CHAPTER 1

The Political Use of Numbers

Lies and Statistics

In 1904, Mark Twain was busy writing his last book, an autobiography, in which he reminisced about his life. At one point, as he recalled writing *Innocents Abroad* in 1868, he became a bit mournful as he remembered that while writing it, he had been able to write 3,000 words a day. He began to feel his age as he remembered that twenty-nine years later, in 1897, he had slowed down from producing 3,000 to only 1,800 words a day, and seven years later, in 1904, his output had declined by an additional 400 words to only 1,400 words a day. But as he began to despair over the limitations of age, Twain realized that what had declined in those years was not how fast he wrote, but how much time he spent writing. He had begun by writing seven to nine hours a day, but ended with only four to five. After he realized that the number of pages had halved when the number of hours spent writing had also halved, he understood that the change in writing was not that his mind had gotten any less sharp; rather, he just didn't work as much. Twain concluded that numbers do not always speak for themselves. And then he repeated a well-known saying of the day, attributing it to Benjamin Disraeli: "There are three kinds of lies: lies, damned lies, and statistics." For Twain,[1] the aphorism was a reminder to be careful about the use of numbers.

In 1954, well-known "How to" author Darrell Huff[2] repeated the refrain when he wrote his best-selling statistics textbook *How to Lie with Statistics.* In it, Huff regales readers with examples of statistics used to delude the public. The interesting thing about his examples is that the lies are perpetrated by the violation of principles taught in all introductory texts. He begins the work in the same place as most statistics texts, with measurement. The ability to perform statistical analyses depends on our ability to attach numbers to concepts. How well we measure those concepts determines how accurately we can describe our world. Huff then proceeds to work his way through all the major topics addressed in any introductory statistics textbook: measures of central tendency, measures of dispersion, measures of association, probability, and control variables. Although his ostensible purpose is to teach the reader how to lie, in reality, the book shows how knowing the basic principles of statistics empowers us to be skilled consumers of data. To that end, he concludes the book with Chapter 10, "How to Talk Back to a Statistic." More than half a century later, that principle still holds: Understanding statistics is a fundamental life skill for being an active citizen in a democracy.

THE POWER OF NUMBERS

The ability to, in Huff's words, "look a phony statistic in the eye and face it down"[3] has societal benefits. Sir Francis Bacon said that knowledge is power. I normally think about that quote at the personal level: The more I know, the more I am able to accomplish. The concept is certainly a valid reason to get an education. But it also applies at the national level. Those political actors with more knowledge are more politically powerful.

Political entities can use knowledge in ways that can be helpful. I'm sure that doctors probably get irritated with filling out the forms required by the Centers for Disease Control and Prevention (CDC). But the CDC's efforts to document the occurrence of various diseases have incredible societal benefits. In 2006, because of requirements that doctors report cases of *E. coli,* the CDC was able to initially recognize that the seventy-one cases in the northeastern United States were not random and so constituted an outbreak with a single cause. Next, they were able to identify the commonality of the affected people having eaten at Taco Bell. They then narrowed the source down to lettuce and so recalled all potentially contaminated produce. The next step was to identify the source of the contamination at the farm of origin. Finally, the CDC was able to enter the policy-making arena in making suggestions about what policies would prevent future outbreaks. A similar process is replicated in all areas of public policy. This careful collection and analysis of data make government agencies very powerful actors in public health as well as in other public policy areas.

Political entities can also use knowledge in harmful ways. Information was one of the major sources of power for the Nazi regime. The Nazis were able to collect detailed information about individuals living within their borders. They kept such good records that genealogists today have a wealth of information that is not available in most countries. The Nazis' goal, however, was not to aid family historians. Rather, they used their knowledge to identify neighborhoods with higher Jewish populations and then enclosed these ghettos in preparation for transporting the residents to concentration camps. They were also able to use records documenting family ties to extend their control over those of mixed family origin. Perhaps it is this type of misuse of knowledge that led Chinese philosopher Lao-Tzu to say, "People are difficult to govern because they have too much knowledge."[4] A monopoly of information is a very powerful resource for a government intent on keeping control of its citizens.

But the power of statistics is a two-way street. In one direction, it allows politicians to gather information to make decisions. Indeed, the word "statistics" originally referred to the political state in the context of governments collecting data to help them govern better.

But access to information can be just as powerful to citizens as it is to policymakers. Statistics allow citizens to see the causes and effects of political decisions more clearly. A wealth of information is available that allows us to keep tabs on the government and potentially check its misdeeds. If you go to the website of virtually any government agency, you can find readily accessible data pertaining to the mission of that agency. Max Weber[5] identified the reliance on written documents as the third basic function of bureaucracies. One of the points of all that paperwork is to provide transparency into the workings of government. And with the advent of the Internet, all of that information is easily available to the citizen who knows enough to use it.

The collection of statistics has expanded dramatically since governments first began collecting demographic data through censuses. In this book, you will see data from a wide variety of sources. From government sources, you will see discrimination data collected by the Equal Employment Opportunity Commission (EEOC), crime data from the FBI's Uniform Crime Reports and demographic data from the Census Bureau. From research organizations, you will see survey data from the Pew Research Center, the National Opinion Research Center, and the American National Election Studies. You will see international data collected by the World Bank, the Center for Systematic Peace, and the United Nations. All of these organizations collect the data for their own purposes, but they are also bureaucracies that rely on written information. The benefit to us is that we have access to all that information. It is up to you to learn how to use it. With it, you have the power to be an effective citizen. And as you share what you learn, you can make the powers that be more responsive to this nation of citizens who gave them the power to begin with.

THE SCIENCE OF POLITICS

The power of information rests on the scientific use of it. I can't count the number of times people have laughed when I told them I am a political scientist: "What's so scientific about politics?" Their confusion is a result of misunderstanding the nature of science. Sometimes, we mistakenly confuse science with technology—test tubes and mass spectrometers. Sometimes, we confuse science with the areas of study that use those technologies—biology and chemistry. But the nature of science is found in neither the tools nor the subject matter. Rather, it is found in the method used to find out how the world works. Depending on the subject matter, we may call it "hard science" or "social science," but the process is the same.

The **scientific method** begins with a question of how the world works. We call this our research question. Sometimes, we ask why a phenomenon occurred the way it did. Sometimes, we ask how an event affects another phenomenon. Depending on our worldview, we may posit different answers to our research question. This allows us to develop a model of how we think the world works.

$$\text{Independent Variable} \rightarrow \text{Dependent Variable}$$

In regular English, we would say that a cause leads to an effect. But in science-speak, we say that the independent variable affects the dependent variable. Model in hand, we are able to articulate a hypothesis detailing the precise causal relationship between the dependent and independent variables. But not all hypotheses are scientific. To earn that appellation, the hypothesis needs to be subject to an **empirical** test that is capable of evaluating whether the evidence confirms or disconfirms the hypothesis.

Notice that nowhere in the description of the scientific method do we limit what aspect of the world is appropriate to study. We could use this approach to study living organisms, in which case we would call ourselves biologists. But we can also use this approach to study political interactions, in which case we call ourselves political scientists.

The description of the scientific method also does not rigidly limit the nature or method of collecting the evidence we use to test our hypotheses. Sometimes, political science uses the same data collection techniques as the hard sciences: experiments, observation, or public records. But due to the nature of our subject, we have additional sources of data not necessarily useful for the hard sciences: content analysis and survey data.

Sometimes, we collect data through **experiments**, or controlled tests where we vary a possible cause in order to measure the resulting change in the effect. Political scientists Daniel Butler and David Broockman[6] decided to use an experimental research design to answer the question of whether politicians discriminate against constituents due to race. Although their experiment did not take place in a laboratory using test tubes, the structure of their experiment is not dissimilar to experiments used by scientists studying the physical world. They decided to write to state legislators using letters that differed only in the names of the senders. Controlling for the party and race of the legislator, they found that the purported race of the constituent did make a difference on whether legislators responded to the letter.

Whereas some scientists collect data by observing the natural world, political scientists can collect data by observing political interactions. Both hard scientists and social scientists can participate in scientific **observation** by noting and recording phenomena in a systematic way. Political scientist Richard Fenno[7] wondered how members of Congress interact with their constituents. By following several representatives around, he was able to observe how they behave in their home districts, as opposed to in Washington, D.C. Fenno found that each representative builds trust with his or her constituency by explaining who he or she is and how that relates to the needs of the district. On the basis of that trust, representatives are able to justify their activities in Washington and so get reelected.

We not only expect scientists to collect original data through experiments and observation, we also expect them to use public data archived for public use. In the same way that scientists who study public health use data collected by the CDC and the World Health Organization, political scientists have access to huge amounts of political data found in archived **public records**, which would include any information gathered and maintained by a governmental body that is openly available. Political scientist Gary Jacobson[8] wondered how campaign spending affects elections. He used campaign spending data collected by the Federal Election Commission in conjunction with publicly collected election data to find that, at least in congressional elections, incumbents who spend the most win by the smallest margin. (Ponder on that. In Chapter 11, I'll be asking you to come up with a possible explanation!)

Although some political scientists use data collection techniques similar to other scientists, others use techniques that are not part of the hard science repertoire. Because much political data are contained in documents, political scientists study political communications systematically using a technique called **content analysis**. Political scientist Daniel Coffey[9] wondered how polarized the two major political parties are in the United States. He collected state party platforms for both the Republican and Democratic parties between 2000 and 2004 and analyzed the difference between the number of liberal and conservative sentences contained in them. Through this quantitative content analysis, he was able to conclude that the state parties are very polarized.

Finally, although popular opinion does not determine scientific reality, public opinion is an important political factor. As a result, **surveys** (questionnaires used to collect the self-reported attitudes of people) can be a very useful tool for political scientists. For example, political scientist Robert Putnam[10] wondered whether Americans have the kind of social resources they need to be good citizens in a democracy. He used data from various sources to measure how social capital in the United States has changed over time. By using the General Social Survey conducted yearly by the National Opinion Research Center, he found that one aspect of social capital, group membership, has been declining.

Experiments, observation, public records, content analysis, and surveys are also used by political decision makers in order to gather the information they need to make informed decisions. For example, in 1954, when the Supreme Court was deliberating *Brown v. Board of Education* and the segregation of races in public schools, it received an amicus brief containing the results of an experiment conducted at Columbia University. In it, psychologists Kenneth and Mamie Clark wondered what influence race had on young children. In the experiment, they showed two dolls, one white and one black, to a series of children. Although they expected to find that white children would be biased, they were surprised to find that black children were also more likely to associate negative attributes with the black doll. The Supreme Court was equally surprised with the results of the experiment and found it sufficient grounds to conclude that separate could never be equal in an educational context.

During the Cold War, data about Soviet internal politics were sorely lacking. As a result, a form of content analysis called Kremlinology was carefully developed in order to draw conclusions from what little data were available—public speeches and photographs of public events. During the Cuban missile crisis, this technique was used to analyze two contradictory messages that had been received, both of which were purportedly signed by Soviet premier Nikita Khrushchev. Because the content of the two letters was so different, U.S. President John F. Kennedy asked who had written each message. Using content analysis, the Kremlinologists were able to conclude that Premier Khrushchev had written the first, more conciliatory message, whereas the second had been written by hardliners in the Kremlin who were trying to take control. President Kennedy was able to use this information to resolve the conflict successfully.

Much more recently, President Obama charged the Department of Defense with studying the feasibility of discontinuing the "Don't Ask, Don't Tell" (DADT) policy toward gay members of the military. One way in which the Department of Defense chose to study the issue was with a survey of military personnel. The survey found that an overwhelming majority of soldiers had no problem serving with gay comrades. On the basis of the survey, Secretary of Defense Robert Gates advised President Obama to repeal DADT and allow gays to serve openly in the military.

It is not an accident that political scientists and political decision makers use the same data-gathering techniques as other scientists. We are all interested in figuring out how the world works. As we collect data and analyze them systematically to see whether they support or disconfirm our hypotheses, we are better able to understand what is happening around us. Sometimes we do so in a quantitative way, collecting and analyzing information as numbers, as in Coffey's study of political parties and the Department of Defense's study of military personnel. Sometimes we do so in a qualitative way, collecting information as descriptions and explanations rather than as numbers, as in Fenno's study of members of Congress and the Kremlinologists' study of the Khrushchev letters. Regardless of whether the data are numerical or not, the scientific method requires that they be analyzed systematically. Those researchers who rely on **qualitative data** have developed various specialized techniques to ensure the validity of their results. Those of us who rely on **quantitative data** use statistical analysis to ensure the validity of our results.

BOX 1.1 Numbers in the News

Since 1939, when FDR first made a request that George Gallup measure public opinion about potential U.S. involvement in the war in Europe, presidents have used public opinion polls in various ways. Late in 2001, in the warm bipartisan aftermath of 9/11, President George W. Bush's press secretary, Ari Fleischer, invited a slew of prior press secretaries to the White House for lunch. At one point, the conversation turned to whether, in the face of vague terrorist threats, it is better to stay quiet or alert the public. Dee Dee Myers, who had been press secretary in the Clinton administration, asked what the polls said on the subject. President Bush replied, "In this White House, Dee Dee, we don't poll on something as important as national security."[11]

Although Bush consciously distanced himself from Clinton's practice of molding policy to fit public opinion, he still availed himself of the valuable information that polls contain. For President Bush, polls were valuable when times were rough. He primarily used them to determine how to frame unpopular policies to seem more appealing to the public. Whether presidents use polls to find popular opinion or to figure out how to mold it, that use reflects their belief that for policies to succeed, they need (at least some) public support.

The purpose of this textbook is to teach you the basics of statistical analysis. This will be useful in many ways to you as a political science student. It will undoubtedly help you as you read articles assigned in your classes because a lion's share of political science research is quantitative. In addition, if your department requires you to conduct original research, understanding statistics will help you in that process. Or, if you choose to pursue graduate studies, this course could form the foundation of your future study of statistical techniques.

But if you are like my students, you are less likely to go to graduate school than you are to get a job in the public arena. In that case, the skills you learn in this course might be the most profitable skills you acquire in college. Whether you work for the government, a political organization, or a public interest group, you will need to know how to both collect and analyze data. Your potential employers will all be impressed when they see you list "SPSS" as one of the computer programs with which you are proficient because all of them need employees who know how to answer questions by systematically analyzing data.

INTRODUCTORY STATISTICS: AN OVERVIEW

Statistics are normally divided into two types: descriptive statistics and inferential statistics. By definition, the difference between them relies on who is being studied, a population or a sample. If you have data about all of the cases you are trying to describe, you are measuring the **population** and so your analysis would be called **descriptive statistics**. You are actually describing the phenomenon. If you can measure only a **sample** or subset of the population, your analysis would be called **inferential statistics**. You are inferring characteristics about the population from the sample.

Frequently, though, we use the term "descriptive statistics" to refer to the techniques with which we analyze population characteristics. Most frequently, we want to describe **univariate statistics**, or statistics that analyze a single variable, such as measures of central tendency and measures of dispersion. By "measures of central tendency" we mean the different statistical ways to measure the average for the single variable in question—the mean, median, and mode. By "measures of dispersion" we mean the way all of the data are distributed around the average. The tighter the distribution, the better the average describes the variable. These measures of dispersion can include the range, the interquartile range, the variance, or (most often used for statistical purposes) the standard deviation.

But because any analysis of a population is by definition descriptive, some relationships can be analyzed using **bivariate statistics**, which would be classified as descriptive in spite of the fact that they analyze the relationship between two variables. In particular, cross tabulations can be considered descriptive statistics if they compare the relative frequency of variables for the full population. Although statistics as a technique for analyzing data began with census data and the process of counting people (a univariate question), it developed as political leaders realized that these numbers could be useful in answering questions about the differences in distributions either between groups or over time (both of which are bivariate questions). The first systematic census taking began in Sweden in 1749 with the Tabellverket. This census did not just count the people of Sweden, it also gathered detailed information about them. As a result, the government learned that many women died in childbirth and that many children died young. By looking at the distributions, the government was able to ask why this was happening. And as it found the answers, it was able to take action. At the heart of answering the questions of why some groups have different distributions than others, or why some factors change over time, are principles of probability. Of course, there will be fluctuations in whatever concept is being measured. The question is whether these fluctuations have a cause or are merely random effects of time and place.

Remember Mark Twain's warning that numbers don't speak for themselves—they need to be placed in context. Probability allows us to find that context. For example, Progressive

Insurance did a study of car accidents and found that 77 percent of policy holders who had been involved in an accident were less than fifteen miles from home. Fifty-two percent of them were less than five miles from home. Car safety advocates uniformly responded by advising drivers to be particularly careful when driving on short trips.[12] This advice reflected two unsupported assumptions: (1) You are more likely to get in an accident if you are driving close to home than if you are on a longer trip, and (2) there must be something about driving close to home that makes it more dangerous. So should we only drive on long-distance trips? Or would we be safer if we parked our car five miles from home and then walked or took mass transit to get to it? What the reported statistics fail to do is put the numbers into context: What percent of driving occurs close to home? If more than 52 percent of the driving occurs within five miles of home, then we might actually be safer, statistically speaking, driving closer rather than further from home. Before we can begin to analyze why a particular event occurs, we must first answer the question of whether anything actually even happened. Perhaps there is no event to explain; perhaps what we thought we observed is only a random product of underlying probabilities. The first half of this textbook will address descriptive statistics and probability.

Our uncertainty about the reality of a causal relationship can be exacerbated by how many cases we can measure. Although the goal of the United States' decennial census is to count all individuals residing within our borders, the Census Bureau obtains only limited information at that time. In off years, however, it gathers a wealth of information about Americans and American businesses through surveys. Although it is able to obtain this information in valid and reliable ways, the process of generalizing from a sample to a population inserts uncertainty into our measures. The result is an increased level of complexity in our statistical analysis of the data.

Our understanding about underlying probabilities allows us to make the jump from descriptive statistics to inferential statistics. With this basic understanding of probability and sampling in hand, we are able to begin thinking about the world in terms of cause and effect. Recall that with the scientific method, we hypothesize that a change in what we think of as a cause (our independent variable) has an impact on the effect (the dependent variable) we are interested in studying. Once we begin to use inferential statistics, we can begin the process of testing hypotheses in the ways required by the scientific method. With descriptive statistics, we simply described the distribution of the population. With inferential statistics, we will analyze data in a three-step process. In the first step, we describe the pattern that we see in a way similar to the process we used for descriptive statistics. But then we'll add on two more steps.

After describing the pattern of the relationship between our dependent variable (the effect) and independent variable (the cause), we will then ask how confident we are that there is actually a relationship between them. This is where probability comes in. With probability, we can describe what we expect to see if there is no relationship (or if what statisticians call the null hypothesis is correct). Of course, with data from a sample we will not necessarily see that precise pattern. Simply due to the random process of the sample selection, we could see a pattern very different from what the null hypothesis would predict. The second step of our analysis, then, is to find the probability that the relationship we thought we observed is only the result of random error. If we find that it is very unlikely to have occurred due to random error, we reject the null hypothesis and say that the pattern we saw is statistically significant. Notice that this use of the word "significant" is different from what you expect. When a non-statistician hears the word "significant," he or she usually interprets it to mean that a strong relationship exists. Not so here. A statistician says a relationship is significant if he or she is confident that a relationship exists, however weak it might be. We'll come back to the non-statistical use of the word "significant" in the third step of the analysis when we examine how strong the relationship appears to be. But for now, you need to be very conscious of the fact that when a statistician uses the word "significant," contrary to your expectation, he or she usually means the relationship is statistically significant, meaning that we can be fairly confident that a relationship exists. Another thing to keep in mind is that the use of probability dictates that we never know with certainty. We may be confident, but we are never sure. The scientific method never proves anything. Hopefully, by the end of this semester you will have broken yourself of the habit of using the dreaded "p" word. Good statisticians (and scientists) do not claim to prove their hypotheses. We speak in terms of confidence and probabilities.

Once we've concluded that there probably is a relationship between the two variables, we can move to the third step in the process of analyzing data: We ask how strong that relationship probably is. I call the strength of the relationship the substantive significance to line up this third step with what non-statisticians mean when they say a relationship is significant: an important or strong relationship between two variables. In this third step of the analysis, we choose the appropriate measure of association to describe how strong the relationship appears to be. There are various measures of substantive significance that we can use to answer that question. Which measure of association we choose depends on the ways the variables are measured. The second half of the book will address several different measures of association and when to use each.

In parallel with our discussion of the various measures of association will be a discussion of how to control for other factors. One of the problems that dogs statisticians is the reality that correlation is not the same as causation. It is possible for two variables to be correlated even if there is no causal relationship between them. No matter how strong the measure of association may be, it is always possible that there is an alternative explanation. In particular, sometimes a third variable is connected to both the independent and dependent variables that is the actual causal factor. If there is a third factor, we call the relationship we originally saw *spurious*, meaning an apparent relationship that is actually false. Although it is always good to begin an analysis by looking at the raw relationship between our dependent and independent variables, it is also important to control for alternative explanatory variables. We will discuss how to do this as we proceed through the second half of the text.

REMOVING THE BARRIERS TO UNDERSTANDING HOW STATISTICS WORKS

Huff's tongue-in-cheek title *How to Lie with Statistics* plays off the common assumption that the math involved in statistical analysis is so esoteric that statisticians can manipulate the numbers with impunity. I believe the opposite is actually true. Most of the statistics that affect politics and the public actually use fairly basic mathematics. After all, policymakers are not usually professional statisticians. If policymakers can understand the numbers well enough to make a decision, as a citizen you should be able to understand the numbers well enough to call them on their decisions. And in fact, if you can do the most basic of algebra, you can learn to do the introductory statistics that make up the preponderance of evidence used in policy making.

I have structured this book to remove the barriers standing between you and a clear understanding of how statistics works. The structure of most statistics textbooks gets in the way of communicating what should be straightforward concepts. Unfortunately, most statistics textbooks are written by mathematicians and so teach the techniques using an equation-centered format. I know from experience that most political science majors take one look at an equation and are paralyzed with the fear of interpreting what looks like a foreign language (and actually is filled with an awful lot of Greek letters). My approach is to take the equations and interpret them into English. In doing that, I clearly describe, in a step-by-step way, how to do the math. This process should make doing the math very straightforward. So when you see an equation, don't panic. That equation will immediately be followed by instructions on how to do the math.

My goal, not unlike Huff's, is to write a "How to" manual, but for me, it is "How to Keep the Numbers Honest." In the concluding chapter of *How to Lie with Statistics,* "How to Talk Back to a Statistic," Huff drops his literary device of teaching you "how to lie" and suggests five questions to ask in order to distinguish good statistics from bad statistics.

1. Who says so?
2. How does he know?
3. What's missing?
4. Did somebody change the subject?
5. Does it make sense?[13]

All of his suggestions are excellent for a novice, but you can do even better. Your best bet is to actually understand the language of statistics, what the different statistical techniques are, and what assumptions they make. All of that is easier if you know the math that is used to calculate the different statistical measures. So in this book, my first goal is to teach you how to calculate the numbers.

A second barrier most students have to learning statistics is that the equation format is very dry. Students decide to major in political science because they love talking, reading, and learning about politics. Most statistics textbooks set the students up for failure because they fail to show how the techniques are relevant to understanding interesting topics. Frequently, their problem sets rely on made-up data instead of the widely available wealth of real-world data. My second goal is to make the study of statistics interesting by using the techniques I want you to learn to analyze real-world data.

A third barrier most students have is in seeing how they will ever use what they learn. I know my students are registered for the course only because it is required. But I firmly believe that of all the classes they take in the Political Science Department, this is the one that teaches them the most marketable skills. This book is set up using an active learning paradigm in order to show the professional relevance of knowing statistics. To that end, I will model how to use statistics and real-world data to answer hypothetical professional assignments. Although the assignment might be hypothetical, the historical situation giving rise to the political question is not. After I model how to answer the question, I will give similar historical assignments that will require the students to use SPSS (the most commonly used statistical package) to analyze real-world data in order to write a memo answering a question. Thus, my third goal is that students will learn to use statistics to answer realistic professional questions.

THE IMPORTANCE OF STATISTICS: THIS BOOK'S APPROACH

Although statistics are, on occasion, used to deceive and manipulate, the honest use of them can make a great deal of difference politically. For political decision makers, the systematic collection and analysis of data can make all the difference in identifying and solving problems. For citizens, the open dissemination of statistics can keep government in line. For political scientists, statistical analysis allows us to better describe how politics works and why.

My goal is to help you become better political scientists, better citizens, and better political decision makers by teaching you the basic statistics that are relevant to understanding politics. This book covers many of the same topics covered in any introductory statistics class. First, you'll learn about descriptive statistics, which describe the average and distribution of data from a population. Because the conclusions we draw are always uncertain, you will then learn some basic principles of probability. We will then turn to inferential statistics, with which we draw conclusions about data based on a sample.

Unlike most introductory statistics textbooks, I will be modeling the process of calculating the various statistical measures using real-world political data. Each chapter will begin with a cultural example of the relevant concept to help you build an understanding about how it works. I will then work through the details of the concept, usually giving you detailed instructions of how to do the appropriate math in a work table. After covering the concept, I'll summarize it, modeling one last time how to do the math. At the end of the chapter, you'll find an exercise section titled "Your Turn," which will have problems structured very similarly to the ones in the chapter. By mirroring the steps I take in the chapter, you can complete the work. Take the time to do these exercises because doing the math really does help you understand what the numbers mean.

USING DATA TO ANSWER A QUESTION

Although doing the math helps you understand the numbers better, in the real world, you will not be doing statistics by hand. In the real world, you will use a statistical package to get a computer to calculate the numbers for you. In the political world, this statistical package will usually be the Statistical Package for the Social Sciences, or SPSS. As a result, although the bulk of each chapter focuses on building your intuition about how to interpret numbers by teaching you the math, after the "Summarizing the Math" section, you'll find a final section describing how to use the concept in order to answer a question using SPSS.

Each "How to Answer a Question" section (with the exception of the current chapter) will begin by describing how to get the relevant statistics using SPSS. The steps will be outlined in summary tables, which you can find both in the chapter and at the end of the book. Following the description of how to use SPSS will come a section with the heading "A Political Application." This section will describe a real-world situation in which policymakers had to analyze data to answer a question. I place myself in that situation as a hypothetical actor who is asked to answer the question. I then model for you how to use SPSS to answer it. I will include screen shots to show how I navigate SPSS's drop-down menus as well as what the resulting output tables look like. Finally, I write a memo summarizing my findings.

Although most of your academic career will be spent writing papers, in your professional career, you are much more likely to be required to write memos. This is your chance to become proficient at that. Although memos have introductions, bodies, and conclusions in common with term papers, they differ in one key respect. Whereas the purpose of writing a term paper is to convince your professor you know a lot, the purpose of writing a memo is to give your boss an answer to a question. As a result, memos need to be brief and clearly written, and give the answer right up front.

The memos you write in this class will have four parts: the heading, the introduction, the body, and the conclusion. You can find a template for writing a memo in Microsoft Word. If you click on "File" and "New," one of the template options is "Memos." There are several options available for you to download. As you begin to write the memo, the heading should include who it is to, who it is from, the date, and the subject. The introduction should state the problem that you have been asked to solve, describe the data you will use, and give a short answer to the question. The body should give evidence to answer the question. The conclusion should summarize the findings, state how they answer the original question, and indicate the broader implications those findings have for the decision maker.

Box 1.2 provides a quick reference on how to write a memo. It is also contained in Appendix 1, "Tips for Professional Writing." The focus here is answering a practical question in a clear and professional way. You do not want to write a long-winded memo—supervisors give assignments to save themselves time. Stay focused on answering the specific question using only relevant information.

BOX 1.2 How to Write a Memorandum

A. The Heading

1. To: Name and title
2. From: Your name
3. Date: Date sent
4. Subject: Keyword the nature of the assignment

B. The Introduction
   1. Describe the assignment briefly
   2. Describe the data you will use
   3. Summarize your results in one sentence

C. The Body
   Analyze the data

D. The Conclusion
   1. Summarize the data
   2. Explain how it answers the original question
   3. Describe any broader implications

My goal in writing each "Political Application" section is to model for you how to use SPSS to analyze data in a real situation. Following the "Your Turn" exercises at the end of each chapter, you will find a final assignment to "Apply It Yourself." This will place you (hypothetically) in a historical situation where you need to use SPSS to answer a question using real-world data. The datasets are found on the textbook's website, http://college.cqpress.com/sites/statspa. You will use SPSS to analyze the data and then write a memo answering the question. We'll start using SPSS in the next chapter. For now, let me walk you through how to write a memo.

A POLITICAL APPLICATION: INDOCTRINATION U.

It is February 2012 and the Republican presidential primaries are in full force. Candidate Rick Santorum has been attacking President Obama as a snob for thinking that all Americans should be able to attend college. Santorum not only argues that there are lots of great Americans who never attended college, but also accuses Obama of only wanting youth to attend college so that liberal professors can indoctrinate them. You are interning in the presidential campaign offices for former governor Mitt Romney, who, like President Obama, attended Harvard. Your boss is Ryan Williams, the press secretary for the campaign. He is afraid that Santorum might make similar attacks on Romney and asks you to take a look at an op-ed piece on Santorum's comments in the *New York Times*[14] and summarize the statistical argument in a memo.

1. Set up the heading.
   The joy of writing a memo is that you do not need to think about how to start. Because memos have a structured heading, I know that I begin by addressing the

memo to its recipient, in this case, Ryan Williams. The memo is from me and dated today. The subject line usually begins with "Re," which is short for "regarding." Because my boss is likely to have given me multiple assignments, I need to be clear here in identifying what this memo is about. I choose the three key words "Liberal University Indoctrination" in order to focus on this specific topic.

2. Write an introduction.
I begin the introduction by summarizing the question I was asked to answer: Does college make students more liberal? I indicate that the data I am going to use come from the four studies described in the op-ed piece. Finally, I answer the question. It is very important to have a clear answer at the end of the introduction because frequently, bosses will stop reading at this point.

3. Analyze the data in the body of the memo.
Here is where I give the evidence that I used to draw my conclusion. I see four relevant points. First, young adults do become more liberal. Second, professors do tend to be more liberal than the American public. Third, among young adults, college students are actually less liberal than non-college students. Fourth, college students are actually more likely to be religious than non-college students.

4. Conclude the memo.
In the conclusion, it is important to return to the original question so that your boss understands how the data answer it. Because students are not more liberal than non-students, I can conclude that young adults are not more liberal because they have been indoctrinated by their professors. But I need to do more than answer the question. After having done my research, I now know more about this particular topic than my boss. What does my boss need to know about the implications of my results? The underlying fear that motivated this question was that Santorum might make similar attacks against Romney, so I choose to address that fear in the conclusion. Although the finding that college students are more religious isn't actually relevant to answering the question, it might be a good hook if Santorum does end up attacking Romney for being a Harvard elitist.

Memo

To: Ryan Williams, Press Secretary
From: T. Marchant-Shapiro, intern
Date: May 30, 2012
Re: Liberal University Indoctrination

You asked me to summarize the statistical arguments made regarding former Senator Santorum's accusation that universities indoctrinate students to become liberal. The *New York Times* op-ed piece[15] refers to four different studies relevant to the statement. These studies find that although professors do tend to be more liberal and less religious than the general population, if you compare students to non-students of the same age, young adults typically become more liberal regardless of education.

To a certain degree, Santorum is correct in his description of university professors. Whereas only 20 percent of Americans consider themselves liberals, fully 50 percent of university professors do. Similarly, professors are less likely to believe in God than the general population: 20 percent of professors are atheists as compared to 4 percent of Americans more broadly. He is also correct that students tend to become more liberal while in school. But the relationship is not causal. If you compare students to non-students, you find that the change observed among students is simply a result of their age. Students do not become any more liberal than non-students of the same age and actually are more likely to retain their religious views than non-students.

Former Senator Santorum is incorrect when he says that universities indoctrinate students. Although professors do tend to be more liberal and less religious than the general public, their beliefs do not rub off on their students. Governor Romney would be justified in extolling the virtue of a college education both for the economic benefits of being better prepared for the job market and for its tendency to build religious faith in its students.

Your Turn: Using Statistics

YT 1.1

Get a copy of a major newspaper or news magazine and thumb through it looking for statistics. Lots of times, the numbers are missing, but you can look for decisions being made based on some study or another. Even if the news story doesn't report the numbers, the numbers contained in the original study will have led to that particular decision. Find one article that reports something about a study in which statistics were used. Usually, such articles will be based on a news release by whoever did the study—track it down. Now that you've done your research, answer Huff's five questions:

1. Who says so?
2. How does he know?
3. What's missing?
4. Did somebody change the subject?
5. Does it make sense?

Evaluate the statistics in your news story on the basis of those five questions to conclude whether it is a good or bad use of statistics. (Be sure to attach a copy of the article!)

YT 1.2

Watch "The Joy of Statistics" online. How are statistics important for politics? Give three examples from the movie. You can find the movie at www.gapminder.org/videos/the-joy-of-stats/.

YT 1.3

In April 2012, the U.S. Supreme Court spent three days hearing oral argument on the constitutionality of the Affordable Care Act. Connected to that argument, the justices had read amicus curiae briefs, many of which contained empirical studies relevant to their decision. A few weeks later, Justice Stephen Breyer spoke at the Midwest Political Science Association meetings and indicated how helpful the empirical evidence was. He also talked about a decision that the justices had made at the time to not allow television cameras in the court. Their fear was that the ability of news reporters to edit the statements of justices would lead viewers to have less respect for the Supreme Court. In this case, no empirical data had been available. Justice Breyer said that he wished that the justices had had empirical evidence of how decisions by states to televise court proceedings had influenced the public views of those courts. As an appeals court, the Supreme Court does not hear new empirical evidence about the facts of the case. Why would empirical studies about the possible impact of the Court's decisions be useful as justices deliberate?

Apply It Yourself: Assess Grants to Political Scientists

It is May 2012. Arizona representative Jeff Flake has successfully passed a bill through the House of Representatives banning the National Science Foundation (NSF) from funding research by political scientists. Flake's accusation is that studying how politics works is a waste of taxpayer money. In response, Rick Wilson, the editor of the *American Journal of Political Science*, one of the premier journals in the field, blogged about some of the articles published in his journal using data funded by the NSF. You are interning at the National Science Foundation in the Division of Legislative Affairs. Because your boss, Anthony Gibson, is going to need to testify in the Senate when it takes up the appropriations bill (S.2323), the speech director, Lee Herring, asks you to write a memo describing how prior grants to political scientists have been used. Choose two of the studies described in Wilson's blog to summarize and explain why the results are useful. Conclude by identifying which of the two studies would best support the argument that political science research is not a waste of taxpayer money. The blog is found at http://themonkeycage.org/blog/2012/05/15/what-has-the-nsf-wrought/.

Key Terms

Bivariate statistics (p. 7)
Content analysis (p. 5)
Descriptive statistics (p. 7)
Empirical (p. 4)
Experiment (p. 4)
Inferential statistics (p. 7)
Observation (p. 4)
Population (p. 7)
Public record (p. 4)
Qualitative data (p. 6)
Quantitative data (p. 6)
Sample (p. 7)
Scientific method (p. 3)
Survey (p. 5)
Univariate statistics (p. 7)

CHAPTER 2

Measurement

Counting the Biggel-Balls

Whenever someone tells me "I love statistics," my gasp of surprise is inevitably followed by the discovery that the individual is a baseball fan. I can understand the enthusiasm. I have fond memories of sitting in Wrigley Field on a warm, sunny afternoon, scorecard in hand, keeping track of the stats for that day's game. But I've never been enough of a devotee to tabulate the numbers that most fans follow during the season: home runs, runs batted in, batting average, slugging average. Apparently, collecting statistics originated with a reporter, Henry Chadwick, who gathered data on outs, runs, home runs, and strikeouts in order to measure how valuable different players were to their teams. It was natural that he would begin collecting the data he calculated. One of his major concerns was that data be collected in a uniform way so that baseball players could be compared legitimately.[1]

Like Chadwick, as a statistician, my question is always "Where do the numbers come from?" In baseball, like most sports, the record keeping is done in accordance with a prescribed set of rules. A runner will never be able to claim the record for the fastest mile on an informal neighborhood sprint. The mile needs to be officially measured, the speed has to be observed by an unbiased timer, and the run has to take place in an approved race. Similarly, baseball statistics are collected in official ways during Major League games using a form very similar to the one devised by Chadwick a hundred years ago.

Although we usually hear numbers reported with an attitude of acceptance that they "just are," actually collecting the information is not always such an obvious process. This was brought home to me when my children were little and I ended each day by reading them *Dr. Seuss's Sleep Book*. I interspersed my reading with many fake yawns because, as everyone knows, "A yawn is quite catching, you see. Like a cough. It just takes one yawn to start other yawns off." As a tired mother, I hoped that as my children "caught" my yawns, they would also assimilate my exhaustion and fall asleep. But as a statistician, the page that always caught my attention pictured an enormous counting device:

Counting up sleepers . . . ?

Just how do we do it . . . ?

Really quite simple. There's nothing much to it.
We find out how many, we learn the amount
By an Audio-Telly-o-Tally-o Count.
On a mountain, halfway between Reno and Rome,
We have a machine in a plexiglass dome
Which listens and looks into everyone's home.
And whenever it sees a new sleeper go flop,
It jiggles and lets a new Biggel-Ball drop.
Our chap counts these balls as they plup in a cup.
And that's how we know who is down and who's up.[2]

As I read that page each night, I'm not sure which I was most amazed by: the device that could listen and look into everyone's home, or the chap who could count up to ninety-nine zillion nine trillion and three each night by the time my children had fallen asleep. But those are precisely the initial tasks of any statistician: first, identifying the cases, and second, counting them.

In this chapter, we will begin our exploration into statistical analysis by looking at measurement. The first step is identifying our cases—who (or what) is it we are actually studying? Second, we need to pinpoint the attribute of those cases that we are interested in studying—how can we measure it? Third, we need to evaluate our measurement of the attribute to make sure that it reflects what we are trying to study. Fourth, we record that measurement before we finally enter it into a database preparatory to analysis. As with all chapters in this book, this chapter will conclude with a description of how to use this concept to answer a political question.

FINDING YOUR CASES

Like the Biggel-Ball chap, the U.S. Census Bureau is tasked with two Herculean feats every ten years: The Constitution requires first that they find every person living on U.S. soil and, second, that they count them. It's really not simple, and there's quite a lot to it.[3] The twofold process of finding and counting U.S. residents has evolved and expanded for the past two centuries. The first stage of the census requires enumerators to identify all U.S. residents. When the first census was taken in 1790, U.S. marshals were sent around their districts, tasked with visiting every home and recording the name of each head of household along with the number of members in the household. They wrote down this information on whatever paper was available to them, made two copies, one of which they sent on to Washington for counting and one of which they posted in a public place in their jurisdiction for all to inspect. It wasn't until 1830 that the government provided the enumerators with standard forms on which to record the required information. The U.S. marshals retained responsibility for filling out these forms until 1880, when enumerators were hired specifically for the purpose of completing the census. Their task of finding each individual became easier in 1890, when the enumerators were issued maps of the roads they were assigned to cover. It became easier still in 1970 after U.S. postal workers were tasked with compiling an address register of all addresses in the United States. That same year, the U.S. Post Office also eased the workload of the enumerators with the institution of mailed questionnaires rather than personal visits. The task has eased further in the past two censuses with the availability of Internet and phone responses. Of course, the increased clarity of the process still doesn't change the fact that the population of the United States has increased from 3.9 million counted in the 1790 census to 308.7 million counted in the 2010 census. That is a lot of people to find, even with simplified processes. Somehow, in 2010, they were able to complete this first part of the process in two months.

The second part of the process, counting all the people identified during the enumeration, is even more time consuming. Originally, all those handwritten lists had to be tallied by hand. As the population of the United States increased, the time it took to count it increased as well. Although standardizing the forms helped the counting process, by 1870, counting the 39.8 million inhabitants by hand proved to be impractical. As a result, the chief clerk of the Census Office invented a basic counting machine to ease the problem. But even with the device, the magnitude of counting to 50.2 million in 1880 took seven months to complete. One of the employees of the 1880 census, Herman Hollerith, left that job to invent a more sophisticated device that could not only count, but also cross tabulate variables. His invention used cards based on those used by the Jacquard loom. Each individual identified in the census had a corresponding card with various holes punched in it to identify characteristics about the individual—age, gender, and so on. In the long run, Hollerith cards, with their eighty-column format, went on to become the basis of modern computers, and Hollerith's company went on to become IBM. But in the short run, Hollerith rented his counting machines to the U.S. Census Bureau for the 1890 census, allowing them to complete their marathon count up to 62.9 million in a record six weeks. The use of the first non-military computer in 1950 further helped the process. And the 1960 advent of optical scanners connected to bubble-coded responses helped further. But even with the sophisticated technology, the 2010 census took nine months to count.

As with the Biggel-Ball chap and the census, the first step of any analysis is to identify the **cases** in question and to find a way to measure them. For example, in the waning days of the 2010 lame duck Congress, President Barack Obama was able to get the Senate to ratify the New START Treaty with Russia. The issues raised in it echoed the issues raised by arms control treaties negotiated beginning in 1969. On the surface, it was relatively easy to negotiate a conceptual agreement that both countries be limited to the same number of nuclear weapons. In practice, though, measuring the number of nuclear weapons proved difficult. Much of the difficulty originated in the Soviet preference for weapons with multiple warheads, in contrast to the U.S. preference for single-headed missiles. As a result, the Soviets had fewer weapons but more total warheads than the United States. Before the two nations could agree on a treaty, negotiators had to agree on a definition of what constitutes a single nuclear weapon. Similarly, as we study political events, we need to clearly define our **unit of analysis**—what or who we are studying.

MEASURE AN ATTRIBUTE

Rarely are we interested only in counting the number of cases. Normally, we want to find the cases in order to measure the variation that those individuals exhibit on a particular attribute. For example, the Constitution mandated the decennial census because two key functions of government were dependent on the population of the states: representation and taxation. The marshals were charged not simply with counting the number of people in each state, but also with distinguishing between the status of those people. In particular, the Three-Fifths Compromise written into the Constitution required that they distinguish between "free Persons," those "bound to Service for a Term of Years," "Indians," and "all other Persons." The authors of this portion of the Constitution found it easier to have a residual ("all other persons") category than to define the concept of slavery, even though that was what they meant. But this concept of slavery had to be measured by the census. So the enabling legislation for the 1790 census instructed the marshals to draw a chart with six columns with the headings: "Names of heads of families"; "Free white males of 16 years and upwards, including heads of families"; "Free white males under 16 years"; "Free white females, including heads of families"; "All other free persons"; and "Slaves." The marshals then operationalized the concept of slavery by asking each head of household how many slaves he or she owned. Based on that operational definition, the marshals were able to identify a total of 694,280 slaves in the United States.[4] The point is that even the Census Bureau, which ostensibly simply wants to count people, also wants to measure attributes about individuals.

Distinguish the Conceptual and Operational Definitions

In statistical analysis, it is essential to identify both the conceptual definition of an attribute and its operational definition. The **conceptual definition** is something like the dictionary definition of a concept: What do we visualize when we use a term? The **operational definition** is the process by which we translate our observations of reality into a measurement. Although we want both definitions to be closely connected, we need to clearly define the concept first so that when we decide how we are going to measure it, we can actually evaluate how well our measure connects to the concept.

If we fail to specify both definitions in a systematic way, we will end up being confusing. In *Through the Looking-Glass,* Alice meets Humpty Dumpty, who begins to use words in ways that don't make sense to Alice. And so Alice challenges him on the definitions. "'When *I* use a word,' Humpty Dumpty said in rather a scornful tone, 'it means just what I choose it to mean—neither more nor less.'"[5] Conflating conceptual and operational definitions is akin to Humpty Dumpty's arrogance. A political example of the failure to distinguish between conceptual and operational definitions can be found in court debate over the definition of pornography. Justice Potter Stewart made many people uncomfortable in his suggestion of how to identify pornography: "I know it when I see it."[6] Not every community involved in a dispute over whether something is pornographic can run it by Justice Stewart for evaluation. So in 1973, the Supreme Court detailed the Miller Test to determine if obscene materials are protected.[7] The Miller Test leaves the definition of obscenity to the state, although it requires that the state's measure be clear. It then says that obscenity is not protected if it fails the SLAPS test—if it lacks Serious Literary, Artistic, Political, or Scientific value. Critics of Miller fear that basing the definition of obscenity in community standards makes too encompassing a definition. But in practice, the equal breadth of the SLAPS test has increased the umbrella of what is protected. The failure to define how to measure what is and what is not obscene has led to the inability of communities to limit it. In general, any time we fail to give clear conceptual and operational definitions, we limit the usefulness of our work. In the end, we have measured what we have measured, but everyone is left confused about what it means.

The meaning of statistical results relies on the use of the scientific method. This requires that we describe our procedures in advance in order to prevent us from being tempted to mold the data to support our theories about the world. Take the example of British doctor Andrew Wakefield, who believes that childhood vaccinations cause autism. He conducted a case study of twelve children, finding that eight of them developed "regressive autism" after receiving an MMR vaccine. A recent study in the *British Medical Journal* suggested, however, that Wakefield was so invested in his theory that he failed to fully define the indicators of autism in advance. The current study reexamined Wakefield's original data and found that several of the children he reported getting autism after an inoculation do not actually fit into the commonly accepted definition of the affliction, whereas others actually were exhibiting symptoms before they received the shot.[8] Proper scientific method requires that we define our concepts and measures clearly so that we can be equally clear about the difference between what it would look like if we are right and if we are wrong. This means that before we ever start collecting data, we need to clearly define the concepts contained in our theory.

Articulate the Operational Measure

A clear measurement requires a description of not only what the attribute in which we are interested looks like, but what the alternatives look like as well. When the United States and the Soviet Union negotiated arms limitation treaties, they not only discussed the definition of a nuclear weapon, they also discussed verification. This meant that there had to be monitors to verify the count of weapons. But the monitors had to be able to examine places where weapons were *not* in order to verify that they were not there. Years later, the role of monitors was key in the invasion of Iraq: How could monitors know if there were no weapons of mass destruction? If that question was not answered before the monitoring ever began, there was no way for Iraq to prove to the United States' satisfaction that our expectations were being met.

Assignment of a Number. In Chapter 1, we learned that the difference between quantitative and qualitative analysis is that quantitative analysis looks at variables numerically and analyzes them on the basis of those numbers. As a result, an essential component of the process of operationalizing a variable entails assigning numbers to the different values of the variable. Sometimes, the numbers seems obvious. For example, if we are discussing government spending for programs, it seems obvious to quantify the variable in terms of the dollars spent. Similarly, if we are talking about age, it seems obvious to quantify it in years.

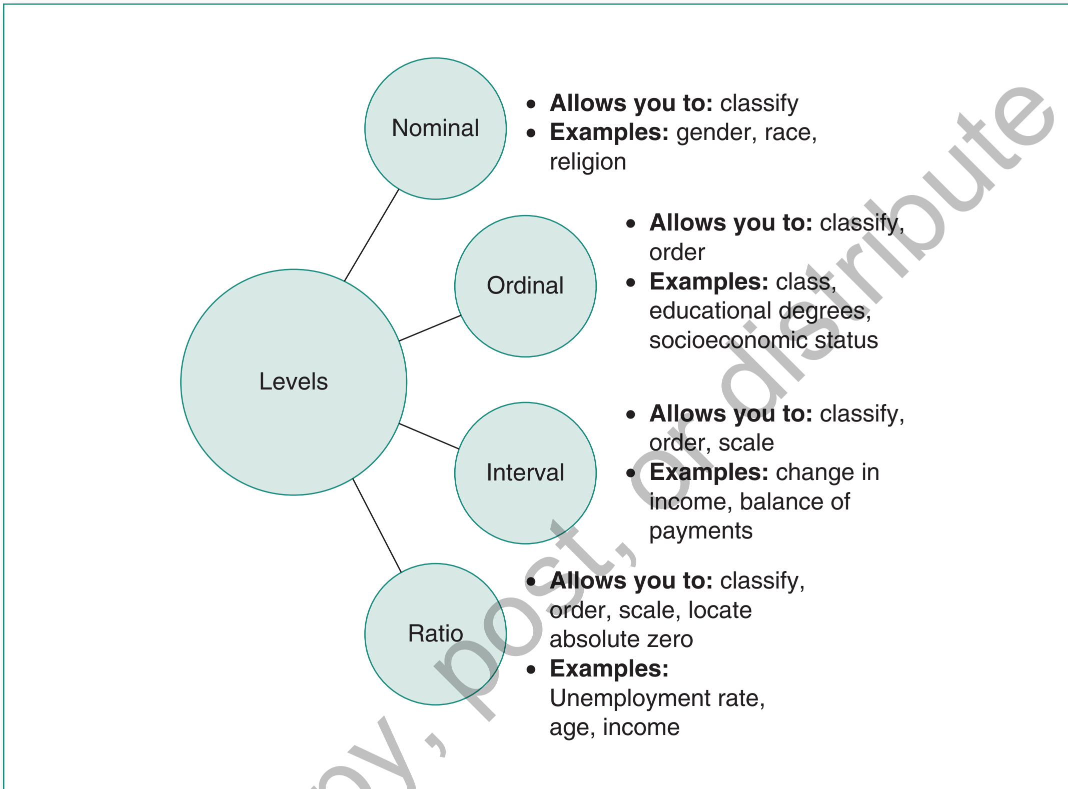
But for other variables, the process of quantifying the values is less obvious. Particularly in the social sciences, some of our most important concepts do not have physical manifestations that are easy to count. We **reify** concepts like democracy—discussing them as if they are real—when they are actually abstractions. Conceptually, we think about democracy as a system of government in which the people rule. But that conceptualization can include multiple factors: competitive elections, political alternatives articulated by multiple parties, social and political freedoms, and even the absence of corruption. Trying to take all those dimensions into consideration while measuring which countries are more (or less) democratic is very difficult. In practical terms, the common practice is to have individuals who are experts about a particular country answer a series of questions about whether the country embodies each of the constituent **indicators**, or characteristics reflecting different aspects of a concept. Each "yes" response increases the final index of democracy by an additional point. This method of measuring democracy is not perfect,[9] but it does allow a number to be assigned to an abstract concept.

Levels of Measurement. Depending on the underlying concept, the numbers associated with its operationalization may have more or less meaning. In statistics, we address three facets of the measure (categorization, order, and scale) and place them in a hierarchy. The more of these dimensions reflected in the numbers associated with a variable, the higher the **level of measurement**. We label the levels of measurement (in order from lowest to highest) nominal, ordinal, interval, and ratio. Figure 2.1 summarizes the characteristics of the different levels of measurement.

The level of measurement about which we have the least amount of information is **nominal**. Nominal-level data have only categories—they do not have a natural order. Some common examples of nominal-level data are race, gender, hair color, and astrological sign. For each of these variables, any individual can be placed in a group, but the variable isn't really an attribute about which an individual can have more or less. You really cannot be more or less gender; you are either a man or a woman. Our computers require us to associate numbers with the categories, but there is no logical reason why I would code men as "0" and women as "1" rather than the reverse. Nominal variables identify only membership in groups.

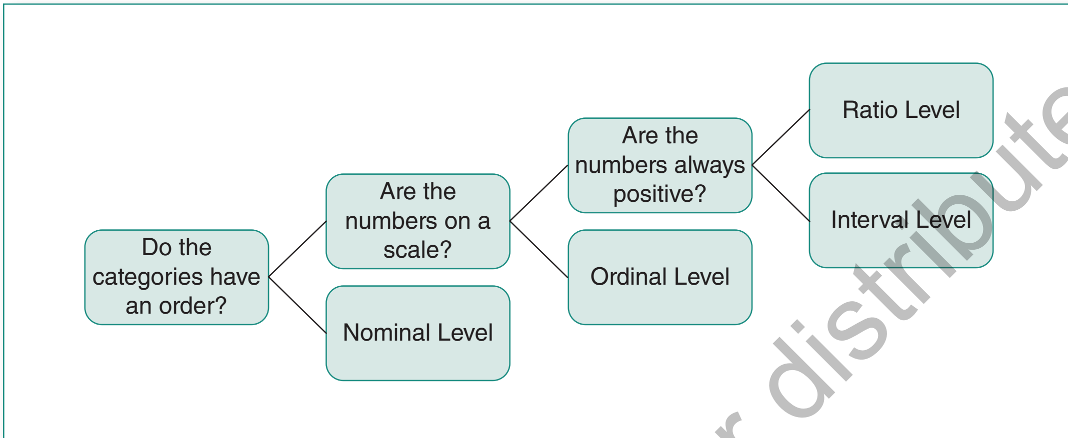
The second level of measurement is **ordinal**. In this case, the categories actually do have an order from less to more. But the numbers we associate with those categories do not reflect a uniform amount of more. Some common examples are class, educational degrees, and socioeconomic status. We can rank degrees in order from high school diploma, associate's, bachelor's, master's, to PhD and know that we have gone from less to more education. But the transition between each category does not reflect an equal amount of education. I can assign each of them a number from zero to five, but the number only puts them in order; it does not reflect a uniform amount of education.

Figure 2.1 Levels of Measurement



The third and fourth levels of measurements are interval and ratio. For both of these, the units not only reflect an order, they also measure a uniform amount. You can say that the numbers have a measurable distance on a scale. As the value of the variable increases by one point, the attribute increases by the same amount regardless of where you are on the scale. If the description of the variable uses the word "scale" or "index," you can be confident it is either interval or ratio. The difference between interval and ratio levels of measurement is that **interval**-level measures can have negative values whereas ratio-level measures have an **absolute zero**. A **ratio**-level measure is a special kind of scale that can have only a positive value. Because zero means a total absence of the attribute being measured, two ratio values can be compared proportionally: If one number is twice as big as another, we can conclude that it has twice as much of that attribute. If the variable is presented either as a percentage or physical measure (size, weight, quantity, or amount), it is probably ratio level.

Figure 2.2 Decision Tree to Determine Level of Measurement



Because the different levels of measurement contain different amounts of information, what you can do with them statistically varies. Statistically, you can conclude more from variables with higher levels of measurement. That is why we spend the second half of this textbook learning about different measures of association. The higher the level of measurement of the variables you are studying, the more you can do in measuring their association. (Except that for statistical purposes, interval- and ratio-level variables get treated the same way mathematically.)

Because identifying the appropriate measure of association depends on the variable's level of measurement, it is essential that you learn to identify the level of measurement correctly. I suggest using a series of three questions to identify the level of measurement. These questions are presented as a decision tree in Figure 2.2. First, ask whether the categories have an order. If they don't, it is a nominal-level variable and you can stop there. If there is an order, then ask if the numbers are on a scale. Does each unit equal the same amount of the attribute you are measuring? If the answer is no, the numbers don't tell you anything more than the order of the categories, then it is an ordinal-level variable and you can stop there. If the answer is yes, then the last question you ask is whether there is an absolute zero, which indicates the total absence of the attribute in question. If the answer is no, you can have negative values, then the variable is interval level. If the answer is yes, the variable can have only positive values, then the variable is ratio level.

Conclusion. After identifying our cases (who we are studying), we need to define what attribute we are studying both conceptually and operationally. The conceptual definition paints a word picture of what we are talking about—along the lines of a dictionary definition. Only after defining the concept do we consider how best to measure it. And that operational definition needs to be clear enough that anyone else could use it and get exactly the same results. Because it is easier statistically to find associations between variables with higher levels of measurement, make that a priority as you are operationalizing your variable.

EVALUATE THE CONCEPTUAL AND OPERATIONAL DEFINITIONS

Once we have defined both our concept and how we are going to measure it, we can evaluate how well the two definitions connect. In practice, no measure is going to quantify a concept perfectly. The 1790 census erred in its count of slaves for at least two reasons. First, it is off by at least sixteen because the sixteen "colored" residents of Vermont were all free but for some reason were included in the final count as slaves.[10] It is also off because in this first census, the relationship between the results and taxes was unclear. As a result, many respondents were cautious about revealing information about their families and "property." As social scientists, we usually look at two aspects in order to evaluate the quality of a measure: validity and reliability.

Valid Measures

Validity means that it makes logical sense to measure our concept in this way. The question here is, How close does the measure come to getting at the underlying concept? When we argue that a particular measure is valid because it makes sense, we call it **face validity**. After the Civil War, the census operationalized race by the identification of the enumerator, but more recently, the census has asked individuals how they identify themselves. It has been valuable for social scientists to realize that physical characteristics are not synonymous with race. Furthermore, this self-identification has led to the realization that "black" and "Hispanic" are not mutually exclusive categories, as was assumed by earlier censuses. As a result, since 2000, the census has used two different questions, the first to identify those of Hispanic descent and the second, those who think of themselves as "black." We can now identify at least some of those who think of themselves as multi-racial. The face validity of self-identification even led to including "negro" as an option in the 2010 census because some older Americans identify themselves that way.[11]

Sometimes, social scientists claim validity because a measure has widespread use—we call this **consensual validity**. Unfortunately, everyone agreeing does not mean that there is a perfect connection between a concept and its measure. Consider, for example, the concept of unemployment. Conceptually, when we use the term "unemployment rate," what we mean is the proportion of the workforce that is without a job at any given point in time. Operationally, though, in the United States unemployment is measured by the Census Bureau, which takes a survey and asks a series of questions without ever using the word "unemployment." Someone in Washington takes those responses and excludes anyone who is jobless but is either not looking for a job or not available for work at that time. Conversely, it counts anyone as employed who can name an employer even if they worked only a minimal amount. And some people who didn't work at all or earn anything are included in the "employed" category if they were ill, on family leave, or involved in an industrial dispute.[12] An alternative method of measuring unemployment would be to use the number of people filing unemployment claims. But that would mean excluding anyone who didn't file for unemployment because, for example, they were being supported by a spouse. It also would exclude all those for whom benefits had expired. Logically speaking, the first operationalization is probably a more valid measure than the second. It has consensual validity because, although it is not perfect, most experts agree that it is the best we can do at measuring what we think of conceptually when we use the term "unemployment."

Alternatively, we can claim **associational validity** if our measure is correlated with other measures connected to the concept about which we are interested. For example, when measuring the partisanship of voters, political scientists will ask two questions, a uniform question first and one of two possible follow-up questions. First, they ask all respondents if they think of themselves as a Democrat or a Republican. If any respondents refuse to associate with one of the two major parties and insist that they think of themselves as Independents, the follow-up question is, "Which party do you lean toward?" Some respondents will continue to insist they are Independents, but many will say they lean toward one of the two major parties. If, however, the respondents allied with one of the two parties in the first question, the follow-up question asks if they think of themselves as a strong or a weak partisan. The possible responses to the questions lead to seven possible categories of partisanship: Strong Democrat, Weak Democrat, Leaning Democrat, Independent, Leaning Republican, Weak Republican, and Strong Republican. In practice, many political scientists will take those seven categories and lump the three categories for each party together. For example, "Democrats" would include Strong Democrats, Weak Democrats, and Leaning Democrats. Thus, those respondents who originally identified themselves as Independents and only when pushed said they leaned toward the Democratic Party will, in the end, get labeled as Democrats. Lumping the leaners with the partisans clearly defies the face validity of self-identification. Political scientists justify it, though, by arguing that if you correlate self-identified partisanship with how these individuals vote, the leaners actually are more likely to vote in a partisan way than the weak partisans. Because we think of partisanship as including how someone votes, political scientists define leaners as partisans on the basis of associational or correlational validity.

We can also claim a measure to be valid if it is a good predictor of an effect we are trying to explain by our concept. We call this **predictive validity**. For example, those who complain about affirmative action say that admission to a good college or getting a job should be based on merit. For them, it is obvious that merit should be operationalized as good grades and high scores on standardized exams. Because both of those measures of merit are good predictors of success in life, opponents of affirmative action claim predictive validity for their operationalization. Face validity, consensual validity, associational validity, and predictive validity are all explanations of why it makes sense to measure a concept in a particular way.

Measuring Variables Reliably

In contrast to validity, **reliability** means that the measurement is consistent—the results don't change with time or researcher. To a large degree, reliability is dependent on how well you define your measure. If you are coding information, you need to be clear from the outset how it will be coded (hence the difference between Dr. Wakefield and his critics in coding "autism"). Usually, this will mean having at least two people coding the same data and reporting the correlation between those two codings as "inter-coder reliability." It would also mean publishing your methods so that other researchers can replicate your findings.

Unclear operationalizations can lead to problems in reliability. For example, if the questions used in surveys don't communicate clearly what they are asking, respondents will tend to respond in random or inconsistent ways. One problem in question wording is being overly general. If you were to ask "Do you support world peace?" any viewer of talent pageants knows that the answer will always be "Yes." Left unmeasured is what policy measures the supporters of peace actually advocate. Do they support intervention in Darfur? Oppose recognition of Myanmar? Support an autonomous Palestinian state? More specific questions would yield more reliable results.

At the other extreme, overly specific question wordings can also yield unreliable results if respondents don't know what the question means. Most Americans could answer reliably whether they support or oppose President Obama's health care act, but it is possible that calling the law by its official name would change the results of the survey: "Do you support or oppose the Patient Protection and Affordable Care Act?" Similarly, getting views on specific aspects of it could be more difficult because of lack of information. "Do you support or oppose the provision that gradually eliminates the donut hole in prescription care coverage?" Someone who is retired would be well informed on the topic and able to take a discerning stand. In contrast, many college students do not know what the prescription donut hole is. Frequently, survey designers will increase the reliability of overly specific questions by giving background information. "Prior to the current health care act, Medicare would cover the first $300 of prescription care, but wouldn't cover any more medicine until the elderly person had paid $6,000 out of pocket for prescriptions. Do you oppose or support the health care act's elimination of this 'donut hole' in prescription coverage?" Respondents need to know what the question means before they can answer in a consistent way.

The honesty of respondents can affect the reliability of data as well. As mentioned earlier, historians of the 1790 census had some reason to fear that property owners were less than forthcoming in their reports because of their fear of taxes. We need to avoid framing our questions in a way that might provoke an emotional response from those answering the questions (e.g., "Would you want an admitted homosexual to be a scout leader for your son?"). But even with neutral terminology, sometimes the political situation leads to unreliable responses. There have been a few elections where the preelection polls overestimated the vote a black candidate would get on Election Day. The "Bradley Effect" is named for the election Los Angeles mayor Tom Bradley unexpectedly lost when he ran for governor of California. The theory is that whites may not report their racial bias when asked a survey question, but that bias erupts in the polling booth. As a result, respondents end up voting differently from how they answered on the survey. If the Bradley Effect is correct, then in elections with one black candidate, survey questions about voting intentions are unreliable indicators of the actual vote. Along a related vein, researchers into the prevalence of workplace dishonesty found that it was less reliable to ask individuals whether they use company property (paper supplies or the company car) for personal use than to ask whether the behavior is typical of most people. In this case, the self-report may seem a more valid measure, but reported perception of peers is more reliable.[13] Sometimes, self-reports are more reliable measures of values than of behavior. For example, more Americans self-report attending church than actually show up for services in any given week. A more reliable measure comes when the survey is conducted on a Monday and the respondents are asked to detail their activities of the previous day (without any mention of church). The self-report of church attendance reflects how much Americans value church attendance. But the Monday survey is a more reliable measure of actual behavior.[14] However valid self-reports may be, the (dis)honesty of respondents can interfere with reliability.

Claims of reliability need to carefully distinguish between unreliable data and attributes that really are changing. The measure of partisanship that lumps leaning Independents with self-identified partisans is considered reliable because the proportion of Americans in each category changes slowly over time. But **cross sectional surveys** can discuss only aggregate-level changes because they interview a different group of respondents every time. When partisanship is asked about in a **panel study**, where the same people are asked the same questions at various points in time, the reliability of this measure of partisanship becomes questionable. Particularly among the self-identified Independents, their response to "toward which party they lean" can be inconsistent.[15] Categorizing the leaners as partisans may then have associational validity, but it is not necessarily a reliable measure of the long-term political attribute we were trying to measure. It could be that Independents have no long-term commitment to a party; they are really only indicating for which party they plan to vote in the next election.

Sometimes, a measure is just too volatile to be considered reliable. For example, many people follow the stock market because it has an important role both in the U.S. economy and in our personal retirement plans. But although the evening news always reports how the stock market did that day, it is rarely reported as a "leading economic indicator" because it fluctuates too much. For example, the "Flash Crash" of May 2010 highlighted the problem of computerized trading. An unusual combination of events combined with computer algorithms for selling led to an unprecedented 1,000-point drop in the stock market in just five minutes. The "crash" did not reflect real changes in the economy. Similarly, when you invest in the stock market, any financial advisor worth his or her pay will tell you to ignore what is going on in the stock market. Watching it too closely will encourage you to sell when you become concerned about your investments and buy when you are pleased that your investments are doing well. But if you want to have a comfortable nest egg when you retire, you actually want to do the reverse: Buy low, sell high. The stock market is an unreliable indicator of both the economy and your retirement because it is too volatile.

After you define the concepts you are interested in studying and carefully detail how you will measure them, you then need to evaluate how well your measure embodies your concept. This discussion should address the validity and reliability of your measure in quantifying your concept. Is your measure getting at the underlying essence of your concept? Is your measure clear enough that it will yield the same results given the same situation?

BOX 2.1 Numbers in the News

In the fall of 2012, the U.S. Bureau of Labor Statistics (BLS) released two contradictory numbers. First, it said that the unemployment rate had increased from 7.8 percent in September to 7.9 percent in October. Second, the BLS said that employers had hired an additional 171,000 workers. If employers were finally starting to hire new employees again, shouldn't the unemployment rate have decreased? The answer to the puzzle is found in the measures. The Bureau of Labor Statistics calculates the unemployment rate based on responses to a survey of households, whereas it bases job growth on a survey of businesses and government agencies. The two measures have two key differences. First, fewer households are surveyed than businesses, so the unemployment rate is less reliable than the estimate of job growth. Second, the estimate of job growth does not include numbers for farmers or self-employed workers, so the estimate of job growth has lower validity than the unemployment rate.[16]

TRANSLATE INFORMATION IN NUMBERS: CODING YOUR DATA

After thinking carefully about your concept and how best to measure it, you can begin the process of actually doing so. We call the process of translating information in numbers **coding**. This involves four steps. First, you create a **coding sheet** on which to record the data. Second, as you collect the information, you actually record the data on the coding sheet. Third, you transfer the data from the coding sheet to a database in preparation for data analysis. Finally, you clean any mistakes you might have made in entering your data.

Create a Coding Sheet

After you have made the decision about how to measure the variable, you then need to actually measure it. This process is easier if you have a coding sheet on which you can enter the data while you collect it. Coding sheets can be elaborate. If you are collecting a lot of data about each case, you would want to have a separate coding sheet for each case. (Remember how in the 1880 census there was a different Hollerith card for each individual?) For now, though, we are only gathering information about one variable, so we can use a single coding sheet to collect information about all of the cases. Set up the coding sheet so that each case has its own line and each variable has its own column. The first column should identify the case; the second, the attribute you are measuring.

As you are collecting the data, it is helpful to have your coding sheet as structured as possible to minimize human error. You will need to write in the identity of each case by hand, but for the attribute, it is easier to have the possibilities already identified. That will mean that you just have to check off which value applies for each case. To do this, you will have to go through the very helpful process of identifying all of the possible categories an attribute can take.

As you are identifying the possible categories that a case can have for a particular attribute, you need to obey two rules. First, the categories need to be **mutually exclusive**. This means that no one case can belong in more than one category. Second, the categories need to be **exhaustive**. This means that every case has to belong in one of the categories. You'll have problems following these rules if you try to measure two different attributes within a single variable. For example, suppose you're categorizing phenotypes and create a variable that has the alternatives "blond haired, blue eyed," "red haired, green eyed," and "brown haired, brown eyed." Even if hair color and eye color tend to be correlated, the correlation is not perfect. I have (had) red hair and brown eyes. How would you code my phenotype? One option is to include a category for each possible combination of hair and eye color. Sometimes, you will want to designate an "other" category to lump together several uncommon possibilities. Alternatively, you could include a "missing" category for the cases that don't fit into your original options as well as the cases for which you have no data. I normally use the number "–9" to indicate that I don't have data for a variable for that case, although if –9 were a possible value, I would make –99 or –999 the missing value. But in the end, every case needs to fall into one and only one of the categories of each variable.

You will not only want to set up your coding sheet to make gathering your data easier, you will also want to set it up to make using the data easier. Just as the Census Bureau can count faster using automated counting machines, in general, for quantitative data, you are going to want to use a statistical package to analyze it. That will entail entering the data electronically into a database. Although computers can keep track of words (we call that "string" or "alphanumeric" data), they can't analyze them very well. Computers (and statistics) are most useful for analyzing numerical data. So, although there isn't a problem just typing in the name of the case identifier (because you won't be analyzing that), for your attribute, you are going to want numbers rather than the labels for your categories. Your life will be much easier if you include next to each category the number you want associated with it. So if my attribute were gender, I would know that my possible categories are to be male and female. If, every time I entered another case, I had to remember to type "0" when I saw "male," I would be very likely to make mistakes along the way. If, however, I put a "0" next to "male" and "1" next to "female," then as I enter the data into a database, I easily type the number I see without risking mistake.

Code Your Data

Once you've set up your coding sheet properly, you are ready to collect your data for use. Your first column should contain the case identification. Begin by writing the name of each case in the first column. The cells of the second column should contain a list of the possible categories of the attribute you are measuring. As you find the information for each case, check off the appropriate category. As you are collecting the data, be sure to keep track of any sources you use to find the information. You will need to give them credit in the end, and it will be easier to keep track of the sources now, rather than to go back and find them later.

GET A FREQUENCY DISTRIBUTION

With only one variable and a limited number of cases, it is easy enough to just count the number of cases that have each of the possible categories of your variable. In general, though, you are going to have more cases and more variables than it is practical to count by hand. Just as the Census Bureau was able to enumerate the population in 1890 much more expeditiously with the incipient IBM computer, you'll find your job much easier if you enter your data in a spreadsheet before you try to analyze it. In the application section, I'll give instructions on how to enter the data in Excel and import it into SPSS for counting.

But even with SPSS to do the counting, the task of producing a professional-looking table is entirely yours. Box 2.2 details how to make a professional-looking frequency table. You'll notice that each of the boxes in this chapter begins with a title. The title is prefaced with an identifying number that is assigned in the order that the boxes are found. In this case, "Box 2.2" refers to the second box in Chapter 2. After labeling the table, you need to give it a descriptive title that includes the attribute you have measured.

The first row of the table should include column headings. The first column is going to contain the categories of the variable, so it should be headed with the name of the variable. The second column should be headed "Frequency" and the third, "Percent." In the second row, place the name of the first category followed by the number of cases in this category along with what percent of all the cases are in this category. In each of the subsequent rows, do the same thing for each of the other categories of the variable. After you finish entering all of the categories, there should be one row left. In the first column, place the word "Total"; in the second, the total number of cases; in the third, add up the percentages in the third column to get "100.0%," plus or minus 0.1%. (This difference is due to rounding error. If it totals less than 99.9% or greater than 100.1%, recheck your math.) If you have any cases for which you were unable to determine their values for this variable, you exclude them completely from the table.

In professional political science papers, each major section of the table is separated by a horizontal line. This is a little different than what you see in this book. For papers, the standard is to place a line at the top of the table (below the title) and at the bottom. You place a third line under the column headings. Begin by highlighting the full table (without the title). When you created the table, Word automatically put lines around all the cells of the table; you'll want to get rid of those first. If you look on the right-hand side of the "Paragraph" section of the "Home" menu of Word, you'll see a square icon—sometimes it is shaped like a box, sometimes it has a grid in it. Right now, it probably has a grid. Click on the arrow next to the icon to get the drop-down menu and choose the "Borders and Shadings" command. Within the "Borders and Shadings" window, click on the setting "None" and the lines in the "View" should vanish. A single line will already be selected as the "Style." In the "Preview," click above and below the table to insert lines there. Once you click on "OK," your table should have single lines at the top and at the bottom. Next, add the line below the headings by highlighting the first line of the table—the one containing the column labels. At this point, the "Borders and Shading" icon will have changed to a box. Click on the arrow to open the drop-down menu and choose "Bottom Border." Your table should now have the three lines: one at the top, one below the labels, and one below the table.

BOX 2.2 How to Create a Professional-Looking Frequency Table in Word

1. Give it a title describing its content.
2. After the title, insert a table of the appropriate size: columns = 3; rows = 2 + the number of categories in your variable.

   >Insert

   >Table

   >Insert Table

   Number of columns = 3

   Number of rows = *2 + number of categories of variable*

3. In the first row, label the three columns: *Variable Label*, "Frequency," "Percent."
4. In the first column, list the categories of the variable.
5. In the second column, give the frequencies of those categories.
6. In the last column, give the valid percentages—do not include the missing cases unless this is actually relevant to your analysis.
7. In the last row, give the totals for the columns: "Total," "*number of valid cases*," "100.0%."
8. Draw the appropriate lines.

   Highlight the entire table:

   >Borders and Shading *(the arrow, not the icon)*

   >Borders and Shadings

   Settings

   >None

   Style

   The single line should be highlighted.

   Preview

   Click above and below the table to add lines there.

   >OK

   Highlight the first line (which has the column labels):

   >Borders and Shading *(the arrow, not the icon)*

   >Bottom Border

9. Below the last line, use a smaller font to add the data source and any other clarifying information.

Finally, you can clean up the table. For example, I like to have a little extra space between the column labels in the first line and the data in the second line. To insert that space, I highlight the second line and use the "Paragraph" command in Word to add "Spacing" "Before." Similarly the "Total" line looks best separated from the data, so I highlight that line and use the "Paragraph" command to add space both before and after it. The table also looks nicer if the columns are centered and the decimals are lined up so I highlight the full table and click on the "Center Text" icon in the "Paragraph" box. In addition, I usually adjust the size of the columns. Because the three columns are stretched all the way across the page, they look too far apart. With the table highlighted, I can move my cursor across a row to find the margin of each column. If I left-click and hold, I can slide the margins over. Finally, after the concluding line, you should give any important information about your variable: where you found your information and how you measured it. If your percentages do not add up to precisely 100.0%, you can indicate here that it is due to rounding error. Use a smaller font for any information after the last line so that it doesn't merge into whatever text follows it.

SUMMARIZING THE PROCESS: MEASUREMENT

This chapter has very practical applications in answering political questions. In this section, I'll describe the general process of measuring an attribute for the purpose of answering a question. This will make more sense in the next section, where I will walk you through an example of actually answering a political question. Box 2.3 outlines the steps we've already covered in this chapter on how to measure a variable. First, you need to identify the cases you are studying. You will want to measure an attribute of those cases that varies. Define that attribute conceptually. Then define it operationally—how will you measure it? Next, evaluate the measure. How well does the measure line up with the concept? Is it valid? Is it reliable? Once you have decided how to measure your variable, you need to do so. Set up a coding sheet for entering the data and fill it out. Then copy the data into a database, cleaning the data if necessary to make sure it is correct.

BOX 2.3 How to Measure a Variable to Answer a Question

1. Identify the unit of analysis. Who or what are the cases you want to identify in order to measure their variation on an attribute?
2. Define your variable conceptually. What is the attribute you are interested in studying?
3. Define your variable operationally. How will you measure that attribute?
4. Evaluate the validity and reliability of your measure. Does it make sense? Is the measurement clear enough that anyone else would get the same result as you?

(Continued)

(Continued)

5. Create a coding sheet for collecting your data, including one column to identify your cases and one column to code your variable. List all the categories for your attribute, make sure they are mutually exclusive and exhaustive, and associate numbers with each category.
6. Code your data.
7. Enter the data into Excel.
8. Import your data into SPSS and clean them. Look for any anomalies (fix anything that looks wrong), identify the missing values, and label each of the values.
9. Request a frequency for your variable.
10. Type up a professional-looking table, indicating both the number of cases and percent for each category.

USE SPSS TO ANSWER A QUESTION WITH MEASUREMENT

If you want to answer a question by measuring a concept, you will go through all of the steps covered in this chapter. You will first define the variable conceptually. You will then translate the concept into a measurement, trying to maximize the validity and reliability of the measure. Next, you will take this operational definition and create a coding sheet, which you will then use to code this variable for each of your cases. Coding sheets in hand, you will then enter your data into an electronic database (either SPSS directly or Excel). From there, you can use SPSS to provide you with a basic frequency of the variable, which you can then use to create a professional-looking frequency for use in answering your question.

Transfer Your Data to an Electronic Database

Once you are done collecting the data, you can either enter them directly into SPSS or into an Excel worksheet. For both SPSS and Excel, the structure of this will be similar to your coding sheet: Each row will contain one case; each column, one variable. The first column will be your case identification. The second column will be the value of the attribute. I personally find it easier to enter data into Excel than SPSS, so that's what I'll describe. But you could also go into the data view window of SPSS and enter the data directly into the cells you see there. Each column is a different variable, originally named "V0001," "V0002," and so on. In each row, you take one case and enter the values of the variables in the corresponding column. If you switch to the "Variable View" of the Data Window (the tab to switch between "Data View" and "Variable View" is found in the lower left-hand corner of the window), you can change the names of the variables to be more descriptive. For example, you can change "V0001" to "Case Identifier."

The convenient aspect of Excel is that in the first row, you can insert labels for your variables. So cell “A1” will identify your unit of analysis, and “B1” will identify the attribute you have measured. When you import the file into SPSS, it is very easy to indicate that these are the names of the variables. After you have indicated the variable labels in the first row of your Excel file, you can start entering data in the second. In the first column, beginning in cell “A2,” you can type in the names of your cases. In the second column, beginning in cell “B2,” you will type the number associated with the category for the attribute you measured. You will end up with a spreadsheet that looks very similar to the coding sheet, except that it has only numbers in the second column. Be sure to save this file. If you haven’t saved it, you won’t be able to find or use it.

Although Excel can do some rudimentary statistics, in this textbook, I will describe how to do data analysis using SPSS (short for Statistical Package for the Social Sciences). This is the statistics package most commonly used by social scientists, whether they are in academic settings or in business or government agencies. As a result, learning to use SPSS can be fairly remunerative. I will focus on SPSS in order to achieve my goal of teaching you a marketable skill. Your university will probably have a site license for SPSS, so you should be able to do all of your work on campus.

Many students, though, prefer to do their work at home. Unfortunately, SPSS is not only popular, it is also expensive—expensive enough that it is normally purchased only by organizations. SPSS was purchased by IBM, which is now producing a student version you can rent for under $100. In its prior incarnation, the student version was not worth purchasing because it could analyze only a limited number of cases and variables. Those limitations no longer hold. The one drawback is that the rental license is only good for a period of time—either six or twelve months. If you search for “SPSS GradPack” with your search engine, you can find various vendors who handle the student rental process. Check all of these because the details vary by length of rental, the number of copies, whether you get disks or download electronically, and (most importantly) the price. Get version 19 or later to ensure compatibility. What they call the “Base” package is quite sufficient for your needs.

After you enter the data into an Excel file, you will need to import it into SPSS. Detailed instructions are found in Box 2.4. (You should notice that this book, focusing as it does on “How to do statistics,” has a lot of “How to” boxes. These are collected at the back of the book in appendixes to help you find them later. All of the “How to Use SPSS” boxes are found in Appendix 2.) You first open SPSS by double-clicking on it, and then you open the Excel file that contains your data. There are two tricks to finding the Excel file. First, you need to be sure that you have saved it. Second, you need to be sure that the file type SPSS is looking for (at the bottom of the window) is set to Excel. By default, SPSS looks for an SPSS data file, so those are the only ones that show up. Under the window that allows you to navigate through your documents to find the data file, you will see two boxes: “File Name” and “Files of Type.” The box next to “Files of Type” will be prefilled with the default file type “SPSS Statistics (*.sav).” Next to that default file type, you’ll see a down arrow. Click on that to get the drop-down menu shown in Figure 2.3. Choose “Excel (*.xls, *.xlsx, *.xlsm).” Once you change the file type to “Excel,” the Excel files will show up as available to you. Click on the Excel file so that its name enters the “File Name” box. Then click on “Continue.”

BOX 2.4 How to Import an Excel Data File into SPSS

Find "IBM SPSS Statistics" in your computer's programs and double click to open it. At this point, your first window opens called "Output [Document]." This is the window where any statistical results that you request will be reported. Superimposed on this window is a box asking if you want to open an existing data source. Since you want to import a new Excel file, you want to cancel out of this box.

What would you like to do?

>Cancel

>File

>Open

>Data

Find the appropriate folder

Files of type = "Excel" (*.xls, *.xlsx, *.xlsm)

File Name = *younamedit.xlsx*

>Open

✓Read variable name from first row of data

>OK

At this point, your second window opens called "Untitled [Dataset]." This window will show your data. But to be able to use it in the future, you will need to save the data into SPSS format. Make sure you are in the Data Window and do the following:

>File

>Save As

File Name = *make sure you are in the right folder and give it a name*

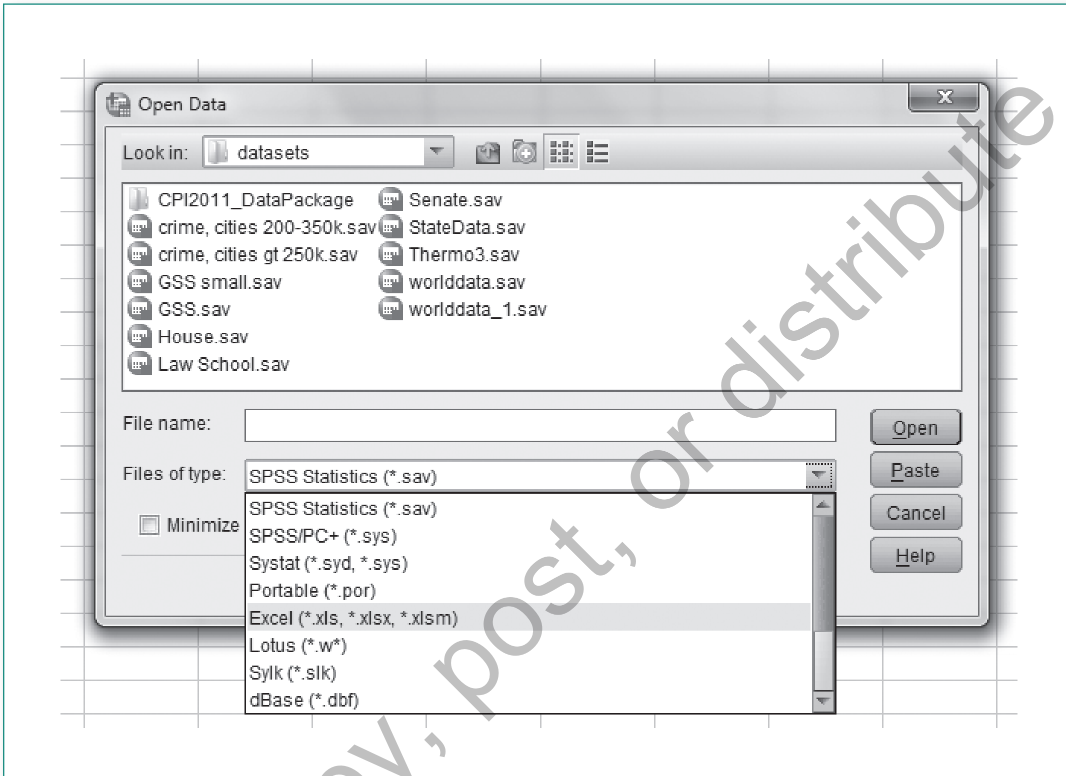
>OK

Once you have opened the data in SPSS, be sure to save it into an SPSS data file. SPSS opens two windows: the Output Window and the Data Window. To save the data, you need to save while you are in the Data Window. You'll know you are there because it will look a lot like an Excel worksheet. Once there, simply use the "File," "Save As" commands in the same way you would with Word.

Clean Your Data

Every time you plan to analyze data, you should always make sure that it has been entered properly. We call the process of correcting any errors **data cleaning**. (Instructions can be

Figure 2.3 How to Find Excel Files to Open in SPSS



found in Box 2.5.) You can sometimes catch errors by looking over the raw data, and you definitely should do this. But you are actually more likely to find errors by looking at how the data are set up. In SPSS, go to the "Data" command and click on "Define Variable Properties." Identify the variables on which you want to work by highlighting them with a click and then clicking on the arrow. Once you have moved the variables over, click on "Continue" to process the request. You will now see a window that has a box on the left listing each of the variables you requested with the first variable highlighted. The box on the right will describe the highlighted variable by listing all the possible values and their frequencies. Look through all the possible values to see if they make sense. If there is a value that you know isn't possible for that variable, you need to fix this. Go into the Data Window, find the column for the variable, and then scan down to find the weird value. Once you find it, look to the left to find the case. With the case identifier, you can check the original coding sheet to find the correct value. Simply type that value into the appropriate cell of the Data Window.

BOX 2.5 Cleaning Your Data in SPSS

>Data

>Define Variable Properties

Click on each relevant variable to highlight it.

→Variables to Scan

>Continue

For each variable, check that the values make sense; if not, fix in the Data Window.

>−9=Missing

Type in the value label in the box next to each number.

>OK

After assuring yourself that all the values make sense, give labels to each of the values. In the "Define Variable Properties" window, one of the columns is labeled "Missing." You'll want to click on the appropriate box to set that value to missing. If you followed my example and made "–9" missing, you'll want to click on the missing box in the row for the value of "–9." Of course, if you do not have missing data for any of your cases, "–9" will not appear as a value and so you will not be able to set a missing value. To the right of the numerical values, you can enter labels to correspond with each number. So, for the gender variable, if you coded men as "0" and women as "1," you can type "men" next to "0" and "women" next to "1." SPSS has some odd quirks with adding in variable labels. Later on, when you are working with more than one variable, you might have problems if you label from the top down. So it is probably a good idea at this point to get in the habit of assigning a label to the last category first and working your way up to the top of the categories. Once you have labeled all of the categories for the first variable, move on to the second. Identify the missing values and then start assigning labels at the bottom and work your way up. Systematically work your way through each of the variables in the box on the left, identifying the missing values and labeling each possible value.

Get a Frequency Table

After you've made your corrections, request a frequency for each of your variables to verify that everything has been fixed. A **frequency** is a table that shows the result of totaling the number of cases in each category of your variable. The command to analyze data is at the top of both the Data Window and the Output Window—it doesn't matter which you use. To get a frequency, the procedure is to click on "Analyze," "Descriptive Statistics," and "Frequencies." At this point, you will see a list of the variables in your dataset in a box on the left. If you click on a variable to highlight it and then click on the arrow, the variable name will move into the "Variable" box. Once you have all your desired variables in the variable box, you can click on "OK" to process your request. The frequency table will then show up in the Output Window. A summary of this process is found in Box 2.6. If everything in the frequency looks correct, go into the Data Window and save your data file. You do this in much the same way as you would in Word—find "File" in the upper left-hand corner and "Save" a file with a name and in a place you will remember in the future.

BOX 2.6 How to Get a Frequency of a Variable in SPSS

After opening your data in SPSS:

>Analyze

>Descriptive Statistics

>Frequencies

Click on the variable to highlight it.

→*To bring the variable over to the "Variable" box*

>OK

At this point, the frequency of the variable will show up in the "Output" window.

These steps are all preparatory to actually answering a question. In order to answer a question, you'll need to translate the final (clean) frequency into a more usable form. First, you're going to want to copy the frequency table from SPSS and paste it into a Word document. Although you can save the Output Window for future reference, it is only readable by SPSS. So if you are going to want to access your frequency at home, I recommend opening a Word document at this point, copying the frequency from the Output Window, and pasting it into the Word document.

The SPSS frequency, however, is not the professional-looking table you will want to use when presenting your findings. Professional-looking tables follow certain protocols you will want to emulate so that your reader takes your results seriously. The SPSS frequency has all the information you need, but it is only a resource from which you will borrow the appropriate numbers as you are creating your own frequency table. Follow the instructions in the chapter (summarized in Box 2.2) to use Word to create a professional-looking frequency table.

Now that you have collected your data in a professional-looking table, you can write a professional memo answering your question. Chapter 1 described the elements of a memo. The box summarizing that process is found in Appendix 1. As with all writing, memos should have an introduction, a body, and a conclusion. Unlike other writing, brevity (combined with clarity) is highly valued in memos, so do your best to keep it short.

An SPSS Application: Europe and the European Union

It is November 2010. Ireland's economy has been tottering, but the government has been refusing pressure from the European Union (EU) to accept a bailout. Prime Minister Brian

Cowen, afraid of a vote of no confidence, has been claiming that an EU loan is unnecessary—that his planned austerity measures will be enough for Ireland to regain its financial footing. But on November 21, Finance Minister Brian Lenihan announces that Ireland has formally applied to the EU and the International Monetary Fund for a bailout package. You are interning in the Congressional Joint Economic Committee, and today you have been charged with dealing with constituent emails. You receive the following query:

To: Joint Economic Committee
From: Alice Concerned
Date: November 22, 2010
Re: European Bailouts

I just heard in the news that the EU is going to bail out Ireland. But last summer, they allowed Iceland to go bankrupt. Why didn't they bail out Iceland, too? I am of Scandinavian descent, and it sure sounds like the EU has a prejudice against northern climes.

When you show it to your supervisor and ask her what to do, she says, "Reply to it, of course." Noticing your look of panic, she gives you a hint: "Why don't you put together a table of how many European nations are members of the European Union?"

1. Identify the unit of analysis.
   I want to identify which countries are in Europe. I could look up Europe in a dictionary, find its boundaries on a world map, and then look for all of the countries within those boundaries. But my dictionary doesn't have an entry for Europe. (Is that weird or what?) And aren't there some European countries the size of cities that wouldn't show up on a map? I would be liable to miss some. So instead, I decide to look up "Europe" on Wikipedia—it is sure to have a list. After doing that, I am glad I went that route because it indicates that even defining the concept of "country" isn't as straightforward as I had assumed it would be. It actually has several categories: "recognized sovereign states," "partially recognized sovereign states," "unrecognized sovereign states," and "dependent territories." I choose to look only at the "recognized sovereign states" because that is the closest to how I had conceptualized a country. I might need to revise that later if it ends up that some EU nations fall into one of the other categories.

2. Conceptually define your variable: What does it means to be a member of the European Union?
   The European Union is a formal association between various countries in Europe based on a signed treaty. Each country retains its political and military autonomy, but commits to a level of economic interrelatedness. Most EU nations share a common currency, which means that each individual country's economy is dependent on the economies of all the other EU nations. As a result, there are prerequisites that a country has to achieve in order to be allowed into the organization. And once it joins, the treaty specifies certain responsibilities that each country needs to uphold.

3. Operationally define what is meant by being a member of the EU: How will you measure it?
A member of the EU is any nation that is currently a signatory of the treaty. This is a valid measure because it is the technical definition—a country cannot be a member without signing.
For ease of data collection, I am going to find the list of member states on Wikipedia. But because I've been told that Wikipedia is not always a reliable source of information, I decide to double-check by looking at the EU's website.

4. Evaluate the validity and reliability of your variable.
Because both sources are publicly available, anyone can follow the same procedures and get the same results. As a result, I can be confident that my coding will be reliable.

5. Create a coding sheet for collecting your data.

Table 2.1 EU Coding Sheet, Blank

Coding Sheet **Which European Countries are Members of the European Union?**	
European Country	*Member of the European Union?*
	☐ 0 No ☐ 1 Yes
	☐ 0 No ☐ 1 Yes
	☐ 0 No ☐ 1 Yes
	☐ 0 No ☐ 1 Yes
	☐ 0 No ☐ 1 Yes
	☐ 0 No ☐ 1 Yes
	☐ 0 No ☐ 1 Yes
	☐ 0 No ☐ 1 Yes
	☐ 0 No ☐ 1 Yes

Sources: "List of Sovereign States and Dependent Territories in Europe," Wikipedia, The Free Encyclopedia. http://en.wikipedia.org/w/index.php?title=List_of_sovereign_states_and_dependent_territories_in_Europe&oldid=494486538. "Member States of the European Union," *Wikipedia, The Free Encyclopedia*, May 23, 2012. "Countries," *European Union*, 2012. http://europa.eu/about-eu/countries/index_en.htm.

Table 2.1 shows a truncated version of my coding sheet. The first column leaves space to write in the names of the European countries identified in (A). The second column contains all the possible categories for EU membership: either yes it is, or no it is not a member. Each of those categories is assigned a numerical value. So "no" it is not a member of the EU will be coded "0," and "yes" it is will be coded "1." There are check boxes next to the number so that as I am coding the data, I can just check the appropriate box. The coding sheet also includes any information a coder might need—in this case, where one should go to find the data.

6. Code your data.
   Table 2.2 shows a truncated version of the coding sheet with the data.

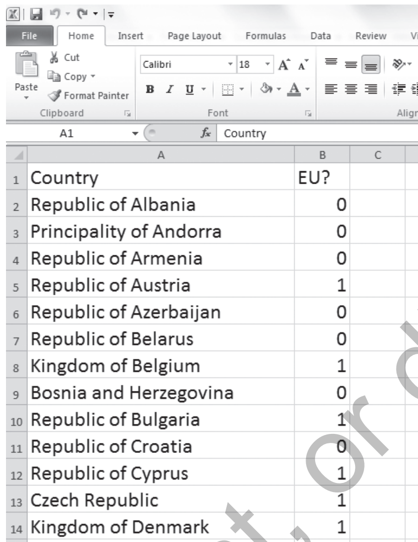
Table 2.2 EU Coding Sheet, Completed

Coding Sheet **Which European Countries Are Members of the European Union?**	
European Country	*Member of the European Union?*
Republic of Albania	☑ 0 No ☐ 1 Yes
Principality of Andorra	☑ 0 No ☐ 1 Yes
Republic of Armenia	☑ 0 No ☐ 1 Yes
Republic of Austria	☐ 0 No ☑ 1 Yes
Republic of Azerbaijan	☑ 0 No ☐ 1 Yes

Sources: "List of Sovereign States and Dependent Territories in Europe," *Wikipedia, The Free Encyclopedia*, May 28, 2012. http://en.wikipedia.org/w/index.php?title=List_of_sovereign_states_and_dependent_territories_in_Europe&oldid=494486538. "Member State of the European Union," *Wikipedia, The Free Encyclopedia*, May 23, 2012. "Countries," *European Union*, 2012. http://europa.eu/about-eu/countries/index_en.htm.

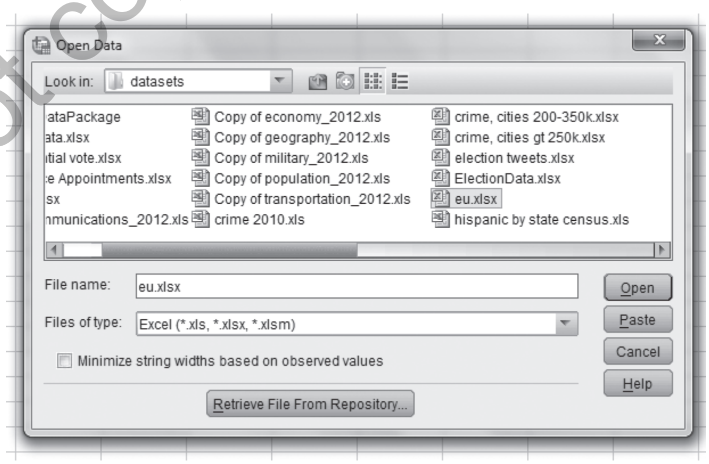
7. Enter your data into Excel.
   Entering data into Excel is really pretty easy. I just need to be sure to make each column corresponds to a different variable. In Excel, the first row in each column should name the variable. The first column is the case identifier; in this instance, it is the name of the European country, so this column is headed "Country." The second column is the variable about European Union membership, so it is headed "EU?" Each line in the file will look at a single case and include all the data for every variable in the appropriate column. Figure 2.4 shows what this file looks like. I am careful to save this Excel file ("File" "Save As") before trying to do anything with it.

Figure 2.4 Excel File of EU Data



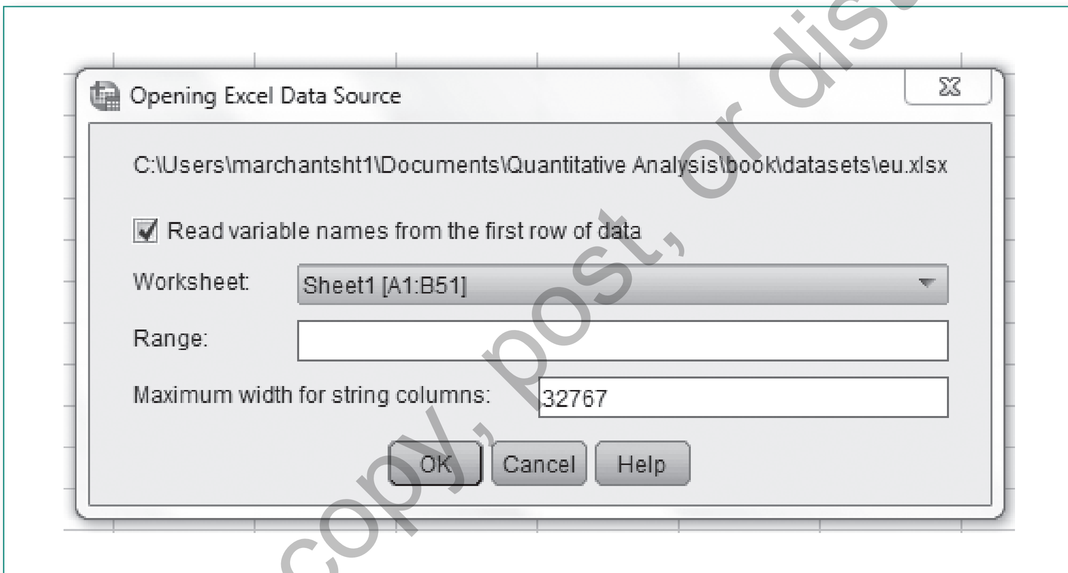
8. Import your data into SPSS and clean it. Look for any anomalies (fix anything that looks wrong), identify the missing values, and label each of the values.

Figure 2.5 Opening an Excel File in SPSS



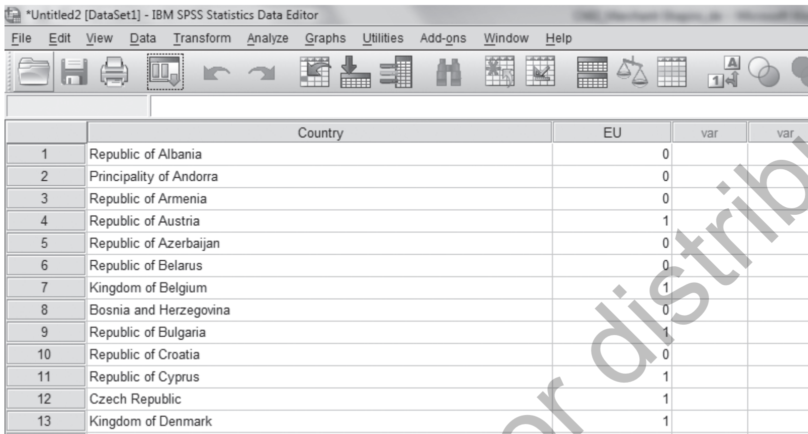
First, I open up SPSS. Once in there, I click on "File," "Open," and "Data" to open a data file. I find the location on my computer where I saved my Excel file. (Note: the file won't show up until you have saved it.) Once I got to the appropriate folder, however, only SPSS/PC+ files showed up because that is the default. At the bottom of the window, I changed the "Files of type" to indicate Excel as shown in Figure 2.5. Now Excel files show up in the "Open Data" box, so I find my file and click on it to enter it into the "File Name" box. I then click on "Open" to actually open the file in SPSS.

Figure 2.6 Indicating the First Line Contains the Variable Names



After opening the right file, a message asks if the first line includes the variable labels. Because the Excel file has the labels "European Country" and "EU?" rather than data in my first line, I leave the box checked and reply "OK" as shown in Figure 2.6. SPSS then opens two windows. In the Data Window, I see the data looking very much like my Excel file. Figure 2.7 shows what the Data Window looks like. While I am in the Data Window, I save my data in SPSS format with the standard "File," "Save As" command. (If I were to check my directory at this point I would notice that a file has appeared named "EU.sav." The ".sav" extension indicates that it is an SPSS data file.) In the Output Window, I see all of my commands along with their results. In the top of both windows are commands to get SPSS to perform statistical functions on my data.

Figure 2.7 SPSS Data Window for EU Data



To clean the data, I use the “Data,” “Define Variable Properties” command as shown in Figure 2.8. After designating my EU variable for scanning, I can assign missing values (which I do not have in this case) and the value labels. Figure 2.9 shows where to enter this information. I am careful to begin at the bottom when entering the value labels.

Figure 2.8 The Define Variable Properties Command, Step 1

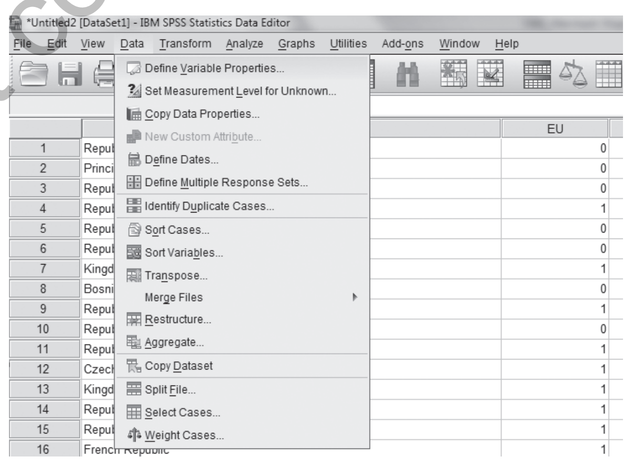
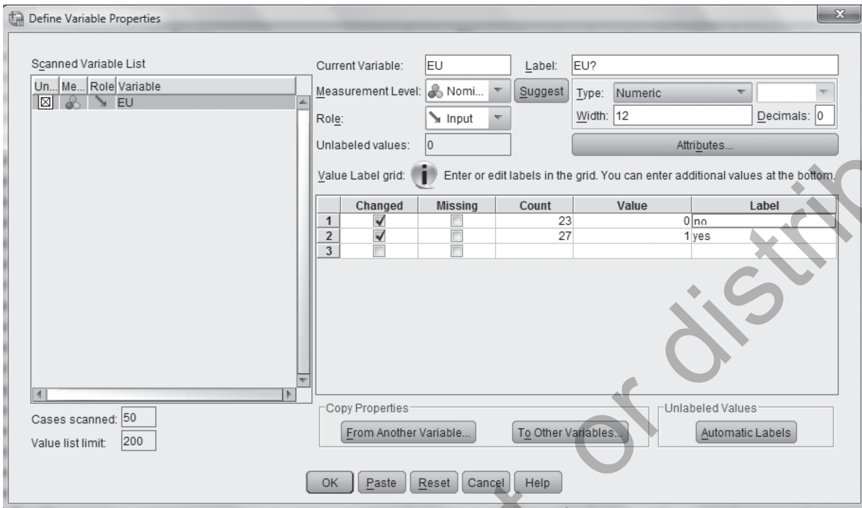


Figure 2.9 The Define Variable Properties Command, Step 2



9. Request a frequency for your variable.
   I want a frequency, so I click on "Analyze" (in either of the windows), choose "Descriptive Statistics," and then choose "Frequencies." At this point, a window opens with two boxes and an arrow between them. The left box has a list of my variables; the right is empty. I click on "EU?" because that is the variable for which I want a frequency. I click on the arrow and "EU?" moves into the right-hand box. At this point, I click on "OK" to get my frequency. It shows up in the "Output" window. Figure 2.10 shows what this looks like.

10. Type up a professional-looking table indicating both the number of cases and percent for each category.
    To have access to the frequency from my computer without SPSS, I copy the components of the Output Window into a Word document. Each component is in a table format, so I copy each component and then paste it into Word. When exiting SPSS, I am sure to save the data in the Data Window as an SPSS file, but do not bother saving the output in the Output Window.
    Because it is not professional to cut and paste the printout of a statistical package into written work, I transfer the information from the SPSS output into a more professional format that looks like Table 2.3. Notice several key features of the table: (1) It has a title describing its content. (2) The first column describes the categories of the variable. (3) The second column gives the frequencies of those categories. (4) The last column gives the percentages. (5) The last row gives the total for the columns. (6) Each section of the table is divided by lines. (7) The data source and any other clarifying information are found after the last line.

Figure 2.10 SPSS Output Window of Frequency for EU Data

Frequencies

[DataSet1]

Statistics

EU?

N	Valid	50
	Missing	0

EU?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	no	23	46.0	46.0	46.0
	yes	27	54.0	54.0	100.0
	Total	50	100.0	100.0	

Table 2.3 Frequency of Membership in the European Union by European Countries

Member of EU?	*Frequency*	*Percent*
No	23	46.0
Yes	27	54.0
Total	50	100.0

Sources: "List of Sovereign States and Dependent Territories in Europe," *Wikipedia, The Free Encyclopedia*. http://en.wikipedia.org/w/index.php?title=List_of_sovereign_states_and_dependent_territories_in_Europe&oldid=494486538. "Member State of the European Union," Wikipedia, The Free Encyclopedia, May 23, 2012. "Countries," *European Union*, 2012. http://europa.eu/about-eu/countries/index_en.htm.

11. Write a professional memo to Ms. Concerned answering her question and include a table at the end with the data you collected.

Memo

To: Alice Concerned
From: T. Marchant-Shapiro, Joint Economic Committee
Date: November 23, 2010
Subject: European Bailouts

We received your email wondering why the European Union agreed to a bailout for Ireland but allowed Iceland to go bankrupt. It is a very good question, especially in the face of the current economic crisis, which has made apparent the interrelated nature of the economies of the nations of the world. To answer your question, I collected data about which European countries are members of the European Union. The EU treats Ireland and Iceland differently because Ireland is a member of the EU and Iceland is not.

The European Union is the result of a treaty between specific European countries. In particular, these nations have forged strong economic ties and a common currency. Because of this interdependence, the EU is very careful about allowing countries to join: They must meet strict preconditions in order to join as well as make commitments about future economic behavior. As the table below shows, slightly more than half of the countries in Europe are actually members of the European Union.

That is the crux of the difference in the cases of Iceland and Ireland. Iceland is not a member of the EU and so it did not have the responsibilities of membership. And the EU did not have the responsibility to keep Iceland's economy afloat. Ireland, on the other hand, is a member of the EU and so, although its government actually did not want to accept the outside loans, it eventually buckled to pressure from its allies to accept the bailout.

You can expect that in the future, as in this instance, the European Union will bail out only its member states. Thank you for your interest in this issue. It is concerned citizens like you that keep our country strong.

Frequency of Membership in the European Union by European Countries

Member of EU?	*Frequency*	*Percent*
No	23	46.0
Yes	27	54.0
Total	50	100.0

Sources: "List of Sovereign States and Dependent Territories in Europe," *Wikipedia, The Free Encyclopedia*, May 28, 2012. http://en.wikipedia.org/w/index.php?title=List_of_sovereign_states_and_dependent_territories_in_Europe&oldid=494486538. "Member State of the European Union," Wikipedia, The Free Encyclopedia, May 23, 2012. "Countries," *European Union*, 2012. http://europa.eu/about-eu/countries/index_en.htm.

Your Turn: Measurement

YT 2.1 After the economic downturn of 2008, one aspect of the public debate was about what to call it.

1. Give a conceptual definition of an economic "recession."
2. In order to know whether or not we are in a recession, the operational definition needs to indicate both what is and what is not a recession. Find a reputable source and give its operational definition of a recession.
3. During this downturn, during what span of time was the United States in a recession according to this measure?

YT 2.2

One of the questions political scientists like to ask is, "Why do countries go to war?" Before you can even begin to answer it, though, you face a sticky measurement problem. How do you decide whether a country is at war or not?

1. Identify your unit of analysis.
2. Describe a way to determine whether a country is at war or not.
3. Evaluate the quality of that measure: Is it valid? Is it reliable?

YT 2.3

Determine the level of measurement (nominal, ordinal, or interval/ratio) for each of the following variables:

1. Education (in years)
2. Level of Education (no high school diploma, high school graduate, some college, college graduate, some graduate school)
3. Graduate Degree (MA/MS, PhD, JD, MBA, MD)
4. Race (White, Black, Hispanic, Asian, Other)
5. Marital Status (married, civil union, widowed, divorced, single)
6. Abortion (number of instances in which abortion should be legal: 0–7)
7. Income (in dollars)
8. Social Class (working class, middle class, upper class)

Apply It Yourself: Measure the Norm for Chief Justice Appointments

The year is 2005. Chief Justice William Rehnquist has passed away and President George W. Bush has nominated John Roberts to replace Rehnquist both for his seat on the bench and as chief justice. You are interning in the Senate Judiciary Committee and your boss, Committee Chair Arlen Specter, knowing that the learning curve is very steep just for becoming a justice, wonders whether it is practical to become chief justice at the same time. But perhaps that's the norm for chief justice appointments? He asks you to find out how many chief justices were appointed simultaneously to being placed on the Court.

1. The unit of analysis for your assignment is Supreme Court chief justices. How will you find all your cases?
2. Your variable is a combination of two concepts: appointment to the Supreme Court and appointment as chief justice. Define these concepts.
3. How will you measure them? Think about any problems you might have and predetermine how you will resolve them. For example, will you include the first chief justice? (By necessity, he must have been appointed simultaneously.) What will be the source of the data?

4. Does this measure give you valid data? How can you check to see if the data are reliable?
5. Create a coding sheet to record your data. You should have one variable for the name of the chief justices (which can be an alphanumeric variable) and one for your appointment variable (which should be a numerical variable). Be sure to indicate the source of the data.
6. Code the data. If you realize that you need to make adjustments in your operationalization, adjust point 3 above. If you do refine your measure, be sure to review your previous codings to make sure that the refinement does not affect how they should be coded.
7. Enter the data into Excel.
8. Import the Excel file into SPSS and clean the data.
9. Get a frequency for your Chief Justice variable.
10. Translate the output into a professional-looking table. How many chief justices had already been on the Supreme Court? How many were appointed simultaneously?
11. Write a brief memo for Senator Specter answering his question. Follow the format described for writing memos, and be sure it contains a properly formatted frequency table at the end.
12. If you were actually writing a memo for an assignment from work, you would not include any SPSS printout with your memo, but this semester you should attach copies of the relevant SPSS tables to your memos for grading purposes. For this assignment, turn in your memo along with your completed coding sheet, a printout of your spreadsheet, and a printout of your SPSS frequency.

Key Terms

Absolute zero (p. 23)
Associational (correlational) validity (p. 26)
Case (p. 19)
Coding (p. 29)
Coding sheet (p. 29)
Conceptual definition (p. 20)
Consensual validity (p. 25)
Cross sectional surveys (p. 28)
Data cleaning (p. 36)
Exhaustive (p. 30)
Face validity (p. 25)
Frequency (p. 38)
Indicator (p. 22)
Interval (p. 23)
Level of measurement (p. 22)
Mutually exclusive (p. 30)
Nominal (p. 22)
Ordinal (p. 22)
Operational definition (p. 20)
Panel study (p. 28)
Predictive validity (p. 26)
Ratio (p. 23)
Reify (p. 22)
Reliability (p. 26)
Unit of analysis (p. 19)
Validity (p. 25)