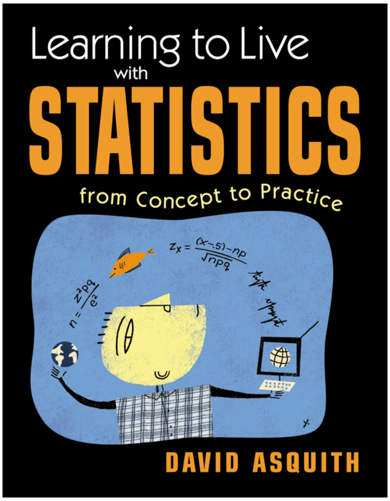
EXCERPTED FROM

# Learning to Live with Statistics: From Concept to Practice

David Asquith





ISBNs: 978-1-58826-524-1 hc
978-1-58826-549-4 pb

LYNNE RIENNER PUBLISHERS
1800 30th Street, Ste. 314
Boulder, CO 80301
USA
telephone 303.444.6684
fax 303.444.0824

# Contents

# CHAPTER 1

# Beginning Concepts

In this chapter, you will learn how to:

- Recognize nominal, ordinal, and interval-ratio levels of measurement
- Explain why the difference between levels of measurement is important in statistics
- Read frequency distributions
- Distinguish between descriptive and inferential statistical analyses

BEFORE GETTING INTO ANY DETAILS, STATISTICS, or formulas, it is worthwhile to consider both the scope of the text and some introductory concepts. This general overview introduces common questions and procedures in statistics and presents the sequence of topics in the text.

Statistics have options. First, there are different kinds of data or information with which we work, and they call for different statistical procedures. **Data*** consist of the measurements and numbers we summarize and analyze, but how do we decide which statistical methods are best? Which statistical procedures are appropriate and which are not? An important consideration in answering these questions is the **level or scale of measurement.** This refers to whether we have actual numbers or merely categories as data. Numbers naturally refer to actual quantities as data: the number of miles you drive per year, your age, how much money you earned last year, your exam scores, and so on. If you work for an immigration lawyer who wishes to know the typical length of time clients have spent in the United States before applying for

---

*Terms presented in bold face in the text appear in the glossary, which begins on page 315.

citizenship, you would have measurements or data points that consist of the numbers of months clients had been in the United States before submitting citizenship applications. Based on those numbers, you could calculate a typical or **average** time to applying for citizenship, and that would not be a complicated procedure. Sometimes, however, your data or measurements will not consist of numbers.

At a different level or scale of measurement, your information may consist simply of designations or categorizations people have made. Examples include checked boxes to indicate sex or gender, academic major, religious preference, and so on. These measurements are clearly not numerical. They are merely descriptive categories: female or male? Major in social science, chemistry, business, Spanish, etc.? Catholic, Buddhist, Protestant, Muslim, Jewish, and so on? These sorts of measurements or data would require different statistical treatments than would numerical information. Therefore, whether we have numerical or categorical data influences any decision as to what statistical procedures are acceptable. Simply put, the typical statistical procedures we use with a set of ages, for instance, do not work if we are asked to analyze data involving sex or religion, and vice versa. A further consideration is whether our purpose is descriptive or inferential statistical analysis.

Briefly, the distinction between **descriptive** and **inferential** statistical analyses refers to how broadly we wish to generalize our statistical results and the subsequent conclusions. If 200 people leaving the voting booth tell us their selections, is this a valid indicator of how the overall election may go? Well, maybe and maybe not. We may do statistical analyses on *any* set of data and simply *describe* that sample of cases. Summarizing how our 200 people voted would be easy enough. However, may we legitimately *infer* something about a whole population of voters from this 200? Are they representative of all voters in that precinct? If we wish to make *inferences* about larger populations, we must be especially careful to analyze truly representative and randomly selected samples from those populations. This is crucial; therefore, probability statements always accompany our inferences. What is the probability our inferences are correct? Incorrect? That makes inferential analyses different from a mere description. We will consider such issues in more depth later. For now, we turn to a text overview.

## A Preview: Text Overview

As the text progresses, we will build upon more elementary concepts. We start with material that is no doubt familiar. Chapter 2 starts with statistics that tell us the **central tendency** in a set of data. These statistics give us a sense of the typical cases or central themes in sets of numbers. **Measures of variation** follow; they tell us how much numbers in a set tend to vary from each other. Are they spread over a wide range or, conversely, do they tend to cluster near the average? As used in Chapter 2, measures of central tendency and of variation are descriptive statistics. They simply summarize sets of available data.

The next topic, covered in Chapter 3, is **probability.** Probability forms a bridge between descriptive statistics and inferential statistics. All statistical inferences include information as to the probability they are correct. Probability, however, is a varied topic in itself. Chapter 3 provides an overview of the field and looks at two types of measurement commonly used in statistics, continuous and discrete binomial variables. **Continuous** variables may be measured in fractions (i.e., in less than just whole units). For example, time or distance may be measured down to thousandths of a minute or a mile if we wish. In contrast, **discrete** measurements exist only in whole numbers. How many courses are you taking? How many TV sets in your home? These measurements, of necessity, are made in whole units or numbers. Moreover, as the prefix *bi* suggests, **binomial events** are those in which only two things or outcomes are possible. Whether a newborn is a girl or boy and whether a coin flip comes up heads or tails are examples of binomial situations, and we have unique procedures for determining such probabilities. Examples include the probability of a woman having two girls and then a boy or the probability of 7 heads in 10 flips of a coin. Whether our variables are continuous or discrete and binomial, however, we may take advantage of **probability distributions** to help us determine, for any data set, what numbers are the most or least likely to occur. One of these distributions is especially useful: the **normal distribution,** or bell-shaped curve. It is a prominent part of Chapter 3.

Following probability, Chapter 4 turns to inferential statistics. It introduces **sampling distributions,** which are special cases of normal distributions that form the theoretical foundation for making inferences. We use sampling distributions for **estimation,** that is, estimating unknown averages or **proportions** (percentages) for large populations based upon samples. For instance, how much does the average student spend on campus per week? What proportion, or percentage, of students favor an ethnic studies course being required for graduation?

Chapters 5 and 6 also use sampling distributions for **hypothesis testing.** Hypotheses are simply statements of what we expect to be true or what we expect our statistical results to show. Chapter 5 examines whether a single sample average or percentage differs from a corresponding population figure. We might test the hypothesis that the average time to graduation in a sample of college athletes does not differ significantly from the campus-wide or population average. Alternatively, we might test whether the percentage of people owning pets is significantly greater among people over age 65 than among adults in general. Chapter 6 looks at two-sample situations. Instead of comparing a sample to a population, we compare one sample to another and ask whether they differ significantly. At graduation, do the grade point averages in samples of transfer and nontransfer students differ? Among drivers under age 18, does a random sample of females receive significantly fewer tickets than a similar sample of males?

Next, while still dealing with hypothesis testing, Chapter 7 shifts gears somewhat. Chapter 7 differs from preceding topics in two respects. First, it compares sample data *not* to population data but rather to other criteria. How does our sample compare to what is expected by **random chance** or expected according to some other stated criterion? We now test hypotheses that our sample statistics do not differ significantly from random chance or from other presumed values. For example, if a campus bookstore sells sweatshirts in three school colors, is there a statistically significant preference for one color over the other two? Random chance says all three colors should be selected equally and that buyers pick their colors at random. But do actual sales vary significantly from what random chance suggests they should be?

Second, Chapter 7 also introduces hypothesis tests for **ranked** and for **categorical data.** Here, we use data that must first be ranked or are already broken into categories. We may compare ranked scores (e.g., from essay tests or opinion scales) in two samples. Categorized measurements may be related to gender (female/male), religious denomination (Catholic/Protestant/Jewish/Muslim/Buddhist), or type of vehicle driven (car/truck/van/motorcycle). In these situations, we look at data in categories and consider the number of cases expected to fall into each category by random chance versus the number of actual cases in that grouping. For vehicles driven, for instance, how many people would be expected to list "car" according to random chance versus how many people actually did list "car?" Are drivers in our sample more (or less) likely to use cars than random chance would suggest?

Chapter 7 also uses categorical data to establish correlations or associations between characteristics. Whereas a hypothesis test may tell us whether an association between two characteristics differs from random chance, other statistics allow us to calculate the approximate strength of that association. Various **measures of association** tell us how closely two characteristics are correlated.

Chapter 8 rounds out hypothesis testing by introducing situations involving three or more samples. For example, suppose we wish to compare the average weight losses for people on four different diet plans. Do the average losses differ significantly? Chapters 5, 6, and 7 deal with one- and two-sample cases. Somewhat different methods are needed when we have more than two samples, however, and these make up Chapter 8.

Finally, in Chapter 9, we return to correlations between sets of numbers. Do the numbers of hours studied per week correlate with grade point averages? Do more years of education translate into (correlate with) higher incomes? Besides such correlations, we will also consider how statistical predictions can be made. Suppose we knew that scores on a first statistics test correlated strongly with scores on the second test. We may then predict scores on that second test based upon students' scores on the first test and also get an idea of how accurate those predictions might be. This is known as **correlation and regression.**

Each chapter concludes with a set of practice exercises. Some are essay questions, but most call for you to diagnose the situations described, pick out the relevant bits of information, and solve the exercise by using that chapter's procedures. The exercises are based on real-world situations, and you are asked to translate those word problems into workable and complete solutions. The task is to identify the nature of a problem and to then use the correct statistical procedures in your analyses—without being told specifically which formulas to use. That is part of the learning process: diagnose a typical situation, consider what you are asked to do, and decide upon the appropriate statistical solution. Your instructor may assign selected questions as homework or as classroom exercises. You are strongly encouraged to try as many of these exercises as your time allows, even if they aren't assigned. There is no one single strategy more conducive to learning statistics than practice, practice, practice. The end-of-chapter problems are designed to cover all the procedures and possible alternatives introduced. If you can do the exercises, you have mastered the chapter.

With this general plan of the text in mind, the remainder of this chapter turns to two issues important in any statistical analysis. First, with which kind or level of data are we working? Second, is our analysis descriptive or inferential in nature?

## The Level of Measurement: Using the Right Tools

In the research process, **measurement** is the first step, preceding any statistical analyses. Measurement is simply a matter of being able to reliably and validly assess and record the status or quantity of some characteristic. For students' academic levels, for instance, we simply record their statuses as freshmen, sophomores, juniors, seniors, or graduates. We often refer to the things we measure as **variables,** that is, characteristics we expect to vary from one person or **element** to the next. Conversely, if something is true of *every* person or element, it is a **constant.** For example, grade point averages (or GPAs) among history majors would certainly vary, but the designation "history major" would be a constant. Constants become parts of our definition for the **population** we are studying, such as all upper-division history majors or all commuting students, and we often do not measure them directly. We assume everyone measures the same on those characteristics. Our measurements of variables, however, require statistical analyses.

We use the upper-case letter ***X*** to denote any single score. For instance, if we are measuring TV viewing, and if respondent number 17 watches 25 hours of TV per week, then $X_{17} = 25$. We need this way of referring to individual measurements. We must have a way to write formulas and express how we are going to treat the individual scores or observations. If we wish to square each measurement, we write "$X^2$." If we wish to multiply each *X* score by its companion *Y* score, it is "*XY*," and so on. "*X*" will appear in many formulas throughout the text, and we will have numerous occasions to refer to the "*X* variable."

One other feature of variables should be briefly noted here. To statisticians, the variables they analyze are obviously measurable and recordable. They have their measurements or data with which to work. Sometimes, however, a researcher assumes the variables reflect more broad or abstract concepts and characteristics. Research often includes more general and theoretical factors that do not lend themselves easily to direct measurement, such as personality, socioeconomic status (someone's location in a prestige or lifestyle hierarchy), employee morale, marital happiness, and so on. In these cases, the actual measurable variables or data become **empirical indicators** for the abstract concepts. For instance, answers to items on personality inventories (e.g., "Would you rather go to a movie with friends or stay home and read a bestselling book?") serve as indicators for broader and more theoretical dimensions of personality makeup. People may be asked about their incomes, educational histories, or occupations rather than directly asking their socioeconomic class levels. Measuring morale or marital happiness may mean asking about recommending one's job to a friend or about whether one has thought about divorce or would remarry the same person. Variables, then, and the resulting measurements will sometimes be obvious and concrete (sex, age), whereas at other times they may be indicators or reflections of more abstract concepts (sociability, mental health). In either case, the statistical analyses measurements used on these variables are of different kinds or **levels.**

The level of measurement involves the quantitative precision of our variables. Some variables naturally lend themselves to precise numerical measurements (e.g., age and income), whereas others do not (e.g., gender, academic major, and political party preference). Still other variables fall between these two extremes. Often, variables have been somewhat subjectively or arbitrarily quantified (e.g., test scores or attitudinal/opinion scores). Generally speaking, the *higher* or more quantitatively precise the level of measurement, the more we can do with the data statistically. The most precise levels are the **interval** and the **ratio.** Interval or ratio level data consist of legitimate and precise numerical measurements. This allows us to choose from a very wide range of statistical tools or operations. **Nominal** or **categorical** data, at the other end of the continuum, are sometimes described as nonquantitative. We are simply labeling or categorizing things (e.g., Democrat versus Republican) and are not measuring quantities. We may still summarize our data or test for multivariate correlations and so on, but our statistical choices are more limited. The basic point is that, with the higher levels of measurement, the more statistical options or possibilities we have. As we proceed through the following chapters, one consideration will be the kinds of statistical treatments appropriate when we have certain levels of data. Or, to think of this the other way around: If we wish to use a certain statistical procedure, what sort of data must we have?

### Creating Categories

The nonquantitative level of measurement is called the nominal level. This is the most simple and elementary level or scale of measurement. A nominal variable's categories or measurements are qualitatively different from each other. The categories—say, female and male—cannot be ranked or put in any natural order or sequence. Designating females as category 1 and males as category 2 would be no more legitimate or correct than making males 1 and females 2. Besides gender, other examples are racial or ethnic background, marital status, academic major, or religious denomination. Considering the latter, we might establish the following categories: Protestant, Catholic, Jewish, Buddhist, Muslim, Other. We would have six major categories, but that is all. We could not go an additional step and rank the categories from highest to lowest. That would be nonsensical; the categories of measurement have no natural or logical sequence to them. They make just a much sense in *any* order: Muslim, Jewish, Catholic, Buddhist, Protestant, Other. Or, alphabetically, we would have Buddhist, Catholic, Jewish, Muslim, Protestant, and finally Other. The same is true if we list major racial/ethnic categories. We could list them alphabetically as African American, Asian American, Caucasian, Latino, Other, but it would be just as valid to list them in a different sequence: Asian American, Latino, African American, Caucasian, Other. Our measurements or categories do not fall into any *one* order or sequence. The measurements differ *qualitatively* from each other, not *quantitatively.* We often refer to these measurements as qualitative or categorical variables. We are measuring differences of type, not differences of amount. For instance, a researcher might ask your racial or ethnic extraction; the researcher would not ask how *much* race or ethnicity you have. Race, ethnicity, and the other examples are simply nominal and not quantitative variables.

Statistical operations with nominal data use the category **counts** or tallies, also called category **frequencies.** For example, with a sample of 100 people ($n = 100$) that includes 53 females and 47 males, we use the 53 and the 47 in any statistical operations. Saying a variable is nonquantitative does not mean we cannot do *any* statistics with the data. It just means we may only use the frequencies, tallies, or counts when we do so. The categories of measurement themselves are nonnumerical, e.g., Latino or Asian American, or drivers of cars versus trucks. Even so, we may at least count *how many* people fall into the respective categories, and it is these latter numbers that we use in our statistical analyses. We will work with frequencies and nominal data when we look at cross-tabulations and measures of association.

Frequency distributions illustrate the different levels of measurement. **Frequency distributions** *show all the measurements recorded for a particular sample or population.* The figures typically include the actual numbers and percentages falling into each measurement category. Researchers typically look at such frequency distributions before doing anything else. The frequencies (or tallies or counts or percentages) represent the most elementary level of analysis and are purely descriptive.

We look first here at frequency distributions in the form of tables. Although not shown, pie charts, line graphs, or bar graphs may be used as well. Each chart or table shows the responses for a single variable, starting with nominal variables. Table 1.1 shows the distribution of marital statuses in a recent survey of students at a large state university. Participants were enrolled in randomly

**Table 1.1 Marital Status Among a Sample of Students at a State University**

| Marital Status | Frequency | Percentage | Valid Percentage |
|---|---|---|---|
| Single, solo | 336 | 45.3 | 46.2 |
| Single, attached | 280 | 37.7 | 38.5 |
| Married | 98 | 13.2 | 13.5 |
| Separated | 2 | .3 | .3 |
| Divorced | 10 | 1.3 | 1.4 |
| Widowed | 2 | .3 | .3 |
| Total | 728 | 98.1 | 100.0 |
| Missing | 14 | 1.9 | |
| Total | 742 | 100.0 | |

*Notes:* "Frequency" refers to the actual number of students giving each answer. Altogether 742 students participated in the survey. "Missing" tallies the unusable responses. Fourteen students either omitted the question or gave unreadable answers. "Percentage" simply expresses all category frequencies as proportions of the total sample, totaling to 100%. "Valid Percentage" discounts the missing cases and recalculates the category percentages based on just those who answered the question. That number, or *n,* comprises 728 actual responses.

selected classes.* Not surprisingly, most students were single, but a distinction is shown between singles who were unattached ("solo") and those who reported being in monogamous relationships ("attached"). The main point, however, is that Table 1.1 illustrates a frequency distribution for a nominal variable. Each category or status is qualitatively different from the others. We do not have more marriage or less marriage. We have a series of different relationship statuses. Moreover, although the categories "Single, Solo" through "Widowed" may appear to be in some sort of logical order, we could actually list them in any sequence we wished. Nominal categories have no inherent order or linear, unidimensional feature to them. They are not different degrees of one thing. They are different things or statuses. Similar attributes of nominal variables are also illustrated in Table 1.2.

Table 1.2 shows the distribution of religious denominations among a recent sample of college students. As in Table 1.1, this table shows the actual frequencies, percentages, and the adjusted valid percentages for each category or answer. As was the case with marital status, the categories here might have been listed in any order. Nominal categories have no inherent order or sequence. Finally, note also that Table 1.2 includes an "Other" category. It would be prohibitive to list all possible religions in our original question, so this option makes the choices exhaustive. Students may answer no matter what their religious denominations. In Table 1.1, however, the choices shown for marital status include all options, and no residual or "Other" category is necessary.

At the next level of measurement, the order or sequence of categories is important. In fact, this feature is reflected in the name, the **ordinal level** or **scale.**

### Comparing Ranks

In contrast to nominal variables, ordinal measurements can be ranked. As the name suggests, there is a natural or logical *order* to them. Common examples of ordinal variables are social class or socioeconomic status (SES), religiosity (how religious one is), one's degree of ethnocentrism or prejudice, political liberalism or conservatism, and scores on essay exams. Not only may we categorize respondents as to their liberalism or conservatism, religiosity, or

*The tables in this chapter—and most in the text—derive from recent campus surveys conducted by the author and students in survey research classes. The tables were prepared using SPSS, a comprehensive statistical program used on virtually all college campuses. Originally developed to aid social science research, its full name was the Statistical Package for the Social Sciences. It proved extremely popular, however, and its use spread to many academic disciplines and to business and other venues. To reflect its broader applications but still retain the familiar acronym, the name was changed to Statistical Products and Service Solutions. To most people, however, it remains simply SPSS.

**Table 1.2 Religious Denomination Among a Sample of Students at a State University**

| Religious Denomination | Frequency | Percentage | Valid Percentage |
|---|---|---|---|
| None/not applicable | 204 | 27.5 | 29.8 |
| Catholic | 232 | 31.3 | 33.9 |
| Other Christian | 125 | 16.8 | 18.2 |
| Buddhist | 66 | 8.9 | 9.6 |
| Hindu | 13 | 1.8 | 1.9 |
| Muslim (Islam) | 10 | 1.3 | 1.5 |
| Sikh | 7 | .9 | 1.0 |
| Jewish | 11 | 1.5 | 1.6 |
| Other | 17 | 2.3 | 2.5 |
| Total | 685 | 92.3 | 100.0 |
| Missing | 57 | 7.7 | |
| Total | 742 | 100.0 | |

test scores, we may also rank the resulting categories or measurements. There is a logical ranking to them, for example: Very Liberal, Liberal, Middle-of-the-Road or Centrist, Conservative, and Very Conservative. We could reverse the order, of course, and put Very Conservative first, but the overall set or list would not make sense in any other sequence. Whenever they are listed, the categories should proceed from one end of the continuum to the other. Similarly, when we display a set of religiosity categories or test scores, we would logically list them from most to least religious, highest to lowest, or vice versa. We do not have qualitative differences between the measurements but rather differences based upon more of something or less of something, that is, upon differences of amount rather than differences of kind or type.

The preceding examples also illustrate the two different types of ordinal measurements we might encounter. On the one hand, we may have a set of *rankable categories*. That is the case with the liberalism-conservatism variable above. Not only do we place people's responses in particular political categories, we may also legitimately *rank* the categories from the most liberal to the most conservative (or least liberal, if you like). We may cross-tabulate sets of rankable categorical measurements with other variables or use measures of association similar to those used with nominal measurements. Tables 1.3 and 1.4 illustrate frequency distributions for sets of ordinal categories.

Please notice two things about Tables 1.3 and 1.4. First, the measurement categories appear in logical sequences. Table 1.3 ranks answers along

**Table 1.3 Responses of a Student Sample to the Question: "Have You Ever Told Minor 'White Lies' or Fibs?"**

| Told Minor Lies | Frequency | Percentage | Valid Percentage | Cumulative Percentage |
|---|---|---|---|---|
| Many times | 172 | 23.2 | 24.3 | 24.3 |
| Occasionally | 361 | 48.7 | 51.0 | 75.3 |
| Rarely | 156 | 21.0 | 22.0 | 97.3 |
| Never | 19 | 2.6 | 2.7 | 100.0 |
| Total | 708 | 95.4 | 100.0 | |
| Missing | 34 | 4.6 | | |
| Total | 742 | 100.0 | | |

*Note:* Percentages may not total 100.0 due to rounding.

**Table 1.4 Responses of a Student Sample to the Statement: "If Asked, It Would Be OK to Help Family Members or Friends with the Answers on Tests."**

| OK to Help Family or Friends on Tests | Frequency | Percentage | Valid Percentage | Cumulative Percentage |
|---|---|---|---|---|
| Strongly agree | 17 | 2.3 | 2.3 | 2.3 |
| Agree | 166 | 22.4 | 22.4 | 24.7 |
| DK/NS | 151 | 20.4 | 20.4 | 45.1 |
| Disagree | 295 | 39.8 | 39.8 | 84.9 |
| Strongly disagree | 112 | 15.1 | 15.1 | 100.0 |
| Total | 741 | 99.9 | 100.0 | |
| Missing | 1 | .1 | | |
| Total | 742 | 100.0 | | |

*Notes:* DK/NS stands for "Don't Know/Not Sure," i.e., generally ambivalent or undecided. Percentages may not total 100.0 due to rounding.

a frequency dimension, "Many Times" through "Never," and Table 1.4 does the same with agreement, "Strongly Agree" through "Strongly Disagree." Second, a new column on the right shows the **Cumulative Percentage.** It shows the previous column, the Valid Percentage, *cumulatively* as one proceeds down the sequence of categories. Percentage-wise, it presents a running total for all the cases as we proceed from the first category to the last. The cumulative percentage makes sense only when we have ordered or rankable measurements because we are counting down the categories and accumulating more and more cases as we proceed, *in sequence,* from one category or measurement or rank to the next. The cumulative percentage has no meaning if the order of our categories is arbitrary or discretionary, as it is with nominal measurements. It would change with every different order in which the categories could be listed, and so it would not tell us anything useful.

In contrast, ordinal data may consist of numerical scores or measurements, that is, not rankable categories, but actual numbers. These might result from, say, grading exams or administering approve/disapprove opinion scales. Each person now has his or her own individual score, and the scores may be ranked from highest to lowest. Such a distribution is illustrated in Table 1.5. In three separate questions, college students were asked to agree or disagree that TV covered local, national, and world events well. Each question was scored with the same 5-point, agree-disagree answer format as in Table 1.4. "Strongly Agree" responses were the most positive and scored as 1, whereas "Strongly Disagree" responses were negative and scored as 5. Adding the scores for the three items yielded cumulative values ranging from 3 to 15. We may put

**Table 1.5 College Students' Scaled Scores Evaluating the Adequacy of TV News Coverage**

| | | Frequency | Percentage | Valid Percentage | Cumulative Percentage |
|---|---|---|---|---|---|
| (+) | 3 | 8 | .6 | .6 | .6 |
| | 4 | 10 | .8 | .8 | 1.4 |
| | 5 | 27 | 2.0 | 2.0 | 3.4 |
| | 6 | 102 | 7.7 | 7.7 | 11.1 |
| | 7 | 102 | 7.7 | 7.7 | 18.9 |
| | 8 | 167 | 12.6 | 12.7 | 31.5 |
| | 9 | 119 | 8.9 | 9.0 | 40.5 |
| | 10 | 149 | 11.2 | 11.3 | 51.8 |
| | 11 | 146 | 11.0 | 11.1 | 62.9 |
| | 12 | 185 | 13.9 | 14.0 | 76.9 |
| | 13 | 114 | 8.6 | 8.6 | 85.5 |
| | 14 | 65 | 4.9 | 4.9 | 90.5 |
| (–) | 15 | 126 | 9.5 | 9.5 | 100.0 |
| | Total | 1320 | 99.2 | 100.0 | |
| | Missing | 10 | .8 | | |
| | Total | 1330 | 100.0 | | |

these combined scores in order, or rank them, but we may not claim, for instance, that a scale score of 6 is exactly twice as positive about TV news coverage as a score of 12. Our measurements of the variables are not that precise. All we may say is that the lower the score, the more positive a student's view of TV news programming.

Another example of a numerical ordinal scale comes from a survey on lying. Regarding work assignments, students were asked whether they had ever called an employer and falsely claimed to be sick and, separately, whether they had ever lied to a professor about the reason for late or missing assignments. Each item was originally scored on a 1 to 4 ("Many Times" to "Never") scale. Adding each student's tallies on the two items yields a possible range of combined scores from 2 through 8 (Table 1.6). As in the previous table, we may not claim that a student scoring 3 is twice as likely to have lied as a student scoring 6, but we may at least rank the *X* values. Our survey, after all, asked students to generally estimate how often they had lied: Many Times, Occasionally, Rarely, or Never. We had not asked for actual and precise numbers of times. We may claim a set of ordered and rankable numerical scores, with the lower score meaning more lying about reasons for not meeting one's obligations, but we cannot claim to have measured those "lying histories" with true mathematical precision.

Ordinal data in numerical form are suitable for statistical procedures using ranks. Realizing the scores do not represent precise mathematical increments of the variable, we simply rank them from highest to lowest or vice versa, and

**Table 1.6 College Students Scaled Responses About Lying to Employers and/or Professors**

| | | Frequency | Percentage | Valid Percentage | Cumulative Percentage |
|---|---|---|---|---|---|
| Most | 2 | 9 | 1.2 | 1.3 | 1.3 |
| | 3 | 21 | 2.8 | 2.9 | 4.2 |
| | 4 | 83 | 11.2 | 11.6 | 15.8 |
| | 5 | 112 | 15.1 | 15.6 | 31.4 |
| | 6 | 204 | 27.5 | 28.5 | 59.8 |
| | 7 | 170 | 22.9 | 23.7 | 83.5 |
| Least | 8 | 118 | 15.9 | 16.5 | 100.0 |
| | Total | 717 | 96.6 | 100.0 | |
| | Missing | 25 | 3.4 | | |
| | Total | 742 | 100.0 | | |

thereafter we ignore the original numbers and use the ranks in our statistical procedures. We will return to this concept in Chapters 7 and 9. For now, there are two remaining levels of measurement.

### When the Numbers Count

The interval and ratio levels of measurement are legitimately accurate and precise measurements with no subjectivity or doubt. These are the kinds of measurements about which there is no ambiguity. An earlier example, age, is such a variable. Other examples would be the number of units you are taking this semester, how many units you have accumulated over your college career, the number of miles you traveled, or the number of blocks you typically walk to campus, how many people live in your household, how much money you earned last year, and so on.

Statistically, the difference between the interval and ratio levels of measurement does not matter. We may treat interval level measurements just as we would ratio level observations. However, there is an important difference between the two. Ratio scales have a true (or legitimate and meaningful) zero point, that is, a complete absence of whatever is being measured, and interval scales do not. Annual income, for example, would constitute a ratio scale. Someone could, at least theoretically, have absolutely no income or even a negative income, so therefore an $X = \$000$ measurement could be legitimate and valid. In contrast, the Fahrenheit temperature scale represents interval level measurements. A reading of zero degrees has no particular meaning because freezing occurs at 32° Fahrenheit. And yet both scales, income and Fahrenheit, are numerical, accurate, precise, and unambiguous. One has a meaningful zero point, however, and the other does not. This is the difference between interval and ratio scales, but we may treat interval and ratio data the same and ignore that difference. It does remain essential, however, to distinguish between nominal, ordinal, and interval-ratio measurements.

As examples, Tables 1.7 and 1.8 show frequency distributions for interval-ratio measurements. As for the previous tables, Table 1.7 shows data from a campus survey and clearly reflects a college population. Notice the comparatively large frequencies for people in their mid-twenties and just single-digit tallies at age 33 and above. This predominance of people under age 30 is also confirmed in the cumulative percentage column. We have accumulated or accounted for fully 89.1% of the sample when everyone up through age 30 is counted. This feature of cumulative percentages also has another name. We sometimes refer to it as the **percentile rank** of a number, defined as *the proportion or percentage of cases that fall at or below a certain point in a distribution.* With 89.1% of the cases falling at age 30 or below, someone exactly 30 would fall at about the 89th percentile rank in this distribution. We could

**Table 1.7 Age at Last Birthday Among a Sample of College Students**

| | Age | Frequency | Percentage | Valid Percentage | Cumulative Percentage |
|---|---|---|---|---|---|
| | 17 | 1 | .1 | .1 | .1 |
| | 18 | 30 | 4.0 | 4.2 | 4.4 |
| | 19 | 49 | 6.6 | 6.9 | 11.3 |
| | 20 | 62 | 8.4 | 8.8 | 20.1 |
| | 21 | 96 | 12.9 | 13.6 | 33.7 |
| | 22 | 115 | 15.5 | 16.3 | 49.9 |
| | 23 | 104 | 14.0 | 14.7 | 64.6 |
| | 24 | 48 | 6.5 | 6.8 | 71.4 |
| | 25 | 47 | 6.3 | 6.6 | 78.1 |
| | 26 | 24 | 3.2 | 3.4 | 81.5 |
| | 27 | 16 | 2.2 | 2.3 | 83.7 |
| | 28 | 20 | 2.7 | 2.8 | 86.6 |
| | 29 | 8 | 1.1 | 1.1 | 87.7 |
| | 30 | 10 | 1.3 | 1.4 | 89.1 |
| | 31 | 10 | 1.3 | 1.4 | 90.5 |
| | 32 | 13 | 1.8 | 1.8 | 92.4 |
| | 33 | 5 | .7 | .7 | 93.1 |
| | 34 | 6 | .8 | .8 | 93.9 |
| | 35 | 2 | .3 | .3 | 94.2 |
| | 36 | 2 | .3 | .3 | 94.5 |
| | 37 | 1 | .1 | .1 | 94.6 |
| | 38 | 1 | .1 | .1 | 94.8 |
| | 39 | 5 | .7 | .7 | 95.5 |
| | 40 | 3 | .4 | .4 | 95.9 |
| | 42 | 3 | .4 | .4 | 96.3 |
| | 43 | 3 | .4 | .4 | 96.7 |
| | 44 | 2 | .3 | .3 | 97.0 |
| | 45 | 3 | .4 | .4 | 97.5 |
| | 46 | 3 | .4 | .4 | 97.9 |
| | 47 | 3 | .4 | .4 | 98.3 |
| | 48 | 1 | .1 | .1 | 98.4 |
| | 50 | 2 | .3 | .3 | 98.7 |
| | 51 | 2 | .3 | .3 | 99.0 |
| | 54 | 1 | .1 | .1 | 99.2 |
| | 55 | 1 | .1 | .1 | 99.3 |
| | 57 | 1 | .1 | .1 | 99.4 |
| | 59 | 1 | .1 | .1 | 99.6 |
| | 67 | 1 | .1 | .1 | 99.7 |
| | 77 | 1 | .1 | .1 | 99.9 |
| | 82 | 1 | .1 | .1 | 100.0 |
| Total | | 707 | 95.3 | 100.0 | |
| Missing | | 35 | 4.7 | | |
| Total | | 742 | 100.0 | | |

*Note:* Percentages may not total 100.0 due to rounding.

**Table 1.8 If an Immigrant: Number of Years in US (College Student Sample)**

| | Years in US | Frequency | Percentage | Valid Percentage | Cumulative Percentage |
|---|---|---|---|---|---|
| | 1 | 7 | 1.0 | 3.4 | 3.4 |
| | 2 | 13 | 1.9 | 6.3 | 9.8 |
| | 3 | 11 | 1.6 | 5.4 | 15.1 |
| | 4 | 8 | 1.2 | 3.9 | 19.0 |
| | 5 | 12 | 1.7 | 5.9 | 24.9 |
| | 6 | 11 | 1.6 | 5.4 | 30.2 |
| | 7 | 5 | .7 | 2.4 | 32.7 |
| | 8 | 7 | 1.0 | 3.4 | 36.1 |
| | 9 | 9 | 1.3 | 4.4 | 40.5 |
| | 10 | 19 | 2.8 | 9.3 | 49.8 |
| | 11 | 12 | 1.7 | 5.9 | 55.6 |
| | 12 | 14 | 2.0 | 6.8 | 62.4 |
| | 13 | 5 | .7 | 2.4 | 64.9 |
| | 14 | 5 | .7 | 2.4 | 67.3 |
| | 15 | 7 | 1.0 | 3.4 | 70.7 |
| | 16 | 8 | 1.2 | 3.9 | 74.6 |
| | 17 | 6 | .9 | 2.9 | 77.6 |
| | 18 | 8 | 1.2 | 3.9 | 81.5 |
| | 19 | 1 | .1 | .5 | 82.0 |
| | 20 | 10 | 1.4 | 4.9 | 86.8 |
| | 21 | 12 | 1.7 | 5.9 | 92.7 |
| | 22 | 1 | .1 | .5 | 93.2 |
| | 23 | 5 | .7 | 2.4 | 95.6 |
| | 24 | 2 | .3 | 1.0 | 96.6 |
| | 25 | 1 | .1 | .5 | 97.1 |
| | 26 | 2 | .3 | 1.0 | 98.0 |
| | 30 | 1 | .1 | .5 | 98.5 |
| | 31 | 1 | .1 | .5 | 99.0 |
| | 32 | 1 | .1 | .5 | 99.5 |
| | 45 | 1 | .1 | .5 | 100.0 |
| Total | | 205 | 29.7 | 100.0 | |
| Missing | | 485 | 70.3 | | |
| Total | | 690 | 100.0 | | |

*Note:* Percentages may not total 100.0 due to rounding.

also say that a 40-year-old student would be at almost the 96th percentile rank. He or she would be as old or older than 96% of all students in the sample. We will use percentile ranks again in Chapter 3, when we discuss the bell-shaped curve.

Table 1.8 shows a similar distribution but from a different survey. It illustrates an interval-ratio variable and distribution from a survey on immigration.

In this case, a questionnaire asked immigrant students how many years they had been in the United States. The **range** of data is extensive, from a low of 1 year (rounded off) to a high of 45 years. According to the cumulative percentage column, about half the immigrant students (49.8%) had been in the United States 10 years or less. Moreover, notice the large number of *Missing* cases in Table 1.8. Most students ($n$ = 485, or 70.3% of all respondents) were *not* immigrants and, of course, did not answer the question. For this question, they were coded as omits or "Missing." Still, fully 29.7% of students in the survey did answer as immigrants, no doubt reflecting recent and changing demographics among young adults in the United States.

Interval-ratio measurements allow us to have the utmost confidence in their mathematical accuracy and precision. Therefore—and this is a key point—any calculations may use the actual $X$ scores or measurements. Unlike nominal or ordinal measurements, the data are not in categories nor must we convert the data values to ranks. To justify their use, however, the original scores must be absolutely reliable, unambiguous, and precise measurements of the variable in question.

When we do have such reliable data, we may use the actual $X$ scores in our calculations. If we need to know the average or typical or usual response, we may add up all the $X$ values and then divide by the total number of cases or values we have, or $n$. If we have a set of ages, the $X$ variable being age, we may do this. If we have a set of essay exam scores, however, we have ordinal-level measurements only and should not calculate an average. We may not be sure how to *precisely* interpret each exam score, so we should not use them in any statistical procedure. Instead of an average, we have alternative statistics available, as discussed in Chapter 2.

Two final considerations are important. First, as noted earlier, when we do have interval-ratio data, we may use a broad range of statistical treatments. This text assumes we have such data for the most part. Therefore, we may look at a full range of introductory statistical tools. We will, however, also look at statistics specifically designed for ordinal and nominal data.

Second, sometimes we are working with more than one level of data at the same time. If we are correlating measurements on two variables, one nominal, say, and the other interval-ratio, what do we do then? A common rule is to use a statistic (or statistics, plural) appropriate for the lower level. If we have both nominal and interval-ratio data, for example, we use statistics suitable for the nominal level. The higher-level variable meets all the assumptions and criteria of the lower level of measurement, but the reverse is obviously not true. Interval-ratio measurements meet all the criteria of nominal measurements (we can distinguish between different categories or scores of the variable), but nominal data would certainly not meet the precise quantitative requirements of the interval-ratio level. Consequently, we should use statistics appropriate for the nominal level. That is, we choose a level of measurement we are sure is met by both our variables. As in most statistical situations, we would probably have *some* choice in exactly how to treat the data, but we must be prepared to sacrifice that higher level of measurement in one of our variables on occasion.

Beyond this, and no matter what level(s) of data we have, there is another matter to consider. What is the purpose of our analysis? Are we simply summarizing data, or do we wish to make inferences about a larger universe or population based upon seeing just a tiny fraction or sample of it? Do we wish to summarize what we have *or* do we want to make educated guesses beyond the data immediately available? These questions lead to the last part of this introduction: the difference between descriptive and inferential kinds of statistical analyses.

## To What End: Description or Inference?

Statistical analysis has two very broad areas: the descriptive and the inferential. The former is the more basic and, as the name suggests, amounts to describing and summarizing data. Given a set of numbers, what is the average, do the numbers vary much or very little from that average, and what percentage of the cases fall below such-and-such a score? Examining the data at hand, descriptive statistics look at variables one at a time and simply summarize the data by calculating various statistical measures (e.g., averages, medians, standard deviations) and by showing frequency distributions. If we have 20 variables, we may easily summarize the scores or measurements for each one and provide a report. We are distilling a lot of raw or original data down into more comprehensible summary measures. Because this involves averages and the like, a good part of descriptive statistics and the material in the following chapter should be familiar to you (or will come back quickly). After that, we begin moving toward the second branch of statistics, inference.

Inferential statistics are a bit more involved and theoretical. The term *inferential* derives from the fact that we are making inferences about larger universes or populations based upon just sample data. And this, in turn, requires that we have **random samples.** Random samples (sometimes called probability samples) are those that give every member of the population a statistically equal chance to be selected. The procedures required for good random samples can be quite involved and are beyond the scope of this text. Nevertheless, we may justify making inferences about the population only if we have random and representative samples. Our discussions of inferential statistical procedures assume we are dealing with random samples.

This second branch of statistics also covers most of what we think of when we use the term "statistics": the normal or bell-shaped curve, procedures known as confidence interval estimates, hypothesis testing, correlation and regression, and so on. But, as noted before, when we look at inferential statistics, we must also consider probability. Statistical inference is based upon probability. Whenever we make that extrapolation or inferential leap from the sample to the population, there is always a chance we are wrong. What is the probability the population average is or is *not* what we have estimated? These probabilities must accompany any inferences we make.

Succeeding chapters look first at descriptive statistics, next at probability, and finally turn to inferential statistics. Before proceeding, however, we must be wary about assuming too much from this introduction: The level of measurement, on the one hand, and the distinction between descriptive and inferential statistics, on the other, are quite independent concepts. They do not necessarily correlate in any way. Descriptive analyses, being the more simple and elementary of the two, do not apply only to the nominal level of measurement. Inferential statistical methods, being more involved and complex, do not apply only to the higher levels of measurement. We may have descriptive statistical summaries involving *any* scale of measurement, nominal to interval-ratio. The same is true of inferential analyses, which may involve nominal, ordinal, or interval-ratio data. It is not a case of a certain level of measurement being appropriate for either descriptive *or* inferential methods. It is a matter of selecting a basic statistical treatment (descriptive, inferential, or both), depending on the study's purpose, and only thereafter tailoring the specific statistics used to the level(s) of measurement involved.

In Chapter 2, we turn to a few descriptive statistics, some of which will be familiar. We first consider averages, more formally known as measures of central tendency, and then look at a popular and valuable measure of variation.

## Exercises

1. What is the level or scale of measurement, and why is it important in statistical analyses?

2. Describe the principal levels of measurement, including the characteristics of each and at least one example of each.

3. How do the levels of measurement differ regarding the statistical procedures possible with each?

4. What is the difference between numerical measurements at (1) the ordinal level and (2) the interval-ratio level?

5. How do descriptive and inferential statistical analyses differ?

6. What is the importance of random sampling to statistical analyses?