Displaying the Order in a Group of Numbers Using Tables and Graphs

Chapter Outline

We imagine you to be as unique as the other students we have known who have taken this course. Some of you are highly scientific sorts; others are more intuitive. Some of you are fond of math; others are less so, or even afraid of it. Whatever your style, we welcome you.

Given that you *can* learn statistics, you still have to decide if you want to make the effort it will require. Why would you want to do that, except to meet a requirement of your major? (Not a very energizing motive.) First, you will be far better equipped

to read research articles in your major. Second, you'll be on your way to being able to do your own research if you so choose. Third, you will improve both your reasoning and your intuition. Fourth, each time you finally grasp something you were struggling with, you will feel great.

Formally, **statistics** is a branch of mathematics that focuses on the organization, analysis, and interpretation of a group of numbers. But really what is statistics? Think of statistics as a tool that extends a basic thinking process employed by every human: You observe a thing; you wonder what it means or what caused it; you have an insight or make an intuitive guess; you observe again, but now in detail, or you try making some little changes in the process to test your intuition. Then you face the eternal problem: was your hunch confirmed or not? What are the chances that what you observed this second time will happen again and again so that you can announce your insight to the world as something probably true?

Statistics is a method of pursuing truth. As a minimum, statistics can tell you the likelihood that your hunch is true in this time and place and with these sorts of people. This pursuit of truth, or at least its future likelihood, is the essence of science. It is also the essence of human evolution and survival. Think of the first research questions: What will the mammoths do next spring? What will happen if I eat this root? It is easy to see how the early accurate "researchers" survived. You are here today because your ancestors exercised brains as well as brawn. Do those who come after you the same favor: Think carefully about outcomes. Statistics is one good way to do that.

Behavioral and social scientists usually use a computer and statistical software to carry out statistical procedures. However, the best way to develop a solid understanding of statistics is to learn how to do the procedures by hand (with the help of a calculator—it's not the multiplying and adding that you learn from, but the going through all the steps). In order to minimize the amount of figuring you have to do, we use relatively small groups of numbers in examples and practice problems. We hope that this will also allow you to focus more on the *underlying principles and logic* of the statistical procedures, rather than on the computations for each practice problem (such as subtracting 3 from 7 and then dividing the result by 2 to give an answer of 2). Having said that, we also recognize the importance of learning how to do statistical procedures on a computer, as you most likely would when conducting your own research. SPSS statistical software is commonly used by behavioral and social scientists to carry out statistical analyses. Check with your instructor to see if you have access to SPSS at your institution.

statistics A branch of mathematics that focuses on the organization, analysis, and interpretation of a group of numbers.

descriptive statistics Procedures for summarizing a group of scores or otherwise making them more understandable.

The Two Branches of Statistical Methods

There are two main branches of statistical methods:

1. **Descriptive statistics:** Behavioral and social scientists use descriptive statistics to summarize and describe a group of numbers from a research study.

2. **Inferential statistics:** Behavioral and social scientists use inferential statistics to draw conclusions and make inferences that are based on the numbers from a research study but that go beyond these numbers. For example, inferential statistics allow researchers to make inferences about a large group of individuals based on a research study in which a much smaller number of individuals took part.

inferential statistics Procedures for drawing conclusions based on the scores collected in a research study but going beyond them.

BOX 1 Math Anxiety, Statistics Anxiety, and You: A Message for Those of You Who Are Truly Worried about This Course

Let's face it: Many of you dread this course, even to the point of having a full-blown case of "statistics anxiety" (Zeidner, 1991). If you become tense the minute you see numbers, we need to talk about that right now.

First, this course is a chance for a fresh start with digits. Your past performance in (or avoidance of) geometry, trigonometry, calculus, or similar horrors need not influence in any way how well you comprehend statistics. This is largely a different subject.

Second, if your worry persists, you need to determine where it is coming from. Math or statistics anxiety, test anxiety, general anxiety, and general lack of confidence each seems to play its own role in students' difficulties with math courses (Cooper & Robinson, 1989; Dwinell & Higbee, 1991).

Is your problem mainly math/statistics anxiety? An Internet search will yield many wonderful books and Web sites to help you. We highly recommend Sheila Tobias's *Overcoming Math Anxiety* (1995). Tobias, a former math avoider herself, suggests that your goal be "math mental health," which she defines as "the willingness to learn the math you need when you need it" (p. 12). (Could it be that this course in statistics is one of those times?)

Tobias (1995) explains that math mental health is usually lost in elementary school, when you are called to the blackboard, your mind goes blank, and you are unable to produce the one right answer to an arithmetic problem. What confidence remained probably faded during timed tests, which you did not realize were difficult for everyone except the most proficient few.

Conquer your math anxiety by thinking back to when it might have started. Before that you were doing as well as others, so this is not about your lack of ability, but about some bad experiences or inadequate teaching that left you feeling inferior to others who were better prepared. Starting now, treat yourself more gently. Do not expect to understand things instantly. You are still recovering.

Start out confident. To do well, you only need a semester of high school algebra, but do review it, perhaps online, so that your math brain is warmed up. Go to class faithfully and do keep yourself alert during it. Do your statistics homework first, before you are tired. Ask for help from teachers, teaching assistants, or anyone you know in the class. Form a study group. If you don't understand something, try explaining it to someone who understands it even less. You may figure it out in the process. Read ahead in a relaxed way, so that what you hear in class is already familiar.

Is your problem test anxiety? Anxiety produces arousal, and one of the best understood relationships in psychology is between arousal and performance. Whereas moderate arousal helps performance, too much or too little dramatically decreases cognitive capacity and working memory. Things you have learned become harder to recall. Your mind starts to race, creating more anxiety, more arousal, and so on. There are many ways to reduce anxiety and arousal in general, such as learning to breathe properly and to take a brief break to relax deeply. Your counseling center should be able to help you or direct you to some good books on the subject. Again, many Web sites deal with reducing anxiety.

To reduce test anxiety specifically, overprepare for the first few tests, so that you go in with the certainty that you cannot possibly fail, no matter how aroused you become. Do this especially for the first test, because there will be no old material to review and a good score on your first test will make you less anxious about the next. Do a practice test, and make it as similar to a real test as possible, making a special effort to duplicate the aspects that bother you most. Once you think you are well prepared, set yourself a time limit for solving some homework problems. Make yourself write out answers fully and legibly. This may be part of what makes you feel slow during a test. If the presence of others bothers you—the sound of their scurrying pencils while yours is frozen in midair—do your practice test with others in your course. Even make it an explicit contest to see who can finish first.

For the test itself, reduce the effects of anxiety by jotting down in the margins any formulas you are afraid you

will forget. Glance through the whole test so that you know how to pace yourself. Read the instructions carefully. Start with questions that you can answer easily, and if some seem impossible, don't worry—you still may pass even if you do not answer them correctly. Try to ignore the people around you. Some people may seem to be working faster, but they may not be working as well as you.

Is your problem a general lack of confidence? Is there something else in your life causing you to worry or feel bad about yourself? Then we suggest that it is time you tried your friendly college counseling center.

Lastly, could you be highly sensitive? High sensitivity, reactivity, or responsiveness is an innate trait found in about 15–20% of humans (and at least a hundred other species; Wolf, Van Doorn, & Weissing, 2008). Individuals with the trait naturally adopt a strategy of reflecting before acting, rather than rushing into things. They process information more deeply and notice subtleties, as seen in brain imaging studies (Jagiellowicz et al., 2010). This makes them especially intuitive, creative, and conscientious (Aron, 1996; Aron & Aron, 1997). They also have stronger emotional reactions and are more affected by their environment and events in their childhood (Aron, Aron, & Davies, 2005; Belsky, Bakermans-Kranenburg, & Van IJzendoorn, 2007; Ellis, Essex, & Boyce, 2005). Being aware of so much, they become overstimulated sooner than others, so that college life itself can be a challenge for them, even though they have the potential to perform exceptionally well when not overaroused.

You might want to find out if you are a highly sensitive person (at *http://www.hsperson.com*). If you are, appreciate the trait's assets and make some allowances for its one disadvantage, this tendency to become easily overaroused. It has to affect your performance on tests. What matters is what you actually know, which is probably quite a bit. This simple act of self-acceptance—that you are *not* less smart but *are* more sensitive—may in itself help ease your arousal when trying to express your statistical knowledge.

So good luck to all of you. We wish you the best both while taking this course and in your lives.

In this chapter, we focus on descriptive statistics. This topic is important in its own right, but it also prepares you to understand inferential statistics.

In this chapter, we introduce you to some basic concepts, then you learn to use tables and graphs to describe a group of numbers. The purpose of descriptive statistics is to make a group of numbers easy to understand. As you will see, tables and graphs help a great deal.

Some Basic Concepts

Variables, Values, and Scores

As part of a larger study (Aron, Paris, & Aron, 1995), researchers gave a questionnaire to 151 students in an introductory statistics class during the first week of class. One question asked was, "How stressed have you been in the last 2½ weeks, on a scale of 0 to 10, with 0 being *not at all stressed* and 10 being *as stressed as possible?*" (How would *you* answer?) In this study, the researchers used a survey to examine students' level of stress. Some other methods that researchers use to study stress include creating stress with laboratory tasks (like having to be videotaped giving a talk for humans or swimming in water for rats) and measuring stress-related hormones or brain changes.

variable Characteristic that can have different values.

value Number or category that a score can have.

score Particular person's value on a variable.

In the current example, level of stress is a **variable,** which can have **values** from 0 to 10, and the value of any particular person's answer is the person's **score.** If you answered 6, your score is 6; your score has a value of 6 on the variable called level of stress.

More formally, a variable is a condition or characteristic that can have different values. In short, it can *vary.* In our example, the variable is level of stress. It can have

Table 1 Some Basic Terminology

Term	Definition	Examples
Variable	Condition or characteristic that can have different values	Stress level; age; gender; religion
Value	Number or category	0, 1, 2, 3, 4; 25, 85; female; Catholic
Score	A particular person's value on a variable	0, 1, 2, 3, 4; 25, 85; female; Catholic

values of 0 through 10. Height is a variable, social class is a variable, score on a creativity test is a variable, number of people absent from work on a given day is a variable, dosage of a medication is a variable, political party preference is a variable, and class size is a variable.

A value is just a number, such as 4, −81, or 367.12. A value can also be a category, such as female or male or the country in which you live (Canada, the United States, Australia, etc.).

Finally, on any variable, each person has a particular number or score that is that person's value on the variable. For example, your score on the stress variable might have a value of 6. Another student's score might have a value of 8.

Behavioral and social science research is about variables, values, and scores (see Table 1). The formal definitions are a bit abstract. In practice, you will find that what we mean when we use these words is usually clear.

Kinds of Variables

Most of the variables behavioral and social scientists use are like those in the stress ratings example: The scores are numbers that tell you how much there is of the thing being measured. In the stress ratings example, the higher the number is, the more stress there is. We call this kind of variable a **numeric variable.** Numeric variables are also called *quantitative variables.*

Behavioral and social scientists use two main kinds of numeric variables: equal-interval variables and rank-order variables. The kind of variable used most often is a variable in which the numbers stand for approximately equal amounts of what is being measured. This is called an **equal-interval variable.** Take grade point average (GPA). This is a roughly equal-interval variable. For example, the difference between a GPA of 2.5 and 2.8 means about as much of a difference as the difference between that of 3.0 and 3.3 (each is a difference of 0.3 of a GPA). Most behavioral and social scientists also consider scales like the 0 to 10 stress ratings as roughly equal interval. So, for example, a difference between stress ratings of 4 and 6 means about as much as the difference between 7 and 9.

The other kind of numeric variable behavioral and social scientists often use is where the numbers only stand for relative rankings. This is called a **rank-order variable.** (Rank-order variables are also called *ordinal variables.*) An example is rank in one's graduating class. Notice that with a rank-order variable, the difference between one number and the next does not always mean the same amount of the underlying thing being measured. For example, the difference between being second and third in your graduating class could be a very unlike amount of difference in underlying GPA than the difference between being eighth and ninth. A rank-order variable provides less information than an equal-interval variable. It is less precise. However, behavioral and social scientists often use rank-order variables because they are the only information available.

numeric variable Variable whose values are numbers. Also called *quantitative variable.*

equal-interval variable A variable in which the numbers stand for approximately equal amounts of what is being measured.

rank-order variable Numeric variable in which the values are ranks, such as class standing or place finished in a race. Also called *ordinal variable.*

Table 2 Levels of Measurement

Level	Definition	Example
Equal-interval	Numeric variable in which differences between values correspond to differences in the underlying thing being measured	Stress level, age
Rank-order	Numeric variable in which values correspond to the relative position of things measured	Class standing, position finished in a race
Nominal	Variable in which the values are categories	Gender, religion

There is also a kind of variable that is not about numbers at all, but that refers just to names or categories. This is called a **nominal variable.** The term *nominal* comes from the idea that its values are names. (Nominal variables are also called *categorical variables* because their values are categories.) For example, for the nominal variable gender, the values are female and male. A person's "score" on the variable gender is one of these two values. Similarly, hair color has values, such as brown, black, and so forth.

These different kinds of variables are based on different **levels of measurement** (see Table 2). Researchers sometimes have to decide how they will measure a particular variable. For example, they might use an equal-interval scale, a rank-order scale, or a nominal scale. The level of measurement selected affects the type of statistics that can be used with a variable. We focus mostly on numeric equal-interval variables. However, rank-order and nominal variables also are fairly common in the behavioral and social sciences.

How are you doing?

1. A father rates his daughter as a 2 on a 7-point scale (from 1 to 7) of crankiness. In this example, (a) what is the variable, (b) what is the score, and (c) what is the range of possible values?
2. What is the difference between a numeric and a nominal variable?
3. Name the kind of variable for each of the following variables: (a) a person's nationality (Mexican, French, Japanese, etc.), (b) a person's score on a standardized IQ test, (c) a person's place on a waiting list (first in line, second in line, etc.).

Answers

1. (a) crankiness, (b) 2, (c) 1 to 7.

2. A numeric variable has values that are numbers that tell you the degree or extent of what the variable measures; a nominal variable has values that are different categories and have no particular numerical order.

3. (a) nominal, (b) equal-interval, (c) rank-order.

nominal variable Variable with values that are categories (that is, they are names rather than numbers). Also called *categorical variable.*

level of measurement Type of underlying numerical information provided by a measure, such as equal-interval, rank-order, and nominal (categorical).

Frequency Tables

An Example

Let's return to the stress rating example. Recall that in this study, students in an introductory statistics class during the first week of the course answered the question "How stressed have you been in the last 2½ weeks, on a scale of 0 to 10, with 0 being *not at*

all stressed and 10 being *as stressed as possible?*" The actual study included scores from 151 students. To ease the learning for this example, we are going to use a representative subset of scores from 30 of these 151 students (this also saves you time if you want to try it for yourself). The 30 students' scores (their ratings on the scale) are:

8, 7, 4, 10, 8, 6, 8, 9, 9, 7, 3, 7, 6, 5, 0, 9, 10, 7, 7, 3, 6, 7, 5, 2, 1, 6, 7, 10, 8, 8.

Looking through all these scores gives some sense of the overall tendencies. But this is hardly an accurate method. One solution is to make a table showing how many students used each of the 11 values that the ratings can have (0, 1, 2, etc., through 10). That is, the number of students who used each particular rating is the *frequency* of that value. We have done this in Table 3. We also figured the percentage each value's frequency is of the total number of scores. Tables like this sometimes give only the raw-number frequencies and not the percentages, or only the percentages and not the raw-number frequencies.[1]

Table 3 is called a **frequency table** because it shows how frequently (how many times) each rating number was used. A frequency table makes the pattern of numbers easy to see. In this example, you can see that most of the students rated their stress around 7 or 8, with few rating it very low.

Table 3 Frequency Table of Number of Students Rating Each Value of the Stress Scale

Stress Rating	Frequency	Percent
0	1	3.3
1	1	3.3
2	1	3.3
3	2	6.7
4	1	3.3
5	2	6.7
6	4	13.3
7	7	23.3
8	5	16.7
9	3	10.0
10	3	10.0

Source: Data based on Aron et al. (1995).

How to Make a Frequency Table

There are four steps in making a frequency table.

1. **Make a list of each possible value down the left edge of a page, starting from the lowest and ending with the highest.** In the stress rating results, the list goes from 0, the lowest possible rating, up to 10, the highest possible rating. Note that even if one of the ratings between 0 and 10 had not been used, you would still include that value in the listing, showing it as having a frequency of 0. For example, if no one gave a stress rating of 2, you still include 2 as one of the values on the frequency table.
2. **Go one by one through the scores, making a mark for each next to its value on your list.** This is shown in Figure 1.
3. **Make a table showing how many times each value on your list was used.** To do this, add up the number of marks beside each value.
4. **Figure the percentage of scores for each value.** To do this, take the frequency for that value, divide it by the total number of scores, and multiply by 100. You usually will need to round off the percentage. We recommend that you round percentages to one decimal place. Note that because of the rounding, your percentages will not usually add up to exactly 100% (but the total should be close).

Frequency Tables for Nominal Variables

The preceding steps assume you are using numeric variables, the most common situation. However, you can also use a frequency table to show the number of scores in each value (that is, for each category) of a nominal variable. For example, researchers

frequency table Ordered listing of the number of individuals having each of the different values for a particular variable.

[1]In addition, some frequency tables include, for each value, the total number of scores with that value and all values preceding it. These are called *cumulative frequencies* because they tell how many scores are accumulated up to this point on the table. If percentages are used, cumulative percentages also may be included. Cumulative percentages give, for each value, the percentage of scores up to and including that value. The cumulative percentage for any given value (or for a score having that value) is also called a *percentile.* Cumulative frequencies and cumulative percentages help you see where a particular score falls in the overall group of scores.

TIP FOR SUCCESS

When doing Step ❷, cross off each score as you mark it on the list. This should help you to avoid mistakes, which are common in this step.

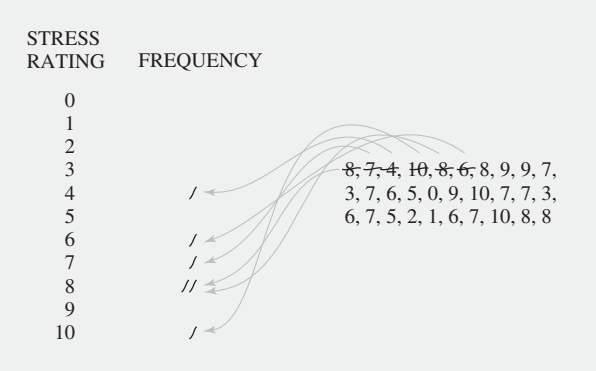


Figure 1 Making a frequency table for the stress ratings scores. (Data based on Aron, Paris, and Aron, 1995.)

Table 4 Frequency Table for a Nominal Variable: Closest Person in Life for 208 Students

Closest Person	Frequency	Percent
Family member	33	15.9
Nonromantic friend	76	36.5
Romantic partner	92	44.2
Other	7	3.4

Source: Data from Aron, Aron, and Smollan (1992).

(Aron, Aron, & Smollan, 1992) asked 208 students to name the closest person in their lives. As shown in Table 4, 33 students selected a family member, 76 a nonromantic friend, 92 a romantic partner, and 7 selected some other person. As you can see in Table 4, the values listed on the left-hand side of the frequency table are the values (the categories) of the variable.

Another Example

McLaughlin-Volpe, Aron, and Reis (2001) had 94 first- and second-year university students keep a diary of their social interactions for a week during the regular semester. Each time a student had a social interaction lasting 10 minutes or longer, the student would fill out a card. The card included questions about who were the other people in the interaction and about various aspects of the conversation. Excluding family and work situations, the number of social interactions of 10 minutes or longer over a week for these 94 students were as follows:

48, 15, 33, 3, 21, 19, 17, 16, 44, 25, 30, 3, 5, 9, 35, 32, 26, 13, 14, 14, 47, 47, 29, 18, 11, 5, 19, 24, 17, 6, 25, 8, 18, 29, 1, 18, 22, 3, 22, 29, 2, 6, 10, 29, 10, 21, 38, 41, 16, 17, 8, 40, 8, 10, 18, 7, 4, 4, 8, 11, 3, 23, 10, 19, 21, 13, 12, 10, 4, 17, 11, 21, 9, 8, 7, 5, 3, 22, 14, 25, 4, 11, 10, 18, 1, 28, 27, 19, 24, 35, 9, 30, 8, 26.

Now, let's follow our four steps for making a frequency table.

❶ **Make a list of each possible value down the left edge of a page, starting from the lowest and ending with the highest.** The lowest possible number of interactions is 0. In this study, the highest number of interactions could be any number. However, the highest actual number in this group was 48; so we can use 48 as the highest value. Thus, the first step is to list these values down a page. (It might be good to use several columns so that you can have all the scores on a single page.)

❷ **Go one by one through the scores, making a mark for each next to its value on your list.** Figure 2 shows the results of this step.

❸ **Make a table showing how many times each value on your list was used.** Table 5 is the result.

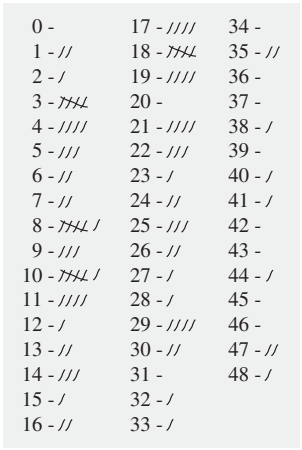


Figure 2 Making a frequency table of students' social interactions over a week. (Data from McLaughlin-Volpe et al., 2001.)

Table 5 Frequency Table for Number of Social Interactions During a Week for 94 College Students

Score	Frequency	Score	Frequency	Score	Frequency
0	0	17	4	34	0
1	2	18	5	35	2
2	1	19	4	36	0
3	5	20	0	37	0
4	4	21	4	38	1
5	3	22	3	39	0
6	2	23	1	40	1
7	2	24	2	41	1
8	6	25	3	42	0
9	3	26	2	43	0
10	6	27	1	44	1
11	4	28	1	45	0
12	1	29	4	46	0
13	2	30	2	47	2
14	3	31	0	48	1
15	1	32	1		
16	2	33	1		

Source: Data from McLaughlin-Volpe et al. (2001).

TIP FOR SUCCESS

Be sure to check your work by adding the frequencies for all of the scores. This sum should equal the total number of scores you started with.

❹ **Figure the percentage of scores for each value.** We have *not* done so in this example because with so many categories, it would not help much for seeing the pattern of scores. However, if you want to check your understanding of this step, the first five percentages would be 0.0%, 2.1%, 1.1%, 5.3%, and 4.3%. (These are the percentages for frequencies of 0, 2, 1, 5, and 4, rounded to one decimal place.)

Grouped Frequency Tables

Sometimes there are so many possible values that a frequency table is too awkward to give a simple picture of the scores. The last example was a bit like that, wasn't it? The solution is to make groupings of values that include all values in a certain range. For example, consider our stress example. Instead of having a separate frequency figure for the group of students who rated their stress as 8 and another for those who rated it as 9, you could have a combined category of 8 and 9. This combined category is a range of values that includes these two values. A combined category like this is called an **interval.** This particular interval of 8 and 9 has a frequency of 8 (the sum of the 5 scores with a value of 8 and the 3 scores with a value of 9).

A frequency table that uses intervals is called a **grouped frequency table.** Table 6 is a grouped frequency table for the stress ratings example. (However, in this example, the full frequency table has only 11 different values. Thus, a grouped frequency table was not really necessary.) Table 7 is a grouped frequency table for the 94 students' numbers of social interactions over a week.

A grouped frequency table can make information even more directly understandable than an ordinary frequency table can. Of course, the greater understandability of a grouped frequency table is at a cost. You lose information about the breakdown of frequencies in each interval.

Table 6 Grouped Frequency Table for Stress Ratings

Stress Rating Interval	Frequency	Percent
0–1	2	6.7
2–3	3	10.0
4–5	3	10.0
6–7	11	36.7
8–9	8	26.7
10–11	3	10.0

Source: Data based on Aron et al. (1995).

interval In a grouped frequency table, the range of values that are grouped together. (For example, if the interval size is 10, one of the intervals might be from 10 to 19.)

grouped frequency table Frequency table in which the number of individuals (frequency) is given for each interval of values.

Table 7 Grouped Frequency Table for Number of Social Interactions During a Week for 94 College Students

Interval	Frequency	Percent
0–4	12	12.8
5–9	16	17.0
10–14	16	17.0
15–19	16	17.0
20–24	10	10.6
25–29	11	11.7
30–34	4	4.3
35–39	3	3.2
40–44	3	3.2
45–49	3	3.2

Source: Data from McLaughlin-Volpe et al. (2001).

When setting up a grouped frequency table, it makes a big difference how many intervals you use. There are guidelines to help researchers with this, but in practice it is done automatically by the researcher's computer (see this chapter's "Using SPSS" section for instructions on how to create frequency tables using statistical software). Thus, we will not focus on it. However, should you have to make a grouped frequency table on your own, the key is to experiment with the interval size until you come up with one that is a round number (such as 2, 3, 5, or 10) and that creates about 5–15 intervals. Then, when actually setting up the table, be sure you set the start of each interval to a multiple of the interval size and the top end of each interval to the number just below the start of the next interval. For example, Table 6 uses six intervals with an interval size of 2. The intervals are 0–1, 2–3, 4–5, 6–7, 8–9, and 10–11. Note that each interval starts with a multiple of 2 (0, 2, 4, 6, 8, 10) and the top end of each interval (1, 3, 5, 7, 9) is the number just below the start of the next interval (2, 4, 6, 8, 10). Table 7 uses 10 intervals with an interval size of 5. The intervals are 0–4, 5–9, 10–14, 15–19, and so on, with a final interval of 45–49. Note that each interval starts with a multiple of 5 (0, 5, 10, 15, etc.) and that the top end of each interval (4, 9, 14, 19, etc.) is the number just below the start of the next interval (5, 10, 15, 20, etc.).

How are you doing?

1. What is a frequency table?
2. Why would a researcher want to make a frequency table?
3. Make a frequency table for the following scores: 5, 7, 4, 5, 6, 5, 4
4. What does a grouped frequency table group?

Answers

1. A frequency table is a systematic listing of the number of scores (the frequency) of each value in the group studied.
2. A frequency table makes it easy to see the pattern in a large group of scores.
3.

Value	Frequency	Percent
4	2	28.6
5	3	42.9
6	1	14.3
7	1	14.3

4. A grouped frequency table groups the frequencies of adjacent values into intervals.

Histograms

histogram Barlike graph of a frequency distribution in which the values are plotted along the horizontal axis and the height of each bar is the frequency of that value; the bars are usually placed next to each other without spaces, giving the appearance of a city skyline.

A graph is another good way to make a large group of scores easy to understand. A picture may be worth a thousand words but it is also sometimes worth a thousand numbers. A straightforward approach is to make a graph of the frequency table. One kind of graph of the information in a frequency table is a kind of bar chart called a **histogram.** In a histogram, the height of each bar is the frequency of each value in the frequency table. Ordinarily, in a histogram all the bars are put next to each other with no space in between. The result is that a histogram looks a bit like a city skyline. Figure 3 shows two histograms based on the stress ratings example, one based on

(a) Frequency Table

❶

STRESS RATING	FREQUENCY
0	1
1	1
2	1
3	2
4	1
5	2
6	4
7	7
8	5
9	3
10	3

(b) Grouped Frequency Table

❶

STRESS RATING INTERVAL	FREQUENCY
0 – 1	2
2 – 3	3
4 – 5	3
6 – 7	11
8 – 9	8
10 – 11	3

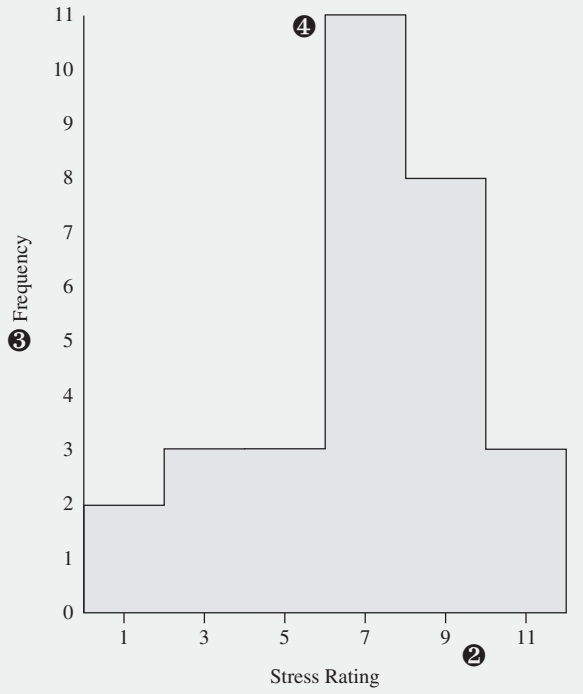


Figure 3 Four steps in making a histogram based on (a) a frequency table and (b) a grouped frequency table for the stress ratings example. (Data based on Aron, Paris, and Aron, 1995.) ❶ Make a frequency table. ❷ Put the values along the bottom of the page. ❸ Make a scale of frequencies along the left edge of the page. ❹ Make a bar for each value.

INTERVAL	FREQUENCY
0 – 4	12
5 – 9	16
10 – 14	16
15 – 19	16
20 – 24	10
25 – 29	11
30 – 34	4
35 – 39	3
40 – 44	3
45 – 49	3

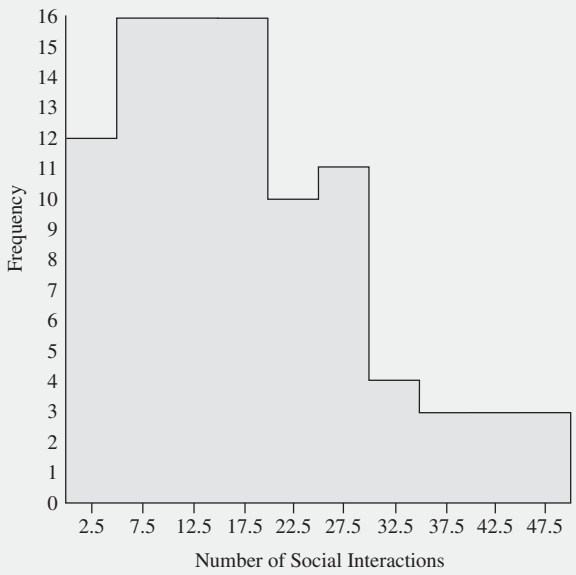


Figure 4 Histogram for number of social interactions during a week for 94 college students, based on grouped frequencies. (Data from McLaughlin-Volpe et al., 2001.)

the ordinary frequency table and one based on the grouped frequency table. Figure 4 shows a histogram based on the grouped frequency table for the example of the numbers of students' social interactions in a week.

How to Make a Histogram There are four steps in making a histogram.

1. **Make a frequency table (or grouped frequency table).**
2. **Put the values along the bottom of the page.** The numbers should go from left to right, from lowest to highest. If you are making a histogram from a grouped frequency table, the values you put along the bottom of the page are the interval midpoints. The midpoint of an interval is halfway between the start of that interval and the start of the next highest interval. So, in Figure 4, the midpoint for the 0–4 interval is 2.5, because 2.5 is halfway between 0 (the start of the interval) and 5 (the start of the next highest interval). For the 5–9 interval, the midpoint is 7.5, because 7.5 is halfway between 5 (the start of the interval) and 10 (the start of the next highest interval). Do this for each interval. When you get to the last interval, find the midpoint between the start of the interval and the start of what would be the next highest interval. So, in Figure 4, the midpoint for the 45–49 interval is halfway between 45 (the start of the interval) and 50 (the start of what would be the next interval), which is 47.5.
3. **Make a scale of frequencies along the left edge of the page.** The scale should go from 0 at the bottom to the highest frequency for any value.
4. **Make a bar for each value.** The height of each bar is the frequency of the value over which it is placed. For each bar, make sure that the middle of the bar is above its value.

TIP FOR SUCCESS

Now try this yourself! Work out the interval midpoints for the grouped frequency table shown in Table 6. Your answers should be the same as the values shown along the bottom of Figure 3b.

TIP FOR SUCCESS

You will probably find it easier to make histograms if you use graph paper.

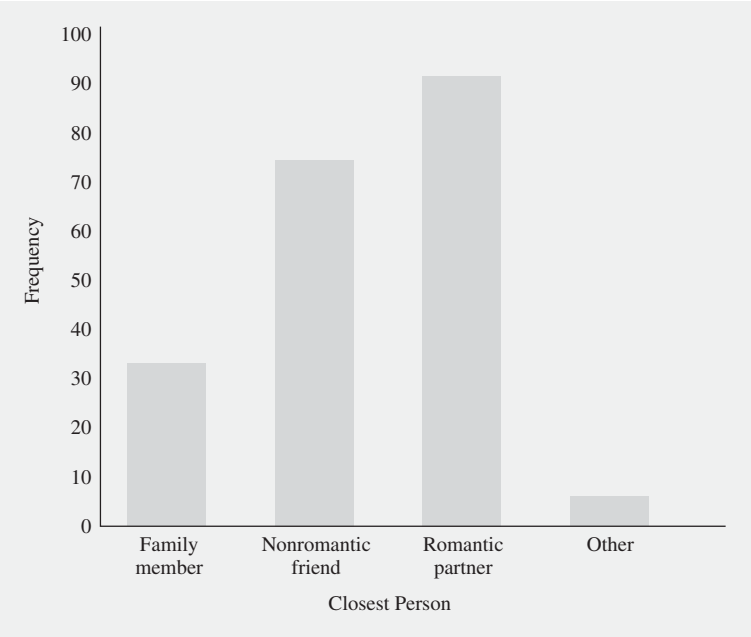


Figure 5 Bar graph for the closest person in life for 208 students (see Table 4). (Data from Aron et al., 1995.)

When you have a nominal variable, the histogram is called a bar graph. Since the values of a nominal variable are not in any particular order, you leave a space in between the bars. Figure 5 shows a bar graph based on the frequency table in Table 4.

How are you doing?

1. Why do researchers make histograms?
2. When making a histogram from a frequency table, (a) what goes along the bottom, (b) what goes along the left edge, and (c) what goes above each value?
3. Make a histogram based on the frequency table below.

Value	Frequency
1	3
2	4
3	8
4	5
5	2

Answers

1. Researchers make histograms to show the pattern visually in a frequency table.
2. (a) The values, from lowest to highest go along the bottom; (b) the frequencies, from 0 at the bottom to the highest frequency of any value at the top, go along the left edge; (c) above each value is a bar with a height of the frequency for that value.
3. See Figure 6.

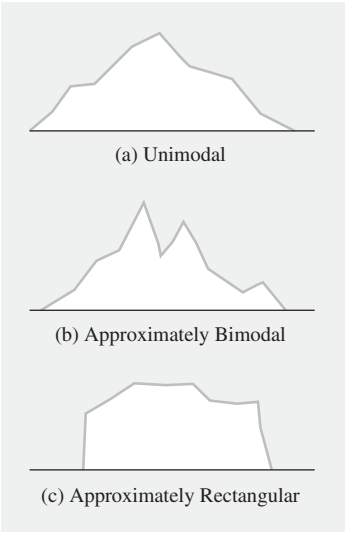


Figure 7 Examples of (a) unimodal, (b) approximately bimodal, and (c) approximately rectangular frequency polygons.

frequency distribution Pattern of frequencies over the various values; what a frequency table or histogram describes.

unimodal distribution Frequency distribution with one value clearly having a larger frequency than any other.

bimodal distribution Frequency distribution with two approximately equal frequencies, each clearly larger than any of the others.

multimodal distribution Frequency distribution with two or more high frequencies separated by a lower frequency; a bimodal distribution is the special case of two high frequencies.

rectangular distribution Frequency distribution in which all values have approximately the same frequency.

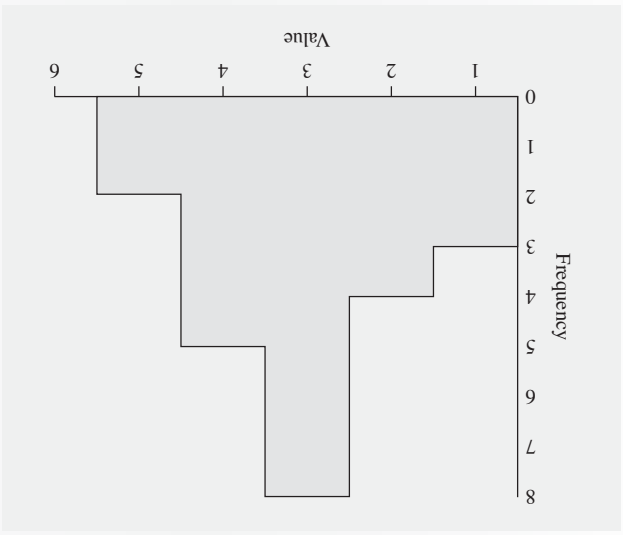


Figure 6 Histogram for "How Are You Doing?" question 3.

Shapes of Frequency Distributions

A **frequency distribution** shows the pattern of frequencies over the various values. A frequency table or histogram describes a frequency distribution because each shows the pattern or shape of how the frequencies are spread out, or "distributed." Behavioral and social scientists also describe this shape in words. Describing the shape of a distribution is important both in the descriptive statistics we focus on in this chapter and also in inferential statistics.

Unimodal and Bimodal Frequency Distributions

One important aspect of a distribution's shape is whether it has only one main high point: one high "tower" in the histogram. For example, in the stress ratings study, the most frequent score is a 7, giving a histogram with only one very high area. This is called a **unimodal distribution.** If a distribution has two fairly equal high points, it is called a **bimodal distribution.** Any distribution with two or more high points is called a **multimodal distribution.** Finally, if all the values have about the same frequency, it is called a **rectangular distribution.** Figure 7 shows examples of these frequency distribution shapes. As you will see, the graphs in Figure 7 are not histograms, but special line graphs called *frequency polygons,* which are another way to graph a frequency table. In a frequency polygon, the lines move from point to point. The height of each point shows the number of scores with that value. This creates a mountain peak skyline.

The information we collect in behavioral and social science research is usually approximately unimodal. Bimodal and other multimodal distributions occasionally turn up. A bimodal example is the distribution of the ages of people in a toddler's play area in a park, who are mostly either toddlers with ages of around 2–4 years or caretakers with ages of 20–40 years or so (with few infants, a few people aged

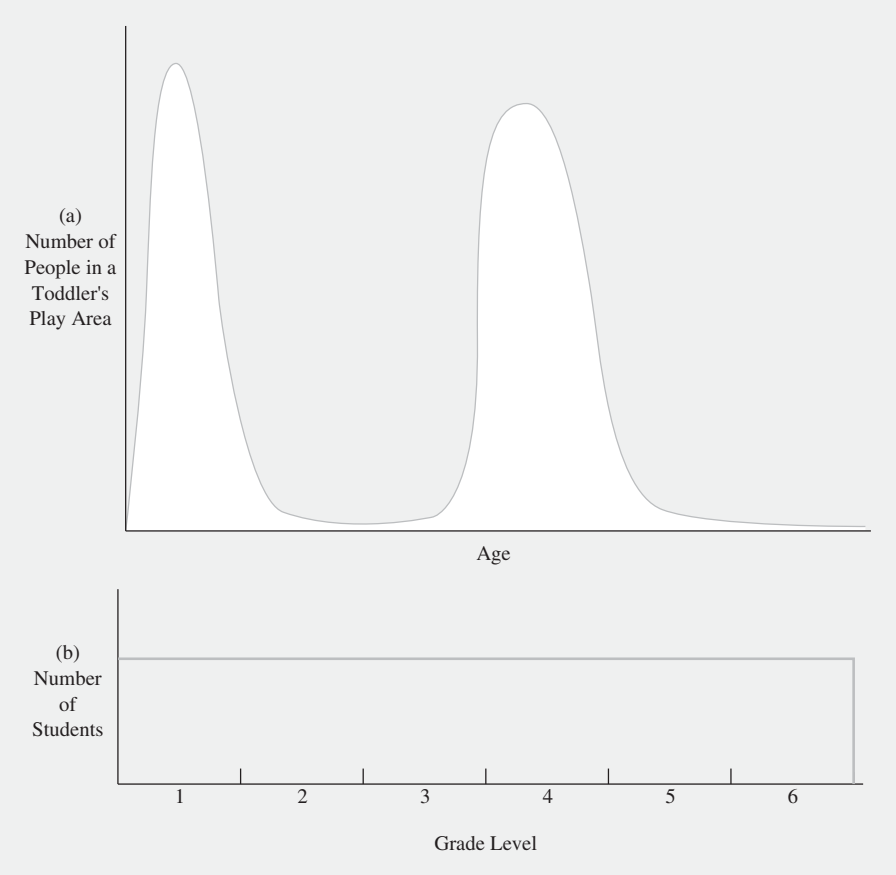


Figure 8 Fictional examples of distributions that are not unimodal: (a) A bimodal distribution showing the possible frequencies for people of different ages in a toddler's play area. (b) A rectangular distribution showing the possible frequencies of students at different grade levels in an elementary school.

5–19 years, and a few above 40). Thus, if you make a frequency distribution of these ages, the large frequencies would be at the values for toddler ages (2–4 years) and for higher ages (20–40 years or so). An example of a rectangular distribution is the number of children at each grade level attending an elementary school: there is about the same number in first grade, second grade, and so on. Figure 8 shows these examples.

Symmetrical and Skewed Distributions

Look again at the histograms of the stress rating example (Figure 3). The distribution is lopsided, with more scores near the high end. This is somewhat unusual. Most things we measure in the behavioral and social sciences have about equal numbers on both sides of the middle. That is, most of the time, the scores follow an approximately **symmetrical distribution** (if you fold the graph of a symmetrical distribution in half, the two halves look the same).

A distribution that is clearly not symmetrical is called a **skewed distribution.** The stress ratings distribution is an example. A skewed distribution has

symmetrical distribution Distribution in which the pattern of frequencies on the left and right side are mirror images of each other.

skewed distribution Distribution in which the scores pile up on one side of the middle and are spread out on the other side; distribution that is not symmetrical.

Figure 9 Examples of frequency polygons of distributions that are (a) approximately symmetrical, (b) skewed to the right (positively skewed), and (c) skewed to the left (negatively skewed).

TIP FOR SUCCESS

You may be interested to know that the word *skew* comes from the French *queue,* which means "line or tail." This should help you remember that the direction of the skew (to the left or right) is the side that has the long line or tail.

one side that is long and spread out, somewhat like a tail. The side with *fewer* scores (the side that looks like a tail) describes the direction of the skew. Thus a distribution with fewer scores left of the peak, like our stress ratings example, is *skewed to the left.* The other example we have examined in this chapter, the distributions of students' numbers of interactions in a week, is *skewed to the right.* Figure 9 shows examples of approximately symmetrical and skewed distributions.

A distribution that is skewed to the right is also called *positively skewed.* A distribution skewed to the left is also called *negatively skewed.*

Strongly skewed distributions come up in the behavioral and social sciences mainly when what is being measured has some lower or upper limit. For example, a family cannot have fewer than zero children. When many scores pile up at the low end because it is impossible to have a lower score, the result is called a **floor effect.** A skewed distribution caused by a lower limit is shown in Figure 10a.

floor effect Situation in which many scores pile up at the low end of a distribution (creating skewness to the right) because it is not possible to have any lower score.

A skewed distribution caused by an upper limit is shown in Figure 10b. This is a distribution of adults' scores on a multiplication table test. This distribution

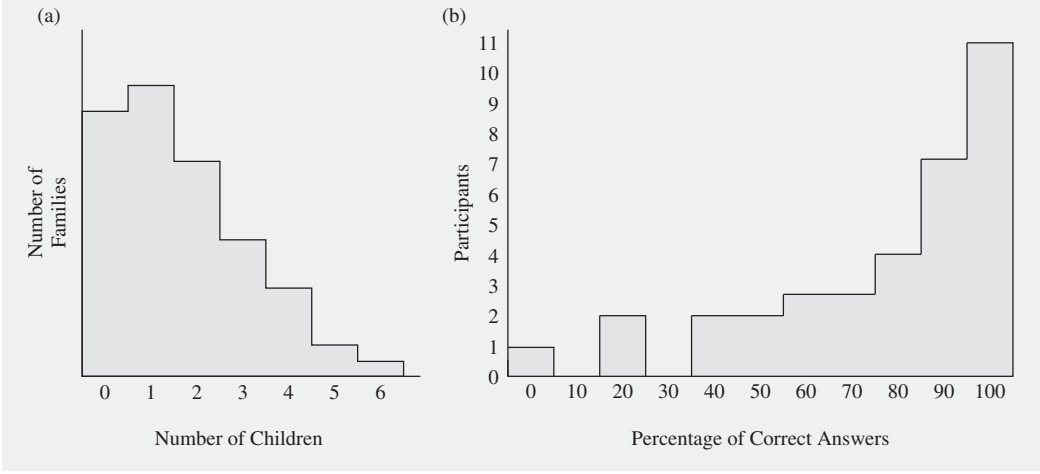


Figure 10 (a) A distribution skewed to the right due to a floor effect: fictional distribution of the number of children in families. (b) A distribution skewed to the left due to a ceiling effect: fictional distribution of adults' scores on a multiplication table test.

is strongly skewed to the left. Most of the scores pile up at the right, the high end (a perfect score). This is an example of a **ceiling effect.** The stress ratings example also shows a mild ceiling effect. It is a ceiling effect because many students had high levels of stress and the maximum rating was 10. It was only a mild ceiling effect, in part, because people often do not like to use ratings right at the maximum.

Normal, Heavy-Tailed, and Light-Tailed Distributions

Behavioral and social scientists also describe a distribution in terms of whether its tails are particularly "heavy" (thick, with many scores in them) or "light" (thin, with few scores in them). These are called **heavy-tailed distributions** and **light-tailed distributions.** (This aspect of the shape of a distribution is also called *kurtosis.*) The standard of comparison is a bell-shaped curve. In behavioral and social science research and in nature generally, distributions often are quite similar to this bell-shaped standard, called the **normal curve.** The normal curve is a unimodal, symmetrical curve with average tails—the sort of bell shape shown in Figure 11a. The stress ratings example in this chapter is very roughly like a normal curve, except that it is somewhat skewed to the left. In our experience, most distributions that result from behavioral and social science research are actually closer to the normal curve than this example.

Figures 11b and 11c show examples of heavy-tailed and light-tailed distributions, with the normal distribution shown as a dashed line. Distributions that have tails that are heavier (thicker) or lighter (thinner) than a normal curve also tend to have a different shape in the middle. Those with heavy tails are usually more peaked than the normal curve (see Figure 11b). It is as if the normal curve got pinched in the middle and some of it went up into a sharp peak and the rest spread out into thick tails. Those with light tails are usually flatter than the normal curve (see Figure 11c). It is as if the tails and the top of the curve both got sucked in to right near the middle on both sides. (An extreme case of a light-tailed distribution would be a rectangular distribution.)

ceiling effect Situation in which many scores pile up at the high end of a distribution (creating skewness to the left) because it is not possible to have a higher score.

heavy-tailed distribution Distribution that differs from a normal curve by being too spread out so that a histogram of the distribution would have too many scores at each of the two extremes ("tails").

light-tailed distribution Distribution that differs from a normal curve by being too peaked or pinched so that a histogram of the distribution would have too few scores at each of the two extremes ("tails").

normal curve Specific, mathematically defined, bell-shaped frequency distribution that is symmetrical and unimodal; distributions observed in nature and in research commonly approximate it.

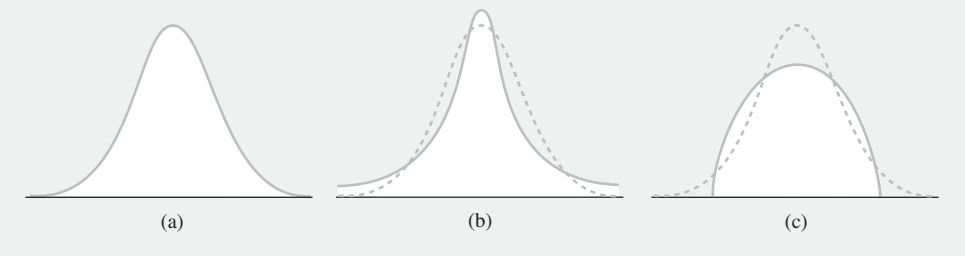


Figure 11 Examples of (a) normal, (b) heavy-tailed, and (c) light-tailed distributions. The normal distribution is shown as a dashed line in (b) and (c). *Source:* Adapted from DeCarlo, T. (1997). On the meaning and use of kurtosis. *Psychological Methods, 3*, 292–307. Copyright © 1997 by the American Psychological Association. Adapted with permission. The use of APA information does not imply endorsement by APA.

BOX 2 Gender, Ethnicity, and Math Performance

Women and minorities often struggle with doubts about their ability to do well academically, especially in math. Many lines of research indicate that prejudices, not genetics, are the probable cause of differences in test scores between women and men or minorities and others. For example, the same difference of 15 IQ points between a dominant and minority group has been found all over the world, even when there is no genetic difference between the groups, and in cases where opportunities for a group have changed, such as when they emigrate, differences have rapidly disappeared (Block, 1995). Hyde and Mertz (2009) found that the more gender equality there is in a society, the more that women are equal to men in their math performance, and the greater likelihood that women are found to have "profound mathematical talent."

However, if you are a woman or a minority and have been made to doubt your mathematical abilities, you are vulnerable to "stereotype threat" (Steele, 1997). That is, if you are in a situation that reminds you of your culturally induced doubts, you will behave more in accordance with that negative stereotype. For example, in a typical experiment creating stereotype threat (Alter, Aronson, Darley, Rodriguez, & Ruble, 2010), fifth- and sixth-grade students were randomly assigned to fill out a form asking for their race before taking their end-of-year mathematics test. African American students who filled out this form answered only about half as many math questions correctly as African American students who did not. Women told before a math test that men generally score better on the test did worse than women told the two genders score equally well (Spencer, Steele, & Quinn, 1999) and African Americans told that there are racial differences in scores on the Graduate Record Exam did worse than those told there are no differences (Steele, 1997).

What Can You Do for Yourself?

If you feel you belong to a group that is expected to do poorly at math, what can you do to get out from under the shadow of stereotype threat as you take this course? In the study above with fifth- and sixth-graders, when the students were made to think of the test as an opportunity rather than a measure of their ability, the effect of the stereotype threat was greatly reduced. These students heard that the test would allow them to "learn a lot of new things" and "be a big help in school because it sharpens the mind" and would help them in their studies. You can do the same for yourself. You are about to learn many new things. Studying for statistics and taking the exams truly will sharpen your mind. Knowing statistics will help you in many other courses and throughout your life. These are the thoughts to keep in mind—not your gender or your race.

In another study (Martens, Johns, Greenberg, & Schimel, 2006), women were told they were taking a test that would diagnose their math abilities (women in the other group were told their input was needed on the test questions to improve them). The effect of the stereotype threat was greatly reduced when women followed a "self-affirmation" procedure in which they listed in order their most valued characteristics, and then wrote about why the highest characteristic on their list was personally important and about a time when it was particularly so. This procedure gave the women a sense of other things about themselves that mattered to them besides the test they were taking. You can do the same thing. Think of what is important to you about yourself and your values. It puts your performance in statistics in perspective, doesn't it?

Finally, we encourage you to fight the activation of prejudices inside or outside of you that could stand in the way of your right to know statistics well. Consider these words from the former president of the Mathematics Association of America:

> The paradox of our times is that as mathematics becomes increasingly powerful, only the powerful seem to benefit from it. The ability to think mathematically—broadly interpreted—is absolutely crucial to advancement in virtually every career. Confidence in dealing with data, skepticism in analyzing arguments, persistence in penetrating complex problems, and literacy in communicating about technical matters are the enabling arts offered by the new mathematical sciences. (Steen, 1987, p. xviii)

Power is another word for *influence.* How would you like to influence the world? See to it that you are enabled by introductory statistics to do all that you wish to do.

How are you doing?

1. Describe the difference between a unimodal and multimodal distribution in terms of (a) a frequency polygon and (b) a frequency table.
2. What does it mean to say that a distribution is skewed to the left?
3. What kind of skew is created by (a) a floor effect and (b) a ceiling effect?
4. When a distribution is described as heavy-tailed or light-tailed, what is it being compared to?

Answers

1. (a) A unimodal distribution has one main high point; a multimodal distribution has more than one main high point. (b) A unimodal distribution has one value with a higher frequency than all the other frequencies; a multimodal distribution has more than one value with large frequencies compared to the values around it.
2. When a distribution is skewed to the left, fewer scores have low values than have high values.
3. (a) A floor effect creates a distribution that is skewed to the right; (b) a ceiling effect creates a distribution that is skewed to the left.
4. The distribution is being compared to the normal curve.

Frequency Tables and Histograms in Research Articles

Frequency tables and histograms are not usually included in research articles. However, they are commonly used by researchers as a first step in more elaborate statistical procedures, particularly when the distribution seems to deviate from normal. The shapes of distributions (normal, skewed, etc.) are occasionally described in words. When frequency tables are included in research articles they are often used to summarize the characteristics of the people in the study. For example, Raskauskas and Stoltz (2007) asked a group of 84 adolescents about their involvement in traditional and electronic bullying. The researchers defined electronic bullying as "... a means of bullying in which peers use electronics [such as text messages, emails, and defaming Web sites] to taunt, threaten, harass, and/or intimidate a peer" (p. 565). Table 8 is a frequency table showing the adolescents' reported incidence of being victims or perpetrators of traditional and electronic bullying. The table shows, for example, that about half (48.8%) of the adolescents reported being the victim of electronic bullying, and the most common vehicle for electronic bullying (experienced by 32.1% of adolescents) was text messaging.

Table 8 Incidence of Traditional and Electronic Bullying and Victimization ($N = 84$)

Form of Bullying	*N*	%
Electronic victims	41	48.8
Text-message victim	27	32.1
Internet victim (Web sites, chatrooms)	13	15.5
Picture-phone victim	8	9.5
Traditional victims	60	71.4
Physical victim	38	45.2
Teasing victim	50	59.5
Rumors victim	32	38.6
Exclusion victim	30	50.0
Electronic bullies	18	21.4
Text-message bully	18	21.4
Internet bully	11	13.1
Traditional bullies	54	64.3
Physical bully	29	34.5
Teasing bully	38	45.2
Rumor bully	22	26.2
Exclusion bully	35	41.7

Source: Raskauskas, J., & Stoltz, A. D. (2007). Involvement in traditional and electronic bullying among adolescents. *Developmental Psychology, 43,* 564–575. Copyright © 2007 by the American Psychological Association. Reproduced with permission. The use of APA information does not imply endorsement by APA.

Learning Aids

Summary

1. Behavioral and social scientists use descriptive statistics to describe and summarize a group of numbers from a research study.
2. A value is a number or category; a variable is a characteristic that can have different values; a score is a particular person's value on the variable.
3. Most variables in the behavioral and social sciences are numeric with approximately equal intervals. However, some numeric variables are rank-ordered (the values are ranks), and some variables are not numeric at all, but are nominal (the values are categories).
4. A frequency table organizes the scores into a table in which each of the possible values is listed from lowest to highest, along with its frequency (number of scores that have that value and percentage).
5. When there are many different values, a grouped frequency table is useful. It is like an ordinary frequency table except that the frequencies are given for intervals that include a range of values.
6. The pattern of frequencies in a distribution can be shown visually with a histogram (or bar graph), in which the height of each bar is the frequency for a particular value.
7. The general shape of a histogram can be unimodal (having a single peak), bimodal (having two peaks), multimodal (having multiple peaks; including bimodal), or

rectangular (having no peak); it can be symmetrical or skewed (having a long tail) to the right or the left; and, compared to the bell-shaped normal curve, it can be light-tailed or heavy-tailed.

8. Frequency tables rarely appear in research articles, but when they do they often summarize the characteristics of the people in the study. Histograms almost never appear in articles, though the shapes of distributions (normal, skewed, etc.) are occasionally described in words.

Key Terms

statistics
descriptive statistics
inferential statistics
variable
value
score
numeric variable
equal-interval variable
rank-order variable
nominal variable
level of measurement
frequency table
interval
grouped frequency table
histogram
frequency distribution
unimodal distribution
bimodal distribution
multimodal distribution
rectangular distribution
symmetrical distribution
skewed distribution
floor effect
ceiling effect
heavy-tailed distribution
light-tailed distribution
normal curve

Example Worked-Out Problems

Ten first-year university students rated their interest in graduate school on a scale from 1 = *no interest at all* to 6 = *high interest*. Their scores were as follows: 2, 4, 5, 5, 1, 3, 6, 3, 6, 6.

Making a Frequency Table

See Figure 12.

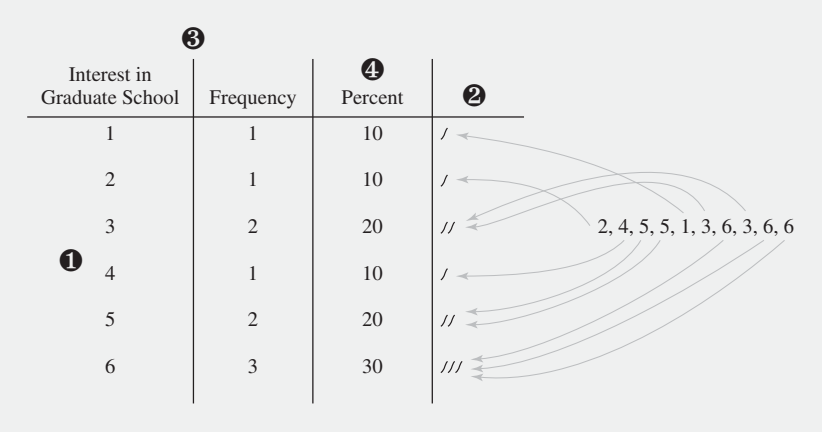


Figure 12 Answer to Example Worked-Out Problem for making a frequency table. ❶ Make a list down the page of each possible value, from lowest to highest. ❷ Go one by one through the scores, making a mark for each next to its value on your list. ❸ Make a table showing how many times each value on your list is used. ❹ Figure the percentage of scores for each value.

Making a Histogram

See Figure 13.

❶

Interest in Graduate School	Frequency
1	1
2	1
3	2
4	1
5	2
6	3

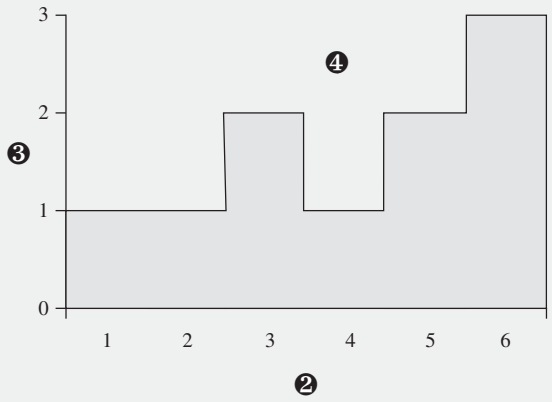


Figure 13 Answer to Example Worked-Out Problem for making a histogram. ❶ Make a frequency table. ❷ Put the values along the bottom of the page (from left to right, from lowest to highest). ❸ Make a scale of frequencies along the left edge of the page (going from 0 at the bottom to the highest frequency for any value). ❹ Make a bar for each value (with a height for the frequency of that value).

Practice Problems

These problems involve tabulation and making graphs. Most real-life statistics problems are done on a computer with special statistical software. Even if you have such software, do these problems by hand to ingrain the method in your mind. To learn how to use a computer to solve statistics problems like those in this chapter, refer to the "Using SPSS" section at the end of this chapter.

All data are fictional unless an actual citation is given.

Set I (for answers, see the end of this chapter)

1. A client rates her satisfaction with her vocational counselor as a "3" on a 4-point scale from 1 = *not at all satisfied* to 4 = *very satisfied.* What are the (a) variable, (b) possible values, and (c) score?
2. You fill out a survey question that asks you to give the likelihood that you will vote in an upcoming election. The question uses a 5-point response scale from 1 = *not at all likely* to 5 = *very likely* and you answer "5." What are the (a) variable, (b) possible values, and (c) score?
3. Name the kind of variable for each of the following: (a) ethnic group to which a person belongs, (b) number of times a mouse makes a wrong turn in a laboratory maze, and (c) position one finishes in a race.
4. A particular block in a suburban neighborhood has 20 households. The number of children in these households is as follows:

   2, 4, 2, 1, 0, 3, 6, 0, 1, 1, 2, 3, 2, 0, 1, 2, 1, 0, 2, 2

Make (a) a frequency table and (b) a histogram. Then (c) describe the general shape of the distribution.

5. Fifty students were asked how many hours they had studied this weekend. Here are their answers:

   11, 2, 0, 13, 5, 7, 1, 8, 12, 11, 7, 8, 9, 10, 7, 4, 6, 10, 4, 7, 8, 6, 7, 10, 7, 3, 11, 18, 2, 9, 7, 3, 8, 7, 3, 13, 9, 8, 7, 7, 10, 4, 15, 3, 5, 6, 9, 7, 10, 6

   Make (a) a frequency table and (b) a histogram. Then (c) describe the general shape of the distribution.

6. Following are the speeds of 40 cars clocked by radar on a particular road in a 35-mph zone on a Sunday afternoon:

   30, 36, 42, 36, 30, 52, 36, 34, 36, 33, 30, 32, 35, 32, 37, 34, 36, 31, 35, 20, 24, 46, 23, 31, 32, 45, 34, 37, 28, 40, 34, 38, 40, 52, 31, 33, 15, 27, 36, 40

   Make (a) a frequency table and (b) a histogram. Then (c) describe the general shape of the distribution.

7. These are the scores on a measure of sensitivity to smell taken by 25 chefs attending a national conference:

   96, 83, 59, 64, 73, 74, 80, 68, 87, 67, 64, 92, 76, 71, 68, 50, 85, 75, 81, 70, 76, 91, 69, 83, 75

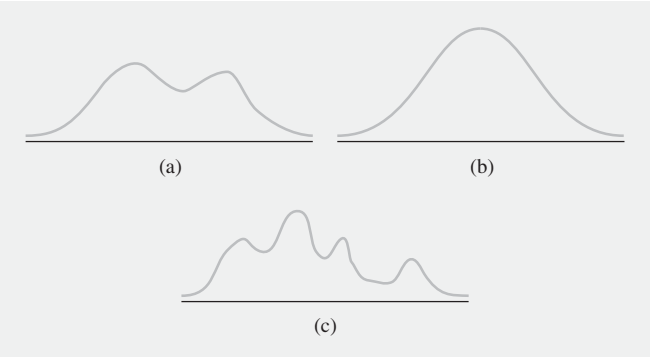
   Make (a) a frequency table and (b) histogram. (c) Make a grouped frequency table using intervals of 50–59, 60–69, 70–79, 80–89, and 90–99. Based on the grouped frequency table, (d) make a histogram and (e) describe the general shape of the distribution.

8. The following data are the number of minutes it took each of 34 10-year-olds to do a series of abstract puzzles:

   24, 83, 36, 22, 81, 39, 60, 62, 38, 66, 38, 36, 45, 20, 20, 67, 41, 87, 41, 82, 35, 82, 28, 80, 80, 68, 40, 27, 43, 80, 31, 89, 83, 24

   Make (a) a frequency table and (b) a grouped frequency table using intervals of 20–29, 30–39, 40–49, 50–59, 60–69, 70–79, and 80–89. Based on the grouped frequency table, (c) make a histogram and (d) describe the general shape of the distribution.

9. Describe the shapes of the three distributions illustrated.



10. Explain to a person who has never had a course in statistics what is meant by (a) a symmetrical unimodal distribution and (b) a negatively skewed unimodal distribution. (Be sure to include in your first answer an explanation of what "distribution" means.)
11. What is a ceiling effect?

12. McKee and Ptacek (2001) asked 90 college students about a time they had "delivered bad news" to someone. Table 9 (their Table 1) shows the results for the type of bad news given. (a) Using this table as an example, explain the idea of a frequency table to a person who has never had a course in statistics. (b) Explain the general meaning of the pattern of results.

Table 9 Descriptive Statistics for the Type of News Given

Category	Frequency	Percentage
1. Relationship with family	19	21.1
2. School	1	1.1
3. Job/work	6	6.7
4. Relationship with actual/potential girlfriend/boyfriend	17	18.9
5. Personal health	1	1.1
6. Finance	1	1.1
7. Relationship with friends	21	23.3
8. Health of family member/friend	23	25.6
9. Other	1	1.1

Source: McKee, T. L. E., & Ptacek, J. T. (2001). I'm afraid I have something bad to tell you: Breaking bad news from the perspective of the giver. *Journal of Applied Social Psychology, 31,* 246–273. Copyright © 2006, 2001 V. H. Winston. Reprinted by permission of Blackwell Publishing.

Set II

13. A participant in a research study is given 50 words to remember and later asked to recall as many of the words as she can. This participant recalls 17. What is the (a) variable, (b) possible values, and (c) score?
14. Explain and give an example for each of the following types of variables: (a) equal-interval, (b) rank-order, and (c) nominal.
15. Here are the number of children in each of 30 classrooms in an elementary school:

    24, 20, 35, 25, 25, 22, 26, 28, 38, 15, 25, 21, 24, 25, 25, 24,
    25, 20, 32, 25, 22, 26, 26, 28, 24, 22, 26, 21, 25, 24

    Make (a) a frequency table and (b) a histogram. Then (c) describe the general shape of the distribution.
16. Pick a book and page number of your choice (select a page with at least 30 lines; *do not pick a textbook or any book with tables or illustrations*). Make a list of the number of words on each line; use that list as your group of scores. Make (a) a frequency table and (b) a histogram. Then (c) describe the general shape of the distribution. (Be sure to give the name, author, publisher, and year of the book you used, along with the page number, with your answer.)
17. An organizational researcher asks 20 employees in a company to rate their job satisfaction on a 5-point scale from 1 = *very unsatisfied* to 5 = *very satisfied.* The ratings were as follows:

    3, 2, 3, 4, 1, 3, 3, 4, 5, 2, 3, 5, 2, 3, 3, 4, 1, 3, 2, 4

    Make (a) a frequency table and (b) a histogram. Then (c) describe the general shape of the distribution.
18. A researcher asked 15 college students how many times they "fell in love" before they were 11 years old. The numbers of times were as follows:

    2, 0, 6, 0, 3, 1, 0, 4, 9, 0, 5, 6, 1, 0, 2

Make (a) a frequency table and (b) a histogram. Then (c) describe the general shape of the distribution.

19. Here are the number of holiday gifts purchased by 25 families randomly interviewed at a local mall at the end of the holiday season:

    22, 18, 22, 26, 19, 14, 23, 27, 2, 18, 28, 28, 11, 16,
    34, 28, 13, 21, 32, 17, 6, 29, 23, 22, 19

    Make (a) a frequency table and (b) a grouped frequency table using intervals of 0–4, 5–9, 10–14, 15–19, 20–24, 25–29, and 30–34. Based on the grouped frequency table, (c) make a histogram and (d) describe the general shape of the distribution.
20. Explain to a person who has never taken a course in statistics the meaning of a grouped frequency table.
21. Nownes (2000) surveyed representatives of interest groups who were registered as lobbyists of three U.S. state legislatures. One of the issues he studied was whether interest groups are in competition with each other. Table 10 (Nownes's Table 1) shows the results for one such question. (a) Using this table as an example, explain the idea of a frequency table to a person who has never had a course in statistics. (b) Explain the general meaning of the pattern of results.

Table 10 Competition for Members and Other Resources

	Question: How much competition does this group face from other groups with similar goals for members and other resources?	
Answer	**Percentage**	**Number**
No competition	20	118
Some competition	58	342
A lot of competition	22	131
Total	100	591

Note: There were no statistically significant differences between states. For full results of significance tests, contact the author.
Source: Nownes, A. J. (2000). Policy conflict and the structure of interest communities. *American Politics Quarterly, 28,* 309–327. Copyright © 2002, 2000 American Psychological Society. Reprinted by permission of Sage Publications, Inc.

22. Mouradian (2001) surveyed college students selected from a screening session to include two groups: (a) "Perpetrators"—students who reported at least one violent act (hitting, shoving, etc.) against their partner in their current or most recent relationship—and (b) "Comparisons"—students who did not report any such uses of violence in any of their last three relationships. At the actual testing session, the students first read a description of an aggressive behavior, such as "Throw something at his or her partner" or "Say something to upset his or her partner." They then were asked to write "as many examples of circumstances of situations as [they could] in which a person might engage in behaviors or acts of this sort with or toward their significant other." Table 11 (Mouradian's Table 3) shows the "Dominant Category of Explanation" (the category a participant used most) for females and males, broken down by comparisons and perpetrators. (a) Using this table as an example, explain the idea of a frequency table to a person who has never had a course in statistics. (b) Explain the general meaning of the pattern of results.

Table 11 Dominant Category of Explanation for Intimate Aggression by Gender and Perpetrator Status

	Group							
	Female				Male			
	Comparisons (n = 36)		Perpetrators (n = 33)		Comparisons (n = 32)		Perpetrators (n = 25)	
Category	***f***	**%**	***f***	**%**	***f***	**%**	***f***	**%**
Self-defense	2	6	3	9	3	9	1	4
Control motives	8	22	9	27	9	28	3	12
Expressive aggression	4	11	3	9	3	9	8	32
Face/self-esteem preservation	1	3	2	6	2	6	3	12
Exculpatory explanations	5	14	3	9	3	9	3	12
Rejection of perpetrator or act	12	33	6	18	10	31	7	28
Prosocial/acceptable explanations	0	0	0	0	0	0	0	0
Tied categories	4	11	7	21	2	6	0	0

Note: f = frequency. % = percentage of respondents in a given group who provided a particular category of explanation.
Source: Mouradian, V. E. (2001). Applying schema theory to intimate aggression: Individual and gender differences in representation of contexts and goals. *Journal of Applied Social Psychology, 31,* 376–408. Copyright © 2006, 2001 V. H. Winston. Reprinted by permission of Blackwell Publishing.

23. Make up and draw an example of each of the following distributions: (a) symmetrical, (b) rectangular, and (c) skewed to the right.
24. Give an example of something having these distribution shapes: (a) bimodal, (b) approximately rectangular, and (c) positively skewed. Do not use an example given in this text or in class.

Using SPSS

The ✍ in the following steps indicates a mouse click. (We used SPSS version 17.0 to carry out these analyses. The steps and output may be slightly different for other versions of SPSS.)

Creating a Frequency Table

❶ Enter the scores from your distribution in one column of the data window.
❷ ✍ *Analyze.*
❸ ✍ *Descriptive statistics.*
❹ ✍ *Frequencies.*
❺ ✍ the variable you want to make a frequency table of and then ✍ the arrow.
❻ ✍ *OK.*

Practice the steps above by creating a frequency table for the social interactions example in this chapter (the scores are listed on previous page). After Step ❺, your screen should look like Figure 14. Your output window (which will appear after you ✍ *OK* in Step ❻) should look like Figure 15. As you will see, SPSS automatically produces a column with the cumulative percent for each value (see footnote 1).

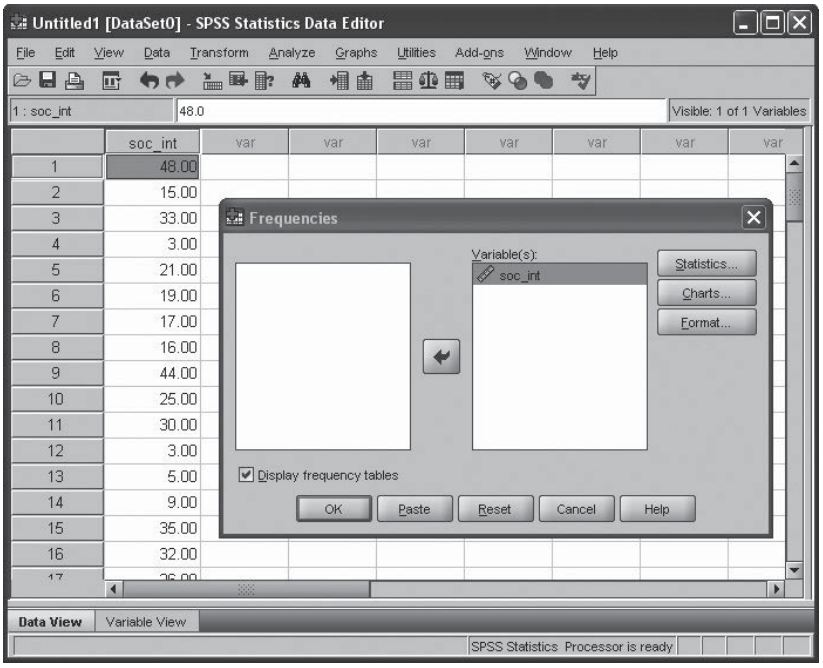
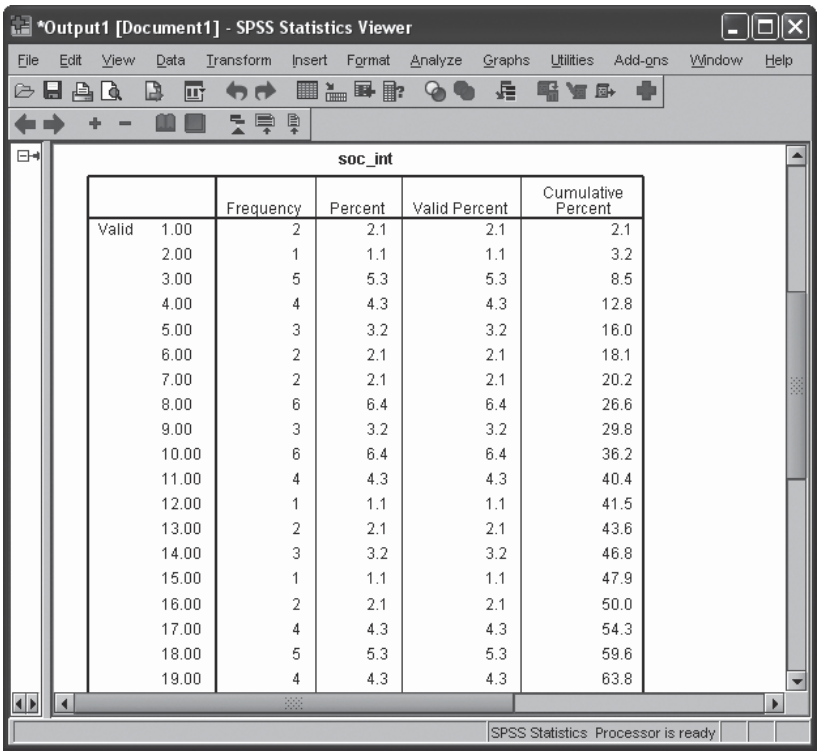


Figure 14 SPSS data window and frequencies window for the social interactions example. (Data from McLaughlin-Volpe et al., 2001.)



soc_int

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	2	2.1	2.1	2.1
	2.00	1	1.1	1.1	3.2
	3.00	5	5.3	5.3	8.5
	4.00	4	4.3	4.3	12.8
	5.00	3	3.2	3.2	16.0
	6.00	2	2.1	2.1	18.1
	7.00	2	2.1	2.1	20.2
	8.00	6	6.4	6.4	26.6
	9.00	3	3.2	3.2	29.8
	10.00	6	6.4	6.4	36.2
	11.00	4	4.3	4.3	40.4
	12.00	1	1.1	1.1	41.5
	13.00	2	2.1	2.1	43.6
	14.00	3	3.2	3.2	46.8
	15.00	1	1.1	1.1	47.9
	16.00	2	2.1	2.1	50.0
	17.00	4	4.3	4.3	54.3
	18.00	5	5.3	5.3	59.6
	19.00	4	4.3	4.3	63.8

Figure 15 SPSS frequency table for the social interactions example. (Data from McLaughlin-Volpe et al., 2001.)

Creating a Histogram

❶ Enter the scores from your distribution in one column of the data window.
❷ ✍ *Analyze.*
❸ ✍ *Descriptive statistics.*
❹ ✍ *Frequencies.*
❺ ✍ on the variable you want to make a frequency table of and then ✍ the arrow.
❻ ✍ *Charts,* ✍ *Histograms,* ✍ *Continue.*
❼ Optional: To instruct SPSS *not* to produce a frequency table, ✍ the box labeled *Display frequency tables* (this *un*checks the box).
❽ ✍ *OK.*

Practice the preceding steps by creating a histogram for the social interactions example in this chapter (the scores are listed under "Another Example," in the section "Frequency Tables"). Your output window should look like Figure 16.

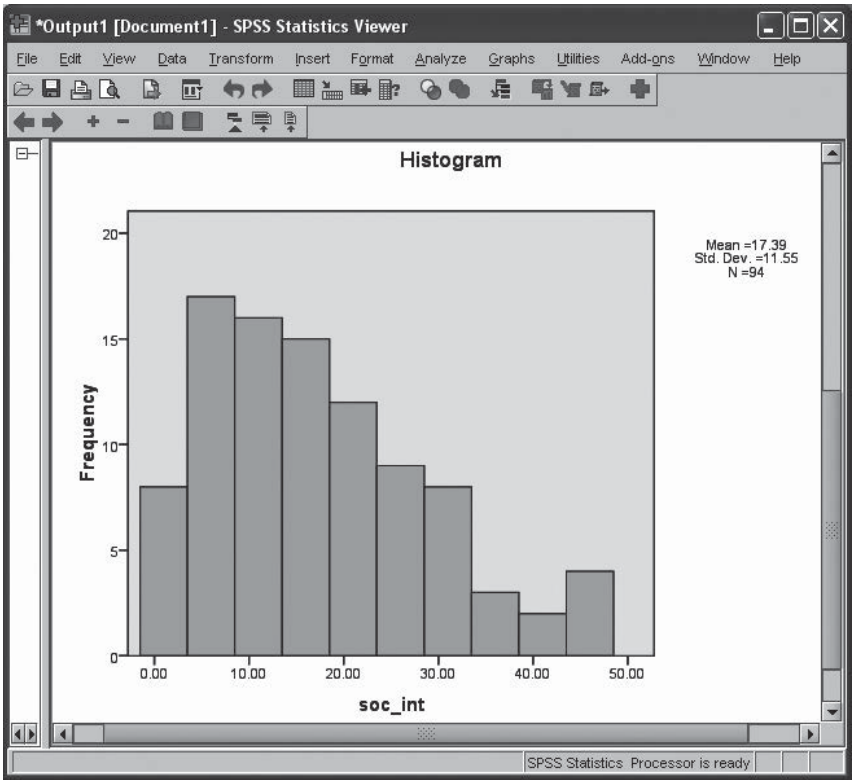


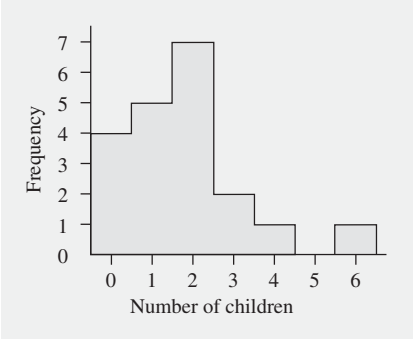
Figure 16 SPSS histogram for the social interactions example. (Data from McLaughlin-Volpe et al., 2001).

Answers to Set I Practice Problems

1. (a) Satisfaction with the vocational counselor; (b) 1, 2, 3, or 4; (c) 3.
2. (a) Likelihood of voting; (b) 1, 2, 3, 4, or 5; (c) 5.
3. (a) Nominal (or categorical); (b) numeric (or quantitative)—more precisely, equal interval; (c) numeric (or quantitative)—more precisely, rank onder (or ordinal).
4. (a) Frequency table.

Number of Children	Frequency	Percent
0	4	20
1	5	25
2	7	35
3	2	10
4	1	5
5	0	0
6	1	5

(b) Histogram.



(c) Unimodal, skewed to the right.

5. (a) Frequency table.

Hours	Frequency	Percent	Hours	Frequency	Percent
0	1	2	10	5	10
1	1	2	11	3	6
2	2	4	12	1	2
3	4	8	13	2	4
4	3	6	14	0	0
5	2	4	15	1	2
6	4	8	16	0	0
7	11	22	17	0	0
8	5	10	18	1	2
9	4	8			

(b) Based on frequency table above. See answer to question 4b for an example.

(c) Approximately unimodal, slightly skewed to the right.

6. (a) Frequency table.

Speed	Frequency	Percent	Speed	Frequency	Percent
15	1	3	24	1	3
16	0	0	25	0	0
17	0	0	26	0	0
18	0	0	27	1	3
19	0	0	28	1	3
20	1	3	29	0	0
21	0	0	30	3	8
22	0	0	31	3	8
23	1	3	32	3	8

(continued)

Speed	Frequency	Percent	Speed	Frequency	Percent
33	3	5	43	0	0
34	4	10	44	0	0
35	2	5	45	1	3
36	6	15	46	1	3
37	2	5	47	0	0
38	1	3	48	0	0
39	0	0	49	0	0
40	3	8	50	0	0
41	0	0	51	0	0
42	1	3	52	2	5

(b) Based on frequency table above. See answer to question 4b for an example.

(c) Unimodal, approximately symmetrical.

7. (a) Frequency table.

Score	Frequency	Percent	Score	Frequency	Percent
50	1	4	74	1	4
51	0	0	75	2	8
52	0	0	76	2	8
53	0	0	77	0	0
54	0	0	78	0	0
55	0	0	79	0	0
56	0	0	80	1	4
57	0	0	81	1	4
58	0	0	82	0	0
59	1	4	83	2	8
60	0	0	84	0	0
61	0	0	85	1	4
62	0	0	86	0	0
63	0	0	87	1	4
64	2	8	88	0	0
65	0	0	89	0	0
66	0	0	90	0	0
67	1	4	91	1	4
68	2	8	92	1	4
69	1	4	93	0	0
70	1	4	94	0	0
71	1	4	95	0	0
72	0	0	96	1	4
73	1	4			

(b) Based on frequency table above. See answer to question 4b for an example.

(c) Grouped frequency table.

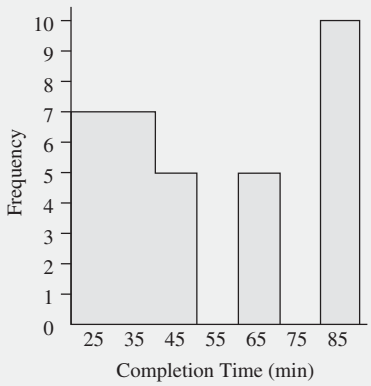
Interval	Frequency	Percent
50–59	2	8
60–69	6	24
70–79	8	32
80–89	6	24
90–99	3	12

(d) Based on frequency table in (c) above. See answer to question 4b for an example.
(e) Unimodal, approximately symmetrical (slightly negatively skewed).

8. (a) Similar to 7a above.
(b) Grouped frequency table.

Interval	Frequency	Percent
20–29	7	20.6
30–39	7	20.6
40–49	5	14.7
50–59	0	0.0
60–69	5	14.7
70–79	0	0.0
80–89	10	29.4

(c) Histogram.



(d) Roughly rectangular.

9. (a) Bimodal; (b) approximately normal (or unimodal or symmetrical); (c) multimodal.
10. (a) A distribution is the way a group of numbers is spread out over the possible values the numbers can have. You can describe such a distribution with a graph, called a histogram—a kind of bar graph with one bar for each possible value with one unit of height for each time its particular value occurs. In a histogram, a symmetrical distribution has a symmetrical shape (the right and left halves are mirror images). A unimodal distribution is one in which this graph has a single high point, with the other values gradually decreasing around it. (b) A negatively skewed unimodal distribution has a single high point, is not symmetrical, and its tail—the long, low side—extends to the left (where the negative scores go on the graph).
11. A ceiling effect describes the situation when many scores pile up at the high end of a distribution because it is not possible to have a higher score. This usually causes the distribution to be negatively skewed (skewed to the left).
12. (a) This is called a frequency table because it lays out how frequently (how many times) each category occurs for nine different categories. A frequency table makes the pattern of numbers easy to see. For example, of the 90 college students in the study, 19 gave bad news about Relationship with family (the first category). The table also gives the percentages. For example, 19 students is 19/90 of the total, or 21.1%. (b) The most bad news is given in four of the nine categories: Relationship with family, Relationship with actual/potential girlfriend/boyfriend, Relationship with friends, and Health of family member/friend. All of these categories had to do with family members or friends, most with relationships; there were few cases in the other categories (which had little directly to do with family or friends).