

Measurement and Numbers

Questions to Ask About Test Scores

Scales of Measurement

Preparation of a Frequency Distribution

Grouped Frequency Distributions

Cumulative Frequency Distributions

Graphic Representation

Measures of Central Tendency

The Mode

The Median

Percentiles

The Arithmetic Mean

Central Tendency and the Shape of the
Distribution

Measures of Variability

The Range

The Semi-Interquartile Range

The Standard Deviation

Interpreting the Standard Deviation

Interpreting the Score of an Individual

Measures of Relationship

Making Predictions

Finding the Regression Line

Summary

Questions and Exercises

Suggested Readings

QUESTIONS TO ASK ABOUT TEST SCORES

Catherine Johnson and Peter Cordero wanted to gather information about achievement levels in their two sixth-grade classes. They gave their students a 45-item reading comprehension test provided in their current reading series, a 65-item review test from the mathematics book, and a dictation spelling test of 80 items based on the words their classes had been studying during the past 6 weeks. They marked the papers, counted the number of correct answers on each, and recorded the scores. Then they made up a joint class list that showed the three scores for each student. In addition, they recorded which class each student was in (1 for Ms. Johnson's class, 2 for Mr. Cordero's) and each student's gender (1 for boys, 2 for girls). With five pieces of information for each of 52 students, they wondered what they should do with all those numbers.

Tests and assessments *do* produce scores, and scores *are* numbers. So, if we are to think about and use test scores, we must be prepared to think about and work with numbers. The numbers that represent test scores can be organized to provide the answers to a range of questions, but first we must know what kinds of questions to ask. Once we have the questions in mind, we can begin to ask how the numbers can be arranged to provide the answers.

Look at the numbers (scores) shown in Table 2-1 for the sixth-graders in the two classes. What kinds of questions might the two teachers ask about this set of numbers?

Table 2-1

Scores for 52 Sixth-Grade Students on Tests of Reading, Spelling, and Mathematics

First Name	Last Name	Gender	Class	Reading (45)*	Spelling (80)	Math (65)
Aaron	Andrews	1	1	32	64	43
Byron	Biggs	1	1	40	64	37
Charles	Cowen	1	1	36	60	38
Donna	Davis	2	1	41	74	40
Erin	Edwards	2	1	36	69	28
Fernando	Franco	1	1	41	67	42
Gail	Galaraga	2	1	40	71	37
Harpo	Henry	1	1	30	51	34
Irrida	Ignacio	2	1	37	68	35
Jack	Johanson	1	1	26	56	26
Kleven	Klipsch	1	1	28	51	25
Laverne	Lappenski	2	1	36	57	53
Mary	Madison	2	1	39	68	37
Nathan	Natts	1	1	22	47	22
Oprah	Oates	2	1	36	59	33
Petula	Peters	2	1	32	64	33
Quadra	Quickly	2	1	21	44	19
Rahim	Roberts	1	1	29	64	43
Salim	Salik	1	1	41	76	33
Thomas	Tank	1	1	35	65	38
Usaka	Urban	2	1	41	65	38
Victor	Vasquez	1	1	37	68	40
Wakana	Watanabe	2	1	25	53	21
Xenum	Xerxes	1	1	25	54	31
Yuan	Young	1	1	32	59	24
Zebulon	Zibberits	1	1	42	73	44
Angela	Ash	2	2	43	64	52
Bellinda	Brown	2	2	33	38	41
Charlotta	Cowen	2	2	33	47	50
Dominik	Dubrow	1	2	39	66	34
Erik	Eriksen	1	2	39	55	47
Francis	French	2	2	38	59	49
Guido	Garcia	1	2	31	52	29
Hillary	Huan	2	2	38	61	48
Igor	Ivanovich	1	2	33	53	43

(continued)

Table 2-1 (Continued)

First Name	Last Name	Gender	Class	Reading (45)*	Spelling (80)	Math (65)
Jill	Johanson	2	2	42	61	45
Kaleen	Knowles	2	2	35	55	51
Larry	Lewis	1	2	29	40	34
Moe	Mastrioni	1	2	36	58	39
Nancy	Nowits	2	2	28	44	44
Orden	Orford	1	2	35	53	38
Petre	Popovich	1	2	36	52	53
Quincy	Quirn	1	2	33	48	33
Rhonda	Rostropovich	2	2	31	50	31
Sally	Stebbens	2	2	33	51	32
Thelma	Thwaites	2	2	38	43	45
Uriah	Urdahl	1	2	42	61	60
Velma	Vauter	2	2	29	49	36
William	Westerbeke	1	2	33	54	33
Xena	Xerxes	2	2	30	57	37
Yannita	Younts	2	2	44	63	49
Zephina	Zoro	2	2	30	47	38

*Maximum score in parentheses.

What questions can you ask? Before reading further, study the sets of scores and jot down the questions that come to your mind in connection with these scores. See how many of the question types that we discuss you can anticipate.

Each student has five numbers assigned to him or her, but we might want to inquire whether each of these numbers actually conveys quantitative information. Do the numbers that are assigned to represent gender and class have the same kind of meaning as the test scores? This is a question of the *scale* the numbers represent, and the answer affects the kinds of operations we can apply to the numbers.

A second, rather general type of question that we might ask is, what is the basic pattern of the set of scores? How do they "run"? What do they "look like"? How can we get a picture of the set of math scores, for example, so that we can get an impression of the group as a whole? To answer this type of question, we will need to consider simple ways of tabulating and graphing a set of scores.

A third type of question that will almost certainly arise is, what is this group like, on the average? In general, have they done as well on the test as some other sixth-grade group? What is the typical level of performance in the group? All these questions call for some single number to represent the group as a whole, some measure of where the *middle* of the group lies. To answer this type of question, we will need to become acquainted with statistics developed to represent the average, or typical, score.

Fourth, to describe the group, we might feel a need to describe the extent to which the scores spread out away from the average value. Have all the students in the group made about the same

progress, or do they show a wide range of achievement? How does this group compare with other classes, with respect to the *spread* of scores, and do the students show the same spread of achievement on all three tests? This type of question calls for a study of measures of variability.

Fifth, we might ask how a particular individual stands on one of the tests. We might want to know whether Aaron Andrews did well or poorly on the mathematics test. And if we decide that his score is a good one, we might want some way of saying just how good it is. We might ask whether Aaron did better in reading or in mathematics. To answer this question, we will need a *common yardstick* on which to express performance in two quite different areas. One need, then, is for some uniform way of expressing and interpreting the performance of an individual, independent of the particular test. How does this person stand relative to the group? We will give some preliminary answers to this question in this chapter and consider it in detail in Chapter 3.

A sixth query is of the following type: To what extent do those who excel in reading also excel in mathematics? To what extent do these two abilities go together in the same individuals? Is the individual who is superior in one area likely to be superior in the other? To express this *association* between two measurements, we will need to become acquainted with indices of correlation.

Test scores are also frequently used to forecast future performance. Therefore, a seventh issue we might want to address is how to make the most accurate prediction of a person's performance either on another test or on some outcome, such as an end-of-the-year achievement assessment. This forecasting function is the primary use that is made of scores on such high-stakes tests as the SAT, and the answer will lead us to a consideration of regression.

Many other questions may arise with respect to a set of scores. The most important ones concern the drawing of general conclusions from data on a limited group. For example, the 26 girls in this group have an average reading score of 35.0, and the 26 boys have an average score of 33.9. These are *descriptive* facts about this testing of these particular girls and boys. But these students might be considered to represent a larger population, such as all students in this school district or the state. From the results in these two classes, we might want to make an estimate or best guess of the average level of reading achievement in the larger group. We also might want to know whether we can safely conclude that the total population of girls from which this sample is drawn would surpass the total population of boys on this same test. These problems are of **inference**. Problems of statistical inference make up the bulk of advanced statistical work. For a detailed description of the principles and applications of statistical inference, see any of the books on statistical methods listed at the end of this chapter. These issues do not enter into the basic interpretation of test scores for an individual or a group, so we will not consider them further here.

The routines developed for organizing numbers to answer these and other questions constitute the field called **statistics**. This name and, in fact, the very prospect of working with numbers seem a bit scary to some people. Fortunately, much of the mechanics of working with numbers can now be performed on a pocket calculator or personal computer, so we can concentrate on the questions to ask and on the ways in which the numbers are arranged to answer them, rather than worry about computational details. As we discuss ways to answer each of the seven types of questions mentioned above, we will introduce easy-to-use computer programs to perform the necessary computations. Two very widely available programs will be presented: the package of statistical programs known as SPSS and the data analysis routines included in the Microsoft Excel® spreadsheet program. Both programs have some shortcomings, but both are relatively easy to use once you get used to them.

SCALES OF MEASUREMENT

One way to define measurement is *the assignment of numbers to objects according to a set of rules*. The set of rules is called a **scale**. Knowing the scale that has been used to assign the numbers is critical to proper interpretation of the measurement. For example, Ms. Johnson and Mr. Cordero assigned the number 1 to their male students and the number 2 to their female students, but what information do these numbers contain? Feminist ideology aside, does the number 2 mean that the girls possess more of the trait of gender than the boys do? Obviously not; in this case the numbers do not convey information about *amount* of anything. The number 1 has been substituted for the label "boy" and the number 2 has been substituted for "girl," but we could just as well have used the numbers 163 and 27. Each number takes on the meaning of a verbal label. When numbers are used in this way, the scale is called a **nominal scale**. The numbers take the place of names. Your university has almost certainly assigned you a student number. This number represents a "score" on a nominal scale, as do the numbers on the backs of athletes' uniforms. When numbers do not contain information about amount of a trait, it is not appropriate to treat them like numbers. In general, you cannot add them or perform any other arithmetic operations and obtain a meaningful result. The only use we ordinarily can make of numbers on a nominal scale is to count how many instances there are of each number. There are 26 ones and 26 twos in Table 2-1, so we know there are an equal number of boys and girls.

Sometimes numbers are assigned to represent the order of individuals on a trait. None of the numbers in Table 2-1 are of this kind, but suppose Mr. Cordero decided to rank order his students on the basis of their spelling test scores with the student who earned the highest score—Dominik Dubrow, getting a rank of 1, second highest, Angela Ash, ranked 2, and so forth. What kind of information would these numbers represent?

A set of ranks conveys information about the order in which the students stand on the trait, but the ranks do not contain information about amount. More importantly, the differences between numbers do not have a constant meaning. The difference between a rank of 1 and a rank of 5 is not necessarily the same as the difference between a rank of 11 and one of 15. Both differ by four units, but four units of rank usually cover a greater portion of the trait at the extremes of a group than in the middle. A scale that tells us the order in which people stand, who has more of the trait, but not how much more, is called an **ordinal scale**. Several common ways of reporting test score information that we will describe in Chapter 3 produce ordinal scales.

Tests that are scored like those in Table 2-1 treat each item as equal to every other item. Getting 10 items correct yields a score of 10, getting 20 items correct yields a score of 20, and getting 30 items correct would give you a score of 30. Because each item is assumed equal in the amount of the trait it measures, equal differences in scores are treated as equal differences in the trait. Although this assumption is somewhat tenuous when measuring human abilities, it represents the same basic kind of measurement as the centigrade or Fahrenheit temperature scales. Equal numerical differences in score represent equal differences in the property being measured. When we can make this assumption, the scale is called an **interval scale**. Most of the computations that are done with test scores require that we assume the scale of measurement is an interval scale.

Suppose Zebulon Zibberits tries his hardest, but still cannot spell any of the words in Mr. Cordero's list correctly. Does this mean that Zebulon has zero spelling ability? Probably not. The scale of spelling ability represented by this test starts at a point well above zero, so

a score of zero on the test does not mean zero ability. Very few scales used in psychology and education are constructed such that a score of zero means exactly none of the trait or property in question. This is one reason Blanton and Jaccard (2006) expressed concern about the arbitrary nature of the metrics used in psychology and education. Exceptions (non-arbitrary metrics) would be scales like those for height, weight, and duration of time. When the scale is constructed so a score of zero means exactly none of the trait, it is called a **ratio scale**. Scales like this allow us to conclude not only that the size of the unit is the same everywhere along the scale, but also that a score that is numerically twice another score means exactly twice as much of the trait. Someone who takes 10 minutes to solve a problem takes twice as long as someone who takes 5 minutes, and the person who takes 20 minutes to solve the problem takes twice as long as the 10-minute person and four times as long as the 5-minute person. Ratio scales allow us to make proportional statements like these. Very few ratio scales exist in the fields of education and psychology, but fortunately interval scales allow us to perform most of the analyses we need. This classification of scales as nominal, ordinal, interval or ratio was suggested by S. S. Stevens in 1951 and has been widely accepted, although some experts in measurement, for example, Torgerson (1958) have suggested alternative classifications.

PREPARATION OF A FREQUENCY DISTRIBUTION

In Table 2-1, we showed a record sheet on which test scores for 52 sixth-graders were listed. Let us look at the scores in the Math column and consider how they can be rearranged to give a clearer picture of how the pupils have performed on the math test. The simplest rearrangement is merely to list the scores in order from highest to lowest, as follows:

60	49	44	41	38	37	33	31	24
53	49	44	40	38	36	33	31	22
53	48	43	40	38	35	33	29	21
52	47	43	39	37	34	33	28	19
51	45	43	38	37	34	33	26	
50	45	42	38	37	34	32	25	

This arrangement gives a somewhat better picture of the way the scores fall than does Table 2-1. We can see the highest (60) and lowest (19) scores at a glance. It is also easy to see that the middle person in the group falls somewhere in the mid-30s. We can see by inspection that most of the scores fall between 30 and 50. But this simple rearrangement of scores still has too much detail for us to see the general pattern clearly. We need to condense the data into a more compact form.

Often, the first step in organizing scores for presentation is to prepare a display called a **frequency distribution**, a table that shows how often each score has occurred. Each score value is listed, and the number of times it occurs is shown. A portion of the frequency distribution for the math test is shown in Table 2-2. However, Table 2-2 is still not a very good form for reporting the scores. The table is too long and spread out. We have shown only part of it; the whole table would take 42 lines, almost as many as the original listing of scores. It would have a number of zero entries, and there would be marked variations in the Frequency column from one score to the next.

Table 2-2

Frequency Distribution of Scores on the Mathematics Test for 52 Students

Score (X)	Frequency
60	1
59	0
58	0
57	0
56	0
55	0
54	0
53	1
52	1
.	.
.	.
.	.
40	2
39	1
38	5
37	4
.	.
.	.
.	.
28	1
27	0
26	1
25	1
24	1
23	0
22	1
21	1
20	0
19	1

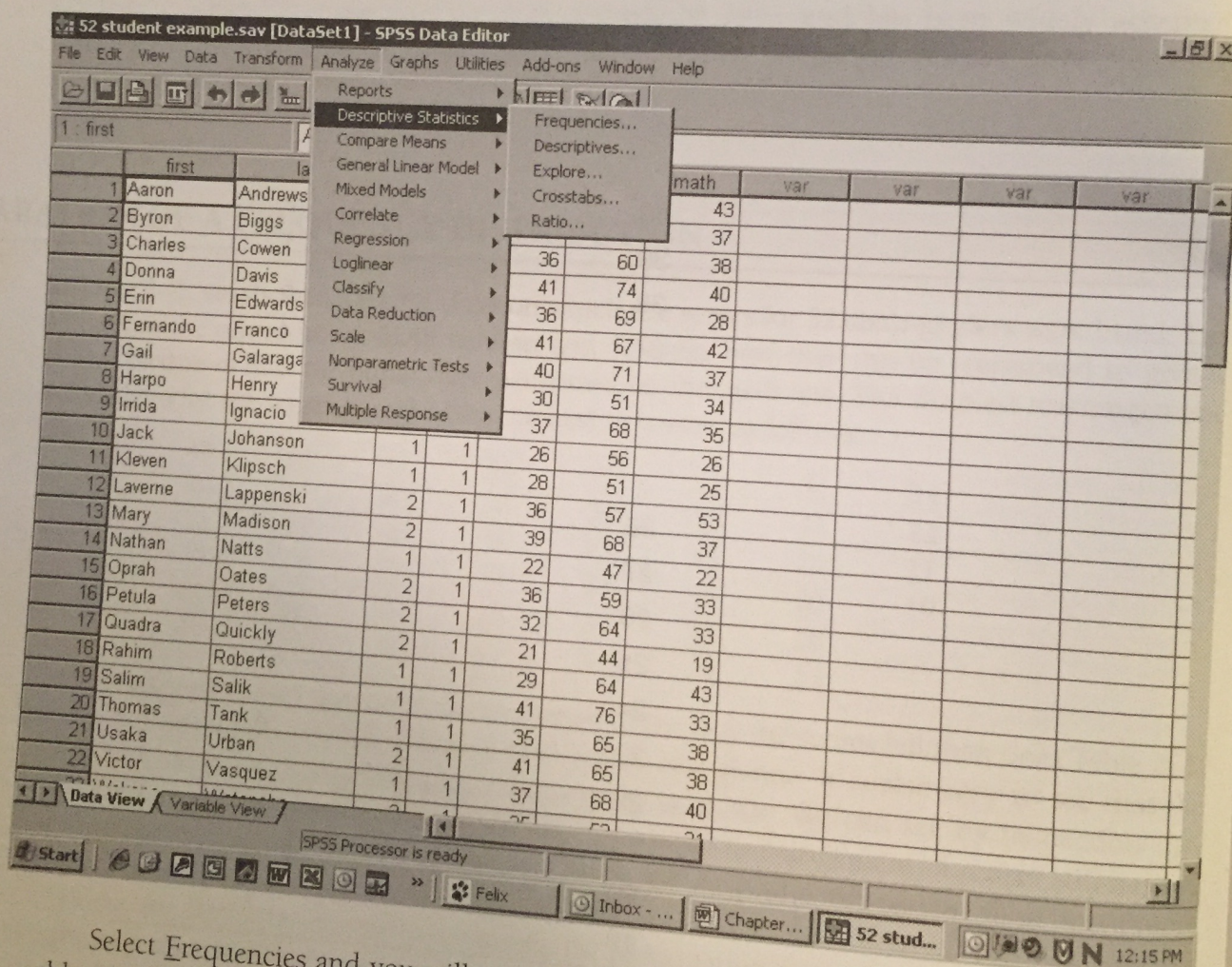
Grouped Frequency Distributions

Scores are often *grouped* into broader categories to further improve the clarity of presentation. We discard some detail in the data to make it easier to grasp the picture presented by the entire set of scores. In our example, we will group three adjacent scores, so that each *grouping interval* includes three points of score. The entire range of scores from 19 to 60 is

MAKING THE COMPUTER DO IT

Frequency Distributions

Excel is not designed to provide a frequency distribution like the one we have just described. Rather, it will prepare a grouped frequency distribution as described in the next section. SPSS will produce a frequency distribution, but you must be careful when reading it because score values with zero frequency are omitted from the list of scores. To prepare a frequency distribution of the data from Table 2-1 using SPSS, start the program, enter the data from Table 2-1 or open the file in which they have been saved, click on the Analyze button in the menu at the top of the screen, and select Descriptive Statistics. The screen should look like the one shown below (the exact layout of the screen will depend on which version of SPSS you are using):



Select Frequencies and you will see a dialogue box in which you can select the variables to be analyzed. Highlight the name of the variable you want by clicking on it, then click on the arrow button to move the variable into the Variables window. If you use the data from Table 2-1 and select the Math variable, you should get the complete frequency distribution, part of which we showed you in Table 2-2. You can generate other reports from this program, but we will wait to describe them until we have covered those topics.

Table 2-3

Grouped Frequency Distribution of Scores from 52 Students on a Math Test Using an Interval of 3

Interval	Frequency
60-62	1
57-59	0
54-56	0
51-53	4
48-50	4
45-47	3
42-44	6
39-41	4
36-38	10
33-35	9
30-32	3
27-29	2
24-26	3
21-23	2
18-20	1

represented by 14 intervals. When this is done, the set of scores is represented as shown in Table 2-3, a fairly compact table illustrating how many people there are in each **score interval**. Thus, for example, we have two people in the interval 19-21. We do not know how many of them got 19s, 20s, or 21s; we have lost this information in the grouping. We assume that they are evenly spread throughout the interval. In most cases, there is no reason to believe that one score will occur more often than any other, and this assumption is a sound one, so the gains in compactness and convenience of presentation more than make up for any slight inaccuracy introduced by the groupings. (In some special applications, such as reports of family income, certain values are more likely than others, for example, \$18,000, \$25,000, \$50,000. Special precautions are required when grouping material of this type. An effort should be made to place the most popular values near the middle of an interval to reduce distortion.)

In practical situations, we always face the problem of deciding how broad the groupings should be, that is, whether to group by 3, 5, 10, or some other number of points of score. The decision is a compromise between (1) losing detail from our data and (2) obtaining a convenient, compact, and smooth representation of the results. The use of broader intervals results in losing more detail, but condenses the data into a more compact picture. A practical rule of thumb is to choose an interval that will divide the total score range into roughly 15 groups. In our example, the highest score is 60, and the lowest is 19. The range of scores is 60 to 19, giving a range of 41 points. Dividing 41 by 15, we get 2.7. The nearest whole number is 3, so we group the data by 3s. In addition to the "rule of 15," we also find that intervals of 5, 10, and

multiples of 10 make convenient groupings. Because the purpose of grouping scores is to arrive at a convenient and clear representation, factors of convenience become a major consideration. It is also conventional to use a multiple of the interval width as the lower limit for each interval, as we have done in Table 2-3.

In cases where graphs are going to be prepared using the scores in their grouped form, it is also convenient to use an interval that includes an odd number of score points, for example, 3, 5, or 7, because it is sometimes necessary to use the midpoint of the score interval to represent all scores in the interval. If the interval has an even number of score values, this midpoint will be halfway between two actual scores, but if an odd number is used for the interval width, the midpoint will be a whole score value, making for a more attractive graph.

Note also that sometimes there is no need to group the data into broader categories. If the original scores cover a range of no more than about 20 points, grouping may not be required. Also, with modern computing equipment it is usually easier to compute the statistical indices, described later in the chapter, from the original set of scores unless the data come to you as a grouped frequency distribution. In that case, most programs require special procedures that are beyond our scope.

MAKING THE COMPUTER DO IT

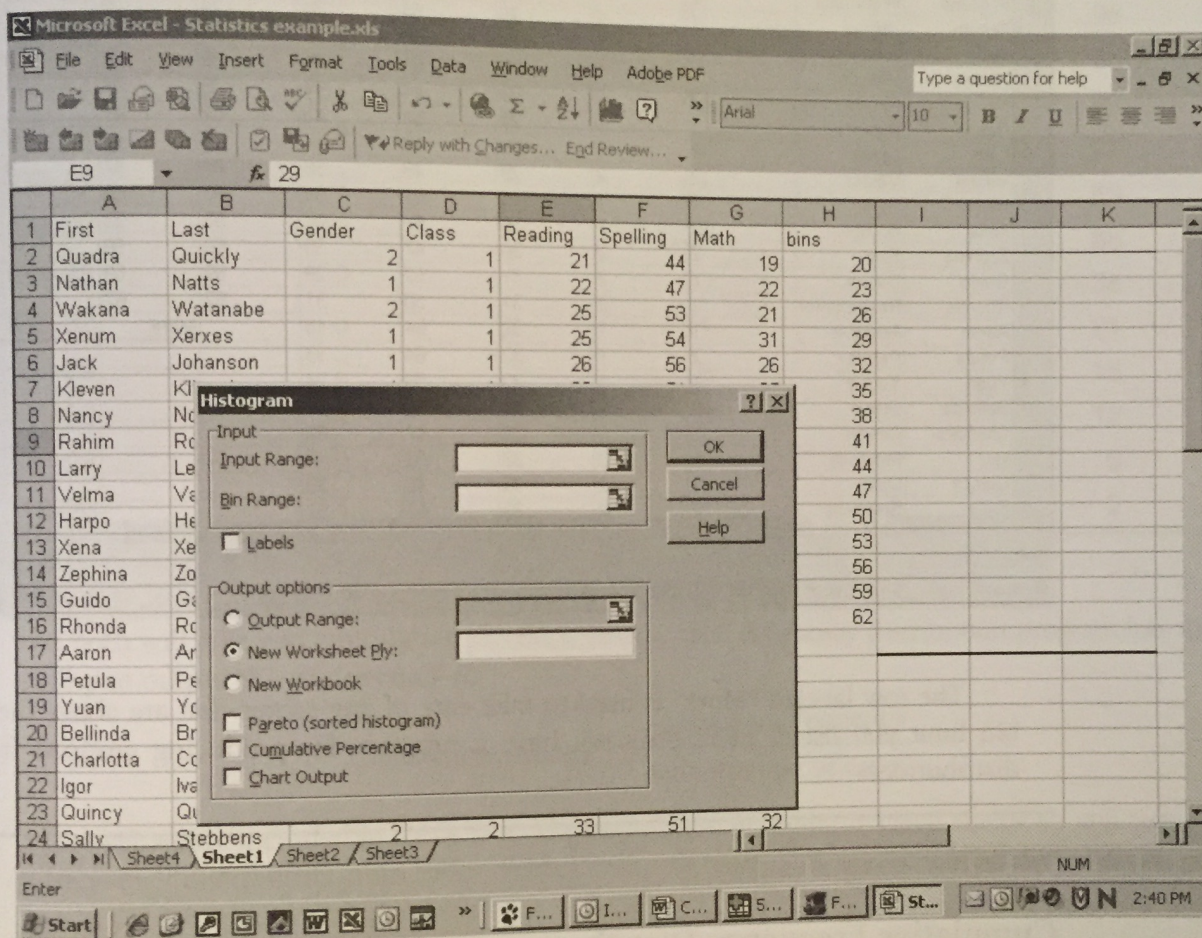
Grouped Frequency Distributions

Excel is designed to produce grouped frequency distributions, but you must follow a particular sequence of steps that can at times be frustrating until you become used to them. After you have opened the file containing the data you wish to analyze, the first step is to decide on the number of intervals. Excel calls these intervals *bins*. Using our set of scores for the math test, the easiest way to determine how many bins to use is to sort the scores to find the range they cover. Click on any score in the Math column, then click on one of the sort icons. The scores will now be in order, so you can determine the highest and lowest. Finding the range to be 41 and applying our "rule of 15," we again decide to use bins of 3. (Note that you could use Excel to prepare a raw frequency distribution like the one we got using SPSS, but you would have to create a bin for every possible score.)

You must now set up your bins. Bins are specified by giving the *highest* score that is to fall in each bin. Therefore, if we want to use a bin size of 3, our lowest bin should have 20 specified as its value (to make the lower limit of each bin a multiple of 3). Excel sorts scores into bins in ascending order. That is, the program looks at all the scores and puts those that are less than or equal to the bin limit in the first bin. If we start with a bin limit that is below the lowest score, it will have a frequency of zero, but if we start with a bin limit that is more than our chosen interval above the lowest score, the lowest interval will be too wide.

After deciding on the first bin limit, set each additional bin limit one interval above the one below it. If we start with a limit of 20, then the other bin limits would be 23, 26, 29, . . . 53, 56, 59, 60. *The bins must be arranged with the lowest bin at the top of its column.* In a column to the right of your data (for instance, column I), type the word *bins* in the first row. Then type in the bin limits, starting with the lowest, in the cells under the "bins" heading.

The Excel frequency distribution routine is part of the Histogram program in the Data Analysis group of programs that are included in the Tools menu. (Histograms are discussed in the next section.) Click on Tools, then on Data Analysis. (Hint: You may have to click on a cell containing data to activate the Data Analysis tool.) A dialogue box will appear that includes many data analysis options. Click on Histogram and click OK. You should see a screen that looks like the one shown here imposed over your data.



Excel works by specifying areas of the data table to be included in the analysis. The scores for the math test are listed in column G with a label in the first row and data in cells G2 to G53. Click on the Input Range window in the Histogram pop-up dialogue box. Then in the main screen highlight the cells containing the math test scores. The cell references \$G\$2 and \$G\$53 will appear in the pop-up box window. Next, click in the Bin Range window of the pop-up box. Then, again on the main screen, highlight the bin limits you placed in the column under "bins." The cell references for these cells will appear in the window. Finally, click the Output Range button on the pop-up box to activate that window and click in a cell where you would like the grouped frequency distribution to start (usually just to the right of your bins column, in this case \$I\$1). Then click OK. If you are using the data from Table 2-1 and have followed the directions as we have outlined them, your screen should look like this:

Microsoft Excel - Statistics example.xls

Type a question for help

File Edit View Insert Format Tools Data Window Help Adobe PDF

Reply with Changes... End Review...

	A	B	C	D	E	F	G	H	Formula Bar	K
	First	Last	Gender	Class	Reading	Spelling	Math	bins	Bin	Frequency
1	First	Last	Gender	Class	Reading	Spelling	Math	bins	20	1
2	Quadra	Quickly		2	21	44	19	20	23	2
3	Nathan	Natts		1	22	47	22	23	26	3
4	Wakana	Watanabe		2	25	53	21	26	29	2
5	Xenum	Xerxes		1	25	54	31	29	32	3
6	Jack	Johanson		1	26	56	26	32	35	9
7	Kleven	Klipsch		1	28	51	25	35	38	10
8	Nancy	Nowits		2	28	44	44	38	41	4
9	Rahim	Roberts		1	29	64	43	41	44	6
10	Larry	Lewis		1	29	40	34	44	47	3
11	Velma	Vauter		2	29	49	36	47	50	4
12	Harpo	Henry		1	30	51	34	50	53	4
13	Xena	Xerxes		2	30	57	37	53	56	0
14	Zephina	Zoro		2	30	47	38	56	59	0
15	Guido	Garcia		1	31	52	29	59	62	1
16	Rhonda	Rostropovich		2	31	50	31	62	More	0
17	Aaron	Andrews		1	32	64	43			
18	Petula	Peters		2	32	64	33			
19	Yuan	Young		1	32	59	24			
20	Bellinda	Brown		2	33	38	41			
21	Charlotta	Cowen		2	33	47	50			
22	Igor	Ivanovich		1	33	53	43			
23	Quincy	Quirn		1	33	48	33			
24	Sally	Stebbens		2	33	51	32			

Ready

NUM

2:45 PM

The row labeled "More" is used to take care of any scores that are above the highest bin limit you listed. SPSS does not have a program for producing grouped frequency distributions.

Cumulative Frequency Distributions

We often want to know how many people got scores below some particular value. The most direct way to answer this question is with a **cumulative frequency distribution**, which lists each score or interval and the number of scores falling in or below the score or interval. A cumulative frequency distribution is easily prepared from the frequency distribution or grouped frequency distribution, as shown in Table 2-4, which presents the **cumulative frequency**, as well as the frequency in each interval. Each entry in the column labeled "Cumulative Frequency" shows the total number of individuals having a score **equal to or less than** the highest score in that interval; that is, there is 1 student scoring at or below 20, $(1 + 2) = 3$ students scoring at or below 23, $(3 + 3) = 6$ scoring at or below 26, $(6 + 2) = 8$ scoring at or below 39, $(8 + 3) = 11$ scoring at or below 32, and so forth. The cumulative frequency distribution is especially useful for determining some expressions of relative position, which we will discuss in Chapter 3. Unfortunately, neither SPSS nor Excel has a routine for providing cumulative frequency distributions.

Table 2-4

Cumulative Frequency Distribution of Scores from 52 Students on a Math Test Using an Interval of 3

Interval	Frequency	Cumulative Frequency	Cumulative Percent
60-62	1	52	100
57-59	0	51	98
54-56	0	51	98
51-53	4	51	98
48-50	4	47	90
45-47	3	43	83
42-44	6	40	77
39-41	4	34	65
36-38	10	30	58
33-35	9	20	38
30-32	3	11	21
27-29	2	8	15
24-26	3	6	12
21-23	2	3	6
18-20	1	1	2

but both will give you cumulative percents; that is, each frequency distribution program has the ability to output the values equal to

$$\text{Cumulative percent} = \frac{\text{cumulative frequency}}{\text{total number of cases}}$$

These values are included in Table 2-4. The cumulative percents are given whenever you request a frequency distribution from SPSS. With Excel you must click on the Cumulative Percentages box in the Histogram dialogue box.

GRAPHIC REPRESENTATION

It is often helpful to translate the facts of a table like Table 2-3 into a pictorial representation. A common type of graphic representation, called a **histogram**, is shown in Figure 2-1. This type of graph can be thought of, somewhat grimly, as "piling up the bodies." The score intervals for the mathematics test scores that we used in Table 2-3 are shown along the baseline (the **abscissa**), and the vertical height of the pile (the **ordinate**) represents the number of people whose scores fall in that interval. The diagram shows that there is one "body" piled up in interval 18-20, two in interval 21-23, and so forth. (Because we used Excel to generate this histogram, the intervals are labeled with the upper bin limits.) This figure gives a clear picture of how the scores pile up, with most of them in the 30 to 45 range and long, low "tails" running out to the extreme low and high scores.

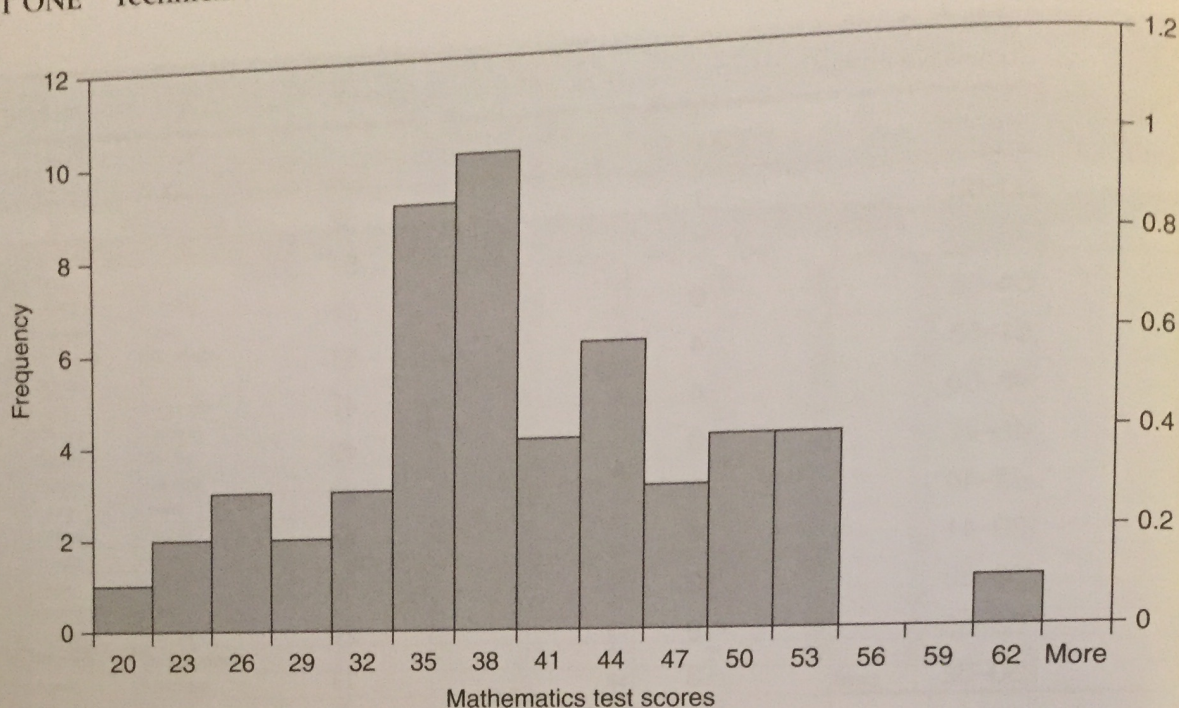


Figure 2-1

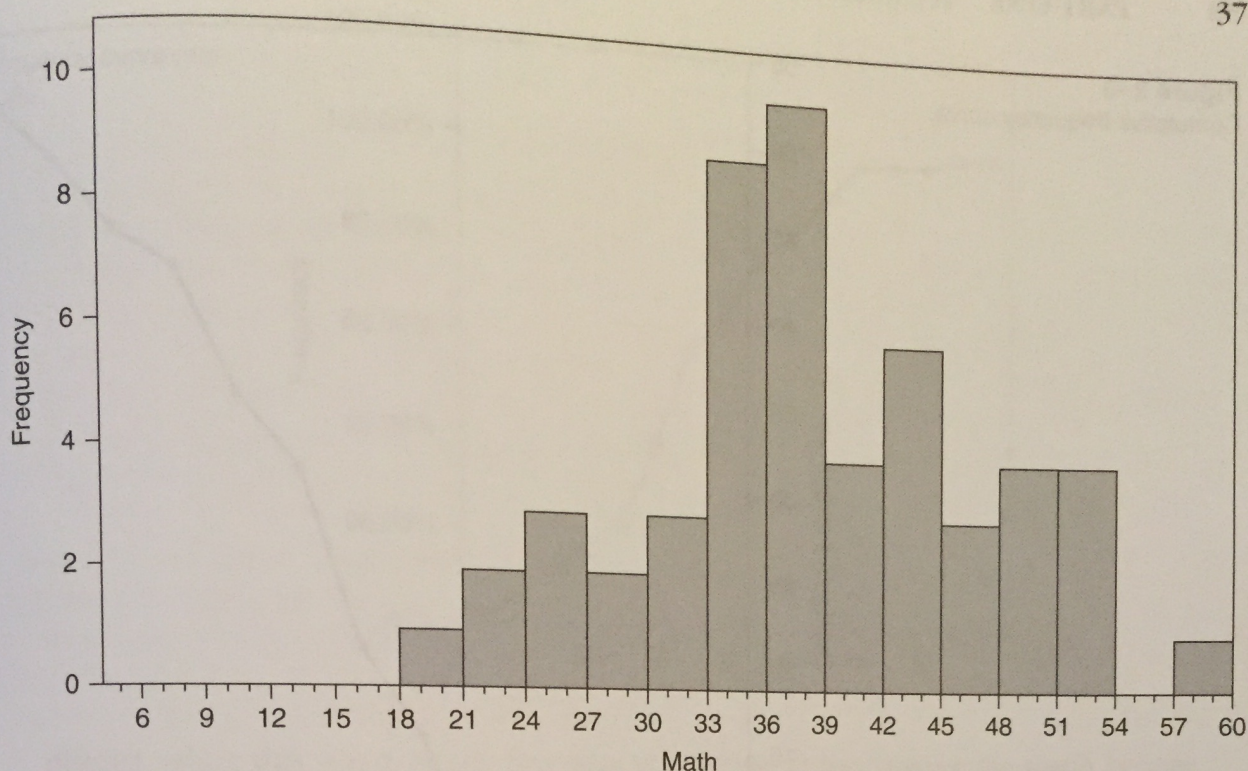
Histogram of 52 mathematics scores.

MAKING THE COMPUTER DO IT

Histograms

Both SPSS and Excel will prepare graphs of grouped frequency distributions, but they do it in very different ways. Excel's procedure is by far the simpler one, but from the technical point of view, the graph Excel's default options produce is properly called a **bar graph**, not a histogram because the bars are separated from each other. This can be corrected by double clicking on any bar, selecting "Options" and changing the "gap width" to zero. Unfortunately, the intervals are labeled with the upper limit of each "bin" rather than the interval midpoint, which may cause problems in graph interpretation if this is not pointed out. The lesson in this example is that you have to look at computer output with caution because the choices made by the programmer who wrote the program may not give you the exact picture you expect or want.

SPSS automatically groups the data to produce a graph with about 15 categories (the default for our example yields 13). The result is a proper histogram (the bars touch each other), but you may want to edit the graph to improve its appearance. If you double-click on the output of the graph, you can edit it. You may want to experiment with the graph editor to see what the editing features produce. The SPSS graph, on page 37, of the mathematics data has been edited to make it as similar as possible, within SPSS editing options, to the graph of our grouped frequency distribution. However, this graph does

**Figure 2-2**

Histogram of math scores produced by SPSS.

not represent the grouped frequency distribution shown in Table 2-4 and pictured in Figure 2-2.

One peculiarity of SPSS is that it treats the scale of the data as strictly continuous, which can create problems. Default interval widths and limits often are not obtainable score values, resulting in interval widths of 3.33, 5.17 and so forth. For the results of psychological and educational measurements where fractional values do not occur, this can produce a histogram that is difficult to interpret. Also, note that in the graph above the axis labeled "Math" has the upper limits of the intervals given and that the interval 57-60 is considered by SPSS to start just above 57 (57.0000001) and go up to and include exactly 60, the highest score. This is different from what we get using "by hand" methods, but it will only occur when the highest value in the data corresponds to the upper limit of the highest interval and is a consequence of the particular way SPSS represents scores (which is different from the traditions of psychological and educational measurement).

We can use a similar procedure to provide a graphic representation of the cumulative frequency distribution. This graph, known as the **cumulative frequency curve** (also sometimes called an **ogive**), is prepared by placing values of the cumulative frequency on the ordinate, and scores or score intervals on the abscissa. A point representing the cumulative frequency for each score or interval is then plotted, and the points are connected, forming a graph such as the one shown in Figure 2-3. Note that the cumulative frequency curve never drops back toward the abscissa.