>Tables>Tally Individual Values,** then double-click TempCodes in the Variables list.
   a) Check the boxes for the desired statistics, such as Counts, Percents, and Cumulative percents.
   b) Click [OK].

   The table will be displayed in the Session Window. Eighteen states have high temperatures between 110 and 114°F. Eighty-two percent of the states have record high temperatures less than or equal to 119°F.

Tally for Discrete Variables: TempCodes

TempCodes	Count	Percent	CumPct
100–104	2	4.00	4.00
105–109	8	16.00	20.00
110–114	18	36.00	56.00
115–119	13	26.00	82.00
120–124	7	14.00	96.00
125–129	1	2.00	98.00
130–134	1	2.00	100.00
N=	50		

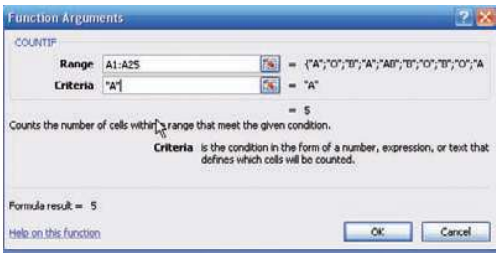
7. Click **File>Save Project As** . . . , and type the name of the project file, **Ch2-2.** This will save the two worksheets and the Session Window.

Excel Step by Step

Categorical Frequency Table (Qualitative or Discrete Data)

1. In an open workbook select cell A1 and type in all the blood types from Example 2–1 down column A.
2. Type in the variable name **Blood Type** in cell B1.
3. Select cell B2 and type in the four different blood types down the column.
4. Type in the name **Count** in cell C1.
5. Select cell C2. From the toolbar, select the Formulas tab on the toolbar.
6. Select the Insert Function icon  , then select the Statistical category in the Insert Function dialog box.
7. Select the Countif function from the function name list.

8. In the dialog box, type **A1:A25** in the **Range** box. Type in the blood type "A" in quotes in the **Criteria** box. The count or frequency of the number of data corresponding to the blood type should appear below the input. Repeat for the remaining blood types.
9. After all the data have been counted, select cell C6 in the worksheet.
10. From the toolbar select Formulas, then AutoSum and type in C2:C5 to insert the total frequency into cell C6.



After entering data or a heading into a worksheet, you can change the width of a column to fit the input. To automatically change the width of a column to fit the data:

1. Select the column or columns that you want to change.
2. On the Home tab, in the Cells group, select Format.
3. Under Cell Size, click Autofit Column Width.

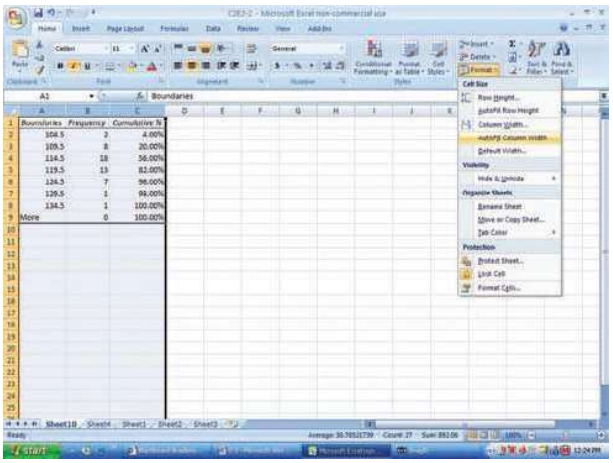
Making a Grouped Frequency Distribution (Quantitative Data)

1. Press **[Ctrl]-N** for a new workbook.
2. Enter the raw data from Example 2–2 in column A, one number per cell.
3. Enter the upper class boundaries in column B.
4. From the toolbar select the Data tab, then click Data Analysis.
5. In the Analysis Tools, select Histogram and click [OK].
6. In the Histogram dialog box, type **A1:A50** in the Input Range box and type **B1:B7** in the Bin Range box.
7. Select New Worksheet Ply, and check the Cumulative Percentage option. Click [OK].
8. You can change the label for the column containing the upper class boundaries and expand the width of the columns automatically after relabeling:

   Select the Home tab from the toolbar.

   Highlight the columns that you want to change.

   Select Format, then AutoFit Column Width.



Note: By leaving the Chart Output unchecked, a new worksheet will display the table only.

2–2

Histograms, Frequency Polygons, and Ogives

Objective 2

Represent data in frequency distributions graphically using histograms, frequency polygons, and ogives.

After you have organized the data into a frequency distribution, you can present them in graphical form. The purpose of graphs in statistics is to convey the data to the viewers in pictorial form. It is easier for most people to comprehend the meaning of data presented graphically than data presented numerically in tables or frequency distributions. This is especially true if the users have little or no statistical knowledge.

Statistical graphs can be used to describe the data set or to analyze it. Graphs are also useful in getting the audience's attention in a publication or a speaking presentation. They can be used to discuss an issue, reinforce a critical point, or summarize a data set. They can also be used to discover a trend or pattern in a situation over a period of time.

The three most commonly used graphs in research are

1. The histogram.
2. The frequency polygon.
3. The cumulative frequency graph, or ogive (pronounced o-jive).

An example of each type of graph is shown in Figure 2–1. The data for each graph are the distribution of the miles that 20 randomly selected runners ran during a given week.

Historical Note

Karl Pearson introduced the histogram in 1891. He used it to show time concepts of various reigns of Prime Ministers.

The Histogram

The **histogram** is a graph that displays the data by using contiguous vertical bars (unless the frequency of a class is 0) of various heights to represent the frequencies of the classes.

Example 2–4

Record High Temperatures

Construct a histogram to represent the data shown for the record high temperatures for each of the 50 states (see Example 2–2).

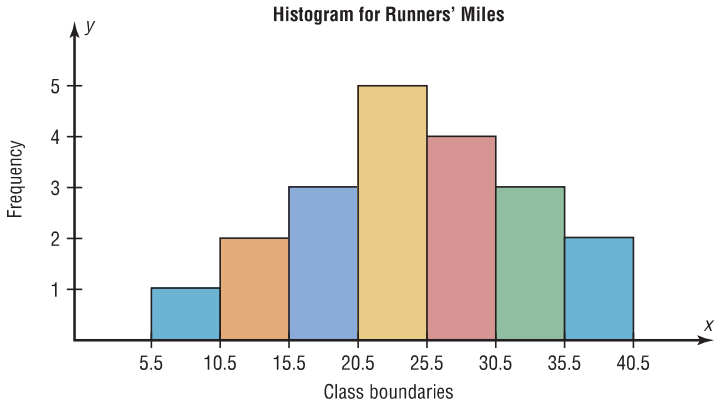
Class boundaries	Frequency
99.5–104.5	2
104.5–109.5	8
109.5–114.5	18
114.5–119.5	13
119.5–124.5	7
124.5–129.5	1
129.5–134.5	1

Solution

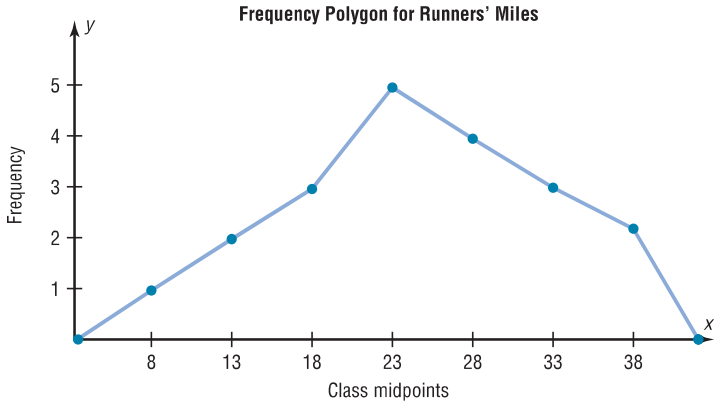
Step 1 Draw and label the x and y axes. The x axis is always the horizontal axis, and the y axis is always the vertical axis.

Figure 2–1

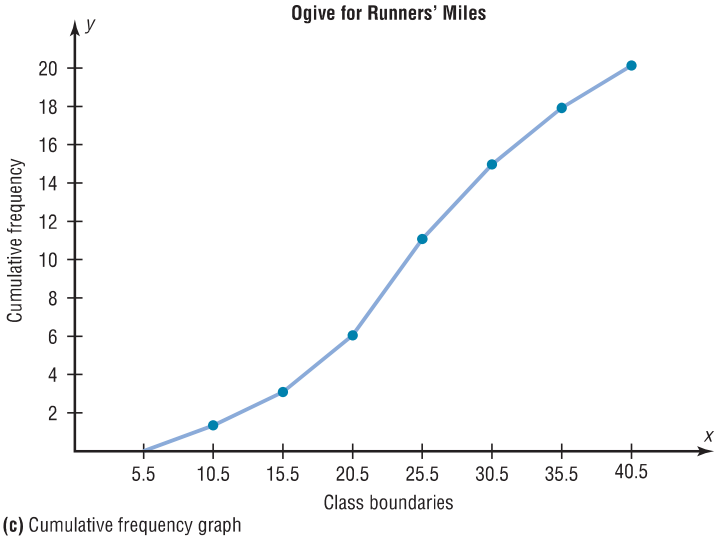
Examples of Commonly Used Graphs



(a) Histogram

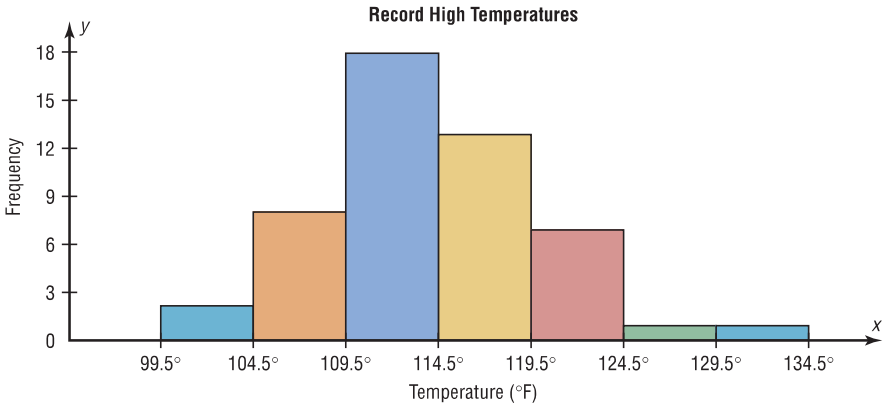


(b) Frequency polygon



(c) Cumulative frequency graph

Figure 2–2
Histogram for Example 2–4



Historical Note

Graphs originated when ancient astronomers drew the position of the stars in the heavens. Roman surveyors also used coordinates to locate landmarks on their maps.

The development of statistical graphs can be traced to William Playfair (1748–1819), an engineer and drafter who used graphs to present economic data pictorially.

Step 2 Represent the frequency on the *y* axis and the class boundaries on the *x* axis.

Step 3 Using the frequencies as the heights, draw vertical bars for each class. See Figure 2–2.

As the histogram shows, the class with the greatest number of data values (18) is 109.5–114.5, followed by 13 for 114.5–119.5. The graph also has one peak with the data clustering around it.

The Frequency Polygon

Another way to represent the same data set is by using a frequency polygon.

The **frequency polygon** is a graph that displays the data by using lines that connect points plotted for the frequencies at the midpoints of the classes. The frequencies are represented by the heights of the points.

Example 2–5 shows the procedure for constructing a frequency polygon.

Example 2–5

Record High Temperatures

Using the frequency distribution given in Example 2–4, construct a frequency polygon.

Solution

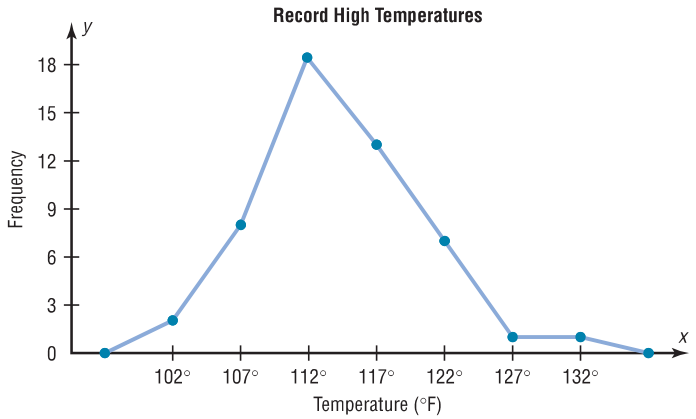
Step 1 Find the midpoints of each class. Recall that midpoints are found by adding the upper and lower boundaries and dividing by 2:

$$\frac{99.5 + 104.5}{2} = 102 \qquad \frac{104.5 + 109.5}{2} = 107$$

and so on. The midpoints are

Class boundaries	Midpoints	Frequency
99.5–104.5	102	2
104.5–109.5	107	8
109.5–114.5	112	18
114.5–119.5	117	13
119.5–124.5	122	7
124.5–129.5	127	1
129.5–134.5	132	1

Figure 2–3
Frequency Polygon for Example 2–5



Step 2 Draw the *x* and *y* axes. Label the *x* axis with the midpoint of each class, and then use a suitable scale on the *y* axis for the frequencies.

Step 3 Using the midpoints for the *x* values and the frequencies as the *y* values, plot the points.

Step 4 Connect adjacent points with line segments. Draw a line back to the *x* axis at the beginning and end of the graph, at the same distance that the previous and next midpoints would be located, as shown in Figure 2–3.

The frequency polygon and the histogram are two different ways to represent the same data set. The choice of which one to use is left to the discretion of the researcher.

The Ogive

The third type of graph that can be used represents the cumulative frequencies for the classes. This type of graph is called the *cumulative frequency graph,* or *ogive.* The **cumulative frequency** is the sum of the frequencies accumulated up to the upper boundary of a class in the distribution.

The **ogive** is a graph that represents the cumulative frequencies for the classes in a frequency distribution.

Example 2–6 shows the procedure for constructing an ogive.

Example 2–6

Record High Temperatures

Construct an ogive for the frequency distribution described in Example 2–4.

Solution

Step 1 Find the cumulative frequency for each class.

	Cumulative frequency
Less than 99.5	0
Less than 104.5	2
Less than 109.5	10
Less than 114.5	28
Less than 119.5	41
Less than 124.5	48
Less than 129.5	49
Less than 134.5	50

Figure 2–4
Plotting the Cumulative Frequency for Example 2–6

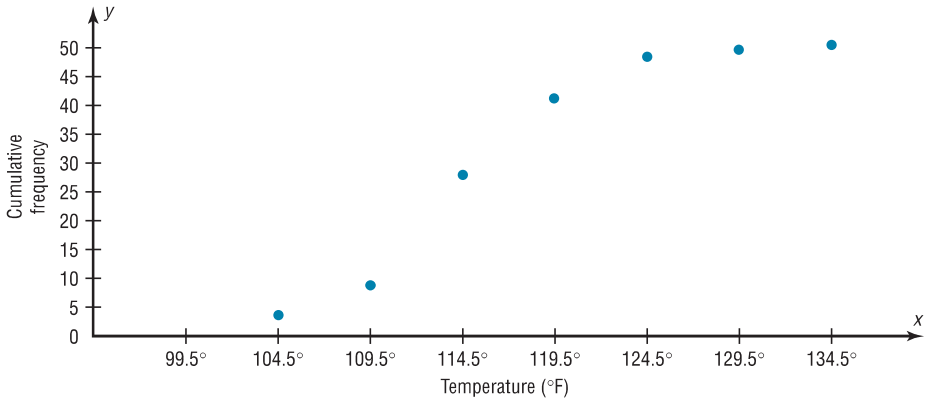
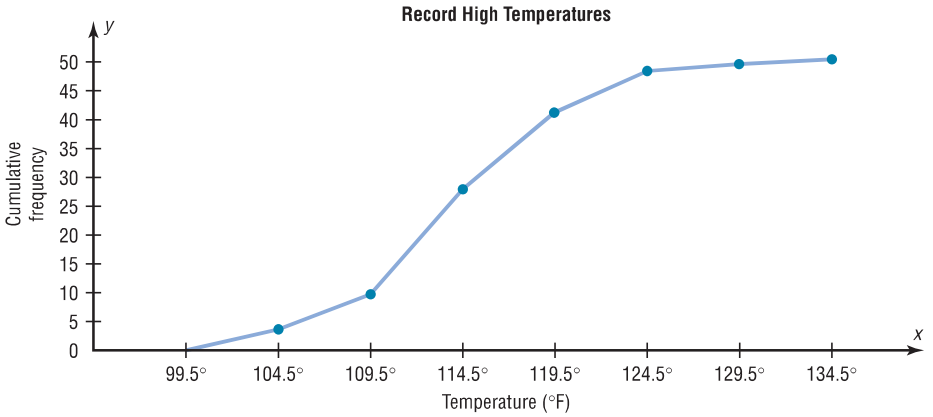


Figure 2–5
Ogive for Example 2–6



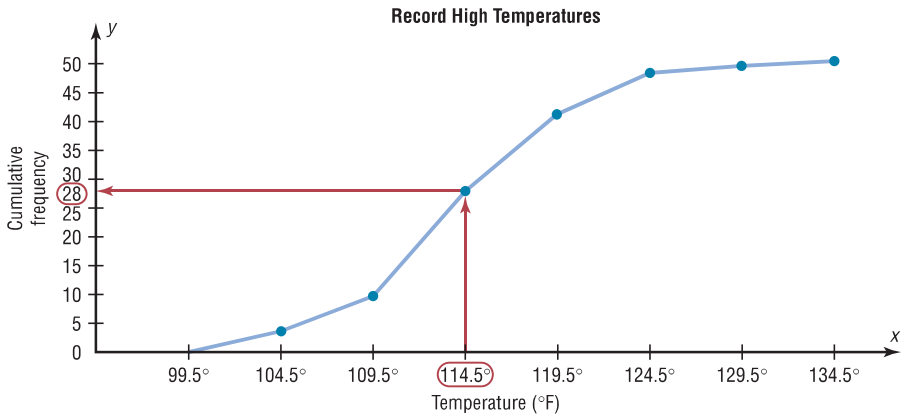
Step 2 Draw the x and y axes. Label the x axis with the class boundaries. Use an appropriate scale for the y axis to represent the cumulative frequencies. (Depending on the numbers in the cumulative frequency columns, scales such as 0, 1, 2, 3, . . . , or 5, 10, 15, 20, . . . , or 1000, 2000, 3000, . . . can be used. Do *not* label the y axis with the numbers in the cumulative frequency column.) In this example, a scale of 0, 5, 10, 15, . . . will be used.

Step 3 Plot the cumulative frequency at each upper class boundary, as shown in Figure 2–4. Upper boundaries are used since the cumulative frequencies represent the number of data values accumulated up to the upper boundary of each class.

Step 4 Starting with the first upper class boundary, 104.5, connect adjacent points with line segments, as shown in Figure 2–5. Then extend the graph to the first lower class boundary, 99.5, on the x axis.

Cumulative frequency graphs are used to visually represent how many values are below a certain upper class boundary. For example, to find out how many record high temperatures are less than 114.5°F, locate 114.5°F on the x axis, draw a vertical line up until it intersects the graph, and then draw a horizontal line at that point to the y axis. The y axis value is 28, as shown in Figure 2–6.

Figure 2–6
Finding a Specific Cumulative Frequency



The steps for drawing these three types of graphs are shown in the following Procedure Table.

Unusual Stat
Twenty-two percent of Americans sleep 6 hours a day or fewer.

Procedure Table

Constructing Statistical Graphs

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

Step 2 Choose a suitable scale for the frequencies or cumulative frequencies, and label it on the *y* axis.

Step 3 Represent the class boundaries for the histogram or ogive, or the midpoint for the frequency polygon, on the *x* axis.

Step 4 Plot the points and then draw the bars or lines.

Relative Frequency Graphs

The histogram, the frequency polygon, and the ogive shown previously were constructed by using frequencies in terms of the raw data. These distributions can be converted to distributions using *proportions* instead of raw data as frequencies. These types of graphs are called **relative frequency graphs.**

Graphs of relative frequencies instead of frequencies are used when the proportion of data values that fall into a given class is more important than the actual number of data values that fall into that class. For example, if you wanted to compare the age distribution of adults in Philadelphia, Pennsylvania, with the age distribution of adults of Erie, Pennsylvania, you would use relative frequency distributions. The reason is that since the population of Philadelphia is 1,478,002 and the population of Erie is 105,270, the bars using the actual data values for Philadelphia would be much taller than those for the same classes for Erie.

To convert a frequency into a proportion or relative frequency, divide the frequency for each class by the total of the frequencies. The sum of the relative frequencies will always be 1. These graphs are similar to the ones that use raw data as frequencies, but the values on the *y* axis are in terms of proportions. Example 2–7 shows the three types of relative frequency graphs.

Example 2–7 Miles Run per Week

Construct a histogram, frequency polygon, and ogive using relative frequencies for the distribution (shown here) of the miles that 20 randomly selected runners ran during a given week.

Class boundaries	Frequency
5.5–10.5	1
10.5–15.5	2
15.5–20.5	3
20.5–25.5	5
25.5–30.5	4
30.5–35.5	3
35.5–40.5	2
	20

Solution

Step 1 Convert each frequency to a proportion or relative frequency by dividing the frequency for each class by the total number of observations.

For class 5.5–10.5, the relative frequency is $\frac{1}{20} = 0.05$; for class 10.5–15.5, the relative frequency is $\frac{2}{20} = 0.10$; for class 15.5–20.5, the relative frequency is $\frac{3}{20} = 0.15$; and so on.

Place these values in the column labeled Relative frequency.

Class boundaries	Midpoints	Relative frequency
5.5–10.5	8	0.05
10.5–15.5	13	0.10
15.5–20.5	18	0.15
20.5–25.5	23	0.25
25.5–30.5	28	0.20
30.5–35.5	33	0.15
35.5–40.5	38	0.10
		1.00

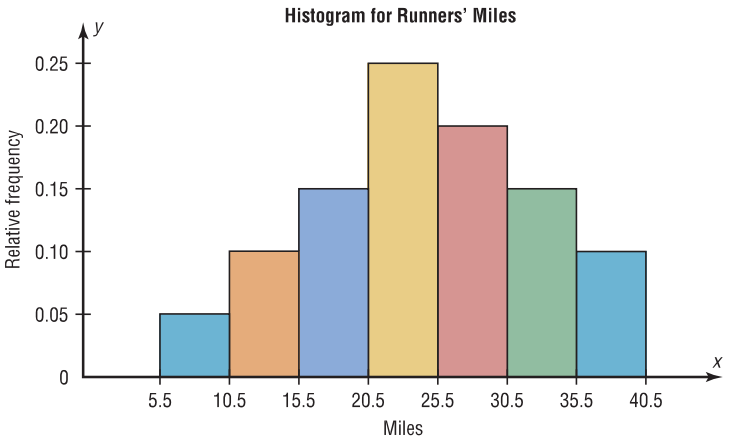
Step 2 Find the cumulative relative frequencies. To do this, add the frequency in each class to the total frequency of the preceding class. In this case, $0 + 0.05 = 0.05$, $0.05 + 0.10 = 0.15$, $0.15 + 0.15 = 0.30$, $0.30 + 0.25 = 0.55$, etc. Place these values in the column labeled Cumulative relative frequency.

An alternative method would be to find the cumulative frequencies and then convert each one to a relative frequency.

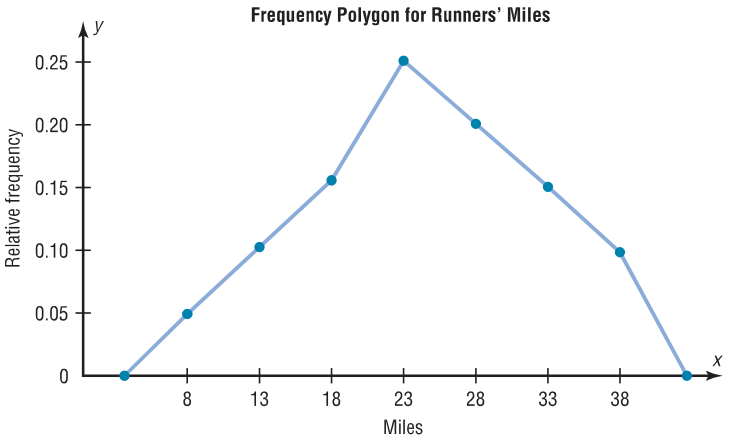
	Cumulative frequency	Cumulative relative frequency
Less than 5.5	0	0.00
Less than 10.5	1	0.05
Less than 15.5	3	0.15
Less than 20.5	6	0.30
Less than 25.5	11	0.55
Less than 30.5	15	0.75
Less than 35.5	18	0.90
Less than 40.5	20	1.00

Step 3 Draw each graph as shown in Figure 2–7. For the histogram and ogive, use the class boundaries along the x axis. For the frequency polygon, use the midpoints on the x axis. The scale on the y axis uses proportions.

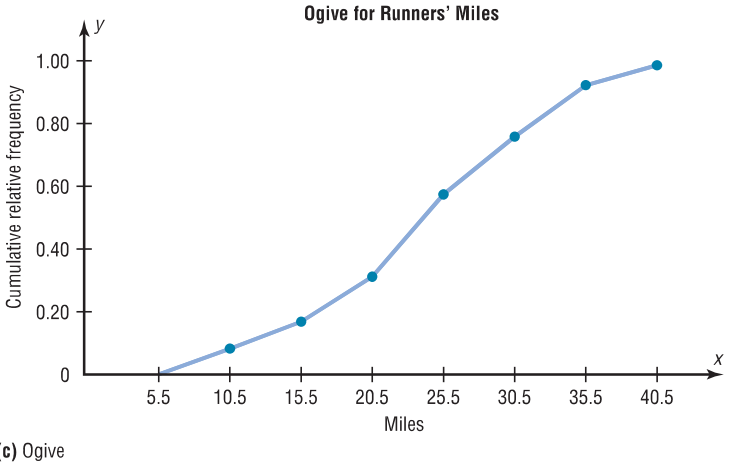
Figure 2–7
Graphs for Example 2–7



(a) Histogram



(b) Frequency polygon



(c) Ogive

Distribution Shapes

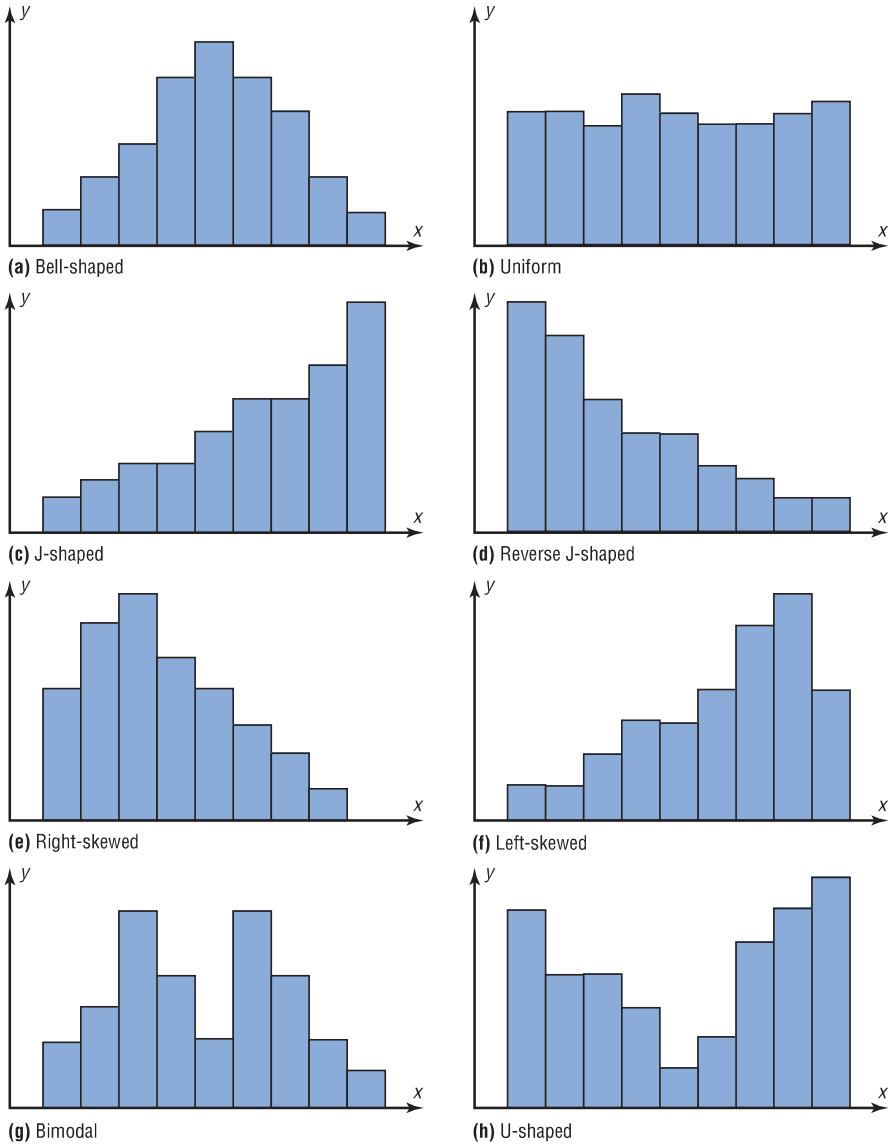
When one is describing data, it is important to be able to recognize the shapes of the distribution values. In later chapters you will see that the shape of a distribution also determines the appropriate statistical methods used to analyze the data.

A distribution can have many shapes, and one method of analyzing a distribution is to draw a histogram or frequency polygon for the distribution. Several of the most common shapes are shown in Figure 2–8: *the bell-shaped or mound-shaped, the uniform-shaped, the J-shaped, the reverse J-shaped, the positively or right-skewed shape, the negatively or left-skewed shape, the bimodal-shaped, and the U-shaped.*

Distributions are most often not perfectly shaped, so it is not necessary to have an exact shape but rather to identify an overall pattern.

A *bell-shaped distribution* shown in Figure 2–8(a) has a single peak and tapers off at either end. It is approximately symmetric; i.e., it is roughly the same on both sides of a line running through the center.

Figure 2–8
Distribution Shapes



A *uniform distribution* is basically flat or rectangular. See Figure 2–8(b).

A *J-shaped distribution* is shown in Figure 2–8(c), and it has a few data values on the left side and increases as one moves to the right. A *reverse J-shaped distribution* is the opposite of the J-shaped distribution. See Figure 2–8(d).

When the peak of a distribution is to the left and the data values taper off to the right, a distribution is said to be *positively or right-skewed.* See Figure 2–8(e). When the data values are clustered to the right and taper off to the left, a distribution is said to be *negatively or left-skewed.* See Figure 2–8(f). Skewness will be explained in detail in Chapter 3. Distributions with one peak, such as those shown in Figure 2–8(a), (e), and (f), are said to be *unimodal.* (The highest peak of a distribution indicates where the mode of the data values is. The mode is the data value that occurs more often than any other data value. Modes are explained in Chapter 3.) When a distribution has two peaks of the same height, it is said to be *bimodal.* See Figure 2–8(g). Finally, the graph shown in Figure 2–8(h) is a *U-shaped* distribution.

Distributions can have other shapes in addition to the ones shown here; however, these are some of the more common ones that you will encounter in analyzing data.

When you are analyzing histograms and frequency polygons, look at the shape of the curve. For example, does it have one peak or two peaks? Is it relatively flat, or is it U-shaped? Are the data values spread out on the graph, or are they clustered around the center? Are there data values in the extreme ends? These may be *outliers.* (See Section 3–3 for an explanation of outliers.) Are there any gaps in the histogram, or does the frequency polygon touch the *x* axis somewhere other than at the ends? Finally, are the data clustered at one end or the other, indicating a *skewed distribution*?

For example, the histogram for the record high temperatures shown in Figure 2–2 shows a single peaked distribution, with the class 109.5–114.5 containing the largest number of temperatures. The distribution has no gaps, and there are fewer temperatures in the highest class than in the lowest class.

Applying the Concepts 2–2

Selling Real Estate

Assume you are a realtor in Bradenton, Florida. You have recently obtained a listing of the selling prices of the homes that have sold in that area in the last 6 months. You wish to organize those data so you will be able to provide potential buyers with useful information. Use the following data to create a histogram, frequency polygon, and cumulative frequency polygon.

142,000	127,000	99,600	162,000	89,000	93,000	99,500
73,800	135,000	119,500	67,900	156,300	104,500	108,650
123,000	91,000	205,000	110,000	156,300	104,000	133,900
179,000	112,000	147,000	321,550	87,900	88,400	180,000
159,400	205,300	144,400	163,000	96,000	81,000	131,000
114,000	119,600	93,000	123,000	187,000	96,000	80,000
231,000	189,500	177,600	83,400	77,000	132,300	166,000

1. What questions could be answered more easily by looking at the histogram rather than the listing of home prices?
2. What different questions could be answered more easily by looking at the frequency polygon rather than the listing of home prices?
3. What different questions could be answered more easily by looking at the cumulative frequency polygon rather than the listing of home prices?
4. Are there any extremely large or extremely small data values compared to the other data values?
5. Which graph displays these extremes the best?
6. Is the distribution skewed?

See page 101 for the answers.

Exercises 2–2

1. Do Students Need Summer Development? For 108 randomly selected college applicants, the following frequency distribution for entrance exam scores was obtained. Construct a histogram, frequency polygon, and ogive for the data. (The data for this exercise will be used for Exercise 13 in this section.)

Class limits	Frequency
90–98	6
99–107	22
108–116	43
117–125	28
126–134	9

Applicants who score above 107 need not enroll in a summer developmental program. In this group, how many students do not have to enroll in the developmental program?

2. Number of College Faculty The number of faculty listed for a variety of private colleges that offer only bachelor's degrees is listed below. Use these data to construct a frequency distribution with 7 classes, a histogram, a frequency polygon, and an ogive. Discuss the shape of this distribution. What proportion of schools have 180 or more faculty?

165	221	218	206	138	135	224	204
70	210	207	154	155	82	120	116
176	162	225	214	93	389	77	135
221	161	128	310				

Source: *World Almanac and Book of Facts.*

3. Counties, Divisions, or Parishes for 50 States The number of counties, divisions, or parishes for each of the 50 states is given below. Use the data to construct a grouped frequency distribution with 6 classes, a histogram, a frequency polygon, and an ogive. Analyze the distribution. (The data in this exercise will be used for Exercise 24 in Section 2–2.)

67	27	15	75	58	64	8	67	159	5
102	44	92	99	105	120	64	16	23	14
83	87	82	114	56	93	16	10	21	33
62	100	53	88	77	36	67	5	46	66
95	254	29	14	95	39	55	72	23	3

Source: *World Almanac and Book of Facts.*

4. NFL Salaries The salaries (in millions of dollars) for 31 NFL teams for a specific season are given in this frequency distribution.

Class limits	Frequency
39.9–42.8	2
42.9–45.8	2
45.9–48.8	5
48.9–51.8	5
51.9–54.8	12
54.9–57.8	5

Source: NFL.com

Construct a histogram, a frequency polygon, and an ogive for the data; and comment on the shape of the distribution.

5. Railroad Crossing Accidents The data show the number of railroad crossing accidents for the 50 states of the United States for a specific year. Construct a histogram, frequency polygon, and ogive for the data. Comment on the skewness of the distribution. (The data in this exercise will be used for Exercise 14 in this section.)

Class limits	Frequency
1–43	24
44–86	17
87–129	3
130–172	4
173–215	1
216–258	0
259–301	0
302–344	1

Source: Federal Railroad Administration.

6. Costs of Utilities The frequency distribution represents the cost (in cents) for the utilities of states that supply much of their own power. Construct a histogram, frequency polygon, and ogive for the data. Is the distribution skewed?

Class limits	Frequency
6–8	12
9–11	16
12–14	3
15–17	1
18–20	0
21–23	0
24–26	1

7. Air Quality Standards The number of days that selected U.S. metropolitan areas failed to meet acceptable air quality standards is shown below for 1998 and 2003. Construct a grouped frequency distribution with 7 classes and a histogram for each set of data, and compare your results.

1998						2003					
43	76	51	14	0	10	10	11	14	20	15	6
20	0	5	17	67	25	17	0	5	19	127	4
38	0	56	8	0	9	31	5	88	1	1	16
14	5	37	14	95	20	14	19	20	9	138	22
23	12	33	0	3	45	13	10	20	20	20	12

Source: *World Almanac.*

8. How Quick Are Dogs? In a study of reaction times of dogs to a specific stimulus, an animal trainer obtained the following data, given in seconds. Construct a histogram, a frequency polygon, and an ogive for the data; analyze the results. (The histogram in this exercise

will be used for Exercise 18 in this section, Exercise 16 in Section 3–1, and Exercise 26 in Section 3–2.)

Class limits	Frequency
2.3–2.9	10
3.0–3.6	12
3.7–4.3	6
4.4–5.0	8
5.1–5.7	4
5.8–6.4	2

9. Quality of Health Care The scores of health care quality as calculated by a professional risk management company are listed for selected states. Use the data to construct a frequency distribution with 6 classes, a histogram, a frequency polygon, and an ogive.

118.2 114.6 113.1 111.9 110.0 108.8 108.3 107.7 107.0 106.7
105.3 103.7 103.2 102.8 101.6 99.8 98.1 96.6 95.7 93.6
92.5 91.0 90.0 87.1 83.1

Source: *New York Times Almanac.*

10. Making the Grade The frequency distributions shown indicate the percentages of public school students in fourth-grade reading and mathematics who performed at or above the required proficiency levels for the 50 states in the United States. Draw histograms for each, and decide if there is any difference in the performance of the students in the subjects.

Class	Reading frequency	Math frequency
17.5–22.5	7	5
22.5–27.5	6	9
27.5–32.5	14	11
32.5–37.5	19	16
37.5–42.5	3	8
42.5–47.5	1	1

Source: *National Center for Educational Statistics.*

11. Construct a histogram, frequency polygon, and ogive for the data in Exercise 16 in Section 2–1 and analyze the results.

12. For the data in Exercise 18 in Section 2–1, construct a histogram for the state gasoline taxes.

13. For the data in Exercise 1 in this section, construct a histogram, a frequency polygon, and an ogive, using relative frequencies. What proportion of the applicants needs to enroll in the summer development program?

14. For the data in Exercise 5 in this section, construct a histogram, frequency polygon, and ogive using relative frequencies. What proportion of the railroad crossing accidents are less than 87?

15. Cereal Calories The number of calories per serving for selected ready-to-eat cereals is listed here. Construct a frequency distribution using 7 classes. Draw a histogram, a frequency polygon, and an ogive for the data, using relative frequencies. Describe the shape of the histogram.

130 190 140 80 100 120 220 220 110 100
210 130 100 90 210 120 200 120 180 120
190 210 120 200 130 180 260 270 100 160
190 240 80 120 90 190 200 210 190 180
115 210 110 225 190 130

Source: *The Doctor's Pocket Calorie, Fat, and Carbohydrate Counter.*

16. Protein Grams in Fast Food The amount of protein (in grams) for a variety of fast-food sandwiches is reported here. Construct a frequency distribution using 6 classes. Draw a histogram, a frequency polygon, and an ogive for the data, using relative frequencies. Describe the shape of the histogram.

23 30 20 27 44 26 35 20 29 29
25 15 18 27 19 22 12 26 34 15
27 35 26 43 35 14 24 12 23 31
40 35 38 57 22 42 24 21 27 33

Source: *The Doctor's Pocket Calorie, Fat, and Carbohydrate Counter.*

17. For the data for year 2003 in Exercise 7 in this section, construct a histogram, a frequency polygon, and an ogive, using relative frequencies.

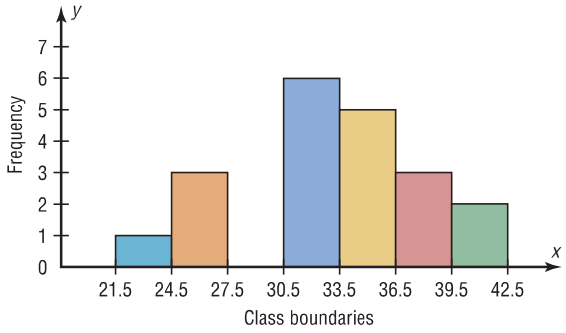
18. How Quick Are Older Dogs? The animal trainer in Exercise 8 in this section selected another group of dogs who were much older than the first group and measured their reaction times to the same stimulus. Construct a histogram, a frequency polygon, and an ogive for the data.

Class limits	Frequency
2.3–2.9	1
3.0–3.6	3
3.7–4.3	4
4.4–5.0	16
5.1–5.7	14
5.8–6.4	4

Analyze the results and compare the histogram for this group with the one obtained in Exercise 8 in this section. Are there any differences in the histograms? (The data in this exercise will be used for Exercise 16 in Section 3–1 and Exercise 26 in Section 3–2.)

Extending the Concepts

19. Using the histogram shown here, do the following.



a. Construct a frequency distribution; include class limits, class frequencies, midpoints, and cumulative frequencies.
b. Construct a frequency polygon.
c. Construct an ogive.

20. Using the results from Exercise 19, answer these questions.

a. How many values are in the class 27.5–30.5? 0
b. How many values fall between 24.5 and 36.5? 14
c. How many values are below 33.5? 10
d. How many values are above 30.5? 16

Technology *Step by Step*

MINITAB Step by Step

Construct a Histogram

1. Enter the data from Example 2–2, the high temperatures for the 50 states.

2. Select **Graph>Histogram.**

3. Select [Simple], then click [OK].

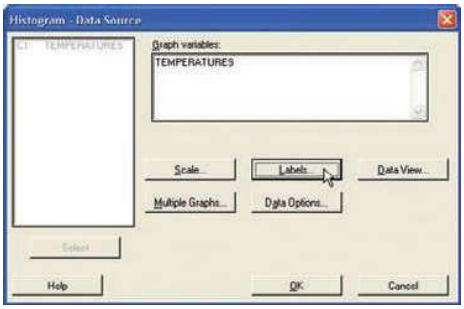
4. Click C1 TEMPERATURES in the Graph variables dialog box.

5. Click [Labels]. There are two tabs, Title/Footnote and Data Labels.

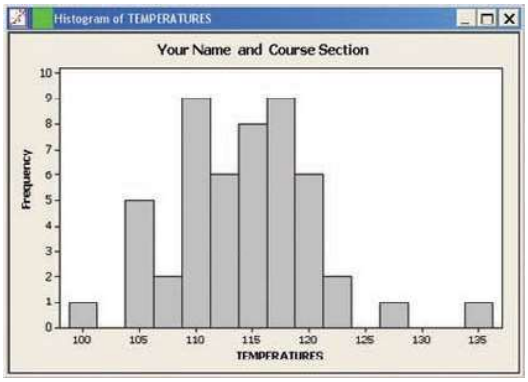
a) Click in the box for Title, and type in Your Name and Course Section.

b) Click [OK]. The Histogram dialog box is still open.

6. Click [OK]. A new graph window containing the histogram will open.



7. Click the **File** menu to print or save the graph.



8. Click **File>Exit.**
9. Save the project as **Ch2-3.mpj.**

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

Constructing a Histogram

To display the graphs on the screen, enter the appropriate values in the calculator, using the WINDOW menu. The default values are $X_{min} = -10$, $X_{max} = +10$, $Y_{min} = -10$, and $Y_{max} = +10$.

The X_{scl} changes the distance between the tick marks on the x axis and can be used to change the class width for the histogram.

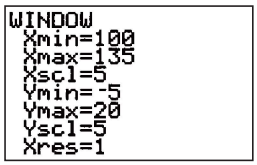
To change the values in the WINDOW:

1. Press **WINDOW.**
2. Move the cursor to the value that needs to be changed. Then type in the desired value and press **ENTER.**
3. Continue until all values are appropriate.
4. Press **[2nd] [QUIT]** to leave the WINDOW menu.

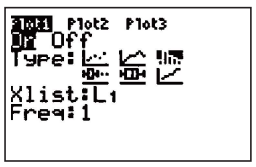
To plot the histogram from raw data:

1. Enter the data in L_1.
2. Make sure WINDOW values are appropriate for the histogram.
3. Press **[2nd] [STAT PLOT] ENTER.**
4. Press **ENTER** to turn the plot on, if necessary.
5. Move cursor to the Histogram symbol and press **ENTER,** if necessary.
6. Make sure Xlist is L_1.
7. Make sure Freq is 1.
8. Press **GRAPH** to display the histogram.
9. To obtain the number of data values in each class, press the **TRACE** key, followed by ◄ or ► keys.

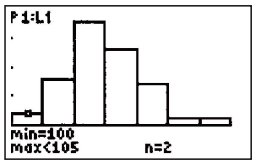
Input



Input



Output



Example TI2–1

Plot a histogram for the following data from Examples 2–2 and 2–4.

112	100	127	120	134	118	105	110	109	112
110	118	117	116	118	122	114	114	105	109
107	112	114	115	118	117	118	122	106	110
116	108	110	121	113	120	119	111	104	111
120	113	120	117	105	110	118	112	114	114

Press **TRACE** and use the arrow keys to determine the number of values in each group.

To graph a histogram from grouped data:

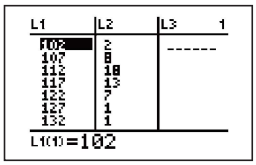
1. Enter the midpoints into L_1.
2. Enter the frequencies into L_2.
3. Make sure WINDOW values are appropriate for the histogram.
4. Press **[2nd] [STAT PLOT] ENTER.**
5. Press **ENTER** to turn the plot on, if necessary.
6. Move cursor to the histogram symbol, and press **ENTER,** if necessary.
7. Make sure Xlist is L_1.
8. Make sure Freq is L_2.
9. Press **GRAPH** to display the histogram.

Example TI2–2

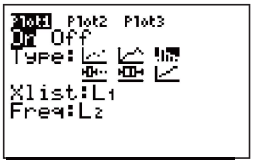
Plot a histogram for the data from Examples 2–4 and 2–5.

Class boundaries	Midpoints	Frequency
99.5–104.5	102	2
104.5–109.5	107	8
109.5–114.5	112	18
114.5–119.5	117	13
119.5–124.5	122	7
124.5–129.5	127	1
129.5–134.5	132	1

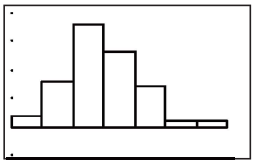
Input



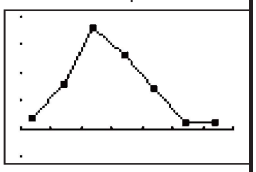
Input



Output

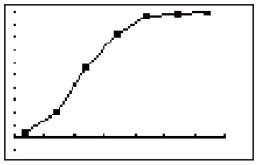


Output



To graph a frequency polygon from grouped data, follow the same steps as for the histogram except change the graph type from histogram (third graph) to a line graph (second graph).

Output



To graph an ogive from grouped data, modify the procedure for the histogram as follows:

1. Enter the upper class boundaries into L_1.
2. Enter the cumulative frequencies into L_2.
3. Change the graph type from histogram (third graph) to line (second graph).
4. Change the Y_{max} from the WINDOW menu to the sample size.

Excel Step by Step

Constructing a Histogram

1. Press **[Ctrl]-N** for a new workbook.
2. Enter the data from Example 2–2 in column A, one number per cell.
3. Enter the upper boundaries into column B.
4. From the toolbar, select the Data tab, then select Data Analysis.
5. In Data Analysis, select Histogram and click [OK].
6. In the Histogram dialog box, type **A1:A50** in the Input Range box and type **B1:B7** in the Bin Range box.