



# A GENTLE INTRODUCTION

## CHAPTER 1 GOALS

- Understand the general purposes of statistics
- Understand the various roles of a statistician
- Learn the differences and similarities between descriptive statistics and inferential statistics
- Understand how the discipline of statistics allows the creation of principled arguments
- Review eight essential questions of any survey or study
- Understand the differences between experimental control and statistical control
- Develop an appreciation for good and bad experimental designs
- Learn four common types of measurement scales
- Recognize and use common statistical symbols

The words “DON’T PANIC” appear on the cover of the book *The Hitchhiker’s Guide to the Galaxy*. Perhaps it may be appropriate for some of you to invoke these words as you embark on a course in statistics. However, be assured that this course and materials are prepared with a different philosophy in mind, which is that statistics should help you understand the world around you and help you make better informed decisions. Unfortunately, statistics may be legendary for driving sane people mad or, less dramatically, causing undue anxiety in the hearts of countless college students. However, statistics is simply the science of the organization and conceptual understanding of groups of numbers and making decisions from them. By converting our world to numbers, statistics helps us understand our world in a quick and efficient way. It also helps us make conceptual sense so that we might

be able to communicate this information about our world to others. More important, just as in the practice of logic, statistics allows us to make arguments based on established and rational principles; thus, it can promote wise and informed decision making. Like any other scientific discipline, statistics also has its own language, and it will be important for you to learn many new terms and symbols and to see how some common words that you already know may have very different meanings in a statistical context.

## HOW MUCH MATH DO I NEED TO DO STATISTICS?

---

Can you add, subtract, multiply, and divide with the help of a calculator? If you answered *yes* (even if you are slow at the calculations), then you can handle statistics. If you answered *no*, then you should brush up on your basic math and simple algebra skills. Statistics is not really about math. In fact, some mathematicians secretly revel in the fact that their science may have little relevance to the real world. The science of statistics, on the other hand, is actually based on decision-making processes, and so statistics must make conceptual sense. Most of the statistical procedures presented in this course can also be performed using specially designed computer software programs such as SPSS ([www.spss.com](http://www.spss.com)). Each chapter in this book has specific SPSS instructions to perform the analyses in that chapter. Programs such as SPSS make it easy to perform statistical calculations, and they do so with blazing speed. However, these programs provide very little or no explanation for the subsequent output, and that is at least one of the purposes of this course.

## THE GENERAL PURPOSE OF STATISTICS: UNDERSTANDING THE WORLD

---

A group of numbers in mathematics is called a set, but in statistics this group is more frequently called *data* (a single number is called a *datum*). Typically, the numbers themselves represent scores on some test, or they might represent the number of people who show up at some event. It is the purpose of statistics to take all these numbers or data and present them in a more efficient way. An even more important use of statistics, contrary to some people's beliefs, is to present these data in a more comprehensible way. People who obfuscate (i.e., bewilder or confuse) with statistics are not really representative of the typical and ethical statistician. This course will also provide you with some protection against the attempts of some people and their statisticians who try to convince you that a product works when it really does not or that a product is superior to another product when it really is not.

### Another Purpose of Statistics: Making an Argument or a Decision

People often use statistics to support their opinions. People concerned with reducing the incidence of lung cancer use statistics to argue that cigarette smoking increases the likelihood of lung cancer. Car and truck makers use statistics to show the reliability of their cars

and trucks. It has been said that being a statistician is like being an honest lawyer. One can use statistics to sell a product, advance an argument, support an opinion, elect a candidate, improve societal conditions, and so on. Because clearly presented statistical arguments can be so powerful—say, compared with someone else’s simple unsupported opinion—the science of statistics can become a very important component of societal change.

## WHAT IS A STATISTICIAN?

Although many of you would probably rather sit on a tack than become a statistician, imagine combining the best aspects of the careers of a curious detective, an honest attorney, and a good storyteller. Well, that is often the job of a statistician, according to contemporary statistician Robert P. Abelson (1928–2005; see Abelson, 1995). Let us examine the critical elements of these roles in detail.



<http://www.clipart.com/en/close-up?o=3982329>

Source: Clipart.com.

### One Role: The Curious Detective

The curious detective knows a crime has been committed and examines clues or evidence at the scene of the crime. Based on this evidence, the detective develops a suspicion about a suspect who may have committed the crime. In a parallel way, a statistician develops suspicions about suspects or causative agents such as which product is best, what causes Alzheimer’s disease, and how music may or may not affect purchasing behavior. In the case of Alzheimer’s disease, the patients’ lives are quantified into numbers (data). The data become the statistician’s clues or evidence, and the experimental design (how the data were collected) is the crime scene. As you already know, evidence without a crime scene is virtually useless, and equally useless are data without knowing how they were collected. Health professionals who can analyze these data statistically can subsequently make decisions about causative agents (e.g., does aluminum in food products cause Alzheimer’s disease?) and choose appropriate interventions and treatments.

<http://www.istockphoto.com/stock-photo-3118029-justice-scale-and-gavel.php>



A good detective is also a skeptic. When other detectives initially share their suspicions about a suspect, good detectives typically reserve judgment until they have reviewed the evidence and observed the scene of the crime. Statisticians are similar in that they are not swayed by popular opinion. They should not be swayed by potential profits or losses. Statisticians examine the data and experimental design and develop their own hypotheses (educated guesses) about the effectiveness or lack of effectiveness of some procedure or product. Statisticians also have the full capability of developing their own research designs (based on established procedures) and testing their own hypotheses. Ethical

statisticians are clearly a majority. It is unusual to find cases of clear fraud in the world of statistics, although they sometimes occur. More often, statisticians are like regular people; they can sometimes be unconsciously swayed by fame, money, loyalties, and prior beliefs. Therefore, it is important to thoroughly learn the fundamental statistical principles in this course. As noted earlier, even if you do not become a producer of statistics, you and your family, friends, and relatives will always be consumers of statistics (e.g., when you purchase any product or prescription drug); thus, it is important to understand how statistics are created and how they may be manipulated intentionally, unintentionally, or fraudulently. There is the cliché that if something seems too good to be true, then it might not be true. This cliché also holds in the world of statistics. Interestingly, however, it might not be the resulting statistics that make something appear too good to be true, but rather it may be how the statistics were gathered, also known as the experimental design. This course will present both the principles of statistics and their accompanying experimental designs. And as you will soon learn, the power of the experimental design is greater than the power of the statistical analysis.

### Another Role: The Honest Attorney

An honest attorney takes the facts of a case and creates a legal argument before a judge and jury. The attorney becomes an advocate for a particular position or a most likely scenario. Frequently, the facts might not form a coherent whole or the facts may have alternative explanations. Statisticians are similar in that they examine data and try to come up with a reasonable or likely explanation for why the data occurred. Ideally, statisticians are not passive people but rather active theoreticians (i.e., scientists). Scientists are curious. They have an idea about the nature of life or reality. They wonder about relationships among variables, for example, what causes a particular disease, why people buy this brand as opposed to that brand, and whether a product is helpful or harmful. These hypotheses are tested through experiments or surveys. The results of an experiment or survey are quantified (turned into numbers). Statisticians then become attorneys when they honestly determine whether they feel the data support their original suspicions or hypotheses. If the data do support the original **hypothesis**, then statisticians argue their case (study) on behalf of their hypothesis before a judge (journal editor, administrator, etc.). In situations where statisticians wish to publish the results of their findings, they select an appropriate journal and send the

article and a letter of justification to the editor of the journal. The journal editor then sends the article to experts in that field of study (also known as peer reviewers). The reviewers suggest changes or modifications if necessary. They are also often asked to decide whether the study is worthy of publication. If the researcher has convinced the editor and the peer reviewers that this hypothesis is the most likely explanation for the data, then the study may be published. The time between submission and the actual publication of the article will typically take between 6 months and 2 years, and this time period is known as publication lag.

There are, of course, unscrupulous or naive attorneys, and sadly too, as noted earlier, there are also unscrupulous or naive statisticians. These types either consciously or unconsciously force the facts or data to fit their hypothesis. In the worst cases, they may ignore other facts not flattering to their case or may even make up their data. In science, although there are a few outright cases of fraud, it is more often that we see data forced into a particular interpretation because of a strong prior belief. In these cases, the role of a skeptical detective comes into play. We may ask ourselves, are these data too good to be true? Are there alternative explanations? Fortunately, in science—and this is where we so strongly differ from a courtroom—our hypothesis will not be decided by one simple study. It is said that we do not “prove” the truth or falsity of any hypothesis. It takes a series of studies, called **replication**, to show the usefulness of a hypothesis. A series of studies that fails to support a hypothesis will have the effect of making the hypothesis fall into disuse. There was an old psychological theory that body type (fat, skinny, or muscular) was associated with specific personality traits (happy, anxious, or assertive), but a vast series of studies found very little support for the original hypothesis. The hypothesis was not disproven in any absolute sense, but it fell into disuse among scientists and in their scientific journals. Ironically, this did not “kill” the scientifically discredited body-type theory given that it still lives in popular but unscientific monthly magazines.

## A Final Role: A Good Storyteller

Storytelling is an art. When we love or hate a book or a movie, we are frequently responding to how well the story was told. I once read a story about the making of steel. Now, steelmaking is not high (or even medium) on my list of interesting topics, but the writer unwound such an interesting and dramatic story that my attention was completely riveted. By parallel, it is not enough for a statistician to be a curious detective and an honest attorney. He or she must also be a good storyteller. A statistician’s hypothesis may be the real one, but the statistician must state his or her case clearly and in a convincing style. Thus, a successful statistician must be able to articulate what was found in an experiment, why this finding is important and to whom, and what the



Source: *The Storyteller* by Mats Rehnman, © 1992. Used with the kind permission of the artist. Find out more about the fine art of storytelling by visiting the artist at [www.Facebook.com/mats.rehnman](http://www.Facebook.com/mats.rehnman).

experiment may mean for the future of the human race (okay, I may be exaggerating on this latter point). Articulation (good storytelling) may be one of the most critical aspects of being a good statistician and scientist. In fact, good storytelling may be one of the most important roles in the history of humankind (see Arsuaga, 2002, and Sugiyama, 2001).

There are many examples of good storytelling throughout science. The origin of the universe makes a very fascinating story. Currently, there are at least two somewhat rival theories (**theories** are bigger and grander than hypotheses, but essentially theories, at their hearts, are no more than educated guesses); in one story a Supreme Being created the universe from nothingness in 6 days, and in the other story, the big bang, the universe started as a singular point that exploded to create the universe. Note that each theory has a fascinating story associated with it. One problem with both stories is that we are trying to explain how something came from nothing, which is a logical contradiction. It has been suggested that the object of a myth is to provide some logical explanation for overcoming a contradiction; however, this may be an impossible task when the contradiction is real. Still, both of these theories remain very popular among laypeople and scientists alike, in part because they make interesting, provocative, and fascinating stories.

Science is replete with interesting stories. Even the reproductive behavior of *Planaria* or the making of steel can be told in an interesting and convincing manner (i.e., any topic has the potential to make a good story). Does our current society reward someone who tells a good story? Of course! For example, who is reportedly the richest woman in England? It is said to be J. K. Rowling, the person who told the world the story of a boy, Harry Potter, who went off to wizardry school.

## LIBERAL AND CONSERVATIVE STATISTICIANS

---

As we proceed with this course in statistics, you may come to realize that there is considerable leeway in the way data are investigated (called the experimental methodology or experimental design) and in the way data are tabulated and reported (called statistical analyses). There appear to be two camps: the liberals and the conservatives. Just as in politics, neither position is entirely correct given that both philosophies have their advantages and disadvantages.

Scientists as a whole are generally conservative, and because statisticians are scientists, they too are conservative as a general rule. Conservative statisticians stick with the tried and true. They prefer conventional rules and regulations. They design experiments in traditional ways, and they interpret and report their interpretations in the same conventional fashion. This position is not as stodgy as it may first appear to be. The conservative position has the advantage of being more readily accepted by the scientific community (including journal editors and peer reviewers). If one sticks to the rules and accepted statistical conventions and one argues successfully according to these same rules and conventions, then there is typically a much greater likelihood that the findings will be accepted and published and will receive attention from the scientific community. Conservative statisticians are very careful in their interpretation of their data. They guard against chance playing a role in their findings by rejecting any findings or treatment effects that are small in nature. Their findings must be very clear, or their treatment effect (e.g., from a new drug) must be very



large, for them to conclude that their data are real, and consequently, there is a very low probability that pure chance could account for their findings.

The disadvantages of the conservative statistical position are that new investigative research methods, creative statistical analyses, and radical conclusions are avoided. For example, in the real world, sometimes new drug treatments are somewhat effective but not on everyone, or the new drug treatment may work on nearly everyone, but the improvement is modest or marginal. By always guarding so strongly against chance, conservative statisticians frequently end up in the position of “throwing out the baby with the bathwater.” They may end up concluding that their findings are simply due to chance when in reality something is actually happening in the data that is not due to chance.

Liberal statisticians are in a freer position. They may apply exciting new methods to investigate a hypothesis and apply new methods of statistically analyzing their data. Liberal statisticians are not afraid to flout the accepted scientific statistical conventions. The drawback to this position is that scientists, as a whole, are like people in general. Many of us initially tend to fear new ways of doing things. Thus, liberal statisticians may have difficulty in getting their results published in standard scientific journals. They will often be criticized on the sole grounds of investigating something in a different way and not for their actual results or conclusions. In addition, there are other real dangers from being statistical liberals. Inherent in this position is that they are more willing than conservative statisticians to view small improvements as real treatment effects. In this way, liberal statisticians may be more likely to discover a new and effective treatment. However, the danger is that they are more likely to call a chance finding a real finding. If scientists are too hasty or too readily jump to conclusions, then the consequences of their actions can be deadly or even worse. What can be worse than deadly? Well, consider a tranquilizer called thalidomide during the 1960s. Although there were no consequences for men, more than 10,000 babies were born to women who took thalidomide during pregnancy, and the babies were born alive but without hands or feet. Thus, scientific liberalism has its advantages (new, innovative, or creative) and its disadvantages (it may be perceived as scary, flashy, or bizarre or have results that are deadly or even worse). Neither position is a completely comfortable one for statisticians. This book will teach you the conservative rules and conventions. It will also encourage you to think of alternative ways of exploring your data. But remember, as in life, no position *is* a position, and any position involves consequences.

## DESCRIPTIVE AND INFERENCE STATISTICS

---

The most crucial aspect of applying statistics consists of analyzing the data in such a way as to obtain a more efficient and comprehensive summary of the overall results. To achieve these goals, statistics is divided into two areas: descriptive and inferential statistics. At the outset, do not worry about the distinction between them too much; the areas they cover overlap, and descriptive statistics may be viewed as building blocks for the more complicated inferential statistics.

**Descriptive statistics** is historically the older of these two areas. It involves measuring data using graphs, tables, and basic descriptions of numbers such as averages and means. These

universally accepted descriptions of numbers are called **parameters**, and the most popular and important of the parameters are the mean and standard deviation (which we will discuss in detail in Chapter 3).

**Inferential statistics** is a relatively newer area that involves making guesses (inferences) about a large group of data (called the **population**) from a smaller group of data (called the **sample**). The population is defined as the entire collection or set of objects, people, or events that we are interested in studying. Interestingly, statisticians rarely deal with the population because, typically, the number of objects that meet the criterion for being in the population is too large (e.g., taxpayers in the United States), it is financially unfeasible to measure such a group, and/or we wish to apply the results to people who are not yet in the population but will be in the future. Thus, statisticians use a sample drawn from the population. For our inferences about our sample to be representative of the population, statisticians have two suggestions. First, the sample should be randomly drawn from the population. The concept of a random sample means that every datum or person in the population has an equal chance of being chosen for the sample. Second, the sample should be large relative to the population. We will see later that the latter suggestion for sample size will change depending on whether we are conducting survey research or controlled experiments.

Statistical studies are often reported as being based on **stratified samples**. A stratified sample means that objects are included in the sample in proportion to their frequency in the subgroups that make up the population. For example, if it was reported that a study was conducted on a religious beliefs stratified sample of the U.S. population, then the ratios of religious beliefs in the sample would be 47% Protestant, 23% no religious beliefs (agnostic, atheist, or none), 21% Catholic, 6% non-Christian (Muslim, Hindu, Jewish, Buddhist), and 3% other Christian churches (Mormon or Jehovah's Witness). Again, large samples, random samples, and stratified samples help us make inferences that are more than likely to be representative of the population that we are studying.

## EXPERIMENTS ARE DESIGNED TO TEST THEORIES AND HYPOTHESES

---

A theory can be considered a group of general propositions that attempts to explain some phenomenon. Typically, theories are also grand; that is, they tend to account for something major or important such as theories of supply and demand, how children acquire language, or how the universe began. A good theory should provoke people to think. A good theory should also be able to generate testable propositions. These propositions are actually guesses about the way things should be if the theory is correct. Hypotheses are specifically stated propositions created from and consistent with the theory that are then tested through experimental research. Theories whose specific hypotheses are frequently supported by research tend to be regarded as useful. Theories whose hypotheses fail to receive support tend to be ignored. It is important to remember that theories and hypotheses are never really proven or disproved in any absolute sense. They either are supported or fail to receive support from research. In the real world, it is also frequently the case that some theories and hypotheses receive mixed results. Sometimes research findings support



the propositions, and other experiments fail to support them. Scientists and statisticians tend to have critical attitudes about theories and hypotheses because, as noted earlier, the repercussions in science for a bad theory can be deadly or even worse. Thus, statisticians like well-designed and well-thought-out research, such that the findings and conclusions are clear, compelling, and unambiguous.

## ODDBALL THEORIES

---

Science is not simply a collection of rigid rules. Scientific knowledge does not always advance smoothly but rather typically moves along in sputters, stops and starts, dead ends, and controversy. Carl Sagan, the late astronomer, said that science requires the mating of two contradictions: a willingness to think about new, unique, strange, or even bizarre explanations coupled with rigorous skepticism and hard evidence. If scientists published only exactly what they thought they would find, new discoveries would be exceedingly rare. Thus, scientists must be willing to take risks and dare to be wrong. The physicist Wolfgang Pauli wrote that being “not even wrong” is even worse than being wrong because that would imply that one’s theory is not even worth contradiction or dispute. The Nobel Prize in Medicine was awarded to an American neurologist, Stanley Prusiner, who proposed that infectious particles called prions do not contain any genes or genetic material yet reproduce and cause “mad cow disease” (a dementia-like disease caused by eating infected meat). For years, Prusiner’s ideas were considered revolutionary or even heretical; however, most research now supports his more than two decades of experimental work. His “oddball” theory was provocative, generated many testable hypotheses, and was supported by his own rigorous testing and skepticism, and scientific knowledge has benefited greatly from his theory. However, remember that there is no shortage of oddballish people creating oddball theories. A good theory is not only interesting and provocative but also must be supported by testable hypotheses and solid evidence.

## BAD SCIENCE AND MYTHS

---

In my Sunday newspaper, I once saw a question in a fitness column concerning adhesive strips that go across the bridge of one’s nose. These same strips seem to be popular among professional athletes. “Do they work?” asked the letter writer. The magazine’s expert says they reduce airflow resistance “as much as 30%.” However, because the “expert” says it is true, does that make it true? Of course not! Nevertheless, we are bombarded daily by a plethora of statistics such as that one in four college-age women have been raped, marriage reduces drug abuse, classical music boosts IQ, people use only 10% of their brains, and left-handers die younger than right-handers. We even pass down many myths as truths such as that Eskimos have 30 words for snow and one should change the oil in one’s car every 3,000 miles. The Internet has even sped up the process of passing on this information. Some of these common myths are also called “urban legends.” Of course, anyone can post any belief or idea on the Internet, and the mere posting or retelling of a “story” does not make it true. For example, Eskimos apparently have just as many descriptions of snow as other

people. Because most of us have been raised to tell the truth, we typically do not challenge the facts or statistics fed to us daily. However, let us be realistic. Many studies are poorly designed, statistics can be misused or manipulated, and many people in our society are not interested in the truth but rather are interested in money and power (that brings more money). The magazine *Consumer Reports* empirically investigated the myth of the 3,000-mile oil change and consistently found no benefits compared with much longer intervals. Interestingly, no reputable scientist ever published the idea that people use only 10% of their brains. If it were true, I would very much hope that my 10% are the ones involved in heart rate, respiration, and blood pressure. A “motivational speaker” probably created this myth, but there’s not a shred of scientific evidence for its truthfulness. It may have persisted because the consequences of the belief (that humans are capable of doing more than they do) are not harmful. But now, let us return to the “nose strips.”

Ultimately, why did someone create nose strips? Probably to make money. Helping people breathe right was perhaps, if we are very lucky, secondary. In fact, it is easy to imagine that many products in the market are not even remotely designed to do what they claim. Their sole purpose is to make money. However, let us give the nose strip creator the benefit of the doubt. Why should we use nose strips? Well, the nose strip people say that they have scientific evidence that nose strips work, and they present the following “facts.” Breathing takes 10% of our total energy, and their nose strips reduce air resistance by up to 30%. With regard to this first fact, I think it would be fair to know how total energy was measured. Isn’t it rather difficult to come up with a single measure of how energy is expended? And even if we could come up with an acceptable measure, did they measure this 10% expenditure when a person was exercising or at rest? If the person was exercising (because I assume that’s when we’d want to use nose strips), how did they measure the total energy expended when the person was running around? Did they remotely or telemetrically send the energy expenditure information? I am not sure that this is even possible. It’s easier for me to imagine the person at rest when he or she was measured for energy expenditure. Therefore, at-rest measures might not be appropriate for someone who is actually running about. Furthermore, who was the person tested? What was that person’s age? Will we be able to generalize to the average person? Is it fair to generalize from just one test subject to all people? Was the participant paid? Some people will say anything for money. An even worse scenario is that some people will even deceive themselves about how well something works if they are paid enough money. (Curiously, Leon Festinger [1919–1989], a social psychologist, found that some people will deceive themselves and act positively about something if they are paid *too little* money.)

The same criticism appears to hold for the second “fact.” How did the nose strip people ever measure this 30% air resistance reduction? Was that 30% over a period of time? Was it a mean score for a large number of participants? Was it a single score? Who did the testing? Was it an independent testing group, or did the nose strip people conduct the research? If the nose strip people did the testing, I think they may have been consciously or unconsciously biased in favor of the product.

A second reason the nose strip people say that we should buy their product is that Jerry Rice, the former great pro football player, used to wear them. Now, we are tempted to use nose strips because of an expert opinion and the prestige associated with that expert. But

does that mean they really work? Of course not, but we rarely challenge the opinion of an “expert.” Many other factors are operating here such as the powerful process of people identifying with the rich and famous. Some people think that if Jerry Rice used them, then they will be as great as he was if they use them. We can see as we closely examine our motivations that many are not based on rational or logical reasoning. They are based, however, on unconscious and powerful forces that cause us to believe, trust, imitate, and follow others, particularly people who we perceive as having more power or status than we have. This willingness to believe on sheer faith that a product or technique works has been called the placebo effect (to be discussed shortly).

In summary, when we closely examine many of our attitudes, beliefs, facts, and “studies,” we find that they often lack any scientific validity. And if we blindly accept facts, figures, cures, and snake oil, we are putting ourselves in a very dangerous position. We may make useless changes in our lives, costly changes for no good reason, or even dangerous changes. We may also unnecessarily subject ourselves to needless stress and worry.

## EIGHT ESSENTIAL QUESTIONS OF ANY SURVEY OR STUDY

---

### 1. Who Was Surveyed or Studied?

Remember, most of the time, we hope that we can use the product or would benefit from a new treatment or technique. A majority of studies still use only men. Would we be able to generalize to women if only men are studied? Would we be able to generalize to children if only adults are studied? Would we be able to generalize to people if only animals are studied? Will the product be safe for people if we do not use animals in our research? If our moral values prevent us from using animals in research, then what other methods are available to ensure the safety of a product?

For a sample to be representative of the larger group (the population), it should be randomly chosen. Did the study in question use random sampling? Random sampling implies that everyone in the population has an equal chance to be in the sample. If we telephone a random sample of voters in our county, is that a random sample? No, because not everyone has a telephone and not everyone has an equal chance to answer the phone if we call at 11 a.m. In reality, random sampling is desirable but is usually impossible or unfeasible. However, scientists recognize that conducting a perfect study without sampling issues is also impossible or unfeasible. In recognition of this restraint, published studies will invariably include a section devoted to the studies’ limitations.

### 2. Why Did the People Participate in the Study?

Did the people volunteer or were they paid? If they were paid, and paid a lot, they might skew the results in favor of the experimenter’s hypothesis. Even if they were not paid, participants have been known to unconsciously bias an experiment in the experimenter’s favor. Why? Because sometimes we are just fond of some people—we’d like to help them out or we’d be embarrassed if they failed in front of us. We might not wish to share their

humiliation. Or it is easy to imagine some participants who would like to see the experimenter fail because they enjoy other people's suffering and humiliation. Ultimately, it is in the experiment's best interest if the experimenter's hypothesis is kept a secret from the participants. Furthermore, the experimenter should not know who received the experimental treatment, so that the experimenter isn't biased when he or she is assessing the results of the experimental treatment or product being tested. If the experimenter doesn't know who was in what group (but someone important does) and the participants do not know what is being tested, the study is said to be a **double-blind experiment**.

### 3. Was There a Control Group, and Did the Control Group Receive a Placebo?

Even if the participants were not paid and were blind to the experimenter's hypothesis, we sometimes wish to help or change a behavior so badly that we do change even if the product or treatment in question does not really work. This is known as the **placebo effect**. Placebo effects are well known throughout the scientific world, and they are real. In other words, some people really do get better when they are given fake substances or sugar pills. Some people really do get better when we shake rattles in front of their faces or wave our fingers back and forth or light incense and chant. Current scientific methods call for a control group when testing some drug, product, or new type of psychotherapy. Furthermore, the new drug or technique should be superior to the improvement we frequently see in the control group that receives only the placebo. Thus, the nose strips should have been tested against a group of participants who received some kind of placebo (e.g., a similar piece of tape on their noses), and both groups should have received the same instructions. Interestingly, the word *placebo* means "I shall please" in Latin. It is also the first word in Catholic evening prayers said for the dead. In about the 12th century, these prayers became known as "placebos." People began to hire professional mourners who would recite the placebos, and relatively quickly the word came to have an unflattering and pejorative meaning that has persisted to modern times (see Brown, 1998, and Frisaldi et al., 2017, for discussions of the placebo effect). It is important to note that the placebo effect appears to be strongest when the treatment outcome is rated by the patients, which suggests that much of the effect is simply the patient's belief that the treatment actually worked. The evidence for a placebo effect is much weaker or even absent when direct physiological evidence is required.

Interestingly, there is an opposite of the placebo effect, which is the **nocebo effect**, where patients think they are being harmed by a drug or treatment. The word *nocebo* means "I shall harm" in Latin, and the term was coined in 1961 to refer to the harmful or noxious effects of a placebo. It is thought that the nocebo effect is primarily of psychological origin; that is, the patient's negative expectations from a drug (even if inert) or treatment (even if a sham) cause harmful effects such as anxiety, nausea, pain, itching, and other somatic problems. One classic example of the nocebo effect is the first of the sensory deprivation experiments, where just prior to being sensorially deprived the participants were told that they might have weird thoughts or even hallucinate, and they did! One increasingly major problem related to the nocebo effect is that known or demonstrable positive effects of a drug or treatment might be avoided or shunned because of social media misconceptions or

falsehoods. Childhood measles is currently on the rise because of the fears of parents about getting their children vaccinated despite the measles (MMR) vaccine being empirically demonstrated to be completely safe and highly effective.

#### 4. How Many People Participated in the Study?

There is no firm set of rules of how many people should be studied. However, the number of people participating can have a profound effect on our conclusions. In surveys, for example, the more, the merrier. The larger the number of people sampled, the more likely that the sample will be representative of the population. Thus, to some extent, the size of our sample will be determined by the size of the population to which we hope to generalize. For example, if we are interested in a study of taxpaying U.S. citizens, then a sample size of 500 might still be considered a small sample if there are 100 million taxpayers. On the other hand, 500 might be considered an unnecessarily large sample if we are attempting to study the population of Colorado Buddhists.

In studies where we employ the scientific method (also known as a controlled experiment) using an experimental group and a control group, large sample sizes may actually be harmful to the truth. This occurs because as we increase the number of participants in each group, there is a peculiar statistical artifact that increases the chances we say the product works when it really does not. Increasing the sample sizes unnecessarily is called an **abuse of power**. **Power** in this context is defined as the ability to detect a real difference in the two groups due to the treatment. Rarely, if ever, will the experimental group be equal to the control group on a response or dependent variable even if the treatment does not work! This occurs because of simple chance differences between the groups. However, if an experimenter increases the sample sizes (to 100 or more in each group), then chance differences will be seen as real differences in most inferential statistical tests. The way to avoid an abuse of power is to perform a statistical test known as **power analysis**, which can help the experimenter choose an appropriate sample size. Power analyses are beyond the scope of this course, but consult an advanced text if you are interested.

There is also the opposite situation where a real difference between two groups might not be detected because of insufficient power or because too few participants are used in each group. In this situation, the treatment may actually work, but we might conclude that it does not because we have not used enough people in each group. Again, a power analysis should help in this situation. Most experimenters use a simple rule of thumb: There should be at least 10 participants in each group or the overall experiment should include at least 30 participants. Yet again, a power analysis is appropriate before the start of any study.

#### 5. How Were the Questions Worded to the Participants in the Study?

Remember, I mentioned earlier the “fact” that one in four college-age women has reported being raped. How did the experimenter actually define the word *rape*? In this case, the experimenter asked 6,159 college students, “Have you ever had sexual intercourse when you didn’t want to because a man gave you alcohol or drugs?” We might argue that this definition of “rape” is too broad. It is possible that because of a woman’s religious beliefs or family values,

she may be ambivalent about sex (some part of her didn't want to have sexual intercourse under ANY circumstances outside of marriage). The experimenter in this situation defended her broad definition of rape by saying that rape has so many different meanings. She took the provocative position that if a student's report met the experimenter's definition of rape, then the student was raped and "whether she realizes it or not is irrelevant." Drugging anyone to have sex without their knowledge or consent is rape, but creating one's own personal definition of rape and then developing "rape facts" is a questionable practice.

How about the "fact" stated on milk cartons that more than 1 million children have been abducted? Here, the debatable issue is the definition of *abducted*. Some missing children organizations define *abducted* as any child not living with his or her legal custodian parent. Thus, it might be argued that some of these abducted children are not really missing if they are living with one of their biological parents (but just not the parent who has legal custody). While this may still be a serious matter, it might not reach the same level of importance as a child who has been taken by someone unknown to the family and the whereabouts of the child are a complete mystery. If we consult current police and Federal Bureau of Investigation (FBI) records, we would find that annually there are about 115 children who meet the latter criterion. The problem with the broad definition of abduction is that if too many children meet the definition of abducted, then we will not have the time and resources to hunt for the truly missing children.

The sexual abuse of children is certainly an important issue and highly worthy of our attention. It has been reported that 25% of all adult males have reported the tendency to sexually abuse children. This statistic might be highly alarming except that the men surveyed were asked, "If you knew you would never get caught or you knew that no one would ever find out, have you ever had the fantasy of sexually fondling a girl under the age of 14?" You might still consider the 25% of men who said yes to this question to be highly despicable, perverts, or sick, yet it could be argued that the question had the word *fantasy* in it. By definition, a fantasy is something not occurring in reality. It seems unfair to ask someone about a fantasy and then assume that it reflects a true tendency in reality. Furthermore, perhaps the men were recalling an old fantasy, one they had when they were under the age of 14 themselves. What would have happened if we had asked the question like this: "Would you ever consider having sexual intercourse with a child, there was a very high probability that you could get caught, you knew you would receive the death penalty, your mother would be the first person to know about it, and she would then commit suicide because you broke her heart?"

A related issue in measurement is how clearly the experimenter has defined the problem he or she is trying to study or solve. Have you heard the "fact" that one in three children goes to bed hungry? How clearly did the experimenter define what he or she meant by *hungry*? Because I go to bed late every night, I nearly always go to bed hungry even after a big meal earlier in the evening. How did the experimenter define the word *hungry*? Did the experimenter ask all the children in the world whether they were hungry before they went to bed? I doubt it. And what if the experimenter defined hungry by asking a child, "Would you like anything else to eat, like a cookie or candy, before you go to sleep?" It is easy to imagine that one definition of hungry is whether a child asks for more to eat, but then, many overweight children would also ask for more to eat and meet the definition of going to bed hungry.



Perhaps you can perceive the monumental task involved in clearly defining any variable, even words as seemingly simple as *hungry*. A child who is really starving is a terrible reality. However, blatant lying, conscious exaggeration, and abusing statistical methods are also abhorrent, even in the cause of ending world hunger.

## 6. Was Causation Assumed From a Correlational Study?

Probably the single most pervasive abuse of statistics occurs when people infer causation from a correlational study. A correlational study is one in which two factors or variables are measured together in a setting. In correlational studies, there is no random assignment of the participants to two groups, there is no treatment applied, and there is no control group. In a correlational study, if the two variables do appear to be related to each other, many people unfortunately assume that one of these variables *causes* the other. While it is possible that there is a causative relationship, a correlational study should never be interpreted as evidence for causation. For example, a correlational study of marriage and drug use found that drug use declined after marriage. The headlines in the newspapers read, “Responsibilities of marriage and family often curb drug use.” One of the problems in this interpretation is that the study was a correlative one, and we must always guard against the inference of causation. It would be equally wrong, but no less absurd, to assume that drug use curbs marriage. A second problem with correlational studies is that many other factors or variables are also operating and may contribute as causes. If a scientist does not measure these other variables, then the scientist cannot assume that it is only a single variable such as marriage that is curbing drug use. In fact, an overwhelming majority of behaviors have multiple and complex causes, and to assume that a single variable such as alcohol use causes crime is not only statistically wrong but wrong in theory as well.

## 7. Who Paid for the Study?

In an ideal world, the investigators in a study should not have any financial ties to or interest in the outcome of a study. However, a few years ago, a study reported that balding men were more likely to have heart attacks than nonbalding men. Who financed the research? A drug company that manufactures a popular hair-growing product! While this financial interest in the study does not automatically invalidate the study, it does tend to cast suspicion on the outcome of the study or the motives of the investigators.

A major controversy began when a study linked a high blood pressure medicine to an increased risk of heart attacks. It was noted that 70 scientists, doctors, and researchers came to the defense of this particular blood pressure medicine in journal articles, reports, and medical publication commentaries. When the 70 were questioned about their defense, 67 of them admitted that they had financial relationships with the manufacturer of the drug, which included travel grants, public speaking fees, research and educational grants, and consulting fees and contracts.

It has been suggested that scientists have been naive about the extent to which financial relationships may affect research. Again, in an ideal world, there should be no financial complications between the investigators and the outcomes of their studies. However, short of the ideal, perhaps scientists should report any appearances of a conflict of interest at the

outset of their studies, either to the journal editors to which the studies are submitted or in a footnote in the published versions of the studies. In this way, the readers of the articles can determine for themselves to what extent they feel the studies may be biased. For example, a prominent medical journal published a meta-analysis (a study summarizing many other studies) of hair-growing medications, and the author concluded that hair-growing medications do indeed work. Only in a subsequent issue of the journal was it revealed that the author had been supported for years by grants from companies that manufacture hair-growing medications.

## 8. Was the Study Published in a Peer-Reviewed Journal?

Nearly all scientific journals use a peer-review procedure where any article submitted to an editor is sent out to other scientists for their opinions on whether the study should be accepted for publication. Some journals reject as many as 90% of all the articles submitted, while some of the “easier” journals may reject only 50%. Even if an article is accepted for publication (a process that could take years), it is rarely accepted without any changes. Typically, a journal editor will summarize the criticisms of the peer reviewers and forward them to the author. The peer reviewers may request additional information about the literature surrounding the issue, more information about the participants, more information about the tests employed or the procedures, additional statistical analyses, or modifications of the conclusions or implications of the study. The first study that ever produced evidence for cold fusion, a process whereby nearly limitless energy could be obtained from a special form of water ( $^2\text{H}_2\text{O}$ ), was not published in a peer-reviewed journal. It was presented by teleconference! Thus, the specific procedures were not revealed in detail, and the study was not subject to any acceptable form of peer review. Is it any wonder that the issue of whether those scientists actually created cold fusion has been vehemently questioned and rejected? It is also important to note that personal or academic politics may play a role in whether a journal article is published. One interesting story of publishing politics occurred when a second-year graduate student in biochemistry determined that half of the matter in the universe was going forward in time but the other was going backward in time. He published his idea in *Nature*, a premier science journal. When physicists finally had time to evaluate his hypothesis, they could find no good evidence for it. The graduate student went back to his specialty, and he discovered how a short piece of DNA could be used to sequence and create much longer pieces of DNA. He called it *polymerase chain reaction* (PCR), and ultimately PCR has been used in diagnosing genetic diseases, solving crimes, tracing human evolution by extracting DNA from bones, and developing treatments and cures for diseases. In 1993, Kary Mullis, the former biochemistry student, was awarded the Nobel Prize in Chemistry. So, what happened when he tried to publish his PCR idea? Because his name was “mud” in the eyes of so many scientists (especially physicists), he could not publish it in any conventional journal! Luckily for Mullis, a guest editor for the semi-obscure journal *Methods in Enzymology* invited him to publish his article in 1987. Within 3 years there were more than 1,000 citations and publications related to PCR, and just 6 years later Mullis received the Nobel Prize.

A relatively new problem with journal article publications is the explosion of predatory journals. Predatory journals are completely for-profit operations (not that there is anything inherently wrong with profits) that will publish virtually any article for money. Predatory journals might not use any review system for submitted articles whatsoever. These journals

usually publish their issues only online (saving money) and call them “open access.” The latter term means that the articles will be freely available on online to readers or institutions. There are legitimate open access journals; however, authors must pay to have their journal articles published regardless of whether the journals are legitimate or predatory. Some of the signs of a predatory journal are that articles are submitted to an email address (rather than an online submission service), the email invitation contains odd uses of English or poor grammar, and the invitation contains the title of a previously published article by the author being solicited but in a different font and size than the email invitation.

What about books? Do they undergo peer review? Typically not. It does vary from publisher to publisher, but there are some publishing companies where anything you write can be published in a book (for a fee). Many scientists will present the results of their research at national or international conventions. Are convention presentations peer-reviewed? Most convention presentations do have some form of peer review, but the standards for convention presentations are much more lax than those for journal articles and the acceptance rates can vary from 50% to 100%. Thus, published journal articles meet the highest standards of scientific scrutiny, convention presentations are a distant second, and some books will meet absolutely no scientific standards. We must consider paid advertisements in newspapers, in magazines, and on the Internet and infomercials on television as the least likely venues for truth or justice.

What about information on the Internet? The Internet is a dicey proposition. There are very legitimate sources of information on the Internet, and reputable scientific journals are increasingly making their studies available on the Internet. There are also quite reputable journals that exclusively publish on the Internet (often called open access journals). However, anyone can post anything he or she likes on the Internet. Thus, one must be able to discern between legitimate Internet sources and illegitimate or questionable ones. I have personally noticed many scientific and pseudoscientific articles posted on the Internet by their authors. The veracity of such studies is completely unknown. These studies may be of value, or they may have been rejected by legitimate journals and may be completely fraudulent. The moral of the Internet story is, let the surfer beware!

## ON MAKING SAMPLES REPRESENTATIVE OF THE POPULATION

---

To ensure that the hypotheses we make from the sample about the population are good ones, remember the two requirements: The sample should be randomly drawn from the population, and the sample should be relatively large. However, these two guidelines do not guarantee that a sample will be representative of the population. There will always be the possibility that a sample, although large and randomly drawn, may still lack some important characteristics that are present in the population. Perhaps this is why single experiments can rarely ever prove a hypothesis or a theory. It takes repeated experiments (replication) by different experimenters before scientists begin to accept that a particular hypothesis may or may not be useful in explaining some phenomenon. In fact, I tell my students that the comedian Father Guido Sarducci (of *Saturday Night Live* fame) created

a five-minute university degree because that is all students will remember 5 years after they graduate. Thus, if my students will remember only one word from my entire statistics course, then I ask them to remember the importance of the word *replication*. Before trusting in a finding, we should see whether the finding has been replicated over a series of studies by different scientists. I have asked my colleagues this same question (i.e., which one word to remember), and one said *measurement*, which I find to be a profound answer because the essence of how data are gathered is absolutely critical to the decisions made in a study. For example, as noted previously, how questions are worded to participants may invoke very high levels of endorsement or very low levels of endorsement. Furthermore, the measurement scale itself can be sensitive or insensitive to measuring a question. For example, an answer scale containing only the responses *true* and *false* may gloss over many of the nuances of behavior (or the variability of a complex behavior) that exist between these two responses.

## EXPERIMENTAL DESIGN AND STATISTICAL ANALYSIS AS CONTROLS

---

To obtain useful and meaningful information from both descriptive and inferential statistics, it is necessary for the data to be collected based on an explicit and appropriate design called the **experimental design**. Essentially, an experimental design is a blueprint for how the data collection will be conducted. At the very beginning of this data collection or experiment, there are two forms of control or power that we have at our disposal: *experimental control* and *statistical control*. Experimental control concerns the design of the experiment—for example, how many people will be used, how many groups there will be, what kind of people will be used, how they will be tested or measured, and so on. The second form of control or power is statistical control, and this is employed after the data collection or the experiment is complete. It is important to remember that although statistical techniques are powerful tools for describing data, experimental control is the more important of the two. One reason for this is that statistics can become meaningless and will communicate no knowledge whatsoever unless we adhere to some basic experimental principles.

For example, a recent TV poll asked viewers to call two different phone numbers. The first phone number was to be called if the viewer thought that a particular university football team should be ranked as the best team in the country, and the second number was to be called if the viewer thought that the same team should not be top ranked. This plan for the experiment is called the experimental design. The statistical analysis consisted of comparing the number of calls for the two phone numbers and seeing which was higher. In this case, the experimental design is so poor that the statistical analysis is rendered meaningless. One of the many problems with this design is that the sample of TV viewers was not random. Therefore, any conclusions that the results reflect popular opinion will be misleading. Another problem is that a viewer could call in more than once and, thus, unduly influence the outcome. ESPN.com once conducted an online survey about which team had the best sports uniform. The Denver Broncos rallied to win the poll with 137,257 votes compared with the University of Michigan's 88,743 votes. On further review, it was revealed that

71,465 Bronco votes came from a single Internet address. Thus, the experimental design is so poor that it is not meaningful to make any conclusions whatsoever about which team has the best sports uniform.

The statistical analyses cannot easily take into account problems like multiple voting. These uncontrolled problem variables in experimental designs are called **confounding variables**. In summary, the major experimental design problem with TV, Internet, and call-in polls is that some people are allowed to participate more than once, and the sample of viewers will not be random. As noted earlier, such studies result in futile, meaningless, and highly untrustworthy data. Often, these polls will cost the viewer some amount of money to participate, and the true purpose of the poll is solely to generate money, not to determine the voters' opinions. However, the experimental design is often so poor that the only truly valid conclusion would be that 137,257 is a larger number than 88,743. It cannot be concluded that more people thought a particular team had a better uniform than another. Any conclusion that gave the implication that a particular team's uniform was the choice of most people would be completely invalid. Subsequently, in the interest of "fairness," the ESPN.com pollsters disqualified all votes for the Broncos and designated the University of Michigan as the winner. In this case, ESPN.com is making a post hoc (after-the-fact) adjustment to the experimental design, further calling its results into question. Again, the experimental design is so poor that no amount of statistical power or analysis can correct the problem. Why do unscientific and unethical telephone surveys persist on TV, on the Internet, and in newspapers? Money! These polls, although completely unscientific, have entertainment value, and entertainment value means money. Also, at 50 cents or more per vote in some telephone polls, the owners of these polls are making a lot of money.

In summary, poor TV, Internet, newspaper, and other surveys demonstrate how you must think carefully about designing your experiment in terms of your statistical analysis. Although statistical control is powerful, it cannot save or make interpretable a shoddy, poorly designed experiment.

## THE LANGUAGE OF STATISTICS

---

Learning about statistics is much like learning a new language. Many of the terms used in statistics may be completely new to you such as *bimodal*. Also, some words you already know will be used in new combinations such as *standard deviation*. Finally, some words will have a new meaning when used in the context of statistics. For example, the word **significance** has a different meaning in a statistical context. When used outside of statistics, significance is used as a value judgment. If something is said to be significant, then we usually mean that it is important or of consequence. The opposite of significant is usually insignificant. However, in statistics, the word *significance* has a different meaning. Significance refers to an effect that has occurred that is not likely due to chance. In statistics, the opposite of significant is nonsignificant, and this means that an effect is likely due to chance. It will be important for you to remember that the word *insignificant* is a value judgment, and it typically has no place in your statistical language.

## ON CONDUCTING SCIENTIFIC EXPERIMENTS

---

Scientific experiments are generally performed to test some hypothesis. Remember, a theory is some grand idea about the way nature (at any level) seems to work. A theory generally should make some predictions in the forms of hypotheses. The hypotheses are then tested through scientific experiments. The typical experiment is a two-group experimental design where there is an experimental group and a control group. Most typically, the experimental group gets or receives some special treatment. The control group usually does not get this treatment. Then, the two groups are measured in some way. These two levels of treatment are aspects of the **independent variable**. The independent variable is the variable that the experimenter manipulates. The experimenter wishes to see whether the independent variable affects the dependent variable. The test that the two groups are measured on is called the **dependent variable** or response variable. Sometimes, it is very important that the participants in each group do not know whether they are receiving some special treatment. If this experiment was to determine whether Vitamin C affects colds, then the control group participants should also be given the same treatment as the experimental group, with the exception of actually receiving Vitamin C. Typically, control groups receive a placebo, which is a pill indistinguishable from the pills that contain the Vitamin C that the experimental groups receive. Remember, the participants should not be aware of who is actually receiving Vitamin C and who is getting the placebo. Also, the experimenter or statistician, who will evaluate the two groups after treatment, should not know who is receiving Vitamin C or the placebo until after the groups have been evaluated because it is a well-known phenomenon that experimenters can have an unconscious role in making their experiments come out in favor of their hypotheses. As noted previously, the procedure is said to be a double-blind experiment. An example of this unconscious influence occurred in a proposed new treatment of childhood autism (an early and severe psychotic disturbance). The treatment was called facilitated communication. The therapists (“facilitators”) would help the highly uncommunicative autistic children form their answers to questions on a keyboard. An experiment was designed to control for the experimenter’s influence, whereby the questions were either the same or different for the autistic children and the facilitators. Despite having received millions of dollars in grants to develop and train people to become facilitators, sadly, in no case could the children answer a question correctly when the questions were different between the children and their facilitators. It appeared that the facilitators were simply consciously or unconsciously answering the questions correctly on behalf of the autistic children.

## THE DEPENDENT VARIABLE AND MEASUREMENT

---

Remember that the test or whatever we use to measure the participants is called the dependent variable. In the Vitamin C experiments, there are a number of different measures that we might use as dependent variables. Some experimenters have focused on the number of new colds in a 1-year period. We could also measure the number of days sick with a cold. The choice of the dependent variable could be very important with regard to the



conclusions of our study. For example, some studies have shown Vitamin C to be ineffective in preventing colds; however, there is some evidence that Vitamin C may reduce the severity or duration of colds.

## OPERATIONAL DEFINITIONS

---

If we are going to count the number of colds each participant gets in a 1-year period, then each participant would somehow be checked for how many colds he or she got. Is there some subjectivity in how you decide whether a person has a cold? Of course there is! To minimize this subjectivity, perhaps the same person (e.g., a medical doctor) would rate all the participants to determine the presence or absence of a cold. However, doctors might vary among themselves about whether a person actually has a cold. Perhaps the diagnoses might vary as a function of what each doctor defines as a cold. For example, one doctor's definition of a cold might be whether the patient claims to have a cold. Another doctor's definition might require the presence of congested sinuses and elevated white blood cells. There is usually no single correct definition; however, whatever definition you do choose should be stated clearly according to some criteria. When you clearly list in your study the criteria that all the participants met, those criteria are said to be your **operational definition**. Thus, in the above example, the operational definition of a cold might be the number of self-reported colds by a participant during a 1-year period, or the operational definition might be the number of colds a participant had during a 1-year period according to a doctor's examination.

## MEASUREMENT ERROR

---

No dependent variable will ever be perfectly measured for each participant. The variation in the dependent variable that is not due to the independent variable is ironically called **measurement error** in statistics. The individual participants contribute the most error to the experiment and to the dependent variable. Other sources of error may come from subtle variations in the testing conditions such as time of day, temperature, extraneous noise levels, humidity, equipment problems, and a plethora of other minor factors. In our Vitamin C study, it is possible that some participants never get colds anyway, and some may frequently get sick. The experimental error associated with the participants is hopefully balanced out when we randomly assign the participants to the two conditions. After all, we could not, for example, assign all our friends to the experimental group and assign all others to the control group. Random assignment of the participants to the groups helps balance out the effects of the participants as a source of error. Using a large number of participants in each group also helps balance out this source of error.

When you hear about or read about experiments or surveys in any form of media, remember to look at the dependent variable closely. Sometimes, the entire interpretation of an experiment may hinge on an adequate dependent variable. How about the results of a study that some city, say "City X," is the best place to live. One critical variable in this study is how they measure "best

city.” Is it in terms of the scenery? What happens if you like or hate ocean views? Typically, these studies try to be “scientific” so that their judges rate the cities on a number of different variables. However, many of these questions may relate to economic growth. Therefore, if a city is not growing rapidly, it might not fare well in the rating system. Would you consider a fast-growing city the best place in which to live? It becomes obvious there are many intangibles in developing a survey to rate the best place in which to live, and any survey that claims to be able to rate the best city is highly questionable. Be sure to examine the operational definitions of critical variables in experimental studies. In this latter example, it would be interesting to see the specific operational definition of “best city.” If the authors do not provide an operational definition, then the study is virtually uninterpretable. Importantly, would you move to Detroit (or Miami) if it was ranked as the best city to live in the United States?

## MEASUREMENT SCALES: THE DIFFERENCE BETWEEN CONTINUOUS AND DISCRETE VARIABLES

---

A variable is typically anything that can change in value, and a variable usually takes on some numeric value. Statisticians most commonly speak of continuous and discrete variables. A **continuous variable** can be measured along a line scale that varies from a smaller number to a larger number. For example, a continuous variable would be the time in seconds in an experiment for a participant to complete a task. A **discrete variable**, for statisticians, typically means that the values are unique and are qualitatively distinct from one another. Speculations are not made between the values. For example, gender could be considered a discrete variable. If the category “man” is assigned a value of 0, “woman” is assigned a value of 1, and “non-binary” is assigned a value of 2, then interpretations will not be appropriate for values between 0 and 1 and between 1 and 2. It should also be noted that when there are only two values of a discrete variable, it is also referred to as a **dichotomous variable**.

## TYPES OF MEASUREMENT SCALES

---

There are various kinds of measurement scales. In any statistical analysis, the type of scale must be identified first so that the appropriate statistical test can be chosen. Some measurement scales supply minimal information about their respective participants; thus, the statistical analysis may be limited. Other types of scales supply a great deal of information; consequently, the statistical possibilities are enhanced. The following are four common types of measurement scales.

### Nominal Scales

**Nominal scales** assign people or objects to qualitatively different categories. Nominal scales are also referred to as *categorical scales* or *qualitative scales*. One example is the assignment of people to categories of self-reported gender (e.g., man, woman, non-binary, other). Thus, when measured on a nominal scale, all the people or objects in a category are the

same on some particular value. Notice also that the people in the category are all considered equal with respect to that value. For example, all the individuals who self-report their gender as “man” are considered equal with respect to the category. Thus, membership in a category does not imply magnitude. Some men in the category are not considered more “manly” than other men in the same category. The frequency of each category can be analyzed, however. For example, it might be noted that 37 men, 44 women, and 3 non-binaries participated in a study. Another example of a nominal scale would be a survey question that required the answer *yes* or *no* or required the answer *yes*, *no*, or *undecided*.

It is also important to note that there are no intermittent values possible on a nominal scale. This means that if, for statistical purposes, a value of 0 is assigned to the man category and 1 is assigned to the woman category, there are no values allowed or assumed between 0 and 1. It is, of course, physiologically possible to have a person who has a mixed biological gender. It is also psychologically possible to have a mixed gender identity. However, it is not possible to have intermittent values if gender is measured on a nominal scale.

### Ordinal Scales

**Ordinal scales** involve ranking people on some variable. The person or object that has the highest value on the variable is ranked “number 1” and so on. The ordinal scale, therefore, requires classification (how much of the value does an individual have) and ranking (where the individual stands relative to all other members of the group).

The ordinal scale has one major limitation—the differences between rankings may appear equal when in reality it is known that they are not. For example, if we rank athletes after a race, the difference in times between the first- and second-place athletes may be huge, while the difference in times between second and third place may be very small. Nevertheless, with an ordinal scale, the appearance is given that the differences among the first, second, and third rankings are all equal.

This limitation is not necessarily bad. Ordinal scales may be sufficient if it is known that the classification variable possesses some arbitrariness. Thus, it may be useful to rank all 50 participants in an experiment with respect to some variable and then compare the top 5 ranked participants with the bottom 5 ranked participants on some other variable. This experimental design is sometimes called an **extreme groups design**.

It has also been assumed in the previous discussion that the people or objects ranked received their respective rankings by a single classification variable. Of course, it is also possible to rank people or objects through more than one classification variable or even nebulous or hazy criteria. For example, movie critics are frequently known for presenting their rankings for the 10 best movies or the 10 worst movies. What are their classification variables? In this case, there are probably many classification or criteria variables, and some of them may even be inexplicable or unconscious. At the very least, this may mean that some types of ordinal scales may be suspect.

### Interval Scales

**Interval scales** probably receive the most statistical attention in all sciences. Interval scales give information about people or objects with respect to their ranking on some classification

variable, and interpretations can be made with regard to how far apart the people or objects are on the variable.

With an interval scale, it is assumed that the difference of a particular size is the same along all points of the scale. For example, on an attitude survey, the difference between the scores of 40 and 41 is the same as the difference between the scores of 10 and 11. On an interval scale, it would also be assumed that a difference of 10 points between two scores would represent the same subjective difference anywhere along the scale. For example, it is assumed that the difference between scores of 40 and 50 on an attitude survey is the same degree of subjective magnitude as between 5 and 15. Obviously, this is a crucial and difficult assumption to meet on any interval scale. Nevertheless, most measurement scales, particularly in the social sciences, are assumed to be interval scales. Examples of interval scales are scores on intelligence tests, scores on attitude surveys, and scores on most personality and psychopathology tests. However, there is some debate about whether these tests should be considered interval scales because it is questionable whether these scales meet the equal magnitude assumption. In practice, however, most statisticians agree that the purported interval scales at least have the property of ordinal scales, and they may at least approximate interval scales. In a more practical evaluation, it is well documented that these purported interval scales yield a host of interpretable findings.

### Ratio Scales

**Ratio scales** have the properties of interval scales and also have some rational zero point. This means that a zero point on a ratio scale has some conceptual meaning. It must be noted that in most social disciplines such as psychology, sociology, and education, ratio scales are rarely used. In economics and business, an income variable could be measured on a ratio scale because it makes sense to talk of “zero” income, while it makes no sense to talk about “zero” intelligence.

Although ratio scales may be thought of as the most sophisticated of the types of scales, they are not necessary to conduct research. Most types of statistical analyses and tests are designed to be used with interval scales. Indeed, one of the primary purposes of the science of statistics is to organize and understand groups of numbers and to make inferences about the nature of the world. Its purpose is not to create the perfect measuring device.

## ROUNDING NUMBERS AND ROUNDING ERROR

---

The general rule for rounding decimal places is if the number to be dropped is 5 or greater, then the remaining number is rounded up. If the number to be dropped is less than 5, then the remaining number is left unchanged. For example, when rounding to the nearest tenth (or rounding to one decimal place), look at just the tenth (first decimal place) and hundredth (second decimal place) decimal places and ignore any places beyond

(e.g., thousandths, which is the third decimal place). So, examine the reasoning underlying the following examples rounded to the tenth or first decimal place:

10.977 would round to 11.0

125.631 would round to 125.6

100.059 would round to 100.1

6.555 would round to 6.6

6.5499 would round to 6.5

Examine the reasoning underlying the next set of examples rounded to the hundredth or second decimal place:

10.977 would round to 10.98

125.631 would round to 125.63

100.059 would round to 100.06

6.555 would round to 6.56

6.5499 would round to 6.55

Examine the reasoning underlying the next set of examples rounded to the thousandth or third decimal place:

10.9774 would round to 10.977

125.6316 would round to 125.632

100.0593 would round to 100.059

6.5557 would round to 6.556

6.54999 would round to 6.550

Anytime a number is rounded off, rounding error is introduced. Income tax forms often ask for rounding to whole numbers so that only the first decimal place would be used to round. For example, when rounding to whole numbers, the following would occur:

\$10.98 would round to \$11

\$125.63 would round to \$126

\$100.06 would round to \$100

\$6.55 would round to \$7

\$6.49 would round to \$6

## Percentages

The word *percentage* is derived from the Latin *per centum*, which means per hundred, so when data are reported as a percentage it is interpreted as parts per hundred. It is important to note at the outset of this discussion that reporting data by percentages is common but not always appropriate. Its inappropriateness becomes evident when any behavior or phenomenon is rare and occurs less than one in a hundred. For example, in the state of Colorado, the definition of driving under the influence (DUI) of alcohol is .08%. Many people wrongly interpret this as 8% or 8 parts of alcohol per 100 parts of blood. However, the *point zero eight percent* actually means 8 parts of alcohol per 10,000 parts of blood. If someone's blood alcohol level was actually 8%, they would be dead in an instant because death from too much alcohol begins to occur at about .35% or 35 parts per 10,000. Thus, the misinterpretation of the DUI definition occurs because the blood alcohol level cannot be measured in a strict whole number percentage. That is, .08% converts to a proportion of .0008, and rounding that number to two decimal places, would mean that the result would be 0%. It is also important to note that percentages are most often reported as two whole numbers with a percentage sign.

Converting proportions to percentages is accomplished by moving the decimal place over two decimal places, examining the third decimal place, and rounding the second decimal place based on the first. Most often, converting proportions to percentages begins with a proportion to three decimal places. In SPSS, the critical output is usually reported to three decimal places. However, the software does round to three decimal places based on the fourth decimal place.

Examine the reasoning underlying this set of examples. The conversion will begin by rounding proportions to three decimal places:

.109 converts to 10.9%, which rounds to 11%

.1256 rounds to .126, which converts to 12.6% or 13%

.1000593 rounds to .100, which converts to 10.0% or 10%

.65557 rounds to .656, which converts to 65.6% or 66%

.654999 rounds to .655, which converts to 65.5% or 66%

## STATISTICAL SYMBOLS

---

Statisticians use symbols to represent various concepts. It will also be important for you to learn most of the common symbols they use. This should not be a very difficult task because the symbols are used repeatedly, so you will have many opportunities to commit them to memory. Two of the most common symbols are  $\bar{x}$  (pronounced *x* bar), which stands for the mean or average of a sample of numbers, and the Greek uppercase letter  $\Sigma$  (**Sigma**), which is used to indicate the sum of a group of numbers. Let us practice with these two symbols. Find the average or  $\bar{x}$  for the following group of numbers:

5, 9, 10



Intuitively, you probably know how to add the numbers up and divide by 3. Therefore, the arithmetic mean (which is commonly used instead of the word *average*) is 8. However, let us use the word *set* instead of group. When you added each of the numbers together, you were taking the sum. Each number in the set of numbers can be represented by  $x$ . Therefore, in statistics, the formula for the mean would look like this:

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

where  $\bar{x}$  = the mean or average of the set of numbers,  $\sum x$  = the sum of all the numbers in the set, and  $N$  = how many numbers are in the set:

$$\bar{x} = \frac{5+9+10}{3}$$

$$\bar{x} = \frac{24}{3}$$

$$\bar{x} = 8$$

Another common symbol is  $\sum x^2$ , which indicates that each number in the set is individually squared and then the products are added together. So, for the original set of numbers 5, 9, 10, obtain  $\sum x^2$ :

$$\sum x^2 = 5^2 + 9^2 + 10^2$$

$$\sum x^2 = 25 + 81 + 100$$

$$\sum x^2 = 206$$

What does  $(\sum x)^2$  indicate? Remember, in mathematics, you must simplify what is in the parentheses first before you begin any other operations. So, for this same set of numbers, add the numbers up first and then square this value:

$$(\sum x)^2 = (5+9+10)^2$$

$$(\sum x)^2 = (24)^2$$

$$(\sum x)^2 = 576$$

Does  $\sum x^2$  equal  $(\sum x)^2$ ? By using the same numbers, you can verify that they are not equal!  $\sum x^2 = 5^2 + 9^2 + 10^2$  or  $\sum x^2 = 25 + 81 + 100$ , which means  $\sum x^2 = 206$ . On the other hand,  $(\sum x)^2 = (5 + 9 + 10)^2$  or  $(24)^2$ , which means  $(\sum x)^2 = 576$ . Thus, you can see the resulting values are very different, and they cannot be used interchangeably. For those of you less confident in your mathematical skills, it may be useful to note the following formulas carefully because you will need them later in this book.

## SUMMARY

For the set of numbers 5, 9, 10:

Sum of the set	$\Sigma x = 5 + 9 + 10$
	$\Sigma x = 24$
Mean	$\bar{x} = \frac{\Sigma x}{N}$
	$\bar{x} = \frac{5+9+10}{3}$
	$\bar{x} = 8$
Sum of each number $x$ in the set squared	$\Sigma x^2 = 5^2 + 9^2 + 10^2$
	$\Sigma x^2 = 25 + 81 + 100$
	$\Sigma x^2 = 206$
Squared sum of the set	$(\Sigma x)^2 = (5+9+10)^2$
	$(\Sigma x)^2 = (24)^2$
	$(\Sigma x)^2 = 576$

Sometimes the symbol  $x_i$  ( $x$  sub  $i$ ) appears.  $x_i$  refers to the  $i$ th number in the set; thus,  $x_1$  is the first number in the set. For the previous set of numbers (5, 9, 10),  $x_1 = 5$ ,  $x_2 = 9$ , and  $x_3 = 10$ . Sometimes you will see  $x_n$  ( $x$  sub  $n$ ), which stands for the  $n$ th number in the set or the last number in the set. Almost as commonly, you will see  $x_k$  ( $x$  sub  $k$ ), which also stands for the last number in the set. In this book, a shorthand notation is used for  $\Sigma x$ .

In reality, the formal notation is  $\sum_{i=1}^k x_i$ , where  $i = 1$  below the  $\Sigma$  and  $k$  above the  $\Sigma$  form a programming loop for the  $x_i$ . Thus,  $x_i$  starts with  $x_1$  and moves progressively through the second number in the set, third number in the set, and so on through  $x_k$  (or the last number in the set).

## HISTORY TRIVIA

## Achenwall to Nightingale

**Florence Nightingale (1820–1920)**

© iStockphoto/TonyBaggett

Source: Public domain.

Although the history of mathematics is ancient, the science of statistics has a much more recent history. It is claimed that the first use of the word *statistics* was by German Professor Gottfried Achenwall (1719–1792) in 1749, and he implied that statistics meant the use of mathematics in the service of the nation. For Achenwall, statistics was another word for state or political arithmetic, which might include counting the population, the size of the army, or the amount and rate of taxation.

One of the earliest uses of the word *statistics* in the English language was by John Sinclair (1754–1835), a Scottish lawyer and writer who published a 21-volume survey of Scotland from 1791 to 1799. In its preface, Sinclair acknowledged the political nature of statistics in Germany; however, he emphasized the use of statistics as a method of inquiry into the

psychological state of a country, for example, to ascertain the level of happiness of its inhabitants and a means by which its future improvements could be made. He specifically thought a new word from the German language, *statistics*, might attract greater public attention.

Concurrent with these developments, mathematicians were making contributions in the area of probability theory that would ultimately form the foundations of inferential statistics. Pierre Simon, the Marquis de LaPlace (1749–1827), better known as “LaPlace,” was a French mathematician who contributed much to early probability theory. Karl Friedrich Gauss (1755–1855) was a German mathematician and astronomer who also made valuable contributions to the foundations of both descriptive and inferential statistics.

During the middle to late 1800s, the application of statistics to social problems became more prominent. Adolph Quetelet (1796–1874), a Belgian mathematician and astronomer, extended some of Gauss’s ideas to the analysis of crime in society. Francis Galton (1822–1911), cousin of Charles Darwin, was an English explorer, meteorologist, and scientist. His book, *Natural Inheritance* (published in 1889), has been recognized as the start of the first great wave of modern statistics. He also profoundly influenced another English person, Karl Pearson (1857–1936), who has been called the founder of the modern science of statistics. Pearson’s intellectually provocative book, *The Grammar of Science*, was published in 1892. In it, he stressed the importance of the scientific method to society and knowledge, and he believed statistical procedures were fundamental to the scientific method.

*(Continued)*

(Continued)

Florence Nightingale (1820–1910) was a contemporary of Pearson and a friend of Galton. She is typically remembered as a nurse and hospital reformer. However, she might just as well be remembered as the mother of descriptive statistics. She trained to become a nurse during the 1850s, and it was her strict observance to sanitation in hospitals that dropped death rates dramatically. She performed not only her nursing duties but also administrative duties. She also founded a school for the training of nurses and established nursing homes in England.

In the course of her administrative work, Nightingale developed a uniform procedure for hospitals to report statistical information about their patients. She is credited with developing the pie chart, which represents portions of the whole as pieces of a pie. She also argued to get statistics in the higher education curriculum. She had suggested to Galton that a professorship be established for the statistical investigation of societal problems, and she pointed out issues to Galton that should be studied under the auspices of this professorship. These issues included crime, education, health, and social services. In 1911, the University of London finally established a Department of Applied Statistics and appointed Pearson its head with the title Galton Professor.

Through the graphic representation of data, called a frequency histogram, Nightingale convinced the queen and prime minister of England to establish a commission on the health and the care of the British army. She did this by showing clearly with graphs that the rate of deaths for the military while at home in England was almost double the rate of deaths for equivalent nonmilitary English males.

One of her biographers argued that Nightingale's interest in statistics transcended her interest in health care and was closely related to her strong religious convictions. She felt the laws governing social phenomena were also laws of moral progress; therefore, they were God's laws and could be revealed by the use of statistics.

## Key Terms

Abuse of power 13

Confounding variable 19

Continuous variable 22

Dependent variable 20

Descriptive statistics 7

Dichotomous variable 22

Discrete variable 22

Double-blind experiment 12

Experimental design 18

Extreme groups design 23

Hypothesis 4

Independent variable 20

Inferential statistics 8

Interval scale 23

Measurement error 21

Nocebo effect 12

Nominal scale 22

Operational definition 21

Ordinal scale 23

Parameters 8

Placebo effect 12

Population 8

Power 13

Power analysis 13

Ratio scale 24

Replication 5

Sample 8

Sigma  $\sum_{i=1}^k x_i$  26

Significance 19

Stratified sample 8

Theory 6

## Chapter 1 Practice Problems

1. Distinguish between a theory and a hypothesis.
2. Name two requirements to help make samples representative of the population.
3. Define *dependent variable* and *independent variable*. Give examples of each and explain how they relate.
4. Based on the following list, identify whether each item represents experimental control or statistical control.
  - a. adding an additional group to an experiment
  - b. increasing the number of participants in the study
  - c. analyzing the data with a different statistical test
  - d. using a double-blind experiment
5. Which of the following are continuous variables?
  - a. stock price
  - b. eye color
  - c. country of origin
  - d. gender
  - e. blood pressure
  - f. height
6. Identify the type of measurement scale that each of the following represents (i.e., nominal, ordinal, interval, or ratio).
  - a. weight
  - b. distance
  - c. birthplace
  - d. heart rate
  - e. IQ score
  - f. first-, second-, and third-place winners in a race
  - g. race
  - h. eye color
  - i. 10 best nursing programs in America
7. Round each number to a whole number.
 

a. 10.999	c. 10.55
b. 10.09	d. 11.399
8. Round each number to one decimal place.
 

a. 12.988	c. 5.555
b. 110.74	d. 55.549
9. Round each number to two decimal places.
 

a. 12.999	c. 6.055
b. 225.433	d. 90.107

10. Round each number to three decimal places.
- a. 12.9994
  - b. 225.4335
  - c. 6.0556
  - d. 90.1073
11. Find  $\bar{x}$  for the following sets of numbers.
- a. 1, 2, 4, 5
  - b. 8, 10, 15
  - c. 7, 14, 22, 35, 40
12. Obtain  $\sum x^2$  for the following sets of numbers.
- a. 6, 9, 12
  - b. 61, 75, 84, 85
  - c. 100, 105, 110, 120
13. Obtain  $(\sum x)^2$  for the following sets of numbers.
- a. 7, 9, 11
  - b. 30, 41, 52, 63
  - c. 225, 245, 255, 275
14. A researcher is interested in determining whether a new “anti-anxiety” drug relieves anxiety in college students studying statistics. Fifteen students were assigned to one of three conditions, where five students received a placebo, five received 50 mg of the new drug, and five received 100 mg of the drug. The students were then placed in a quiet room and were given 15 statistics problems to do. They were administered by the researcher’s assistant. After completing the problems, the tests were scored and the number of errors made by each student was used to determine his or her level of anxiety.
- Based on the information provided above:
- a. Identify the dependent and independent variables.
  - b. Determine the type of measurement scale used (interval, ratio, ordinal, or nominal).
  - c. Explain why the sample used is or is not representative of the population.
  - d. Using only the information above, give an operational definition of anxiety.
  - e. Define confounding variable and cite some examples that could create problems in this experiment.
15. The governor of the state of Georgia proposed that the state should provide classical music (in some format) to the parents of every Georgia baby to raise the child’s intelligence. The governor cited studies that have shown that college students who listen to classical music have higher IQ scores. What is the most essential experimental design problem in the governor’s proposal? What are some potential confounding variables in the studies of college students?

## Chapter 1 Test Yourself Questions

1. Which of the following is NOT a role of the statistician according to Abelson?
- a. curious detective
  - b. honest attorney



- c. good storyteller
  - d. moral advocate
2. A series of studies is used to demonstrate the usefulness of a theory. This process is called
  - a. replication
  - b. duplication
  - c. inferential testing
  - d. demonstrativeness
3. True or false: Theories are often proven or disproven in science.
  - a. true
  - b. false
4. Historically, which is the oldest form of statistics?
  - a. descriptive
  - b. inferential
  - c. parametric
  - d. nonparametric
5. Which of the following is NOT a characteristic of a good theory?
  - a. provable
  - b. testable hypotheses
  - c. provocative
  - d. grandness
  - e. all of the above
6. Who wrote that “not [ever] being wrong” was even worse than being wrong in science?
  - a. Nightingale
  - b. Pauli
  - c. Prusiner
  - d. Abelson
7. Anxiety was measured in a group of anxious elderly patients who were divided into two groups. One group received a new anti-anxiety drug and the other group received a placebo. The operational definition of the dependent variable is
  - a. anxiety
  - b. not specified
  - c. anxious elderly patients
  - d. all of the above
  - e. none of the above
8. The word *placebo* comes from Latin and means
  - a. untrue
  - b. fake belief
  - c. I shall please
  - d. placid

9. Some patients in a control group will get well because of their belief that they are receiving the real treatment. This phenomenon is known as the
- experimenter effect
  - placebo effect
  - Hawthorne effect
  - belief effect
  - all of the above
10. If we randomly choose names from the phone book, call them, and determine whether they are registered voters, does this process result in a random sample of voters?
- yes
  - no
11. Which of the following results in an abuse of statistical power?
- too large a sample
  - too small a sample
  - a nonrandom sample
  - a correlative experimental design
  - all of the above
12. Which of the following was NOT presented as one of the eight questions of any study or survey?
- Who was surveyed?
  - Were the participants paid?
  - How many people participated?
  - Did the survey have too many questions or items?
  - Was causation assumed from a correlational study?
13. A study linked a high blood pressure medicine to an increased risk of heart attacks. How many of the 70 published defenders of the medicine had financial ties to the manufacturer of the drug?
- |       |                |
|-------|----------------|
| a. 24 | d. 67          |
| b. 47 | e. all of them |
| c. 51 |                |
14. Which of the following orders is the correct one for the level of scientific scrutiny of a research study (from highest to lowest scrutiny)?
- journal article, convention paper, book
  - book, convention paper, journal article
  - convention paper, book, journal article
  - newspaper, convention paper, book
  - none has the correct order
15. Which of the following is the most important and powerful form of scientific control in an experiment?
- the design of the experiment
  - the statistical analyses
  - a correlational design
  - a large sample

16. After a normative study of blood pressure in people older than age 50, the researchers realized that they did not record the people's ages. The lack of knowledge about their ages is considered a potential
- confounding variable
  - dependent variable
  - independent variable
  - response variable
  - diagnostic variable
17. Which of the following words typically has no place in a statistical report?
- significant
  - insignificant
  - nonsignificant
18. When the experimenter evaluates the outcome of a study without knowing which participants were in the experimental and control groups and the participants are also unaware, the experiment is said to be
- retrospective
  - archival
  - retroblind
  - archivally blind
  - double blind
19. In a study of autistic children being taught facilitated communication, what percentage of the autistic kids could answer the questions correctly without any help from the facilitators after their training?
- |        |        |
|--------|--------|
| a. 0%  | c. 55% |
| b. 25% | d. 93% |
20. An experimenter defines "consumer assertion" as the total number of products that a consumer has returned for his or her money back in a 1-year period. This definition in statistics is
- a confounding variable
  - a definitional periodicity
  - an operational definition
  - a continuous but discrete definition
  - measurement "error"
21. Which of the following scales requires a rational zero point?
- ordinal
  - nominal
  - ratio
  - rational
  - interdenominational

22. The uppercase Greek letter  $\Sigma$  indicates that the reader should \_\_\_\_\_ a set of numbers.
- add
  - subtract
  - multiply
  - divide
  - calculate
23. For the set of numbers 2, 3, 5, the value of  $\Sigma x^2$  would be equal to
- 38
  - 100
  - 10
24. For the set of numbers 4, 5, 6, 9, the value of  $(\Sigma x)^2$  would be equal to
- 24
  - 158
  - 576
25. For the set of numbers 2, 3, 5, 6, 9, the value of  $\Sigma x^2 - \frac{(\Sigma x)^2}{N}$  would be equal to
- |        |        |
|--------|--------|
| a. -94 | c. 94  |
| b. 30  | d. 125 |
26. Convert the proportion .1356 into a whole number percentage.
- |          |        |
|----------|--------|
| a. 13.6% | c. 13% |
| b. 12%   | d. 14% |
27. Convert the proportion .1999 into a whole number percentage.
- |          |          |
|----------|----------|
| a. 19.9% | c. 20.0% |
| b. 20%   | d. 19%   |
28. Convert the proportion .5555 into a whole number percentage.
- |        |          |
|--------|----------|
| a. 56% | c. 55.6% |
| b. 55% | d. 55.5% |
29. Convert 56% into a proportion.
- |        |        |
|--------|--------|
| a. .57 | c. .56 |
| b. .55 | d. 5.5 |
30. Convert 99% into a proportion.
- |        |         |
|--------|---------|
| a. .99 | c. 1.00 |
| b. .90 | d. .9   |

## SPSS Lesson 1

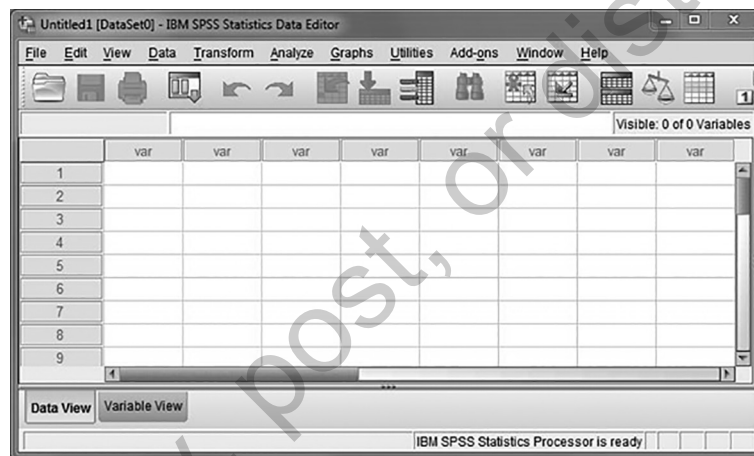
SPSS, as noted previously, can conduct statistical analyses at blindingly fast speed once the data have been entered properly. Each SPSS lesson in this textbook is designed to show you how to use and interpret

the SPSS statistical analyses associated with each chapter of the textbook. Student versions of the SPSS program are usually available at college bookstores at reduced rates. Many schools often have them installed on computers on campus, so purchasing SPSS might not be necessary.

### Opening the Program

1. Click **Start** → **All Programs** → **IBM SPSS Statistics** → **IBM SPSS Statistics 25** to open the program.
2. Close the initial “What would you like to do?” dialog if it appears.

The initial *Data Editor* dialog displays the current data view. Notice the highlighted *Data View* tab in the lower left-hand corner of the screen. To indicate that a file has not yet been created, the *Data Editor* displays *Untitled1 [DataSet0]* in the upper left-hand corner of the screen.



### Creating a Data File

To create a data file, we will start by inputting some data. Imagine an experiment where international college students were given a simple three-item depression questionnaire. They were also asked their names, gender, age, and nationality. The depression items were answered on a 7-point Likert-type scale where 1 indicated *strongly false*, 4 indicated *sometimes false, sometimes true*, and 7 indicated *strongly true*. The first item (Q1) dealt with the prevalence of sad moods, the second (Q2) dealt with suicidal ideation, and the third (Q3) dealt with the ability to get depressed easily. Table 1.1 lists the data.

TABLE 1.1 ■ DEPRESSION STUDY DATA

Name	Gender	Age	Nation	Q1	Q2	Q3
Jim Smith	Man	19	American	4	6	5
Vijay Singh	Man	21	Indian	1	2	1
Teruko Musashi	Woman	17	Japanese	5	2	6
Kenyatta Walker	Woman	20	South African	3	5	6

First, assign identification numbers to the participants of the study to protect their privacy. In my own research, I never input the names of my participants in the SPSS file. It is important to be able to link the names of the participants in your study to their original data. This is particularly helpful if there appears to be a problem with the data; thus, the original participant's testing file can be located. However, I do keep the names and respective identification numbers (IDs) in a secure place (a locked filing cabinet) but not in my SPSS files.

Jim Smith's ID = 1

Vijay Singh's ID = 2

Teruko Musashi's ID = 3

Kenyatta Walker's ID = 4

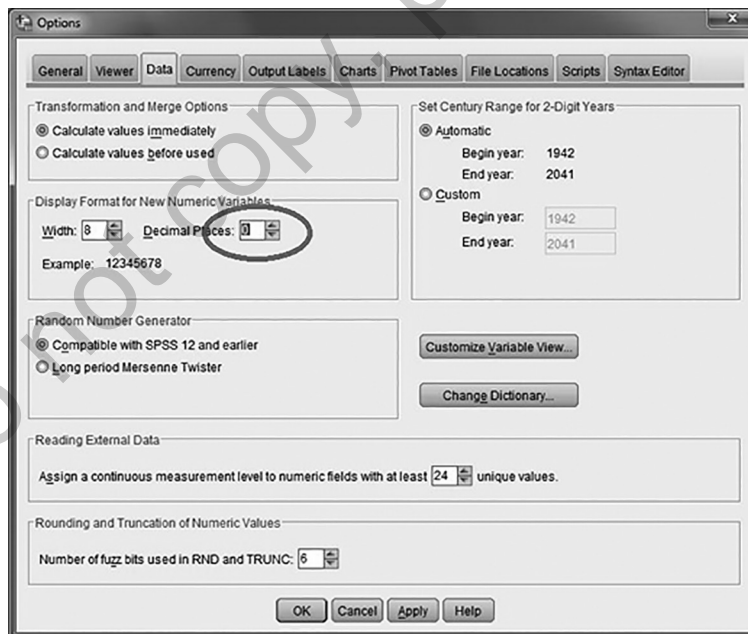
## Changing the Display Format for New Numeric Variables

By default, SPSS displays numeric variables to eight whole digits with two decimal places. The current example requires whole numbers. To change the default display format for new numeric variables, do the following:

1. Click **Edit** → **Options** and select the **Data** tab.
2. In the *Decimal Places* field, enter or select 0 and click **Apply** → **OK**.

This will generate the *Statistics Viewer*, or the output dialog. Minimize the *Statistics Viewer* for now. All your variable values will now be rounded to whole numbers.

*Note:* To change the format of a number for a single cell, activate that cell, click the **Variable View** tab, and change the value in the corresponding *Decimals* table cell.

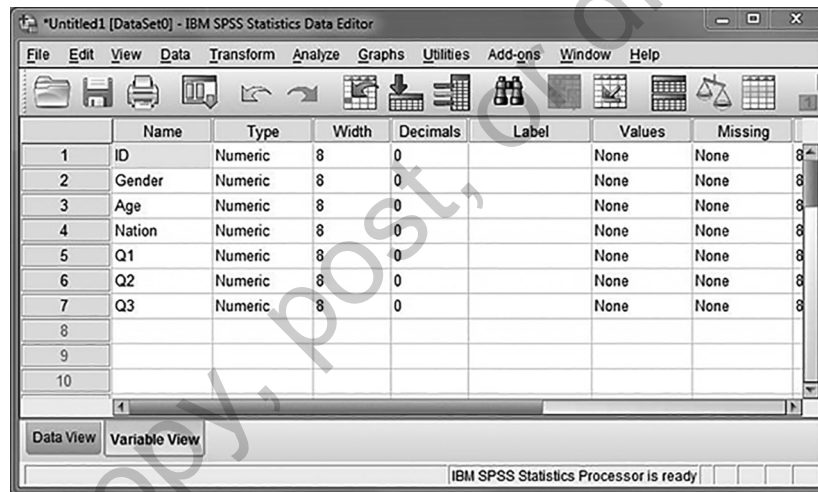




### Entering Variable Names

You will see that the case numbers for each participant are already numbered in a column down the left side of the opening data screen. It is still useful to input ID numbers. The first column in this screen is labeled *var*. We will use this column for our ID numbers. You will use the variable view to enter the variable information for each of the variables ID, gender, age, nationality, Q1, Q2, and Q3.

1. Click the **Variable View** tab in the lower left-hand corner.
2. Select Row 1 *Name* table cell.
3. Enter ID for the first variable name. The variable name is not case sensitive, so it does not matter whether you choose uppercase or lowercase (but be consistent).
4. Continuing downward in the *Name* column, enter Gender in Row 2.
5. Enter Age in Row 3.



6. Enter Nation in Row 4.
7. Enter Q1 in Row 5.
8. Enter Q2 in Row 6.
9. Enter Q3 in Row 7.

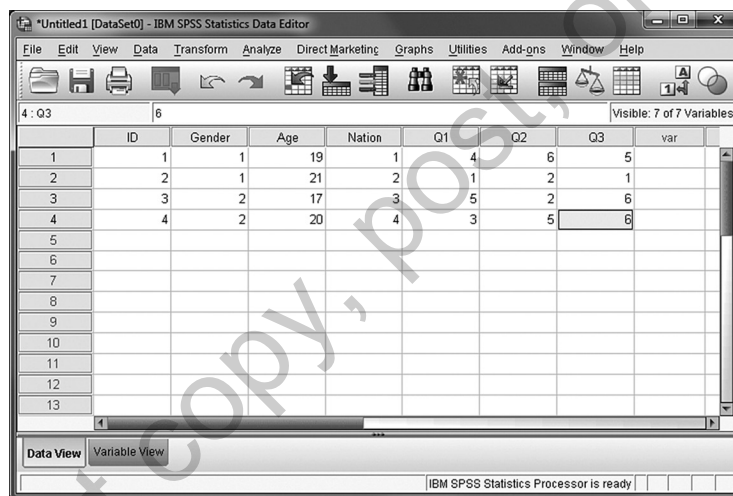
### Entering Data

1. Click the **Data View** tab to return to the data view.
2. Select the Row 1 *ID* table cell. When you click on a table cell, the active cell is highlighted.
3. Enter 1 as Jim Smith's ID value.
4. Use the values listed in Table 1.2 to complete the data entry. You will substitute the whole numbers listed in the following table for *ID*, *Gender*, and *Nation*. Do NOT enter the parenthetical phrases.

TABLE 1.2 NUMERIC VALUE SUBSTITUTIONS FOR DEPRESSION STUDY DATA

ID	Gender	Age	Nation	Q1	Q2	Q3
1 (for Jim Smith)	1 (for male)	19	1 (for American)	4	6	5
2 (for Vijay Singh)	1 (for male)	21	2 (for Indian)	1	2	1
3 (for Teruko Musashi)	2 (for female)	17	3 (for Japanese)	5	2	6
4 (for Kenyatta Walker)	2 (for female)	20	4 (for South African)	3	5	6

As you will quickly see, you may use the cursor or the arrow keys to activate other cells. If you make a mistake, you can use the Delete or Backspace keys to remove any mistakes. If you discover a mistake after you have activated another cell, just reactivate the cell with the mistake. When you are finished entering your data, your *Data Editor* will look like this:



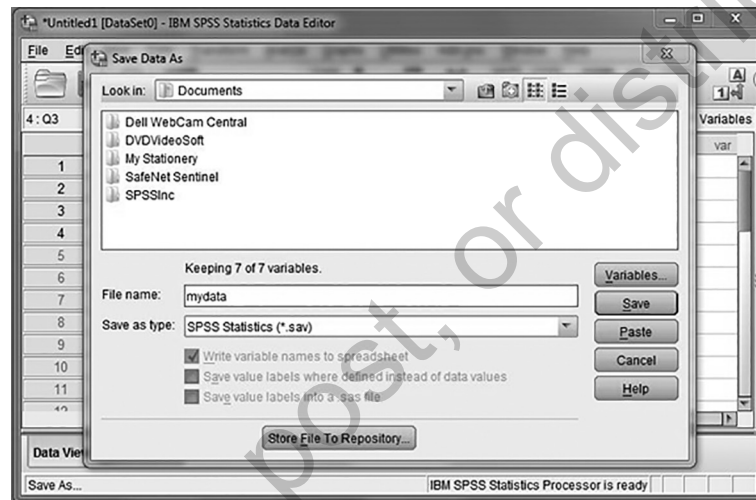
### Numeric Variables Versus String Variables

A numeric variable is entered as a number. A string variable is entered as a *string* of letters. It is almost always better to save variables as numeric values. There will be many more statistical options available if you have saved your variables as numbers. For example, the gender variable could be saved as a string variable, where you entered either "man" or "woman" for each case. However, it is usually better to define the gender variable as a numeric variable and enter a numeric value for being a man and a different numeric value for being a woman (or another category such as *non-binary, other, or prefer not to say*). In the *Variable View* tab, described below, there is a variable attribute called *Values*, and that is where you can enter the information that 1 = man and 2 = woman. Note that the numbers representing the genders are completely arbitrary, so you could also enter 0 = man and 2 = woman.

If in the future you would rather define a string variable, complete the following steps:

1. Click **Variable View**.
2. Click the current *Type* value for the variable you would like to change to string, then click the **ellipsis** button [ . . . ] that appears next to the value. This opens the *Variable Type* dialog.
3. Select **String** and click **OK**.

### Saving Your Data

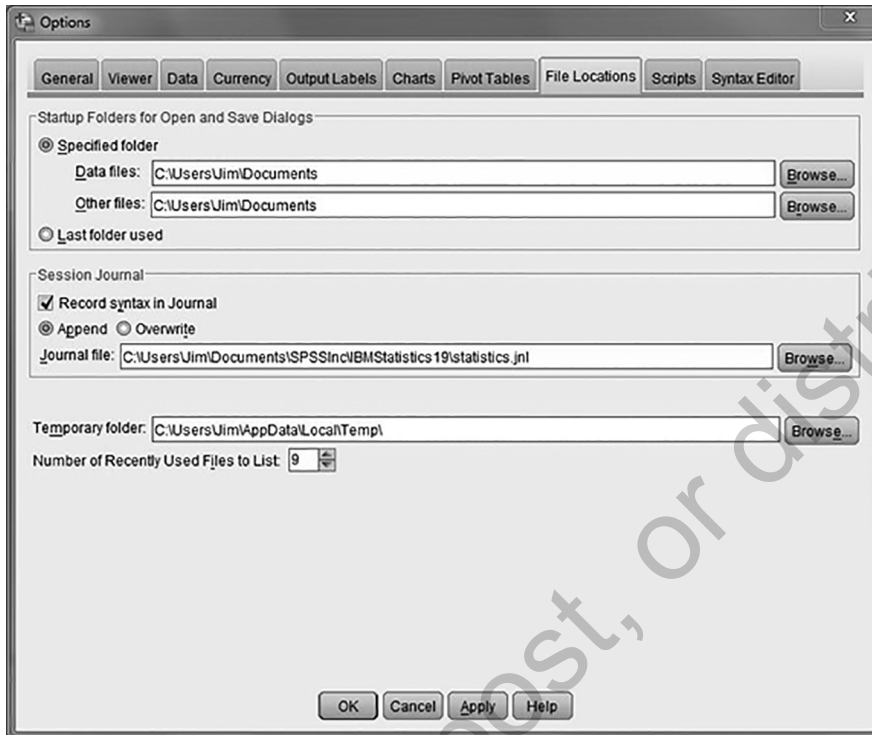


To save your data for future access:

1. Click **File → Save**.
2. Enter *mydata* in the *file name* field and click **OK**. Your data will be saved as *mydata.dat* in the specified location.

### Changing the Folder Where Your Data Are Saved

1. If you would like to change the folder where your data are saved, click **Edit → Options → File Locations**.
2. Use the fields under the *File Locations* tab to specify the folder where your data are stored.



It is wise to save your work periodically while entering data to guard against losing everything you have entered should your computer lock up or in case of a power outage.

To exit the program, click **File** → **Exit**. A dialog will appear asking if you wish to save your currently open file. Click **OK** to save your file and close the program.

### Opening and Saving Your Data Files

Once you have saved a data set, it will be available to you in the future.

To open your data set, click **File** → **Open**, choose the desired file name, and click **OK** to open your file in the *Data Editor*.



Visit [edge.sagepub.com/coolidge4e](https://edge.sagepub.com/coolidge4e) to help you accomplish your coursework goals in an easy-to-use learning environment.