You have probably heard that “a picture is worth a thousand words.” The same can be said about statistical graphs because they summarize hundreds or thousands of numbers. Many people are intimidated by statistical information presented in frequency distributions or in other tabular forms, but they find the same information to be readable and understandable when presented graphically. Graphs tell a story in “pictures” rather than in words or numbers. They are supposed to make us think about the substance rather than the technical details of the presentation.

In this chapter, you will learn about some of the most commonly used graphical techniques. We concentrate less on the technical details of how to create graphs and more on how to choose the appropriate graphs to make statistical information coherent. We also focus on how to interpret information presented graphically.

As we introduce the various graphical techniques, we also show you how to use graphs to tell a “story.” The particular story we tell in this chapter is that of the elderly in the United States. The different types of graphs introduced in this chapter demonstrate the many facets of the aging of society over the next four decades. People have tended to talk about seniors as if they were a homogeneous group, but the different graphical techniques we illustrate here dramatize the wide variations in economic characteristics, living arrangements, and family status among people aged 65 and older. Most of the statistical information presented in this chapter is based on reports prepared by statisticians from the U.S. Census Bureau and other government agencies that gather information about the elderly in the United States and internationally.

Numerous graphing techniques are available to you, but here we focus on just a few of the most widely used ones in the social sciences. The first two, the pie chart and bar graph, are appropriate for...
nominal and ordinal variables. The next two, histograms and line graphs, are used with interval-ratio variables. We also discuss statistical maps and time-series charts. The statistical map is most often used with interval-ratio data. Finally, time-series charts are used to show how some variables change over time.

**THE PIE CHART: RACE AND ETHNICITY OF THE ELDERLY**

The elderly population of the United States is racially heterogeneous. As the data in Table 3.1 show, of the total elderly population (defined as persons 65 years and older) in 2008, about 33.7 million were white, about 3.3 million black, 225,000 American Indian, 1.3 million Asian American, 36,000 Native Hawaiian or Pacific Islander, and 261,000 were two or more races combined.

<table>
<thead>
<tr>
<th>Race</th>
<th>Frequency (f)</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>White alone</td>
<td>33,737,000</td>
<td>86.8</td>
</tr>
<tr>
<td>Black alone</td>
<td>3,314,000</td>
<td>8.5</td>
</tr>
<tr>
<td>American Indian alone</td>
<td>225,000</td>
<td>0.6</td>
</tr>
<tr>
<td>Asian alone</td>
<td>1,294,000</td>
<td>3.3</td>
</tr>
<tr>
<td>Native Hawaiian or Pacific Islander</td>
<td>36,000</td>
<td>0.1</td>
</tr>
<tr>
<td>Two or more races combined</td>
<td>261,000</td>
<td>0.7</td>
</tr>
</tbody>
</table>

*Source: U.S. Census Bureau, The 2010 Statistical Abstract, Table 10.*

A pie chart shows the differences in frequencies or percentages among the categories of a nominal or an ordinal variable. The categories are displayed as segments of a circle whose pieces add up to 100% of the total frequencies. The pie chart shown in Figure 3.1 displays the same information that Table 3.1 presents. Although you can inspect these data in Table 3.1, you can interpret the information more easily by seeing it presented in the pie chart in Figure 3.1. It shows that the elderly population is predominantly white (86.8%), followed by black (8.5%).

Notice that the pie chart contains all the information presented in the frequency distribution. Like the frequency distribution, charts have an identifying number, a title that describes the content of the figure, and a reference to a source. The frequency or percentage is represented both visually and in numbers.
Figure 3.1  Annual Estimates of the U.S. Population 65 Years and Over by Race, 2008


Pie chart A graph showing the differences in frequencies or percentages among the categories of a nominal or an ordinal variable. The categories are displayed as segments of a circle whose pieces add up to 100% of the total frequencies.

Note that the percentages for several of the racial groups are about 3% or less. It might be better to combine categories—American Indian, Asian, Native Hawaiian—into an “other races” category. This will leave us with three distinct categories: White, Black, and Other and Two or More Races. The revised pie chart is presented in Figure 3.2. Confirm for yourself how the percentages are derived from Table 3.1. We can highlight the diversity of the elderly population by “exploding” the pie chart, moving the segments representing these groups slightly outward to draw them to the viewer’s attention.

Note that we “exploded” the segment of the pies representing the black and other population groups in order to highlight the proportion of whites.

THE BAR GRAPH: LIVING ARRANGEMENTS AND LABOR FORCE PARTICIPATION OF THE ELDERLY

The bar graph provides an alternative way to present nominal or ordinal data graphically. It shows the differences in frequencies or percentages among categories of a nominal or an ordinal variable. The
Bar graph  A graph showing the differences in frequencies or percentages among the categories of a nominal or an ordinal variable. The categories are displayed as rectangles of equal width with their height proportional to the frequency or percentage of the category.

Let’s illustrate the bar graph with an overview of the living arrangements of the elderly. Living arrangements change considerably with advancing age—an increasing number of the elderly live alone or with other relatives. Figure 3.3 is a bar graph displaying the percentage distribution of elderly males by living arrangements in 2009. This chart is interpreted similar to a pie chart except that the categories of the variable are arrayed along the horizontal axis (sometimes referred to as the X-axis) and the percentages along the vertical axis (sometimes referred to as the Y-axis). This bar graph is easily interpreted: It shows that in 2009, 18.7% of elderly males lived alone, 72% were married and living with their spouses, 6.1% were living with other relatives, and the remaining 3.2% were living with nonrelatives.
Construct a bar graph by first labeling the categories of the variables along the horizontal axis. For these categories, construct rectangles of equal width, with the height of each proportional to the frequency or percentage of the category. Note that a space separates each of the categories to make clear that they are nominal categories.

Bar graphs are often used to compare one or more categories of a variable among different groups. For example, there is an increasing likelihood that women will live alone as they age. The longevity of women is the major factor in the gender differences in living arrangements. In addition, elderly widowed men are more likely to remarry than elderly widowed women. Also, it has been noted that the current generation of elderly women has developed more protective social networks and interests.

Suppose we want to show how the patterns in living arrangements differ between men and women. Figure 3.4 compares the percentage of women and men 65 years and older who lived with others or alone in 2009. It clearly shows that elderly women are more likely than elderly men to live alone.

We can also construct bar graphs horizontally, with the categories of the variable arrayed along the vertical axis and the percentages or frequencies displayed on the horizontal axis. This format is illustrated in Figure 3.5, which compares the percentage of men and women 55 years and over in the civilian labor force for 2008. We see that for all age categories, men were more likely to be employed than women.

THE STATISTICAL MAP: THE GEOGRAPHIC DISTRIBUTION OF THE ELDERLY

Since the 1960s, the elderly have been relocating to the South and the West of the United States. It is projected that by 2020 these regions will increase their elderly population by as much as 80% (though recent census data reveal that the recession that began in 2008 has halted this dominant immigration
Figure 3.4  Living Arrangement of U.S. Elderly (65 and older) by Gender, 2009


Figure 3.5  Percentage of Men and Women 55 Years and Over Who Are Employed, 2008

trend). We can display these dramatic geographical changes in American society by using a statistical map. Maps are especially useful for describing geographical variations in variables, such as population distribution, voting patterns, crime rates, or labor force composition.

Let’s look at Figure 3.6. It presents a statistical map, by state, of the percentage of the population 65 years and over for 2008. The variable percentage of the population 65 years and over in 2008 has four categories: less than 10%, 10% to 11.9%, 12% to 13.9%, and 14% or more. Each category is represented by a different shading (or color code), and the states are shaded depending on their classification into the different categories. To make it easier to read a map that you construct and to identify its patterns, keep the number of categories relatively small—say, not more than five.

**Figure 3.6** Percentage of the Population 65 Years and Over, 2008

Maps can also display geographical variations on the level of cities, counties, city blocks, census tracts, and other units. Your choice of whether to display variations on the state level or for smaller units will depend on the research question you wish to explore.

*Can you think of a few other examples of data that could be described using a statistical map? What type of data are organized at the state level?*
The **histogram** is used to show the differences in frequencies or percentages among categories of an interval-ratio variable. The categories are displayed as contiguous bars, with width proportional to the width of the category and height proportional to the frequency or percentage of that category. A histogram looks very similar to a bar chart except that the bars are contiguous to each other (touching) and may not be of equal width. In a bar chart, the spaces between the bars visually indicate that the categories are separate. Examples of variables with separate categories are *marital status* (married, single), *gender* (male, female), and *employment status* (employed, unemployed). In a histogram, the touching bars indicate that the categories or intervals are ordered from low to high in a meaningful way. For example, the categories of the variables *hours spent studying*, *age*, and *years of school completed* are contiguous, ordered intervals.

---

**Histogram**  
A graph showing the differences in frequencies or percentages among the categories of an interval-ratio variable. The categories are displayed as contiguous bars, with width proportional to the width of the category and height proportional to the frequency or percentage of that category.

---

Figure 3.7 is a histogram displaying the percentage distribution of the population 65 years and over by age. The data on which the histogram is based are presented in Table 3.2. To construct the histogram of Figure 3.7, arrange the age intervals along the horizontal axis and the percentages (or frequencies) along the vertical axis. For each age category, construct a bar with the height corresponding to the percentage of the elderly in the population in that age category. The width of each bar corresponds to the number of years that the age interval represents. The area that each bar occupies tells us the proportion of the population that falls into a given age interval. The histogram is drawn with the bars touching each other to indicate that the categories are contiguous.

---

**Figure 3.7**  
Age Distribution of U.S. Population 65 Years and Over, 2008

![Histogram of Age Distribution](image)

Table 3.2  Percentage Distribution of U.S. Population 65 Year and Over by Age, 2008

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>65–69</td>
<td>29.2</td>
</tr>
<tr>
<td>70–74</td>
<td>22.6</td>
</tr>
<tr>
<td>75–79</td>
<td>18.7</td>
</tr>
<tr>
<td>80–84</td>
<td>14.8</td>
</tr>
<tr>
<td>85–89</td>
<td>9.2</td>
</tr>
<tr>
<td>90–94</td>
<td>4.0</td>
</tr>
<tr>
<td>95+</td>
<td>1.5</td>
</tr>
<tr>
<td>Total (N)</td>
<td>100.0 (38,870,000)</td>
</tr>
</tbody>
</table>

When bar charts or histograms are used to display the frequencies of the categories of a single variable, the categories are shown on the X-axis and the frequencies on the Y-axis. In a horizontal bar chart or histogram, this is reversed.

STATISTICS IN PRACTICE: THE “GRAYING” OF AMERICA

We can also use the histogram to depict more complex trends, as, for instance, the “graying” of America. Let’s consider for a moment some of these trends: The elderly population today is 10 times larger than it was in 1900, and it will more than double by the year 2030. Indeed, as a journalist has pointed out, if the automobile had existed in Colonial times, half the residents of the New Land couldn’t have taken a spin: One of every two people was under age 16. Most didn’t live long enough to reach old age. Today, the population too young to drive has dropped to one in four while adults 65 and over account for one in eight.4

The histogram can give us a visual impression of these demographic trends. For an illustration, let’s look at Figures 3.8 and 3.9. Both are applications of the histogram. They examine, by gender, age distribution patterns in the U.S. population for 1955 and 2010 (projected). Note that in both figures, age groups are arranged along the vertical axis, whereas the frequencies (in millions of people) are along the horizontal axis. Each age group is classified by males on the left and females on the right. Because this type of histogram reflects age distribution by gender, it is also called an age-sex pyramid.

Visually compare the different pieces of data presented in these graphs. By observing where age groups are concentrated, you can discern major patterns in age distribution over time. Note the different shapes of Figures 3.8 and 3.9. Whereas in 1955 the largest group in the population was 0 to 9 years old, in 2010 the largest group will be 45 to 54 years old. These dramatic changes reflect the “graying” of the baby boom (born 1946–1965) generation. Almost 84 million babies were born in the United States from 1946 to 1965, which is 60% more than were born during the preceding two decades. By 2010, as the baby boom generation reaches 45 to 64, the number of middle-aged and elderly Americans will increase dramatically.
Observe the differences in the number of men and women as age increases. These differences are especially noticeable in Figure 3.9. For example, between ages 70 and 74, women outnumber men 5:4.2; for those 85 years and over, women outnumber men almost 2:1. These differences reflect the fact that at every age male mortality exceeds female mortality.

Although the “graying” of America is a fact today, the age projections for the near future reveal that the younger generations, and particularly the so-called Generation Y (born 1980–1995), will have outnumbered the baby boom generation as early as by 2015. The population projections for 2015 are presented in Figure 3.10. A major reason behind this phenomenon is “the large waves of migrants—legal and illegal—arriving [in the United States] since 1975,” whose children have joined Generation Y and contribute to the younger population’s challenge to the graying baby boomers.  

Note that when we want to use the histogram to compare groups, we must show a histogram for each group (see Figures 3.8, 3.9, and 3.10). When we compare groups on the bar chart, we are able to compare two or more groups on the same bar chart (see Figure 3.4).
Figure 3.9  U.S. Population by Gender and Age, 2010 (in millions)


THE LINE GRAPH

Numerical growth of the elderly population is taking place worldwide, occurring in both developed and developing countries. In 1994, 30 nations had elderly populations of at least 2 million; demographic projections indicate that there will be 55 such nations by 2020. Japan is one of the nations experiencing dramatic growth of its elderly population. Figure 3.11 is a line graph displaying the elderly population of Japan by age.

The line graph is another way to display interval-ratio distributions; it shows the differences in frequencies or percentages among categories of an interval-ratio variable. Points representing the frequencies of each category are placed above the midpoint of the category and are joined by a straight line. Notice that in Figure 3.11 the age intervals are arranged on the horizontal axis and the frequencies along the vertical axis. Instead of using bars to represent the frequencies, however, points representing the frequencies of each interval are placed above the midpoint of the intervals. Adjacent points are then joined by straight lines.
Figure 3.10  U.S. Population by Gender and Age, 2015 (in millions)


Figure 3.11  Population of Japan, Age 55 and Above, 2009

**Line graph**  A graph showing the differences in frequencies or percentages among categories of an interval-ratio variable. Points representing the frequencies of each category are placed above the midpoint of the category and are joined by a straight line.

Both the histogram and the line graph can be used to depict distributions and trends of interval-ratio variables. How do you choose which one to use? To some extent, the choice is a matter of individual preference, but in general, line graphs are better suited for comparing how a variable is distributed across two or more groups or across two or more time periods. For example, Figure 3.12 compares the elderly population in Japan for 2000 with the projected elderly population for the years 2010 and 2020.

Let’s examine this line graph. It shows that Japan’s population of age 65 and over is expected to grow dramatically in the coming decades. According to projections, Japan’s oldest-old population, those 80 years or older, is also projected to grow rapidly, from about 4.8 million (less than 4% of the total population) to 10.8 million (8.9%) by 2020. This projected rise has already led to a reduction in retirement benefits and other adjustments to prepare for the economic and social impact of a rapidly aging society.6

**Figure 3.12**  Population of Japan, Age 55 and Over, 2000, 2010, and 2020

*Source:* Adapted from U.S. Census Bureau, Center for International Research, International Database, 2007.
Look closely at the line graph shown in Figure 3.12, comparing 2010 and 2020 data. How would you characterize the population increase among the Japanese elderly?

TIME-SERIES CHARTS

We are often interested in examining how some variables change over time. For example, we may be interested in showing changes in the labor force participation of Latinas over the past decade, changes in the public’s attitude toward abortion rights, or changes in divorce and marriage rates. A time-series chart displays changes in a variable at different points in time. It involves two variables: (1) time, which is labeled across the horizontal axis, and (2) another variable of interest whose values (frequencies, percentages, or rates) are labeled along the vertical axis. To construct a time-series chart, use a series of dots to mark the value of the variable at each time interval, and then join the dots by a series of straight lines.

Time-series chart A graph displaying changes in a variable at different points in time. It shows time (measured in units such as years or months) on the horizontal axis and the frequencies (percentages or rates) of another variable on the vertical axis.

Figure 3.13 shows a time series from 1900 to 2050 of the percentage of the total U.S. population that is 65 years or older (the figures for the years 2000 through 2050 are projections made by the Social Security Administration, as reported by the U.S. Census Bureau). This time series enables us to see clearly the dramatic increase in the elderly population. The number of elderly increased from a little less than 5% in 1900 to about 12.4% in 2000. The rate is expected to increase to 20% of the total population. This dramatic increase in the elderly population, especially beginning in the year 2010, is associated with the “graying” of the baby boom generation. This group, which was 0 to 9 years old in 1955 (see the age pyramid in Figure 3.8), will be 55 to 64 years old in the year 2010.

The implications of these demographic changes are enormous. To cite just a few, there will be more pressure on the health care system and on private and public pension systems. In addition, because the voting patterns of the elderly differ from those of younger people, the “graying” of America will have major political effects.

Often, we are interested in comparing changes over time for two or more groups. Let’s examine Figure 3.14, which charts the trends in the percentage of divorced elderly from 1960 to 2050 for men and women. This time-series graph shows that the percentage of divorced elderly men and elderly women was about the same until 2000. For both groups, the percentage increased from less than 2% in 1960 to about 5% in 1990. According to projections, however, there will be significant increases in the percentage of men and especially women who are divorced: from 5% of all the elderly in 1990 to 8.4% of all elderly men and 13.6% of all elderly women by the year 2050. This sharp upturn and the gender divergence are clearly emphasized in Figure 3.14.
Figure 3.13  Percentage of Total U.S. Population 65 Years and Above, 1900–2050


Figure 3.14  Percentage Currently Divorced Among U.S. Population 65 Years and Over, by Gender, 1960–2040


How does the time-series chart differ from a line graph? The difference is that line graphs display frequency distributions of a single variable, whereas time-series charts display two variables. In addition, time is always one of the variables displayed in a time-series chart.
In this chapter, we have seen that statistical graphs can give us a quick sense of the main patterns in the data. However, graphs can not only quickly inform us, they can also quickly deceive us. Because we are often more interested in general impressions than in detailed analyses of the numbers, we are more vulnerable to being swayed by distorted graphs. Edward Tufte in his 1983 book *The Visual Display of Quantitative Information* not only demonstrates the advantages of working with graphs but also offers a detailed discussion of some of the pitfalls in the application and interpretation of graphics.

Probably the most common distortions in graphical representations occur when the distance along the vertical or horizontal axis is altered either by not using 0 as the baseline (as demonstrated in Figures 3.15a and b) or in relation to the other axis. Axes may be stretched or shrunk to create any desired result to exaggerate or disguise a pattern in the data. In Figures 3.15a and b, 2009 international data on female representation in national parliaments are presented. Without altering the data in any way, notice how the difference between the countries is exaggerated by using 30 as a baseline (as in Figure 3.15b).

Remember: Always interpret the graph in the context of the numerical information the graph represents.

**Figure 3.15** Female Representation in National Parliaments, 2009: (a) Using 0 as the Baseline and (b) Using 30 as the Baseline
We now illustrate some additional ways in which graphics can be used to highlight diversity visually. In particular, we show how graphs can help us (a) explore the differences and similarities among the many social groups coexisting within American society and (b) emphasize the rapidly changing composition of the U.S. population. Indeed, because of the heterogeneity of American society, the most basic question to ask when you look at data is “compared with what?” This question is not only at the heart of quantitative thinking but underlies inclusive thinking as well.

Three types of graphs—the bar chart, the line graph, and the time-series chart—are particularly suitable for making comparisons among groups. Let’s begin with the bar chart displayed in Figure 3.16. It compares the college degree attainment of those 35 years and over by gender. It shows that the percentage of men with a bachelor’s degree or higher is greater than women in the older age groups (45 to 64 and 65+), whereas the percentage of women having at least a bachelor’s degree exceeds that of men in the younger (35 to 44) age group. The smallest gap is among those who are 35 to 44 years, of whom 29.2% of men had a bachelor’s degree or higher compared with 32.4% of women.

The line graph provides another way of looking at differences based on gender, race/ethnicity, or other attributes such as class, age, or sexual orientation. For example, Figure 3.17 compares years of school completed by black Americans aged 25 to 64 and 65 years and older with that of all Americans in the same age groups.

The data illustrate that in the United States the percentage of Americans who have completed only 8 years of education has declined dramatically from about 10.7% among Americans 65 years and older to 4.2% for those 25 to 64 years old. The decline for black Americans is even more dramatic, from 17% of the black elderly to about 2.2% for those 25 to 64 years old. The corresponding trend illustrated in Figure 3.17 is the increase in the percentage of Americans (all races as well as black Americans) who have completed 13 to 15 years of schooling, or 16 years or more. For example, about 17.3% of black Americans 65 years or older completed 13 to 15 years of schooling compared with 31.3% of those 25 to 64 years.

The trends shown in Figure 3.17 reflect the development of mass education in the United States during the past 50 years. The percentage of Americans who have completed 4 years of high school or more has risen from about 40% in 1940 to 86.6% in 2008 (record high according to the U.S. Census Bureau). Similarly, in 1940 only about 5% of Americans completed 4 years or more of college, compared with 29.4% in 2008.9

Source: U.S. Census Bureau, American Community Survey, 2008, Table B15001.
Figure 3.17  Educational Attainment in the United States by Race and Age, 2008

Figure 3.17 illustrates that overall, younger Americans (25–64 years old) are better educated than elderly Americans. However, despite these overall trends, there are differences between the number of years of schooling completed by “blacks” and “all races.” Examine Figure 3.17, and find these differences. What do they tell you about schooling in America?

Finally, Figure 3.18 is a time-series chart showing changes over time in the teenage birthrate (number of births per 1,000 women) among white, black, and Hispanic women. It shows that between 1990 and 2006, birthrates for teenagers declined; however, rates remain highest for Hispanic and non-Hispanic black women and lowest for non-Hispanic white women.

To conclude, the three examples of graphs in this section as well as other examples throughout this chapter have illustrated how graphical techniques can portray the complexities of the social world by emphasizing the distinct characteristics of age, gender, and ethnic groups. By depicting similarities and differences, graphs help us better grasp the richness and complexities of the social world.

Source: U.S. Census Bureau, The 2010 Statistical Abstract, Table 84.
Note: Rates per 1,000 women in specified group.

Main Points

- A pie chart shows the differences in frequencies or percentages among categories of nominal or ordinal variable. The categories of the variable are segments of a circle whose pieces add up to 100% of the total frequencies.

- A bar graph shows the differences in frequencies or percentages among categories of a nominal or an ordinal variable. The categories are displayed as rectangles of equal width with their height proportional to the frequency or percentage of the category.

- Histograms display the differences in frequencies or percentages among categories of interval-ratio variables. The categories are displayed as contiguous bars with their width proportional to the width of the category and height proportional to the frequency or percentage of that category.

- A line graph shows the differences in frequencies or percentages among categories of an interval-ratio variable. Points representing the frequencies of each category are placed above the midpoint of the category (interval). Adjacent points are then joined by a straight line.

- A time-series chart displays changes in a variable at different points in time. It displays two variables: (1) time, which is labeled across the horizontal axis, and (2) another variable of interest whose values (e.g., frequencies, percentages, or rates) are labeled along the vertical axis.
KEY TERMS

bar graph  
line graph  
histogram  
pie chart  
time-series chart

ON YOUR OWN

Log on to the web-based student study site at www.pineforge.com/ssds6e for additional study questions, quizzes, web resources, and links to social science journal articles reflecting the statistics used in this chapter.

SPSS DEMONSTRATIONS

[HINTS2007]

Demonstration 1: Producing a Bar Chart

SPSS for Windows greatly simplifies and improves the production of graphics. The program offers a separate choice from the main menu bar, Graphs, which lists more than a dozen types of graphs that SPSS can create. We will use HINTS2007 for this demonstration.

Under the Graphs menu select Legacy Dialogs, and then Bar, which will produce various types of bar charts. We will use bar charts to display the distribution of the nominal variable HD06MARITALSTATU (marital status of respondent). After clicking on Bar, you will be presented with the initial dialog box shown in Figure 3.19.

Figure 3.19  Bar Charts Dialog Box
Almost all graphics procedures in SPSS begin with a dialog box that allows you to choose exactly the type of chart you want to construct. Many graph types can display more than one variable (the Clustered or Stacked choices). We will keep things simple here, so click on Simple, then on Define. When you do so, the main dialog box for simple bar charts opens (Figure 3.20).

**Figure 3.20   Simple Bar Charts Dialog Box**

The variable HD06MARITALSTATU (marital status) should be placed in the box labeled “Category Axis.” In the “Bars Represent” box, click on the “% of cases” radio button. This choice changes the default statistic from the number of cases to percentages, which are normally more useful for comparison purposes. Click on **OK** to submit your request. (You should note that SPSS automatically excludes missing values. You can change this by clicking on **Options**. Click in the box labeled “Display groups defined by missing values” to turn on this choice. Then click on **Continue**, then on **OK** to submit your request to SPSS.)

The bar chart for HD06MARITALSTATU is presented in an output window labeled SPSS Viewer. You can see in Figure 3.21 that the bar chart for HD06MARITALSTATU has six bars because the only valid responses
SPSS graphs can be edited by selecting Edit, then Edit Content, In Separate Window, which moves the graph to its own window (Chart Editor) and displays various editing tools and choices.

Demonstration 2: Producing a Histogram

Histograms are used to display interval or ratio variables. We'll use the variable CS19WHENDIAG_AGE (age diagnosed with cancer) from the 2007 HINTS file. Under the Graphs–Legacy Dialogs menu in SPSS, select the Histogram. Click on these choices and you will see the dialog box shown in Figure 3.22.

Histograms are created for one variable at a time (that's why there was no opening dialog box as for bar charts). You simply insert (drag) the variable you want to display in the first empty box. You don't need to worry about missing values in histograms; unlike the bar chart default, SPSS automatically deletes them from the display. Notice that SPSS includes icons to indicate the level of measurement for each variable. Interval-ratio variables (or scale variables as SPSS refers to them) is matched with a ruler icon. Click on the OK button (on the bottom left-hand corner) to process this request. The resulting histogram is shown in Figure 3.23.
SPSS automatically decided the appropriate width for each interval based on the range of the variable CS19WHENDIAG_AGE and the optimal number of bars to be displayed on a screen. The number displayed under each bar is the midpoint of that interval, so, for example, the bar for 20 years of age includes everyone from 20 to 25 years of age when first diagnosed with cancer. The histogram also includes the calculation for mean and standard deviation, which will be discussed in Chapters 4 and 5.

**SPSS PROBLEMS**

[HINTS2007]

1. You've decided to examine the differences in the age of male and female subjects in HINTS2007 data set. You examine AGEGRPA (age group of respondent).
   a. Construct a bar graph for AGEGRPA *(Hint: From the SPSS menu, choose Graphs–Legacy Dialogs–Bar.)*
   b. Construct bar graphs separately for men and women (insert the variable GENDER in the Panel by/Columns box.)
   c. Briefly describe overall age distribution and the difference in age groups between men and women.
2. Are individuals with higher educational attainment more likely to look for information about health or medical topics from any source? Construct separate pie charts for HC01SEEKHEALTHIN (does respondent look for information about health or medical topics?). You will need to select Pie under Graphs–Legacy Dialogs. In the first dialog box, select “Summaries for groups of cases.” Then, make sure you select % of cases under Slices Represent. In the box for Define slices by, insert HC01SEEKHEALTHIN and in the Panel by/Columns box insert HD07EDUCATION (degree). Compare the pie charts. What difference in seeking for information about health exists between the different educational (degree) groups?

3. Examine if there is a difference in responses between men and women in terms of their distress level (DISTRESSSCORE) and their occupational status (HD05OCCUPATIONST) and whether they have ever been diagnosed as having cancer (CS17EVERHADCANCE). Based on the level of measurement for each variable, determine the appropriate graphic display. Produce separate graphs for men and women. What differences, if any, are evident in the data?

4. Determine how best to represent the following variables graphically:
   HD16RENTOROWN—whether respondent rents or owns her/his house
   HD09RACE1—race of respondent
Exercises

1. The time-series chart shown in Figure 3.24 displays trends in birthrates for teenagers from 1980 to 2006.
   a. Write a 60-word report describing the variation in teen birthrates for the three age categories.
   b. What do you think causes some of the age groups to increase or decrease their birthrates across the years? Are there any important years or decades where the change in birthrates is dramatic?

2. We selected a sample of people from the International Social Survey Program (ISSP) 2000. Raw data are presented for their sex (SEX), social class (CLASS), and number of household members (HOMPOP). CLASS is a subjective measure, with respondents indicating L = lower, W = working, M = middle, and U = upper.

Figure 3.24  Birthrates for Teenagers by Age, 1980–2006

a. Construct a pie chart depicting the percentage distribution of sex. (*Hint:* Remember to include a title, percentages, and appropriate labels.)

b. Construct a pie chart showing the percentage distribution of social class.

c. Construct a graph with two pie charts comparing the percentage distribution of social class membership by sex.

3. We continue our analysis of the U.S. elderly population, this time examining the number of elderly from 1900 to 2000. A bar graph is presented in Figure 3.25. Write a brief statement describing the data.

**Figure 3.25**  Population Aged 65 and Over, 1900–2000 (in millions)

4. Using the data from Exercise 2, construct bar graphs showing percentage distributions for sex and class. Remember to include appropriate titles, percentages, and labels.

5. Suppose you want to compare the number of household members for women and men (based on the ISSP data in Exercise 2).
   a. Construct a grouped bar graph (similar to Figure 3.4) to show the percentage distribution of the number of household members by sex.
   b. Which group reported the largest family size?
   c. Why shouldn’t you construct a grouped bar chart showing the frequencies rather than the percentages?

6. During the 2008 presidential election, health care and health insurance were identified by voters as important issues. Policy analysts have noted that the number of uninsured is increasing in the United States. Data from the National Center for Health Statistics are presented in Figure 3.26. What can be said about who did not have health insurance during 2005? How does the percentage of those without health insurance vary by ethnicity/race, age, marital, and poverty status?

**Figure 3.26** The Uninsured Population Below 65 Years of Age by Selected Characteristics, 2005

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
</tr>
<tr>
<td>Below 18 years</td>
<td>16.2</td>
</tr>
<tr>
<td>18–44 Years</td>
<td>61.6</td>
</tr>
<tr>
<td>45–64 Years</td>
<td>22.2</td>
</tr>
<tr>
<td><strong>Poverty</strong></td>
<td></td>
</tr>
<tr>
<td>Below 100%</td>
<td>25.1</td>
</tr>
<tr>
<td>100 to less than 200%</td>
<td>31.6</td>
</tr>
<tr>
<td>200 to less than 400%</td>
<td>28.7</td>
</tr>
<tr>
<td>400% or more</td>
<td>14.6</td>
</tr>
<tr>
<td><strong>Race and Hispanic origin</strong></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>31.3</td>
</tr>
<tr>
<td>White only</td>
<td>48.0</td>
</tr>
<tr>
<td>Black</td>
<td>13.9</td>
</tr>
<tr>
<td>Asian</td>
<td>4.3</td>
</tr>
<tr>
<td>Other</td>
<td>2.4</td>
</tr>
<tr>
<td><strong>Marital status</strong></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>40.2</td>
</tr>
<tr>
<td>Married</td>
<td>41.8</td>
</tr>
<tr>
<td>Separated/divorced</td>
<td>16.3</td>
</tr>
<tr>
<td>Widowed</td>
<td>1.7</td>
</tr>
</tbody>
</table>
7. The MTF 2008 asked students whether they approved trying marijuana. A summary table is provided separating responses for black, white, and Hispanic students. Construct a chart or graph that best displays this information (it could also be separated by racial/ethnic groups).

<table>
<thead>
<tr>
<th>Attitude About Trying Marijuana</th>
<th>Black</th>
<th>White</th>
<th>Hispanic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Don’t disapprove</td>
<td>73</td>
<td>434</td>
<td>69</td>
</tr>
<tr>
<td>Disapprove</td>
<td>44</td>
<td>190</td>
<td>54</td>
</tr>
<tr>
<td>Strongly disapprove</td>
<td>68</td>
<td>242</td>
<td>76</td>
</tr>
<tr>
<td>Total (N)</td>
<td>185</td>
<td>866</td>
<td>199</td>
</tr>
</tbody>
</table>

8. You are writing a research paper on grandparents who had one or more of their grandchildren living with them. In 2000, 2.4 million grandparents were defined as caregivers by the U.S. Census, meaning that they had primary responsibility for raising their grandchildren below the age of 18. You discover the following
information from the U.S. Census Report, “Grandparents Living With Grandchildren: 2000” (C2KBR-31, October 2003): Among grandparent caregivers, 12% cared for a grandchild for less than 6 months, 11% for 6 to 11 months, 23% for 1 to 2 years, 15% for 3 to 4 years, and 39% for 5 or more years.

a. Construct a graph or chart that best displays this information on how long grandparents care for their grandchildren.

b. Explain why the graph you selected is appropriate.

9. Use the data on educational level in Chapter 2, Exercise 4, for this problem.

a. What level of measurement is “years of education”? Why can you use a histogram to graph the distribution of education, in addition to a bar chart?

b. Construct a histogram for years of education, using equal-spaced intervals of 4 years. Don't use percentages in this chart.

10. The 2008 General Social Survey (GSS) data on educational level can be further broken down by race as follows:

a. Construct two histograms for education, one for whites (\(N = 636\)) and one for blacks (\(N = 136\)).

b. Now use the two graphs to describe the differences in educational attainment by race.

<table>
<thead>
<tr>
<th>Years of Education</th>
<th>Whites</th>
<th>Blacks</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td>17</td>
<td>8</td>
</tr>
<tr>
<td>11</td>
<td>27</td>
<td>6</td>
</tr>
<tr>
<td>12</td>
<td>180</td>
<td>36</td>
</tr>
<tr>
<td>13</td>
<td>56</td>
<td>18</td>
</tr>
<tr>
<td>14</td>
<td>101</td>
<td>33</td>
</tr>
<tr>
<td>15</td>
<td>27</td>
<td>5</td>
</tr>
<tr>
<td>16</td>
<td>103</td>
<td>13</td>
</tr>
<tr>
<td>17</td>
<td>32</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>32</td>
<td>2</td>
</tr>
<tr>
<td>19</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>13</td>
<td>4</td>
</tr>
</tbody>
</table>

11. Use the data on minority rights attitudes from Chapter 2, Exercise 11. What would be the most appropriate graphic presentation for the data? Explain the reason for your answer.

12. Examine the bar chart representing the percentage of people speaking a language other than English at home as shown in Figure 3.27.

a. Overall, which age group had the lowest percentage speaking a language other than English at home? Which age group had the highest?

b. Describe the differences between the 2002 and 2008 percentages.

13. Based on educational attainment data (by gender and race) from Chapter 2, Exercise 7, decide how best to graphically represent those data in an easily understood format.

a. Would you choose to use bar charts or pie charts? Explain the reason for your answer.

b. Construct bar or pie charts (depending on your answer) to represent all the data. Remember to include appropriate titles, percentages, and labels.
14. As reported by Catherine Freeman (2004), females have more success in post-secondary education than male students. They are more likely to enroll in college immediately after high school and have higher college graduate rates than males. In her report, Freeman provides the following time-series chart (Figure 3.28), documenting the percentage of women enrolled in undergraduate, graduate, and professional programs. Prepare a brief statement on the enrollment trends from 1970 to 2000.

Figure 3.28  Females as a Percentage of Total Enrollment in Undergraduate, Graduate, and First-Professional Education: Various Years, Fall 1970 to Fall 2000