

ograms. Do you agree with the specific
at, in general, the study has useful

udy, how did you rate the introduction
ew of the literature? Was the approach
you could tell where the study fit into
the study clear?

you enough information about the
s, and analyses to replicate the study
statistics)?

study in your teaching situation?
hat you can at least understand what

sions sections give a nontechnical ex-
y relate to the original purpose of the
given of why the results turned out as
estions for further research? Are the

Are there any unnecessary ones?
rate the study? Are there any serious
, in the scaling, in the definition and
he control of extraneous variables, or
to doubt the results of the study?
e to you and your teaching? To your
or why not?

ve as a reader of research, read any
and critique it.

6 The group and the individuals

In the long run, a study should help you picture how the subjects behaved or performed on each measure, either through graphic representations of the performances or through descriptive statistics that indicate the central tendency and dispersion of the scores. To begin with, then, it is useful to understand how data can be visually represented.

Frequency

If I were to ask you how frequently you brush your teeth, you would probably answer something like two times per day. Likewise, if I were to ask you how frequent a score of 67 was in Table 6.1, the obvious answer would be "three people received 67." *Frequency* is just this commonsense idea. It indicates how many people did the same thing or performed a certain task in the same way. When you figured out how many people received a score of 67, you counted up, or tallied, the number of people at 67. To calculate the frequency at each score level, then, the researcher simply tallies them up and records the result, as shown in Table 6.1. But why would anyone bother to do this? Fre-

TABLE 6.1. TALLYING FREQUENCIES

<i>Students</i>	<i>Score</i>	<i>Tally</i>	<i>Frequency</i>
Robert	65	/	1
Randy	67	///	3
Henk	67		
Shenan	67		
Jeanne	69	////	4
Corky	69		
Millie	69		
Archie	69		
Dean	70	//	2
Elisabeth	70		
Monique	72	/	1
Iliana	73	/	1
Bill	74	/	1

TABLE 6.2. FREQUENCY DISTRIBUTION

Score value	Frequency
65	1
66	0
67	3
68	0
69	4
70	2
71	0
72	1
73	1

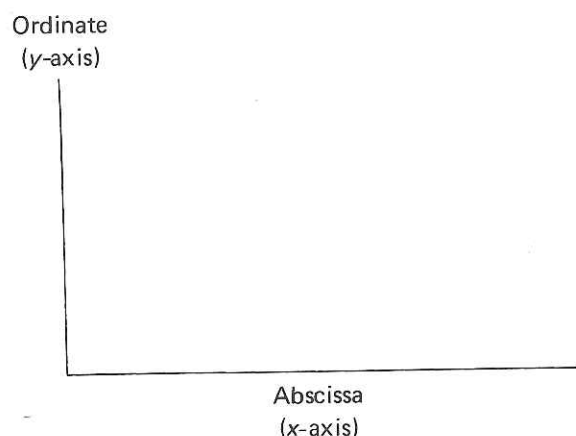


Figure 6.1 Abscissa and ordinate

frequencies are useful because they can be used to collapse, or summarize, information about scores received by a group, as shown in Table 6.2. This arrangement of score values from high to low and the frequency of each score value in a table is called a *frequency distribution*.

Graphic representation of frequencies

There are, however, more appealing ways to show the frequency distribution. You will generally encounter such graphic representations in one of three forms: a histogram, a bar graph, or a frequency polygon. All three are drawn on two axes: a horizontal line (also called the *abscissa*, or *x-axis*) and a vertical line (*ordinate*, or *y-axis*). These are shown in Figure 6.1.

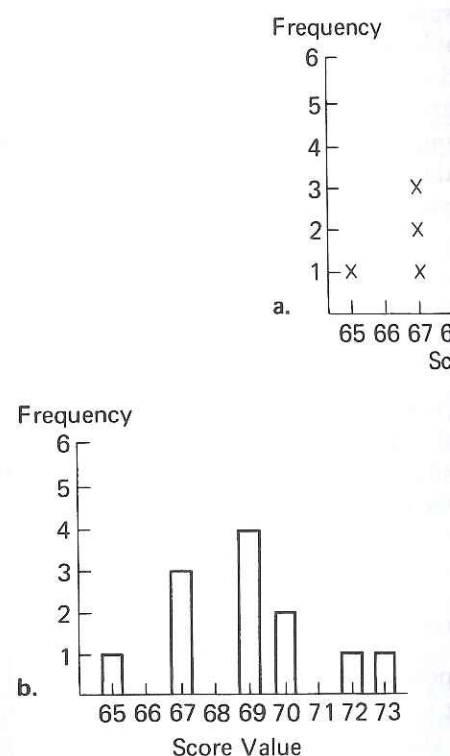


Figure 6.2 Graphic representation of frequency distribution. a: Frequency polygon. b: Bar graph. c: Histogram.

A *histogram* of a frequency distribution is drawn by placing the score values to the horizontal line, the frequencies to the vertical line, and marking an asterisk for each student who received each score, as shown in Figure 6.2a. If used instead of asterisks to represent the frequency of each score, looking at a *bar graph*, as shown in Figure 6.2b, where the top asterisk would be at the top of the bar, connected by lines, you have a *frequency polygon*, as shown in Figure 6.2c. All three types of graphs are important because they represent information that is useful in understanding the groups performed. Another excellent reason for using these concepts is that they are simple and do not distort information graphically (see Figure 6.2).

Studies in our field often omit these concepts, but they usually provide some form of descriptive representation of exactly how each

Frequency
1
0
3
0
4
2
0
1
1

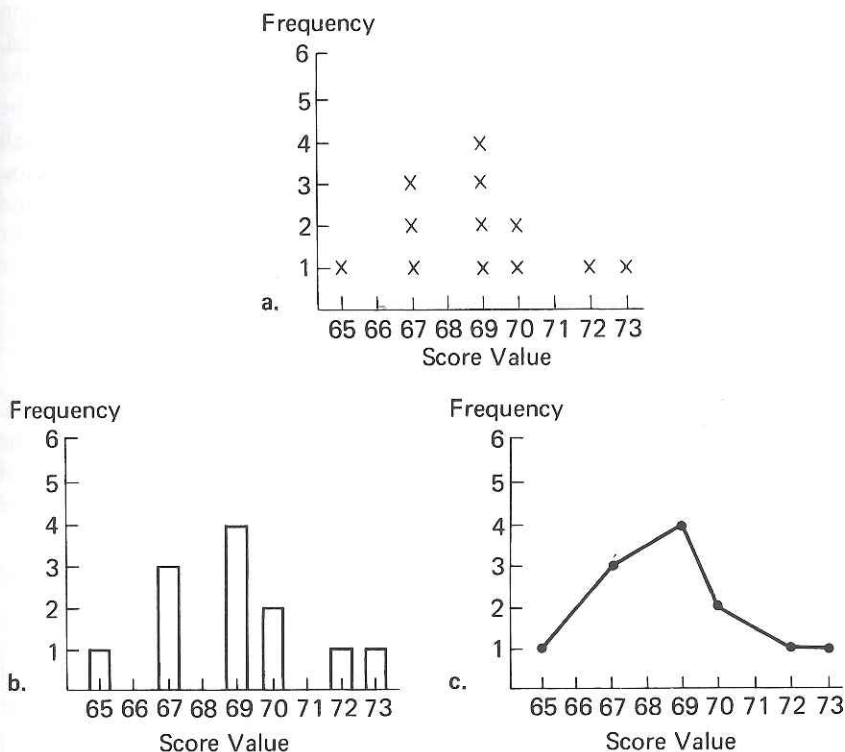


Figure 6.2 Graphic representation of frequency distributions. a: Histogram. b: Bar graph. c: Frequency polygon.

A *histogram* of a frequency distribution is usually created by assigning score values to the horizontal line, putting potential frequency values on the vertical line, and marking an asterisk or "X" to represent each student who received each score, as shown in Figure 6.2a. If bars are used instead of asterisks to represent the same information, you are looking at a *bar graph*, as shown in Figure 6.2b. When dots are placed where the top asterisk would be at each score value and the dots are connected by lines, you have a *frequency polygon*, as shown in Figure 6.2c. All three types of graphs are important in reading statistical studies because they represent information that you should have about the way the groups performed. Another excellent reason for getting a good grip on these concepts is that they are sometimes used to misrepresent or distort information graphically (see Huff and Geis 1954).

Studies in our field often omit these useful forms of graphs. However, they usually provide some form of *descriptive statistics*, that is, numerical representations of exactly how each group performed on the interval

scale measures. It is the reader's responsibility to look at this information and try to create a mental picture of how the group or groups performed. To do so, you must keep in mind two aspects of group behavior: the center and the individuals. Both are important because you should be able to visualize the middle (or typical) behavior of the group, as well as the performance of those individuals who varied from it. In statistics, these two aspects of group behavior are called *central tendency* and *dispersion*.

Central tendency

The first thing to look for, then, is the central tendency in a set of data. Notice that I have carefully avoided using the word *average*. I have done so because the common concept of average is closely related to *one* of three different indicators that are used to look at central tendency: the mean, the mode, and the median.

Mean

The *mean* is probably the single most commonly reported indicator of central tendency. It is virtually the same as the arithmetic average that you may calculate when grading classroom tests. It is symbolized in writing by $\bar{X}$ (said "ex bar"), and the formula for this statistic is as follows:

$$\bar{X} = \frac{\sum X}{N}$$

where $\bar{X}$ = mean, X = scores, N = number of scores, Σ = sum (or add).

For future reference, let's take a brief look at this formula. It simply says, to get the mean ($\bar{X}$), you add up (Σ) the scores (X) and divide by the number (N) of scores that you have. This formula is shown more graphically in Table 6.3. To find this mean in the example, you (1) add up the scores, (2) find the number of scores, and (3) divide the totaled scores by that number. So the mean in the example would be 70.

Even though the focus of this book is on reading and interpreting statistics, rather than doing them, it is often necessary to understand the symbols and how they work. Hence, a little exposure to this simple formula will not hurt. In addition, some of the concepts explained in this book will be easier to introduce with formulas. Of course, the concepts will always be explained in "real" language and examples will be provided. In any case, the formula for the mean was not so difficult.

TABLE 6.3. CALCULATING THE MEAN

Students (Group J)	Scores (X)	Calculations
Jan	40	1 ΣX = sum of scores
Jakob	67	$70 + 76 + 80 + 90$
Jaime	67	2 N = number of scores
Jimmy	70	3 $\bar{X} = \frac{\Sigma X}{N} = \frac{490}{7} = 70$
Juanita	76	
Jean	80	
Jacques	90	

It is just another way of saying something — much in the same way that we use symbols to use the terms *grammar* and *pronunciation* in mathematics as part of learning to understand them.

Mode

The *mode* is the score that occurs most often. In Table 6.3, what would be the mode? The score received by two students. To keep the term simple, I associate the term with its more common use (à la mode). Thus, the mode is the score that is received by the most students. There is a straightforward idea. It should be noted that there may be two or more modes in a set of scores. Such a set of scores would be termed *bimodal*, *trimodal*, etc.

Median

The *median* is defined as the middle point in a set of scores below which 50 percent of the scores fall and above which 50 percent fall. In Table 6.3, you will find three scores below it (40, 67, and 67) and two scores above it (80 and 90). So the median in this example is 67. For the following set of scores: 10, 25, 30, 35, 40, it is clear to you that it is 40.

There are cases, however, when there is an even number of scores, as in the example above. What is the median for these scores? There is an even number of scores, so the median is taken to be midway between the two middle scores. If the two middle scores are 13 and 14, should the median make sense? If so, what is the median?

responsibility to look at this information of how the group or groups performed. and two aspects of group behavior: the are important because you should be (typical) behavior of the group, as well individuals who varied from it. In statistics, behavior are called *central tendency* and

is the central tendency in a set of data. ted using the word *average*. I have done of average is closely related to *one* of e used to look at central tendency: the a.

most commonly reported indicator of he same as the arithmetic average that g classroom tests. It is symbolized in and the formula for this statistic is as

N = number of scores, Σ = sum (or

a brief look at this formula. It simply dd up (Σ) the scores (X) and divide by ou have. This formula is shown more this mean in the example, you (1) add er of scores, and (3) divide the totaled ean in the example would be 70.

s book is on reading and interpreting it is often necessary to understand the ence, a little exposure to this simple on, some of the concepts explained in oduce with formulas. Of course, the d in "real" language and examples will mula for the mean was not so difficult.

TABLE 6.3. CALCULATING THE MEAN

Students (Group J)	Scores (X)	Calculations
Jan	40	1 ΣX = sum of scores = 40 + 67 + 67 + 70 + 76 + 80 + 90 = 490
Jakob	67	
Jaime	67	
Jimmy	70	
Juanita	76	2 N = number of scores = 7
Jean	80	
Jacques	90	
		3 $\bar{X} = \frac{\Sigma X}{N} = \frac{490}{7} = 70$

It is just another way of saying something you already know how to do – much in the same way that we use *syntax* and *phonology* when others use the terms *grammar* and *pronunciation*. Try to view these simple mathematics as part of learning to understand statistics.

Mode

The *mode* is the score that occurs most frequently in a set of scores. In Table 6.3, what would be the mode? The mode would be 67, the only score received by two students. To keep the mode straight in my mind, I associate the term with its more common meaning of fashionable (as in à la mode). Thus, the mode is the score that is most fashionable, or is received by the most students. There is no statistical formula for this straightforward idea. It should be noted, however, that, in some cases, there may be two or more modes in a set of scores. Such a distribution of scores would be termed *bimodal*, *trimodal*, and so on.

Median

The *median* is defined as the middle point in a distribution, or that point below which 50 percent of the scores fall and, logically, above which 50 percent fall. In Table 6.3, you will see that Jimmy's score of 70 has three scores below it (40, 67, and 67) and three scores above it (76, 80, and 90). So the median in this example is 70. What, then, is the median for the following set of scores: 10, 25, 30, 40, 52, 60, 76? It should be clear to you that it is 40.

There are cases, however, when the median is not so clear. For example, what is the median for these scores: 8, 10, 13, 14, 16, 25? When there is an even number of scores, as in this example, the median is taken to be midway between the two middle scores. In this example, the two middle scores are 13 and 14, so the median is 13.5. Does that make sense? If so, what is the median for these scores: 10, 25, 37, 40,

50, 62, 71, 98? Your answer should have been 45, because that is the midpoint between the two middle scores, 40 and 50.

There are other cases in which there may be more than one score at the median, such as in the following: 45, 47, 48, 50, 50, 50, 67, 74, 87. Here, the midpoint is clearly 50, because there is an odd number of like scores at the median separating equal numbers of scores on either side.

Still other situations arise in which fractions other than .5 occur in the median, but the important thing to remember is that the median is the point that divides the scores 50/50, much like the median in a highway divides the road into two equal parts. In statistics, however, the median may turn out to be a fraction.

To review briefly, in Table 6.3, the mean, or arithmetic average, was 70; the mode, or most-often-received score, was 67; and the median, the score that divided the scores 50/50, was 70. These are all indicators of central tendency, and each has its strengths and weaknesses. None is necessarily better than the others, although the mean is most commonly reported. The measures simply serve different purposes and are appropriate in different situations.

Dispersion

Now that you understand how to interpret the typical behavior of a group in the form of central tendency, let's look at how the performance of individuals may vary from that typical behavior. There are two commonly reported indicators of the dispersion of a set of scores: the range and the standard deviation.

Range

You may already be familiar with the notion of range from correcting your classroom tests. *Range* is defined here as the number of points between the highest score on a measure and the lowest score plus one (plus one because it is viewed as including the scores at both ends). Thus, in Table 6.3, where the highest score is 90 and the lowest score is 40, the range is 51 points ($90 - 40 + 1 = 51$). The range provides some idea of how individuals vary from the central tendency. But it represents only the outer edges of that variation and, as a result, is strongly affected by behavior that may not be truly representative of the group as a whole. For instance, if Jan in Table 6.3 had already decided to drop the course involved and, therefore, had just guessed on the test, the range of 51 would not really represent the behavior of the group, because that behavior was strongly affected by something (Jan's personal problem) extraneous to the students' performances on the measure itself. Never-

theless, the range is often reported as it is and should be interpreted as just what it is: the difference between the high and low scores, including the low score.

Standard deviation

The *standard deviation* provides a sort of average of all scores from the mean. To understand it, go back again at the formula. Recall that $\bar{X}$ was the mean, X symbolized the scores, that Σ indicated summation, added up, and that N was the number of scores. The standard deviation (*SD*) is

$$\sqrt{\frac{\Sigma (X - \bar{X})^2}{N}}$$

Starting from the inside and working out, the formula first requires you to subtract the mean from each score ($X - \bar{X}$); square each of these differences; add them up $\Sigma (X - \bar{X})^2$. This sum is then divided by N , and the square root of the result is the standard deviation.

$$\sqrt{\frac{\Sigma (X - \bar{X})^2}{N}}$$

is the standard deviation. Table 6.4 makes this calculation.

Remember that the mean in Table 6.3 was 70. To find the mean for Table 6.4, you must line up the scores and then subtract the mean from each score to get the "differences" from the mean and add them up. Then you insert all this information into the formula. The result for this example is 14.41. Let's now review the steps.

I claimed that the standard deviation is a measure of the squaring and the square root, noticing that the formula is up and then dividing by N – similar to the formula for the mean) of the differences of all scores from the mean. Thus, the standard deviation is a measure of the differences of all scores from the mean. These differences are their *deviations* from the mean – hence the name. What this sort of average tells you will be discussed in the next chapter. Just keep in mind that it is a measure of dispersion. It is often better than the range because, by averaging, the effects of any extreme scores are reduced.

ould have been 45, because that is the
e scores, 40 and 50.

there may be more than one score at
ing: 45, 47, 48, 50, 50, 50, 67, 74, 87.
because there is an odd number of like
equal numbers of scores on either side.
which fractions other than .5 occur in
ing to remember is that the median is
50/50, much like the median in a high-
equal parts. In statistics, however, the
ction.

, the mean, or arithmetic average, was
eived score, was 67; and the median,
50/50, was 70. These are all indicators
s its strengths and weaknesses. None is
, although the mean is most commonly
erve different purposes and are appro-

to interpret the typical behavior of a
ency, let's look at how the performance
t typical behavior. There are two com-
dispersion of a set of scores: the range

th the notion of range from correcting
defined here as the number of points
measure and the lowest score plus one
ncluding the scores at both ends). Thus,
score is 90 and the lowest score is 40,
+ 1 = 51). The range provides some
n the central tendency. But it represents
tion and, as a result, is strongly affected
r representative of the group as a whole.
had already decided to drop the course
t guessed on the test, the range of 51
behavior of the group, because that be-
y something (Jan's personal problem)
ormances on the measure itself. Never-

theless, the range is often reported as one indicator of the dispersion
and should be interpreted as just what it is – the number of points
between the high and low scores, including both of them.

Standard deviation

The *standard deviation* provides a sort of average of the differences of
all scores from the mean. To understand what this means, let's look
again at the formula. Recall that $\bar{X}$ was the symbol for the mean, that
 X symbolized the scores, that Σ indicated that something should be
added up, and that N was the number of subjects. The formula for the
standard deviation (*SD*) is

$$\sqrt{\frac{\Sigma (X - \bar{X})^2}{N}}$$

Starting from the inside and working outward, you can see that the
formula first requires you to subtract the mean, already calculated, from
each score ($X - \bar{X}$); square each of these values ($(X - \bar{X})^2$); and add
them up $\Sigma (X - \bar{X})^2$. This sum is then divided by the number of scores
 $\Sigma (X - \bar{X})^2/N$, and the square root of the result

$$\sqrt{\frac{\Sigma (X - \bar{X})^2}{N}}$$

is the standard deviation. Table 6.4 makes this formula clearer.

Remember that the mean in Table 6.3 was 70. Using the same scores
and mean for Table 6.4, you must line up each score with the mean and
then subtract the mean from each score. Next, you must square each of
the "differences" from the mean and add up the squared values. When
you insert all this information into the formula, you find that the result
for this example is 14.41. Let's now return to the definition.

I claimed that the standard deviation is a sort of average (ignoring
the squaring and the square root, notice that you are adding something
up and then dividing by N – similar to what you did to calculate the
mean) of the differences of all scores from the mean (so it turns out that
what you are averaging is the difference of each student's score from
the mean). Thus, the standard deviation is a sort of average of the
differences of all scores from the mean. These differences are also called
their *deviations* from the mean – hence the name standard deviation.
What this sort of average tells you will be explained more fully in the
next chapter. Just keep in mind that it is a good indicator of dispersion.
It is often better than the range because it is an averaging process. By
averaging, the effects of any extreme scores that are not attributable to

TABLE 6.4. THE STANDARD DEVIATION

Students (Group J)	Score (X)	Mean ($\bar{X}$)	Difference ($X - \bar{X}$)	Difference squared ($X - \bar{X}$) ²
Jan	40	- 70	= - 30	900
Jakob	67	- 70	= - 3	9
Jaime	67	- 70	= - 3	9
Jimmy	70	- 70	= 0	0
Juanita	76	- 70	= + 6	36
Jean	80	- 70	= + 10	100
Jacques	90	- 70	= + 20	400
$\Sigma (X - \bar{X})^2$				= 1454
$SD = \sqrt{\frac{\Sigma (X - \bar{X})^2}{N}} = \sqrt{\frac{1454}{7}}$ $= 207.71 = 14.41$				

performance on the test (e.g., Jan's personal problem) are generally minimized.

Putting the picture together

To review briefly, then, two aspects of group behavior normally are described: the central tendency and dispersion. Central tendency indicates what the middle, or typical, behavior is on a given measure in the form of the mean (arithmetic average), mode (the most-often-received score), and the median (the score that splits the group 50/50). Which of these indicators is presented in a particular study will depend on the type of data and analyses in that study. But, in most cases, you will find at least one of them.

Likewise, you should look for an indication of the dispersion of scores, or the way individuals varied from the typical behavior of the group. This variation may be presented as the range (the difference between the highest and lowest scores, including both) or the standard deviation (a sort of average of how far individuals varied from the mean). Again, this information, in one form or another, is usually presented, and you should look for it so you can picture how the group or groups performed.

Such information is particularly useful in comparing the behavior of several groups on a particular measure. In Table 6.5, for instance, two groups - J and K - are clearly involved, and fairly complete information is presented on the typical behavior within each group and on how the individuals varied from that typical behavior. You should now be able

TABLE 6.5. GROUPS J AND K ON MY E

Group	Central tendency		
	$\bar{X}$	Mode	Median
J	70	67	70
K	60	55	62

to answer the following questions a groups.

1. Which group did generally better o
2. Which group had the single lowest
3. Which group had the widest spread
4. Which group was the most homog
5. Which group was the most intellig

To answer Question 1, you should ha of the central tendency. Clearly, Group and, therefore, performed better as a Question 2, you should have looked bering that they are probably single s lowest score (40). For Question 3, you for Group J was considerably higher th tell you that Group J had the widest Question 4 will be found by looking at for Group K is considerably lower, w the group were more tightly grouped Group J. Thus Group K is considerab look at the standard deviations rather extreme scores can occur for reasons o group on the measure. Since the standar process, such aberrations affect it less. T of the relative homogeneity of a group

You should have seen through Quest idea what the test in this example is whether it is measuring intelligence at is more intelligent than Group K? This should *not* be misinterpreted by you c numbers are just numbers, and it is yo and determine what they mean. This is to do. But it should become even easier because you will see more precisely how relate, how they can go awry, and how tester, and reader of language studies.

TION

Difference ($X - \bar{X}$)	Difference squared ($X - \bar{X}$) ²
30	900
-3	9
-3	9
0	0
+6	36
10	100
20	400

1454

$$\frac{\sum (X - \bar{X})^2}{N} = \frac{1454}{7} = 14.41$$

an's personal problem) are generally

ects of group behavior normally are
nd dispersion. Central tendency indi-
behavior is on a given measure in the
(average), mode (the most-often-received
e that splits the group 50/50). Which
a particular study will depend on the
study. But, in most cases, you will find

a indication of the dispersion of scores,
m the typical behavior of the group.
as the range (the difference between
cluding both) or the standard deviation
ividuals varied from the mean). Again,
another, is usually presented, and you
e how the group or groups performed.
y useful in comparing the behavior of
asure. In Table 6.5, for instance, two
olved, and fairly complete information
or within each group and on how the
al behavior. You should now be able

TABLE 6.5. GROUPS J AND K ON MY EXAM

Group	Central tendency			Dispersion			
	$\bar{X}$	Mode	Median	Low	High	Range	SD
J	70	67	70	40	90	51	14
K	60	55	62	55	68	14	2

to answer the following questions about the behavior of these two groups.

1. Which group did generally better on the exam?
2. Which group had the single lowest score?
3. Which group had the widest spread of scores?
4. Which group was the most homogeneous on the test?
5. Which group was the most intelligent?

To answer Question 1, you should have looked at all three indicators of the central tendency. Clearly, Group J is higher on all three indicators and, therefore, performed better as a group on the examination. In Question 2, you should have looked at the low-high figures (remembering that they are probably single scores in this case) and found the lowest score (40). For Question 3, you should have noted that the range for Group J was considerably higher than that for Group K. This would tell you that Group J had the widest spread of scores. The answer to Question 4 will be found by looking at the standard deviations. The one for Group K is considerably lower, which indicates that the scores of the group were more tightly grouped around the mean than those of Group J. Thus Group K is considerably more homogeneous. But why look at the standard deviations rather than the ranges? Remember that extreme scores can occur for reasons other than the performance of the group on the measure. Since the standard deviation is a sort of averaging process, such aberrations affect it less. Therefore, it is the better indicator of the relative homogeneity of a group.

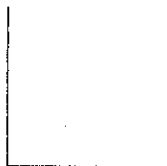
You should have seen through Question 5 immediately. You have no idea what the test in this example is like. How, then, can you decide whether it is measuring intelligence at all, much less whether Group J is more intelligent than Group K? This is an example of how the numbers should *not* be misinterpreted by you or by the authors you read. The numbers are just numbers, and it is your responsibility to look at them and determine what they mean. This is a task you should now be able to do. But it should become even easier and clearer in the next chapter because you will see more precisely how central tendency and dispersion relate, how they can go awry, and how they can help you – the teacher, tester, and reader of language studies.

Terms and symbols

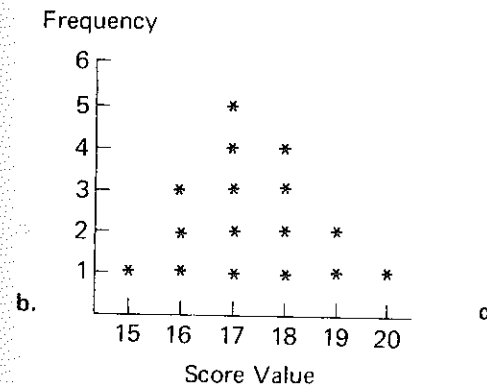
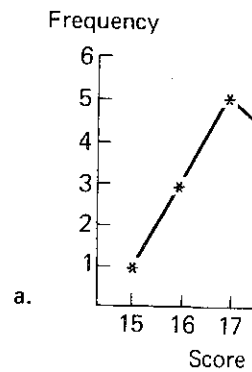
abscissa	mean ($\bar{X}$)
bar graph	median
bimodal distribution, trimodal distribution, and so forth	mode
central tendency	N
descriptive statistics	ordinate
deviations	range
dispersion	standard deviation (SD)
frequency	X
frequency distribution	x -axis
frequency polygon	y -axis
histogram	Σ

Review questions

1. What is central tendency? What are three indexes of central tendency? Which index is the best indicator?
2. What is dispersion? What two indexes are usually used to estimate dispersion? Which index is the best indicator?
3. Why should group behavior on a measure usually be described in terms of both the central tendency and dispersion?
4. What is a frequency distribution? And what is the advantage of a frequency distribution?
5. Can you label the abscissa and ordinate on the axes below?



6. Which of the three graphic representations at the top of page 73 is a histogram? A bar graph? A frequency polygon?



Application

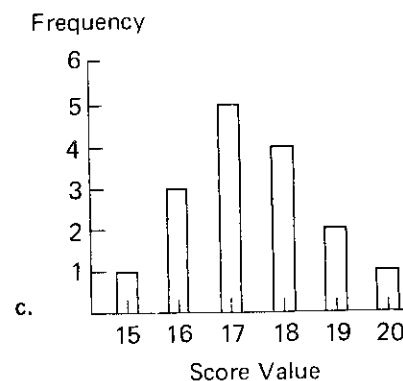
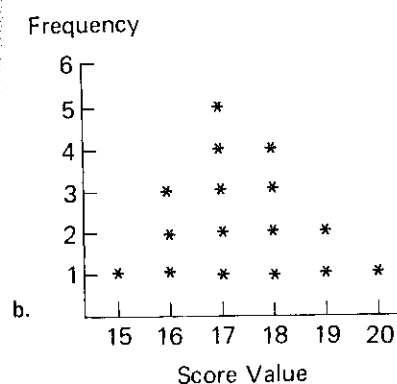
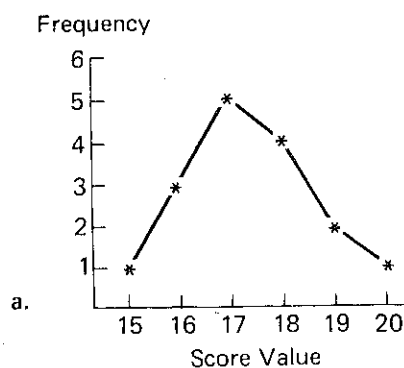
The table on page 74 is adapted from F over, then answer the questions that fol

mean ($\bar{X}$)
 median
 mode
 N
 ordinate
 range
 standard deviation (SD)
 X
 x -axis
 y -axis
 Σ

are three indexes of central tendency?
 or?
 indexes are usually used to estimate
 best indicator?
 n a measure usually be described in
 ncy and dispersion?
 on? And what is the advantage of a
 ordinate on the axes below?

representations at the top of page 73 is
 frequency polygon?

The group and the individuals



Application

The table on page 74 is adapted from Figure 5.7 in Chapter 5. Look it over, then answer the questions that follow.

DESCRIPTIVE STATISTICS FOR FALL 1977 SAMPLE

Group	Measure	$\bar{X}$	SD
Placed Students ($n = 133$)	Course grade	2.99	.62
	Final examination	67.83	9.89
	Cloze test	22.97	4.56
Continuing Students ($n = 31$)	Course grade	2.04	1.04
	Final examination	55.31	9.66
	Cloze test	15.87	4.57

1. Which group was larger?
2. Which group typically had higher grades?
3. Which measure probably had more possible points on it? What additional piece of information would have helped to answer this question?
4. Which group performed better on the cloze test? How do you know that?
5. Which group performed more heterogeneously on the final exam? How do you know that?
6. What additional information would you have liked to see in this table to help you interpret the groups' behaviors?

7 Patterns in human

Whether you believe that events are p essentially random, you will probably certain patterns that occur in your life. Descriptive statistics, in addition to help in terms of central tendency and dispersion, these patterns and how individuals respond to them. To get a grip on how these patterns work, I will discuss concepts and how they interrelate: probability, distribution, and standardized scores.

Probability

Short of taking you on a group trip to Monte Carlo, the best way to demonstrate probability is to take a coin and start flipping it. What did you guess? You probably recognize that there was a 50 percent chance of getting heads. You know that? You probably observed a fairly predictable pattern to events like this.

In more formal terms, the *probability* of an event is the number of *expected outcomes* by which the event can occur (heads or tails). Thus, one expected outcome out of the two possibilities (of heads or tails) is a 50 percent probability of getting heads.

What, then, is the probability of getting a six-sided die? Again, you need only divide the number of possible outcomes (one) by the total number of possible outcomes (six). One divided by six is .166, or a 16.6 percent probability of getting a five.

Why go through all this arithmetic? There are three good reasons. First, it is important to know that the probability of a given event is the ratio of the expected outcomes to the total possible outcomes. This ratio, which is usually stated in percentages. Second, and perhaps most important, this information will help if you are ever asked