1

Why the Social Researcher Uses Statistics

A little of the social scientist can be found in all of us. Almost daily, we take educated guesses concerning the future events in our lives in order to plan for new situations or experiences. As these situations occur, we are sometimes able to confirm or support our ideas; other times, however, we are not so lucky and must face the sometimes unpleasant consequences.

Consider some familiar examples: We might invest in the stock market, vote for a political candidate who promises to solve domestic problems, play the horses, take medicine to reduce the discomfort of a cold, throw dice in a gambling casino, try to psych out our instructors regarding a midterm, or accept a blind date on the word of a friend.

Sometimes we win; sometimes we lose. Thus, we might make a sound investment in the stock market, but be sorry about our voting decision; win money at the craps table, but discover we have taken the wrong medicine for our illness; do well on a midterm, but have a miserable blind date; and so on. It is unfortunately true that not all of our everyday predictions will be supported by experience.

The Nature of Social Research

Similar to our everyday approach to the world, social scientists attempt to explain and predict human behavior. They also take "educated guesses" about the nature of social reality, although in a far more precise and structured manner. In the process, social scientists examine characteristics of human behavior called *variables*—characteristics that differ or vary from one individual to another (for example, age, social class, and attitude) or from one point in time to another (for example, unemployment, crime rate, and population).

Not all human characteristics vary. It is a fact of life, for example, that the gender of the person who gave birth to you is female. Therefore, in any group of individuals, gender of mother is the *constant* "female." A biology text would spend considerable time discussing why only females give birth and the conditions under which birth is possible, but a social scientist would consider the mother's gender a given, one that is not worthy of study because it never varies. It could not be used to explain differences in the mental health of children because all of their mothers are females. In contrast, a mother's age, race, and mental health are variables: In any group of individuals, they will differ from person to person and can be the key to a greater understanding of the development of the child. A researcher therefore might study differences in the mental health of children depending on the age, race, and mental health of their mothers.

In addition to specifying variables, the social researcher must also determine the *unit of observation* for the research. Usually, social scientists collect data on individual persons. For example, a researcher might conduct interviews to determine if the elderly are victimized by crime more often than younger respondents. In this case, an individual respondent is the unit to be observed by the social scientist.

However, researchers sometimes focus their research on *aggregates*—that is, on the way in which measures vary across entire collections of people. For example, a researcher might study the relationship between the average age of the population and the crime rate in various metropolitan areas. In this study, the units of observation are metropolitan areas rather than individuals.

Whether focusing on individuals or aggregates, the ideas that social scientists have concerning the nature of social reality are called *hypotheses*. These hypotheses are frequently expressed in a statement of the relationship between two or more variables: at minimum, an *independent variable* (or presumed cause) and a *dependent variable* (or presumed effect). For example, a researcher might hypothesize that socially isolated children watch more television than children who are well integrated into their peer groups, and he or she might conduct a survey in which both socially isolated and well-integrated children are asked questions regarding the time they spend watching television (social isolation would be the independent variable; TV-viewing behavior would be the dependent variable). Or a researcher might hypothesize that the one-parent family structure generates greater delinquency than the two-parent family structure and might proceed to interview samples of delinquents and nondelinquents to determine whether one or both parents were present in their family backgrounds (family structure would be the independent variable; delinquency would be the dependent variable).

Thus, not unlike their counterparts in the physical sciences, social researchers often conduct research to increase their understanding of the problems and issues in their field. Social research takes many forms and can be used to investigate a wide range of problems. Among the most useful research methods employed by social researchers for testing their hypotheses are the experiment, the survey, content analysis, participant observation, and secondary analysis. For example, a researcher may conduct an experiment to determine if arresting a wife batterer will deter this behavior in the future, a sample survey to investigate political opinions, a content analysis of values in youth magazines, a participant observation of an extremist political group, or a secondary analysis of government statistics on unemployment.

Why Test Hypotheses?

Social science is often referred to, quite unfairly, as the study of the obvious. However, it is desirable, if not necessary, to test hypotheses about the nature of social reality, even those that seem logical and self-evident. Our everyday commonsense observations are generally based on narrow, often biased preconceptions and personal experiences. These can lead us to accept without criticism invalid assumptions about the characteristics of social phenomena and behavior.

To demonstrate how we can be so easily misled by our preconceptions and stereotypes, consider what we "know" about mass murderers—those individuals who simultaneously kill at least four victims. According to popular thinking (and media portrayals), mass murderers are typically insane individuals who go berserk or run amok, expressing their anger in a spontaneous and impulsive outpouring of aggression. Moreover, they are usually regarded as total strangers to their victims, who are unlucky enough to be in the wrong place at the wrong time—at a shopping mall, on a commuter train, or in a fast-food restaurant.

The foregoing conception of mass murderers may seem clear-cut and obvious. Yet, compiling detailed information from FBI reports about 697 mass killers over the period from 1976 to 1995, Fox and Levin found instead that mass murderers are rarely insane and spontaneous—they know exactly what they are doing and are not driven to kill by voices of demons. Random shootings in a public place are the exceptions; most mass murders occur within families or among acquaintances. Typically, mass murderers target spouses and all of their children, or bosses and their co-workers. Far from being impulsive, most mass killers are methodical and selective. They usually plan their attacks and are quite selective as to the victims they choose to kill. In an office massacre, for example, a mass killer might choose to murder only those co-workers and supervisors whom the murderer blames for losing an important promotion or getting fired.

Until recently, even criminologists all but ignored mass killings, perhaps believing that mass murder was merely a special case of homicide (albeit, by definition, yielding a larger body count), explainable by the same theories applied to single-victim incidents and therefore not deserving of special treatment. From this point of view, mass murder occurs in the same places, under the same circumstances, and for the same reasons as single-victim murder.

Comparing FBI reports of single-victim homicides with mass murders reveals quite a different pattern. The location of mass murder differs sharply from that of homicides in which a single victim is slain. First, mass murders do not tend to cluster in large cities as do single-victim crimes; rather, mass killings are more likely to occur in small-town or rural settings. Moreover, while the South (and the deep South in particular) is known for its high rates of murder, this does not hold for mass murder. In comparison to single-victim murder, which is highly concentrated in urban inner-city neighborhoods and in the deep South where arguments are often settled through gunfire, mass murder more or less reflects the general population distribution.

Not surprisingly, the firearm is the weapon of choice in mass-murder incidents, even more than in single-victim crimes. Clearly, a handgun or rifle is the most effective means of mass destruction. By contrast, it is difficult to kill large numbers of people simultaneously with physical force or even a knife or blunt object. Furthermore, although an explosive device can potentially cause the death of large numbers of people (as in the 1995 bombing of the Oklahoma City federal building), its unpredictability would be unacceptable for most mass killers who target their victims selectively. In addition, far fewer Americans are proficient in the use of explosives, as compared with guns.

The findings regarding victim–offender relationships are perhaps as counterintuitive as the weapon-use results may be obvious. Contrary to popular belief, mass murderers infrequently attack strangers who just happen to be in the wrong place at the wrong time. In fact, almost 40% of these crimes are committed against family members, and almost as many involve other victims acquainted with the perpetrator (for example, co-workers). It is well known that murder often involves family members, but this is especially pronounced among massacres.

The differences in circumstance underlying these crimes are quite dramatic. Although more than half of all single-victim homicides occur during an argument between the victim and the offender, it is relatively rare for a heated dispute to escalate into mass murder.

Some of the most notable differences between homicide types emerge in the offender data. Compared to those offenders who kill but one, mass murderers are especially likely to be male, are far more likely to be white, and are somewhat older (middle-aged). Typically, the single-victim offender is a young male and slightly more often black than white.

Victim characteristics are, of course, largely a function of the offender characteristics, indicating that mass killers generally do not select their victims on a random basis. For example, the victims of mass murder are usually white simply because the perpetrators to whom they are related or with whom they associate are white. Similarly, the youthfulness and greater representation of females among the victims of mass murder, as compared to single-victim homicide, stem from the fact that a typical mass killing involves the breadwinner of the household who annihilates the entire family—his wife and his children.

The Stages of Social Research

Systematically testing our ideas about the nature of social reality often demands carefully planned and executed research in which the following occur:

1. The problem to be studied is reduced to a testable hypothesis (for example, "one-parent families generate more delinquency than two-parent families").
2. An appropriate set of instruments is developed (for example, a questionnaire or an interview schedule).
3. The data are collected (that is, the researcher might go into the field and conduct a poll or a survey).
4. The data are analyzed for their bearing on the initial hypotheses.
5. Results of the analysis are interpreted and communicated to an audience (for example, by means of a lecture, journal article, or press release).

As we shall see in subsequent chapters, the material presented in this book is most closely tied to the data-analysis stage of research (see number 4 earlier), in which the data collected or gathered by the researcher are analyzed for their bearing on the initial hypotheses. It is in this stage of research that the raw data are tabulated, calculated, counted, summarized, rearranged, compared, or, in a word, *organized,* so that the accuracy or validity of the hypotheses can be tested.

Using Series of Numbers to Do Social Research

Anyone who has conducted social research knows that problems in data analysis must be confronted in the planning stages of a research project, because they have a bearing on the nature of decisions at all other stages. Such problems often affect aspects of the research design and even the types of instruments employed in collecting the data. For this reason, we constantly seek techniques or methods for enhancing the quality of data analysis.

Most researchers would agree on the importance of *measurement* in analyzing data. When some characteristic is measured, researchers are able to assign to it a series of numbers according to a set of rules. Social researchers have developed measures of a wide range of phenomena, including occupational prestige, political attitudes, authoritarianism, alienation, anomie, delinquency, social class, prejudice, dogmatism, conformity, achievement, ethnocentrism, neighborliness, religiosity, marital adjustment, occupational mobility, urbanization, sociometric status, and fertility.

Numbers have at least three important functions for social researchers, depending on the particular *level of measurement* that they employ. Specifically, series of numbers can be used to

1. *classify* or *categorize* at the nominal level of measurement,
2. *rank* or *order* at the ordinal level of measurement, and
3. assign a *score* at the interval/ratio level of measurement.

The Nominal Level

The *nominal level of measurement* involves naming or labeling—that is, placing cases into categories and counting their frequency of occurrence. To illustrate, we might use a nominal-level measure to indicate whether each respondent is prejudiced or tolerant toward Latinos. As shown in Table 1.1, we might question the 10 students in a given class and determine that 5 can be regarded as (1) prejudiced and 5 can be considered (2) tolerant.

TABLE 1.1 ***Attitudes of 10 College Students toward Latinos: Nominal Data***

Attitude toward Latinos	Frequency
1 = prejudiced	5
2 = tolerant	5
Total	10

Other nominal-level measures in social research are sex (male versus female), welfare status (recipient versus nonrecipient), political party (Republican, Democrat, and Libertarian), social character (inner-directed, other-directed, and tradition-directed), mode of adaptation (conformity, innovation, ritualism, retreatism, and rebellion), and time orientation (present, past, and future), to mention only a few.

When dealing with nominal data, we must keep in mind that *every case must be placed in one, and only one, category.* This requirement indicates that the categories must be nonoverlapping, or *mutually exclusive.* Thus, a respondent's race classified as white cannot also be classified as black; any respondent labeled male cannot also be labeled female. The requirement also indicates that the categories must be *exhaustive*—there must be a place for every case that arises. For illustrative purposes, imagine a study in which all respondents are interviewed and categorized by race as either black or white. Where would we categorize a Chinese respondent if he or she were to appear? In this case, it might be necessary to expand the original category system to include Asians or, assuming that most respondents will be white or black, to include an "other" category in which such exceptions can be placed.

The reader should note that nominal data are not graded, ranked, or scaled for qualities, such as better or worse, higher or lower, more or less. Clearly, then, a nominal measure of sex does not signify whether males are superior or inferior to females. Nominal data are merely labeled, sometimes by name (male versus female or prejudiced versus tolerant), other times by number (1 versus 2), but always for the purpose of grouping the cases into separate categories to indicate sameness or differentness with respect to a given quality or characteristic. Thus, even when a number is used to label a category (for example, 1 = white, 2 = black, 3 = other), a quantity is not implied.

The Ordinal Level

When the researcher goes beyond the nominal level of measurement and seeks to order his or her cases in terms of the degree to which they have any given characteristic, he or she is working at the *ordinal level of measurement*. The nature of the relationship among ordinal categories depends on that characteristic the researcher seeks to measure. To take a familiar example, one might classify individuals with respect to socioeconomic status as lower class, middle class, or upper class. Or, rather than categorize the students in a given classroom as *either* prejudiced *or* tolerant, the researcher might rank them according to their degree of prejudice against Latinos, as indicated in Table 1.2.

The ordinal level of measurement yields information about the ordering of categories, but does not indicate the *magnitude of differences* between numbers. For instance, the social researcher who employs an ordinal-level measure to study prejudice toward Latinos *does not know how much more prejudiced one respondent is than another.* In the example given in Table 1.2, it is not possible to determine how much more prejudiced Joyce is than Paul or how much less prejudiced Ben is than Linda or Ernie. This is because the intervals between the points or ranks on an ordinal scale are not known or meaningful. Therefore, it is not possible to assign *scores* to cases located at points along the scale.

TABLE 1.2 ***Attitudes of 10 College Students toward Latinos: Ordinal Data***

Student	Rank
Joyce	1 = most prejudiced
Paul	2 = second
Cathy	3 = third
Mike	4 = fourth
Judy	5 = fifth
Joe	6 = sixth
Kelly	7 = seventh
Ernie	8 = eighth
Linda	9 = ninth
Ben	10 = least prejudiced

The Interval/Ratio Level

By contrast to the ordinal level, the *interval and ratio levels of measurement* not only indicate the ordering of categories but also the exact distance between them. Interval and ratio measures employ constant units of measurement (for example, dollars or cents, Fahrenheit or Celsius, yards or feet, minutes or seconds), which yield equal intervals between points on the scale.

Some variables in their natural form are interval/ratio level—for example, how many pounds you weigh, how many siblings you have, or how long it takes a student to complete an exam. In the social sciences, naturally formed interval/ratio measures might include the length of a prison sentence, the number of children in a family, or the amount of time—in minutes and hours—an individual spends on the job.

Other variables are interval/ratio because of how we scale them. Typically, an interval/ratio measure that we construct generates a set of scores that can be compared with one another. As currently used by social scientists, for example, a well-known measure of job satisfaction, employed by Tom W. Smith who directs the General Social Survey at the National Opinion Research Center, is treated as an interval variable. In this process, respondents are asked to indicate how satisfied they are with the work they do on a four-point rating scale consisting of 1 for someone who is "very dissatisfied," 2 for someone who is "a little dissatisfied," 3 for someone who is "moderately satisfied," and 4 for someone who is "very satisfied." The occupations are then placed in a hierarchy from lowest to highest, depending on the overall evaluations—the mean satisfaction score—they receive from a group of respondents who hold the jobs they are asked to judge. In one recent study, for example, the job title *clergy* received a rating of 3.79 (almost at the "very satisfied" level), whereas *waiters* received a 2.85 (close to the "moderately satisfied" level); *physical therapists* got a score of 3.72, whereas *roofers* received a 2.84.

TABLE 1.3 ***Satisfaction Scores of Eight Jobs: Interval Data***

Job	Satisfaction Score
Clergy	3.79
Teachers	3.61
Authors	3.61
Psychologists	3.59
Butchers	2.97
Cashiers	2.94
Bartenders	2.88
Roofers	2.84

As depicted in Table 1.3, we are able to order a group of eight occupations in terms of their degree of satisfaction and, in addition, determine the exact distances separating one from another. This requires making the assumption that our measure of job satisfaction uses a constant unit of measurement (one satisfaction point). Thus, we can say that the job of clergy is the most satisfying on the list because it received the highest score on the measure. We can also say that authors are only slightly more satisfied than psychologists, but much more satisfied than bartenders and roofers, both of which received extremely low scores. Depending on the purpose for which a study is designed, such information might be important to determine, but is not available at the ordinal level of measurement.

The ratio level is the same as the interval level, but in addition presumes the existence of an absolute or true zero point. In contrast, an interval level variable may have an artificial zero value or even none at all.

For example, age meets the condition for the ratio level, because a zero represents birth, or the complete absence of age. In contrast, the Fahrenheit scale of temperature possesses an artificial zero point, because "zero degrees" does not represent the total absence of heat, even though it does not feel particularly warm. Similarly, the IQ scale has no zero point at all—that is, there is no such thing as a zero IQ—and therefore qualifies only as an interval scale. Thus, we cannot say that a person with an IQ of 150 is 50% more intelligent than someone with an average 100 IQ.

Similarly, a score of zero on a scale of occupational satisfaction, if it existed, would indicate a total absence of any satisfaction at all ("complete dissatisfaction"), and therefore potentially represents a ratio scale. As constructed by the author, however, the scale of occupational prestige illustrated previously has not been given a score of zero (a score of "1" indicates "very" but not complete dissatisfaction) and is therefore at the interval, not the ratio, level.

When it comes right down to it, it makes little practical difference whether a variable is interval or ratio level. There are many important statistical techniques that assume a standard distance between scale points (that is, an interval scale), but there are very few that require valid ratios between scale points (that is, a ratio scale). Thus, throughout the remainder of the book, we shall indicate whether a technique requires the nominal level, the ordinal level, or the interval level.

Different Ways to Measure the Same Variable

As noted earlier, the level of measurement of certain naturally occurring variables like gender or hair color is clear-cut, while others are constructed by how the social researcher defines them. In fact, the same variable can be measured at different levels of precision depending on the research objectives.

Figure 1.1, for example, illustrates several ways in which the variable "pain" might be measured by a researcher interested in health issues. At the lowest level—nominal—respondents could be classified as being in pain or pain-free, or they could be asked to indicate what type of pain they are experiencing. Of course, the extent of pain could be measured in an ordinal scale ranging from none to severe or respondents could indicate whether their pain was better, worse, or the same. More precisely, the degree of pain could be reported numerically from 0 to 10, reflecting an interval/ratio level scale. Alternative

Nominal Level

Question: Are you currently in pain? Yes or No

Question: How would you characterize the type of pain? Sharp, Dull, Throbbing

Ordinal Level

Question: How bad is the pain right now? None, Mild, Moderate, Severe

Question: Compared with yesterday, is the pain less severe, about the same, or more severe?

Interval/Ratio Level

0–10 Numerical Scale

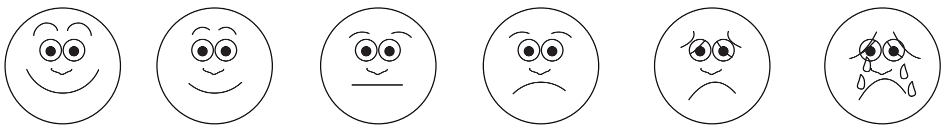
0	1	2	3	4	5	6	7	8	9	10
No pain										Worst pain imaginable

Visual Analog Scale

No pain — Worst pain

Ask patient to indicate on the line where the pain is in relation to the two extremes. Quantification is only approximate; for example, a midpoint mark would indicate that the pain is approximately half of the worst possible pain.

Pain Faces Scale



0	2	4	6	8	10
Very happy, no hurt	Hurts just a little bit	Hurts a little more	Hurts even more	Hurts a whole lot	Hurts as much as you can imagine

FIGURE 1.1 Different ways to measure pain

TABLE 1.4 ***Attitudes of 10 College Students toward Latinos: Interval Data***

Student	Score[a]
Joyce	98
Paul	96
Cathy	95
Mike	94
Judy	22
Joe	21
Kelly	20
Ernie	15
Linda	11
Ben	6

[a]Higher scores indicate greater prejudice against Latinos.

ways of measuring pain at the interval/ratio level include having respondents indicate their degree of pain by marking on a continuum, from no pain to worst pain. Finally, the Pain Faces Scale might be used for children with limited verbal skills or foreign-speaking adults.

Treating Some Ordinal Variables as Interval

As we have seen, levels of measurement vary in terms of their degree of sophistication or refinement, from simple classification (nominal), to ranking (ordinal), to scoring (interval). Table 1.4 gives an example using interval data.

At this point, the distinction between the nominal and ordinal levels should be quite clear. It would be difficult to confuse the level of measurement attained by the variable "color of hair" (blond, redhead, brunette, and black), which is nominal, with that of the variable "condition of hair" (dry, normal, oily), which is ordinal.

The distinction between ordinal and interval, however, is not always clear-cut. Often, variables that in the strictest sense are ordinal may be treated as if they were interval when the ordered categories are fairly evenly spaced. Actually, an earlier example—the measure of job satisfaction—can be used to make this point. To treat this measure as interval rather than ordinal, it is necessary to assume that the distance between "very dissatisfied" and "a little dissatisfied" is roughly the same as the distance between "a little dissatisfied" and "moderately satisfied" and between "moderately satisfied" and "very satisfied." If we are unable to make the assumption of equal intervals between the points on the scale, then the satisfaction measure should be treated as an ordinal scale.

To take another example, the following two variables (*rank of professor* and *attitude toward professor*) are both ordinal:

Scale Value	**Rank of Professor**	**Attitude toward Professor**
1	Distinguished professor	Very favorable
2	Full professor	Favorable
3	Associate professor	Somewhat favorable
4	Assistant professor	Neutral
5	Instructor	Somewhat unfavorable
6	Lecturer	Unfavorable
7	Teaching assistant	Very unfavorable

The *rank-of-professor* variable could hardly be mistaken for interval. The difference between *instructor* (5) and *lecturer* (6) is minimal in terms of prestige, salary, or qualifications, whereas the difference between *instructor* (5) and *assistant professor* (4) is substantial, with the latter generally requiring a doctorate and receiving a much higher salary. By contrast, the *attitude-toward-professor* variable has scale values that are roughly evenly spaced. The difference between *somewhat unfavorable* (5) and *unfavorable* (6) *appears* to be virtually the same as the difference between *somewhat unfavorable* (5) and *neutral* (4). In fact, this is true of most attitude scales ranging from *strongly agree* to *strongly disagree*.

Rather than split hairs, many researchers make a practical decision. Whenever possible, they choose to treat ordinal variables as interval, but only when it is reasonable to assume that the scale has roughly equal intervals. Thus, they would treat the *attitude-toward-professor* variable as if it were interval, but they would never treat the *rank-of-professor* variable as anything other than ordinal. As you will see later in the text, treating ordinal variables that have nearly evenly spaced values as if they were interval allows researchers to use more powerful statistical procedures.

Further Measurement Issues

Whether a variable is measured at the nominal, ordinal, or interval level is sometimes a natural feature of the characteristic itself, and not at all influenced by the decisions that the social researcher makes in defining and collecting data. Hair color (black, brown, blonde, gray, and so on), race (black, white, Asian), and region of residence (Northeast, Mid-Atlantic, South, Midwest, Mountain, and West) are, for example, unquestionably nominal-level variables. A researcher, however, can still expand the meaning of basic characteristics like these in an attempt to increase the precision and power of his or her data. Hair color, for example, can be redefined in terms of shades (for example, from dark brown to platinum blonde) to elevate the level of measurement to ordinal status. Similarly, for the purpose of measuring geographic proximity to Southern culture, an ordinal-level "Southerness scale" might be developed to distinguish Mississippi and Alabama at one extreme, Kentucky and Tennessee next, followed by Maryland and Delaware, and then Connecticut and Vermont at the other extreme. Although it may be somewhat stretching the point, a researcher could also develop an interval-level Southerness scale, using the number of miles a state's center lies above or below the Mason–Dixon line.

More commonly, there are situations in which variables must be downgraded in their level of measurement, even though this might reduce their precision. To increase the response rate, for example, a telephone interviewer might redefine age, an interval-level variable, into ordinal categories such as toddler, child, teenager, young adult, middle-aged, and senior.

Another important measurement distinction that social researchers confront is between discrete and continuous variables. Discrete data take on only certain specific values. For example, family size can be expressed only in whole numbers from 1 on up (there is no such thing as 3.47 people in a family; it's either 1, 2, 3, 4, or more members). Family size therefore represents a discrete interval-level measure. Moreover, nominal variables (such as *New England states:* Massachusetts, Connecticut, Rhode Island, Vermont, Maine, and New Hampshire; *gender:* female and male; *religion:* Protestant, Catholic, Jewish, Muslim, Hindu), by virtue of their categorical nature, are always discrete.

Continuous variables, on the other hand, present an infinite range of possible values, although the manner in which we measure them may appear to be discrete. Body weight, for example, can take on any number of values, including 143.4154 pounds. Some bathroom scales may measure this weight to the nearest whole pound (143 pounds), and others may measure weight to the nearest half pound (143.5), and some even to the nearest tenth of a pound (143.4). Underlying whatever measuring device we use, however, is a natural continuum. Similarly, age is a continuous variable and theoretically could be measured in nanoseconds from birth on. Yet it is customary to use whole numbers (years for adults, weeks for infants) in recording this variable. As shown earlier, it is also a common practice arbitrarily to divide the continuum of age into categories such as toddler, child, teenager, young adult, middle-aged, and senior.

The Functions of Statistics

When researchers use numbers—they *quantify* their data at the nominal, ordinal, or interval level of measurement—they are likely to employ statistics as a tool of (1) *description* or (2) *decision making*. Let us now take a closer look at these important functions of statistics.

Description

To arrive at conclusions or obtain results, a social researcher often studies hundreds, thousands, or even larger numbers of persons or groups. As an extreme case, the U.S. Bureau of the Census conducts a complete enumeration of the U.S. population, in which millions of individuals are contacted. Despite the aid of numerous sophisticated procedures, it is always a formidable task to describe and summarize the mass of data generated from projects in social research.

To take a familiar example, the examination grades of 80 students have been listed in Table 1.5. Do you see any patterns in these grades? Can you describe these grades in a few words? In a few sentences? Can you tell if they are particularly high or low on the whole?

TABLE 1.5 ***Examination Grades for 80 Students***

72	49	81	52	31
38	81	58	68	73
43	56	45	54	40
81	60	52	52	38
79	83	63	58	59
71	89	73	77	60
65	60	69	88	75
59	52	75	70	93
90	62	91	61	53
83	32	49	39	57
39	28	67	74	61
42	39	76	68	65
58	49	72	29	70
56	48	60	36	79
72	65	40	49	37
63	72	58	62	46

Your answer to these questions should be "no." However, using even the most basic principles of descriptive statistics, it is possible to characterize the distribution of the examination grades in Table 1.5 with a good deal of clarity and precision, so that overall tendencies or group characteristics can be quickly discovered and easily communicated to almost anyone. First, the grades can be rearranged in consecutive order (from highest to lowest) and grouped into a much smaller number of categories. As shown in Table 1.6, this *grouped frequency distribution* (to be discussed in detail in Chapter 2) presents the grades within broader categories along with the number or *frequency* (f) of students whose grades fell into these categories. It can be readily seen, for example, that 17 students received grades between 60 and 69; only 2 students received grades between 20 and 29.

TABLE 1.6 ***Examination Grades for 80 Students: A Grouped Frequency Distribution***

Grades	f
90–99	3
80–89	7
70–79	16
60–69	17
50–59	15
40–49	11
30–39	9
20–29	2

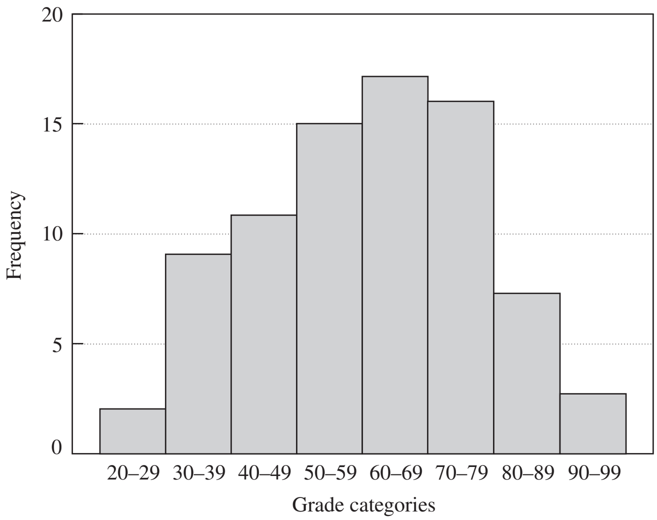


FIGURE 1.2 Graph of examination grades for 80 students

Another useful procedure (explained in Chapter 2) rearranges the grades graphically. As shown in Figure 1.2, the categories of grades are placed (from 20–29 to 90–99) along one line of a graph (that is, the *horizontal base line*) and their numbers or frequencies along another line (that is, the *vertical axis*). This arrangement results in a rather easily visualized graphic representation in which we can see that most grades fall between 50 and 80, whereas relatively few grades are much higher or lower.

As elaborated on in Chapter 3, a particularly convenient and useful statistical method—one with which you are already more or less familiar—is to ask: What is the grade of the *average* person in this group of 80 students? The arithmetic average (or *mean*), which can be obtained by adding the entire list of grades and dividing this sum by the number of students, gives us a clearer picture of the overall group tendency or class performance. The arithmetic average in this example happens to be 60.5, a rather low grade compared against the class averages with which most students may be familiar. Apparently, this group of 80 students did relatively poorly as a whole.

Thus, with the help of statistical devices, such as grouped frequency distributions, graphs, and the arithmetic average, it is possible to detect and describe patterns or tendencies in distributions of scores (for example, the grades in Table 1.5) that might otherwise have gone unnoticed by the casual observer. In the present context, then, statistics may be defined as *a set of techniques for the reduction of quantitative data (that is, a series of numbers) to a small number of more convenient and easily communicated descriptive terms.*

Decision Making

For purposes of testing hypotheses, it is frequently necessary to go beyond mere description. It is often also necessary to make inferences—that is, to make decisions based on data collected on only a small portion or *sample* of the larger group we have in mind to study. Factors such as cost, time, and need for adequate supervision many times preclude taking a complete enumeration or poll of the entire group (social researchers call this larger group from which the sample was drawn a *population* or *universe*).

As we shall see in Chapter 6, every time social researchers test hypotheses on a sample, they must decide whether it is indeed accurate to generalize the findings to the entire population from which they were drawn. Error inevitably results from sampling, even sampling that has been properly conceived and executed. This is the problem of generalizing or *drawing inferences* from the sample to the population.[2]

Statistics can be useful for purposes of generalizing findings, with a high degree of confidence, from small samples to larger populations. To understand better this decision-making purpose of statistics and the concept of generalizing from samples to populations, let us examine the results of a hypothetical study that was conducted to test the following hypothesis:

Hypothesis: *Male college students are more likely than female college students to have tried marijuana.*

The researchers in this study decided to test their hypothesis at an urban university in which some 20,000 students (10,000 males and 10,000 females) were enrolled. Due to cost and time factors, they were not able to interview every student on campus, but did obtain from the registrar's office a complete listing of university students. From this list, every one-hundredth student (one-half of them male, one-half of them female) was selected for the sample and subsequently interviewed by members of the research staff. The interviewers asked each of the 200 members of the sample whether he or she had ever tried marijuana and then recorded the student's gender as either male or female. After all interviews had been completed and returned to the staff office, the responses on the marijuana question were tabulated by gender, as presented in Table 1.7.

TABLE 1.7 ***Marijuana Use by Gender of Respondent: Case I***

	Gender of Respondent	
Marijuana Use	*Male*	*Female*
Number who have tried it	35	15
Number who have not tried it	65	85
Total	100	100

[2]The concept of *sampling error* is discussed in greater detail in Chapter 6. However, to understand the inevitability of error when sampling from a larger group, you may now wish to conduct the following demonstration. Refer to Table 1.5, which contains the grades for a population of 80 students. At random (for example, by closing your eyes and pointing), select a sample of five grades from the entire list. Find the average grade by adding the five scores and dividing by 5, the total number of grades. It has already been pointed out that the average grade for the entire class of 80 students was 60.5. To what extent does your sample average differ from the class average, 60.5? Try this demonstration on several more samples of a few grades randomly selected from the larger group. With great consistency, you should find that your sample mean will almost always differ at least slightly from that obtained from the entire class of 80 students. That is what we mean by *sampling error*.

TABLE 1.8 ***Marijuana Use by Gender of Respondent: Case II***

	Gender of Respondent	
Marijuana Use	*Male*	*Female*
Number who have tried it	30	20
Number who have not tried it	70	80
Total	100	100

Notice that results obtained from this sample of 200 students, as presented in Table 1.7, are in the hypothesized direction: 35 out of 100 males reported having tried marijuana, whereas only 15 out of 100 females reported having tried marijuana. Clearly, in this small sample, males were more likely than females to have tried marijuana. For our purposes, however, the more important question is whether these gender differences in marijuana use are large enough to generalize them confidently to the much larger university population of 20,000 students. Do these results represent true population differences? Or have we obtained chance differences between males and females due strictly to sampling error—the error that occurs every time we take a small group from a larger group?

To illuminate the problem of generalizing results from samples to larger populations, imagine that the researchers had, instead, obtained the results shown in Table 1.8. Notice that these results are still in the predicted direction: 30 males as opposed to only 20 females have tried marijuana. But, are we still willing to generalize these results to the larger university population? Is it not likely that a difference of this magnitude (10 more males than females) would have happened simply by chance? Or can we confidently say that such relatively small differences reflect a real difference between males and females at that particular university?

Let us carry out the illustration a step further. Suppose that the social researchers had obtained the data shown in Table 1.9. Differences between males and females shown in the table could not be much smaller and still be in the hypothesized direction: 26 males in contrast to 24 females tried marijuana—only 2 more males than females. How many of

TABLE 1.9 ***Marijuana Use by Gender of Respondent: Case III***

	Gender of Respondent	
Marijuana Use	*Male*	*Female*
Number who have tried it	26	24
Number who have not tried it	74	76
Total	100	100

us would be willing to call *this* finding a true population difference between males and females rather than a product of chance or sampling error?

Where do we draw the line? At what point does a sample difference become large enough so that we are willing to treat it as significant or real? With the aid of statistics, we can readily, and with a high degree of confidence, make such decisions about the relationship between samples and populations. To illustrate, had we used one of the statistical tests of significance discussed later in this text (for example, chi-square; see Chapter 9), we would already have known that *only those results* reported in Table 1.7 can be generalized to the population of 20,000 university students—that 35 out of 100 males but only 15 out of 100 females have tried marijuana is a finding substantial enough to be applied to the entire population with a high degree of confidence and is therefore referred to as a *statistically significant difference.* Our statistical test tells us there are only 5 chances out of 100 that we are wrong! By contrast, application of the same statistical criterion shows the results reported in Tables 1.8 and 1.9 are *statistically nonsignificant,* probably being the product of sampling error rather than real gender differences in the use of marijuana.

In the present context, then, statistics is *a set of decision-making techniques that aid researchers in drawing inferences from samples to populations and, hence, in testing hypotheses regarding the nature of social reality.*

An Important Note about Rounding

If you are like most students, the issue of rounding can be confusing. It is always a pleasure, of course, when an answer comes out to be a whole number because rounding is not needed. For those other times, however, when you confront a number such as 34.233333 or 7.126534, determining just how many digits to use in rounding becomes problematic.

For occasions when you need to round, the following rule can be applied: *Round a final answer to two more decimal digits than contained in the original scores.* If the original scores are all whole numbers (for example, 3, 6, 9, and 12), then round your final answer to two decimal places (for example, 4.45). If the original scores contain one decimal place (for example, 3.3, 6.0, 9.5, and 12.8), then round your answer to three decimal places (for example, 4.456). A discussion of *how* to round is given in Appendix B.

Many problems in this book require a number of intermediate steps before arriving at the final answer. When using a calculator, it is usually not necessary to round off calculations done along the way (that is, for intermediate steps). Your calculator will often carry many more digits than you will eventually need. As a general rule for intermediate steps, do not round until it comes time to determine your final answer.

Rules of thumb, of course, must be used with some degree of good judgment. As an extreme example, you would not want to round only to two decimal places in calculating the trajectory or thrust needed to send a missile to the moon; even a slight imprecision might lead to disaster. In doing problems for your statistics class, on the other hand, the precision of your answer is less important than learning the method itself. There may be times when your answer will differ slightly from that of your classmate or that contained in this book. For example, you may obtain the answer 5.55, whereas your classmate may get 5.56, yet you both may be correct. The difference is trivial and could easily have resulted from using two calculators with different memory capacities or from doing calculations in a different sequence.

In this text, we have generally followed this rule of thumb. In some illustrations, however, we rounded intermediate steps for the sake of clarity—but only to an extent that would not invalidate the final answer.

Summary

In the first chapter, we linked our everyday predictions about the course of future events with the experiences of social researchers who use statistics as an aid in testing their hypotheses about the nature of social reality. Almost daily, ordinary people take educated guesses about the future events in their lives. Unlike haphazard and biased everyday observations, however, researchers seek to collect *systematic* evidence in support of their ideas. Depending on the particular level of measurement, series of numbers are often employed by social researchers to categorize (nominal level), rank (ordinal level), or score (interval/ratio level) their data. Finally, social researchers are able to take advantage of two major functions of statistics in the data-analysis stage of social research: description (that is, reducing quantitative data to a smaller number of more convenient descriptive terms) and decision making (that is, drawing inferences from samples to populations).

Questions and Problems

1. Someone who ranks a list of cities from slowest to fastest pace of life is operating at the ___________ level of measurement.
a. nominal
b. ordinal
c. interval

2. A researcher who scores a set of respondents (from 0 to 10) in terms of their degree of empathy for accident victims is working at the ___________ level of measurement.
a. nominal
b. ordinal
c. interval

3. Identify the level of measurement—nominal, ordinal, or interval—represented in each of the following questionnaire items:
a. Your gender:
1. _____ Female
2. _____ Male

b. Your age:
1. _____ Younger than 20
2. _____ 20–29
3. _____ 30–39
4. _____ 40–49
5. _____ 50–59
6. _____ 60–69
7. _____ 70 or older

c. How many people are in your immediate family? ___________
d. Specify the highest level of education achieved by your mother:
1. _____ None
2. _____ Elementary school
3. _____ Some high school
4. _____ Graduated high school
5. _____ Some college
6. _____ Graduated college
7. _____ Graduate school

e. Your annual income from all sources: ___________ (specify)
f. Your religious preference:
1. _____ Protestant
2. _____ Catholic
3. _____ Jewish
4. _____ Muslim
5. _____ Other ___________ (specify)

g. The social class to which your parents belong:
1. _____ Upper
2. _____ Upper-middle
3. _____ Middle-middle
4. _____ Lower-middle
5. _____ Lower

h. In which of the following regions do your parents presently live?
1. _____ Northeast
2. _____ South
3. _____ Midwest
4. _____ West
5. _____ Other ___________ (specify)

i. Indicate your political orientation by placing an X in the appropriate space:
LIBERAL ___ :___ :___ :___ :___ CONSERVATIVE
1 2 3 4 5

4. For each of the following items, indicate the level of measurement—nominal, ordinal, or interval:
a. A tailor uses a tape measure to determine exactly where to cut a piece of cloth.
b. The speed of runners in a race is timed in seconds by a judge with a stopwatch.
c. Based on attendance figures, a ranking of the Top 10 rock concerts for the year is compiled by the editors of a music magazine.
d. A zoologist counts the number of tigers, lions, and elephants she sees in a designated wildlife conservation area.
e. A convenience store clerk is asked to take an inventory of all items still on the shelves at the end of the month.
f. The student life director at a small college counts the number of freshmen, sophomores, juniors, and seniors living in residence halls on campus.
g. Using a yardstick, a parent measures the growth of his child on a yearly basis.
h. In a track meet, runners in a half-mile race were ranked first, second, and third place.

5. A researcher who ranks a list of countries according to how much they have depleted their natural resources is working at the ___________ level of measurement.
   a. nominal
   b. ordinal
   c. interval

6. Governments can be divided into three different types—unitary governments, federal governments, and confederations—depending on where the concentration of power is located. This would be considered which level of measurement?
   a. Nominal
   b. Ordinal
   c. Interval

7. A sociologist conducts a survey to determine the effects of family size on various aspects of life. For each of the following questionnaire items, identify the level of measurement (nominal, ordinal, or interval):
   a. Does family size affect school performance? Students are asked to circle their letter grade (A, B, C, D, or F) in various school subjects.
   b. Does family size differ by socioeconomic status? Parents are asked to provide their yearly income in dollars.
   c. Does parental health differ by family size? Parents are asked to rate their overall health on a scale from 1 to 5, with 1 being in very good health and 5 being in very poor health.
   d. Do the effects of family size differ with race and ethnicity? Respondents are asked to indicate if they are Black, White, Hispanic, Asian, or other.

8. Identify the level of measurement (nominal, ordinal, or interval) in each of the following items:
   a. American psychologist William Sheldon developed the idea that there are three major body types: ectomorph, endomorph, and mesomorph.
   b. In a study of short-term memory, a psychologist measures in seconds the time it takes for participants to remember words and numbers that were told to them an hour earlier.
   c. The same psychologist then groups the participants according to how good their short-term memory is, distributing them into five categories that range from "Very good short-term memory" to "Very poor short-term memory."
   d. Participants in a study about eating disorders are asked how many times they eat per day.
   e. Based on blood pressure readings, a psychologist ranks the stressfulness of various activities on a scale of 1 to 10, with 1 being the least stressful and 10 being the most stressful.
   f. In a study on color blindness, a psychologist counts the number of times that participants are able to identify the colors red, yellow, and blue in order to categorize them as either color blind or not color blind.
   g. A researcher interested in family relations focuses on the birth order of siblings.

9. For a very small group of his clients, a psychologist conducts a survey and determines that the most common phobia in the group is acrophobia (fear of heights). In this case, statistics are being used as a tool to perform which function?
   a. Ranking
   b. Population
   c. Description
   d. Decision making
   e. Generalization

Looking at the Larger Picture: A Student Survey

The chapters of this textbook each examine particular topics at close range. At the same time, it is important, as they say, to "see the forest through the trees." Thus, at the close of each major part of the text, we shall apply the most useful statistical procedures to the same set of data drawn from a hypothetical survey. This continuing journey should demonstrate the process by which the social researcher travels from having abstract ideas to confirming or rejecting hypotheses about human behavior. Keep in mind both here and in later parts of the book that "Looking at the Larger Picture" is not an exercise for you to carry out, but a summary illustration of how social research is done in practice.

For many reasons, surveys have long been the most common data-collection strategy employed by social researchers. Through the careful design of a survey instrument—a questionnaire filled out by survey respondents or an interview schedule administered over the telephone or in person—a researcher can elicit responses tailored to his or her particular interests. The adage "straight from the horse's mouth" is as true for informing social researchers and pollsters as it is for handicapping the Kentucky Derby.

A rather simple yet realistic survey instrument designed to study smoking and drinking among high school students follows. The ultimate purpose is to understand not just the extent to which these students smoke cigarettes and drink alcohol, but the factors that explain why some students smoke or drink while others do not. Later in this book, we will apply statistical procedures to make sense of the survey results. But for now, it is useful to anticipate the kind of information that we can expect to analyze.

Suppose that this brief survey will be filled out by a group of 250 students, grades 9 through 12, in a hypothetical (but typical) urban high school. In this chapter, we introduced levels of measurement. Note that many of the variables in this survey are nominal—whether the respondent smokes or has consumed alcohol within the past month—as well as respondent characteristics, such as race and sex. Other variables are measured at the ordinal level—specifically, the extent of respondent's peer-group involvement, his or her participation in sports—exercise, as well as academic performance. Finally, still other variables are measured at the interval level—in particular, daily consumption of cigarettes as well as age and grade in school.

To experience firsthand the way that data are collected, you may decide to distribute this survey or something similar on your own. But just like on those television cooking shows, for our purposes here, we will provide at the end of each part of the text, "precooked" statistical results to illustrate the power of these techniques in understanding behavior. As always, it is important not to get caught up in details, but to see the larger picture.

Student Survey

Answer the following questions as honestly as possible. Do not place your name on the form so that your responses will remain completely private and anonymous.

1. What school grade are you currently in? _________
2. How would you classify your academic performance? Are you
   _____ an excellent, mostly A's student
   _____ a good, mostly B's student
   _____ an average, mostly C's student
   _____ a below average, mostly D's student
3. Within the past month, have you smoked any cigarettes?
   _____ Yes
   _____ No
4. If you are a smoker, how many cigarettes do you tend to smoke on an average day?
   _____ per day
5. Within the past month, have you had any beer, wine, or hard liquor?
   _____ Yes
   _____ No
6. If you have had beer, wine, or hard liquor in the past month, on how many separate occasions?
   _____ times

7. In terms of your circle of friends, which of the following would best describe you?
   _____ I have lots of close friends.
   _____ I have a few close friends.
   _____ I have one close friend.
   _____ I do not have any really close friends.
8. Does either of your parents smoke?
   _____ Yes
   _____ No
9. To what extent do you participate in athletics or exercise?
   _____ Very frequently
   _____ Often
   _____ Seldom
   _____ Never
10. What is your current age? _____ years old
11. Are you _____ Male _____ Female
12. How would you identify your race or ethnicity?
    _____ White
    _____ Black
    _____ Latino
    _____ Asian
    _____ Other