Chapter 1: Introduction
Note: There are no exercises in Chapter 1

Chapter 2: The Empirical Approach to Political Science

Exercise 2-1.
Students can simply list or briefly describe these characteristics. The idea is for them to have the ideas at their (mental) fingertips. I find this exercise useful as an in-class exercise to prime students for discussing the components of the scientific method. There are of course many components of the scientific method. Student responses may include, for example:
- Empirical verification
- Nonnormative
- Transmissible
- Generalization
- Explanatory
- Search for causal relationships
- Predictive
- Reliance on probabilistic explanation
- Theoretical
- Research question/theory/hypothesis

Exercise 2-2.
a. Empirical. This statement makes a claim that can be checked against historical records and can be supported or falsified.

b. Normative. This statement expresses an opinion that is not susceptible to any empirical verification.

c. Combination. This statement contains an empirical component (the claim about early voting favoring Democrats) and a normative component with the use of the word “should.”

d. Either normative or empirical. This statement could be empirical if a specific measure of representation is proposed. Without that measure, the statement is largely normative as it is an opinion. It could go either way.


g. Normative. I regard “better off” as making a value judgment. The exception is if better off were defined with a specific variable, then the statement could be empirical.

h. Normative. In the spirit of the discussion of the scientific method in Chapter 2, I don’t see how rights have “scientific” status. Many people might agree with the statement. But how does it meet the verifiability criterion?

i. Normative.

j. Empirical.

k. Normative. “Too many” is the key phrase.

l. Normative. The premise is empirical. But, again, we think students should be asked to think about this idea: do empirically established conditions logically lead to the truth or desirability of a preference or value?

Exercise 2-3.

a. (1) Map makers ought to “pack” districts with as many like-minded partisans as possible. (2) The House of Representatives does not have enough competitive districts currently.

b. (1) Packing of districts leads to the elections of more extreme representatives (e.g., very liberal or very conservative). (2) Increasing the competitiveness of districts will increase the number of voters dissatisfied with their representatives. (3) Increasing the competitiveness of districts will increase turnover in Congress. (4) Increasing the competitiveness of districts will enhance representation. This statement could go either way. It is ambiguous because the phrase “enhance representation” could be interpreted as meaning making it better, which is a normative statement. But it could also be defined empirically.

Exercise 2-4.

We don’t need the scientific method to make reliable predictions about other people’s activity. All we need is the assumption that “actions” (goal-guided behaviors) flow from beliefs and desires. This supposition allows us to lead orderly, predictable lives. (If we couldn’t predict behavior, would anyone show up in class at the appointed hour to take a test?) Moreover, we see examples of predictions of human behavior all the time in polls. (One could point out that pollsters have become quite good at predicting election results.) And there are countless other examples. What is true, a student, might reply, is that empirically based predictions are invariably probabilistic. They don’t “work” 100% of the time. And these “forecasts” haven’t reached the status of many scientific predictions. All of that argument may be true; social scientists have successfully used adaptations of
the scientific method to develop and test generalizations that seem to have a great deal of support behind them.

Exercise 2-5.
- Historical: too concerned with the history of subjects.
- Legalistic: too concerned with laws and legal analysis.
- Institutional: too concerned with formal institutions.
- Descriptive: described institutions and events but did not explain behavior.

Exercises 2-6.
Answers will vary as there are more than one right answer to each question. The point of the exercise is to get students thinking about how to verify empirical statements. Encourage students to avoid writing “look it up on the internet” for every question. They should be more specific and creative.

   a. Contact the Pennsylvania Secretary of State
   b. Look at the British Parliament webpage and count the number of parties
   c. Look up the federal poverty level at Healthcare.gov, Medicaid, or others
   d. Survey a random sample of respondents and ask if they would support raising taxes expressly to combat climate change
   e. Compare voter identification laws across the states
   f. Consult reports from the World Health Organization

Exercise 2-7.
Students will have great variation in their answers. They should identify two different projects in part a. Part b should include references to components of the scientific method and students should see clear differences across different kinds of projects.

Chapter 3: Beginning the Research Process

Exercise 3-1.
Answers will vary with articles chosen.

Exercise 3-2.
Answers will vary with articles chosen.

Exercise 3-3.
Niven identifies four major problems:
1. Researchers disagree over the effects. Some researchers conclude that negativity depresses turnout while others conclude the effect is slight. Some research argues that negativity in advertising affects not only how recipients of such advertising view the target of the advertising but also how they view the sponsors of the ad and more general feelings toward the political system. Others have found that the effect of negative ads on
recipients depends on the content of the ad (whether it presents unsubstantiated claims or contains issue-related attacks). A meta-analysis of research studies indicates that there are as many studies concluding that negative ads decrease turnout as there are studies concluding that negative ads increase turnout.

2. Previous research has used research designs that may limit the value of the results. Much of the previous research has been conducted in laboratories. Therefore, external validity is a problem because the experiments do not measure the impact of negative ads in real elections. Furthermore, the effect measured has been intention to vote rather than actual voting behavior. Plus, people tend to overreport voting or intention to vote, so studies that use these ways of measuring voting may encounter measurement error. Other studies use survey research in which respondents are asked if they recall seeing negative political advertising: researchers have no control over actual exposure to ads, which makes it difficult to demonstrate a causal link between ad exposure and voting behavior.

3. There have been differences in the type of treatment. Some researchers have used televised ads while others have used mailed material. Also the content of the ads has varied from one study to another.

4. Differences in the selection of subjects: some have used randomly selected participants, others have not.

Exercise 3-4.
The point of this simple assignment is to make sure students get a feel for some of the points made in the chapter. In particular, the searches they are asked to conduct will probably lead to far too many results and will need to be refined. It is also likely that they will turn up lots of irrelevant or useless sources while finding few that would be acceptable for a college paper. Finally, we think it is important for students to give a more complete citation than the simple “www.somewhere.com/topic” that frequently turns up in student writing. Stress the importance of noting the title and date accessed.

a. High Priority; An academic book written by an economist.
b. Low priority; an interview will have interesting material, but we’d rather read his book.
c. High Priority; A book review can be a good place to start as a critical examination will quickly distill the book’s approach and quality.
d. High Priority; Center reports can be useful analyses of data
e. High Priority; Academic journal articles are exactly what is needed
f. Low Priority; popular magazine articles can be interesting but do not have the same rigor as peer reviewed journals.
g. Low priority; The little analysis that is likely to be included will be too narrowly focused or draw on other’s work (Like a research center report)
h. Low priority; normative, not empirical

Exercise 3-5.
This assignment provides a quick and painless introduction to scholarly literature. If used in conjunction with other assignments, students should begin to appreciate the different types of sources. This will also be a good time to make sure that they follow an acceptable citation format.

**Exercise 3-6.**
Answers will vary.

**Exercise 3-7.**
Although this assignment seems straightforward, it might present problems to students embarking on their first major college-level research project. Consequently, it might make a good weekend assignment. One of us had students distribute their bibliographies in class, partly to make sure they were working independently, and partly to provide rapid and general feedback. Doing so helped clarify a lot of misunderstandings about what would and would not be useful sources.

**Exercise 3-8.**
Answers will vary.

**Exercise 3-9.**
Answers will vary.

**Exercise 3-10.**
Students will find web pages with publication information for all of the authors listed in the previous exercises.

**Chapter 4: The Building Blocks of Social Scientific Research: Hypotheses, Concepts, and Variables**

**Exercise 4-1.**
*Content of answer will vary but should include six hypotheses.*

**Exercise 4-2.**
  a. IV: US interests at stake; DV: Support for military force; UA: US public/individuals
  b. IV: # NGOs; DV: form of government; UA: countries
  c. IV: # counties; DV: law enforcement spending; UA: States
  d. IV: amount of aid; DV: public health; UA: countries
  e. IV: party id; DV: likelihood of voting; UA: individual
  f. IV: temperature; DV: turnout; UA: elections

  b. *Possible rewrite:* Are people who live in countries that are more vulnerable to severe weather more concerned about global climate change than are people who live in less vulnerable countries?  
  Or: Is there a relationship between a country’s being affected by global climate change and the percentage of the public who are concerned about climate change?
Hypothesis: People who live in countries most likely to be negatively affected by global climate change are more concerned about global climate change than people who live in countries less likely to be affected.
Or: Countries likely to be affected by global climate change will have a higher percentage of the public who are concerned about climate change than will countries less likely to be affected.
A null hypothesis could also be proposed: No matter what the country, a majority of people are concerned about global warming.

Independent variable: whether or not one lives in a country affected by global warming; Dependent variable: concern about global warming; Unit of analysis: individual.

Or, if students develop a hypothesis that applies to a country, then:
Independent variable: whether or not the country is affected by global climate change; Dependent variable: percentage of public concerned about global climate change; Unit of analysis: country.

c. Possible rewrite: Will primary voters who say the political experience of presidential primary candidates is important more likely to prefer candidates with experience (e.g., Hillary Clinton) than are voters for whom political experience is not important?
Hypothesis: States with public funding for candidates have more competitive elections than states without public funding.
Independent variable: importance of political experience; Dependent variable: amount of experience of preferred candidate; Unit of analysis: voter in primary.

d. Question okay as is.
Hypothesis: States in which both chambers of the legislature are controlled by the same party pass more bills than states in which control is divided.
Independent variable: party control of state legislature: unified or split; Dependent variable: number of bills passed; Unit of analysis: state.

e. Question okay as is.
Hypothesis: States with public funding for candidates have more competitive elections than states without public funding.
Independent variable: public funding; Dependent variable: competitiveness of elections; Unit of analysis: state.

f. Question okay as is.
Hypothesis: The welfare policies in countries with high turnout rates respond more to changes in market based inequalities than in countries with low turnout rates.
Independent variable: turnout rate; Dependent variable: responsiveness to inequalities; Unit of analysis: country.

Exercise 4-3.
a. Hypothesis needs to state the direction of the relationship. Possible rewrite: The president will experience lower approval ratings when the price of gasoline is high.
b. Hypothesis does not specify a relationship. Possible rewrite: An increase in the availability of drug treatment programs in prisons causes the recidivism rate to decrease.

c. Hypothesis is not general. Possible rewrite: Turnout rates in legislative districts in which candidates run unopposed are lower than in districts in which there is more than one candidate.

d. Hypothesis is normative. Possible rewrite: Replace “are better” with some feature that can be measured empirically such as “have less air pollution,” “use less energy,” or “are more compact.”

e. Hypothesis is not plausible.

f. Hypothesis could be a tautology if the definition of “active in politics” includes contributing money to campaigns. Possible rewrite: People who volunteer their time to work on election campaigns are more likely to contribute money to campaigns than people who do not volunteer.

**Exercise 4-4.**
*
*Students have a wide degree of latitude in how they combine the variables. Possible hypotheses:*

a. An increase in interest in politics causes an increase in the likelihood of voting

b. An increase in the strength of gun laws causes a decrease in the number of gun related deaths

c. An increase in the number of teen pregnancies causes an increase in the money spent on abstinence only programs

d. As the president devotes more time to civil rights in the SOU more civil rights bills are introduced in Congress

e. As the corporate income tax rate increases the number of new businesses decreases

**Exercise 4-5.**

Primary caregiver for children (primary caregiver, not); support for Family Medical Leave Law (thermometer scale for support); gender (female, male).

a. Primary caregivers for children will be more likely to support the Family Medical Leave Law than those who are not primary caregivers.

*Independent variable: primary caregiver; Dependent variable: support.*

I expect that the third variable, gender, will affect the hypothesized relationship between primary caregiver and support for the law as an antecedent variable. I theorize that women are more likely to be the primary caregiver for children so gender affects caregiver status as well as support for the law.
gender → primary caregiver → support for law

The third variable is antecedant.

b. People who have a greater general interest in politics are more likely to turn out to vote than are people who are less interested.

*Independent variable:* interest in politics; *Dependent variable:* turn out to vote.

Controlling for the predicted outcome of the election will cause the difference between those with greater or less interest in politics to get smaller. People are more likely to turn out to vote in elections that are too close to call, and are less likely to turn out to vote if the election is predicted to be a lopsided victory.

Interest in politics → Predicted outcome of election → Turn out to vote

The third variable is intervening.

c. The greater the difference in the cost of light bulbs, the more likely consumers will choose the less expensive, regular bulbs.

*Independent variable:* difference in cost of light bulbs; *Dependent variable:* percentage of consumers choosing regular bulbs.

One might hypothesize that concern about global climate change makes no difference and that consumer choice depends on the relative cost of the bulbs.

One might also hypothesize that concern about global climate change does have an impact. For those who are concerned about climate change, the difference in cost is not as important. They will tend to purchase the more expensive, energy-efficient bulbs. For those not concerned about climate change, cost will be the important factor affecting choice.

Difference in cost → choice of bulb

Concern about global climate change

The third variable is alternative.

**Exercise 4-6**
Content of answers will vary. Answers should reflect the appropriate unit of analysis.
Chapter 5: The Building Blocks of Social Scientific Research: Measurement

Exercise 5-1.

a. nominal; dichotomous, so could be used as ordinal

b. ordinal; the likely permanence of the ban decreases across the categories

c. nominal

d. ratio

e. ordinal

f. ordinal as tax policy becomes more progressive at least in theory; could be treated as nominal if one is unwilling to make this claim

g. nominal

h. ordinal

i. ordinal

j. ratio

k. ratio

l. interval

m. ordinal

n. ordinal

o. ordinal

p. nominal

Exercise 5-2.

Measure #1 is ordinal. Measure #2 is ratio. #2 has more information because it captures the number of years. We could create an ordinal scale from the number of years variable. We could not create a ratio variable from the ordinal scale. More information means that we have learned more about the respondent so the second measure is generally better.

Exercise 5-3.
The answers are a bit subjective much like validity. Students should be able to support their answers about validity with reasonable arguments.

**Exercise 5-4.**

Positions on gay marriage, stem cell research, abortion, parental control over access of minors to birth control information, teaching of creationism or alternatives to evolution in schools, school prayer, display of religious symbols such as the Ten Commandments on public property, support for traditional family values, belief that newer lifestyles are contributing to the breakdown of society, and tolerance of alternative lifestyles are likely to be listed. Other possibilities include whether or not war is just, concern about economic and social inequality, and concern about global warming.

Operationalization should match the conceptualization.

**Exercise 5-5.**

Students have some latitude in their responses. The point is that their operationalization of each term should capture the full range of the definition provided.

a. Respondent self-identify as male or female
b. Respondent answers question that provides ordinal categories from which to choose to identify combined income of everyone in the household.
c. Respondent self-identify from a list of Asian, Black, Hispanic, White, Other
d. Respondent answers and index question that includes items on level of support for government programs in food support, healthcare and housing.
e. Respondent indicates how much they feel their voice matters in inter-personal communication with elected officials.

**Exercise 5-6.**

One way is to divide the scores into supportive and unsupportive categories with scores 0–49 placed in the unsupportive category and scores 50–100 in the supportive category. This would make theoretical sense and also results in roughly equally sized groups.

Another way is to divide the scores so that they fall into equally sized groups. Since 50% of the scores are 29 or below, these scores could be put into the unsupportive group. Scores 46–100 would constitute the supportive group. Note that there is a natural break between the scores 29 and 46, so the latter score is much closer to the next score of 50 than it is to 26.

**Exercise 5-7.**

Scores could be divided into the following groups: 0–24, 25–49, 50–74, and 75 and over, which creates equal numbers of theoretical scores in each group (except for the last group, which has one extra score.) This would result in 26% of the cases in the lowest group, 32% in the next highest group, 20% in the next group, and 22% in the highest group. Alternatively scores could be divided into roughly equally sized groups by combining scores as follows: 0–22, 25–39, 40–62, and 67 and over.

**Exercise 5-8.**
For this exercise the independent and dependent variables are easy enough to identify. How the students measure each will vary.

a. IV: religiosity; DV: ideology  
b. IV: distance to polling station; DV: likelihood of voting  
c. IV: economic development; DV: participation in peacekeeping missions  
d. IV: nation at war; DV: frequency of presidential addresses

**Chapter 6: Research Design**

**Exercise 6-1.**  
a. **Internal validity** means that the research procedure demonstrated a true cause-and-effect relationship that was not created by spurious factors. **External validity**, the extent to which the results of a study can be generalized across populations, times, and settings, is the touchstone for natural and social scientists alike. The important difference is that one describes causation while the other describes generalizability.

b. **Causation** is when a change in the state of one thing brings about (in full or in part) a change in the state of another. **Correlation** is simply a statement that two things are systematically related. The important difference is that causation carries with it claims about time order and the elimination of alternative explanations for the observed relationship while correlation does not.

c. The **experimental effect** measures the impact of the independent variable on the dependent variable and, consequently, is a main focus of experimental research. An **experimental group** is a group of subjects that receives or is exposed to an experimental treatment or test factor. The important difference is that the effect is something we measure while the group is a physical group of people participating in the experiment.

d. The **test factor** is the independent variable introduced and controlled by an investigator in order to assess its effects on a response or dependent variable. The **pretest** is the measurement of the dependent variable prior to the administration of the experimental treatment or manipulation of the independent variable. It can also refer to the measurement of many variables. The important difference is that the test factor is a stimulus of some sort while the pretest is a measurement of variables.

e. An **experiment** is a research design in which the researcher controls exposure to the test factor or independent variable, the assignment of subjects to groups, and the measurement of responses. A **field experiment** is an experimental design applied in a natural setting where the researcher usually has less control over the test factor or independent variable, the assignment of subjects to groups, and the measurement of responses. The important difference is the location (lab / field) and the control over over the test factor or independent variable, the assignment of subjects to groups, and the measurement of responses.
Exercise 6-2.
Here are some possibilities:

Extra-legal factors are those mentioned in the chapter: strategic behavior, outside influence, president’s approval rating, and party, for example.

Exercise 6-3.
A good response ought to contain the ideas of internal and external validity. Since the investigator theoretically controls so many facets of the experiment, the potential for the detection of causal effects is high. Still, we listed several problems in experimentation, not the least of which is the level of generalization that one can make. Students should
also mention (in a way that indicates understanding) history, maturation, testing effects on attitudes and behavior, selection bias, experimental mortality, instrument decay, and demand characteristics.

Exercise 6-4.

a.

i. Yes. This is the only plausible yes answer. (Note that the introduction to the question mentioned a random sample.)

ii. No. There is no reason to think that even if the experiment had been done properly the results would apply just to Nowhere’s men.

iii. No. This possibility doesn’t seem appropriate because it implies all people who receive the treatment will be affected, whereas the study was confined to Nowhere.

iv. No. Strictly speaking, we don’t know how the messages would affect people in other places. Students could argue that if Nowhere is a typical community (perhaps a random sample of one from cities and towns throughout the United States), the results would hold for the population as a whole. But the sampling design makes this inference tenuous.

b.

i. Incorrect. Besides not being able to generalize to the U.S. population, it is not possible to generalize to advertisements in general. (The experiment involved only one topic.)

ii. Incorrect. Even assuming the design has no faults, the results cannot be generalized beyond Nowhere.

iii. Correct.

iv. Incorrect. See “i” above.

c. See the previous question. There is random sampling but not randomization. We want students to understand fully that random assignment to treatments ensures that the average composition of the groups is the same at the beginning of the experiment. That way any differences in opinions at the end can presumably be attributed to the experimental variable: exposure. But in this study subjects were assigned more or less on a first-come, first-served basis. There are two control groups, although they are really not comparable to the experimental groups in regard to experimental conditions. There is no way to know of course, but it is likely that the study involved experimental mortality. Subjects with strong pro-choice views might have dropped out, and the people on the other side of the issue might have just become more adamant in their opinions. Differences in pre- and post-experimental choice would then not necessarily reflect the efficacy of extended viewing. (The manner in which the controls were treated clouds the issue, because it is not likely any of them would have dropped out of the study.) It is also possible that demand characteristics were at work, since the extended exposure to the treatment powerfully suggests a “desired” response.

Exercise 6-5.

There are many plausible research designs that would represent correct answers. In general terms, we are looking for some sort of “interrupted” time series analysis that uses,
for example, Gallup polls to track presidential approval ratings. (By *interrupted* we mean a punctuated or broken series, the cause of which may be an event, new policy, or other “shock.”) Other publicly available survey data would allow students to investigate changes in support for the health care law. Although students at this level cannot conduct a formal intervention analysis, they might be able to imagine that after a particular date trends in attitudes begin to change noticeably. They should be able to draw a hypothetical diagram like the one shown in Figure 6-6 in the textbook. Equally important, you might have to remind most students that they are dealing with a quasi-experiment that cannot control for possible confounding factors. Support for the health care law may sharply decline after the criticisms start. But is this effect due to those criticisms or to some other factor? Any research design that is able to capture this dynamic would work.

**Exercise 6-6.**
There are many possible correct answers here but the heart of the answer should revolve around some sort of experimental design that involves a pre and post-test of student political knowledge before and after watching the *Daily Show*.

**Exercise 6-7.**
Design choices will vary. There is no single way to “best” examine the topics. The answers should be reasonable.

**Exercise 6-8.**
Students have considerable leeway in answering the questions. If they follow the hints they should be able to come up with a survey that asks about climate change issues. Perhaps they might want to survey citizens in different countries about their willingness to pay higher taxes to deal with the causes and effects of climate change.

If they follow the hint for the second answer, students may want to test climate change control measures like power plant stack scrubbers or fuel efficiency laws and how they affect CO₂ levels.

**Exercise 6-9.**
a. There are many good choices here. One could use a survey to very good use by surveying district residents or members of Congress. Interviews or focus groups would also be good choices.
b. This answer will depend entirely on the research design chosen. The most important issues to discuss are likely to be internal and external validity.
c. A home-style could be operationalized in many different ways. For example, one could quantify a home-style by counting any number of actions or activities like attending parties, shaking hands, kissing babies, outfits worn, etc.

**Exercise 6-10.**
a. The research question is whether attack advertising demobilizes the electorate—or—whether negative advertising affects turnout.
b. The predominate theory in the literature is that campaigns have a positive impact on the electorate, mobilizing people to vote. Campaigns have changed radically; heavy
reliance on broadcast ads. Campaigns used to tell the voter the positive things about their own candidate. Now, the focus is on the negative aspects of the opponent. Campaigns have turned “hostile and ugly.” Attack ads have become the norm rather than the exception. Given these changes, the authors must re-evaluate the effects of campaigns on the electorate. They think that campaigns are demobilizing forces in politics.

c. Negative advertising causes lower turnout.

d. Previous research on the topic faced many limitations:
   - used students instead of representative sample
   - used fictitious candidates that did not resonate with respondents
   - administered treatment in classroom setting
   - no ongoing campaign at time of experiment
   - used tiny samples
   - respondents could not vote for candidates in treatment ads

e. The authors used many design features to address internal validity including the use of two identical living rooms that served as the lab for exposing participants to the treatment. The living rooms were meant to better approximate actual television watching conditions. Participants could sit on comfortable furniture, eat snacks, get and move around and talk with each other. This is similar to what they might do at home. This was much better than using a classroom as in previous work.

The treatment ads were created by the researchers allowing full control over the content to make sure that there was only one difference in each ad—the tone—so causation could be established. The treatment ads used real candidates for office during an election period. This made it more realistic for participants.

This study used a sample of voters of the surrounding community rather than students.

f. The authors made a concerted effort to create a sample that was similar demographically to the greater Los Angeles area population. They did this by recruiting participants through newspaper ads, flyers, employer announcements and calling people from voter registration lists. They also offered $15 in compensation for participation.

This was important because by bringing in people from different places through different recruiting efforts, the sample was better—it was pretty representative of the community population. Therefore, the results from the study should be more generalizable than if the authors had relied on a less representative sample—students for example.

**Chapter 7: Sampling**

**Exercise 7-1.**

a. A **population parameter** is a characteristic or an attribute in a population (not a sample) that can be quantified. An **estimator** is a statistic based on sample observations that is used to estimate the numerical value of an unknown
population parameter. The difference is that an estimator “estimates” the value of the population parameter.

b. A **sampling frame** is the population from which a sample is drawn, ideally the same as the total population of interest to a study. A **sampling unit** is the entity listed in a sampling frame. It may be the same as an element, or it may be a group or cluster of elements. The difference is that a sampling unit is one subset of a sampling frame.

c. An **element** is a particular case or entity about which information is collected; the unit of analysis. A **stratum** is a subgroup of a population that shares one or more characteristics. The difference is that a stratum is made up of a number of individual elements that share a characteristic.

d. A **Probability sample** is a sample for which each element in the total population has a known probability of being selected. A **nonprobability sample** is a sample for which each element in the total population has an unknown probability of being selected. The difference is whether the sample has a known probability of being selected.

e. A **Proportionate sample** is a probability sample that draws elements from a stratified population at a rate proportional to the size of the samples. A **Disproportionate sample** is a stratified sample in which elements sharing a characteristic are underrepresented or overrepresented in the sample. The difference is that a disproportionate sample will not reflect the same proportions of characteristics as the population.

**Exercise 7-2.**

- a. Systematic random
- b. Quota
- c. Snowball
- d. Cluster
- e. Simple random
- f. Convenience
- g. Stratified
- h. purposive

**Exercise 7-3.**

a. This article is an example of a nonprobability sample of people subjected to offensive public speech.

b. Nielsen sampled people from public places where they were potential targets for offensive public speech. On the assumption that not all people frequent such public places, her selection process ensured that her subjects were, in fact, likely to have been targets.

c. She used systematic procedures to include different types of people and to minimize researcher bias in the selection of potential subjects. She randomly selected a side of a location and used a die to determine which persons to approach. If the die came up “3,” she approached every third person to ask if he or she would participate in her study. She
continued with her selection procedures until she achieved numerical goals for respondents with certain racial and gender characteristics. She used a variety of locations in three communities to ensure representation across race, socioeconomic status, and gender. Her personal contact with subjects likely increased their willingness to agree to a subsequent interview.

**Exercise 7-4.**

a. *Probability sample.* Systematic but not necessarily random. It would be preferable to select a random starting place. The population is students in South High and their parents.

b. *Nonprobability, convenience, or hit-or-miss.* It would be hard to identify any specific population because among other things some who get the questionnaires might not even be students at South High. And, of course, those leaving at that time might not be representative of the student body.

c. *Random probability.* The target population, South High parents and children, might itself be representative of similar individuals in the wider community or conceivably in the state or nation. But there is no way to know without further investigation. The same thing could be said of Question 7-2a. Those who return the questionnaire may or may not be representative, so presumably the population is South High students and parents who return questionnaires. We assume the office uses a truly random mechanism for selecting the names.

d. *Quota sample.* The plan calls for 50 in each group no matter what its overall proportion. Moreover, there is no mention of how the members of each group are to be selected. But if a student mentions in his or her answer that he or she assumes that the guidance counselor selects these students at random, then this could be disproportionate stratified sampling.

e. *Stratified probability sample.* If the classes had unequal numbers of students, this would be a disproportionate stratified probability sample because the plan calls for 50 students from each class.

**Exercise 7-5.**

<table>
<thead>
<tr>
<th>Maryland Firms, by Race, 2002</th>
<th>Expected Numbers for Samples of</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>Population</td>
</tr>
<tr>
<td>White</td>
<td>329,107</td>
</tr>
<tr>
<td>Black</td>
<td>69,192</td>
</tr>
<tr>
<td>American Indian/Alaska Native</td>
<td>3,548</td>
</tr>
<tr>
<td>Asian</td>
<td>26,169</td>
</tr>
<tr>
<td>Hispanic</td>
<td>15,524</td>
</tr>
</tbody>
</table>
c. For a sample size of 200, the expected frequency of American Indian/Alaska Native—, Asian—, and Hispanic-owned firms is too small for statistical analysis. There is a chance that a simple random sample might yield no firms from these groups. The expected frequency of black-owned firms is above 30, but here too there is a probability that a random sample might not yield enough black-owned firms. Increasing the sample size to 1,000 improves the situation, but there still are too few American Indian/Alaska Native— owned firms.

d. Assuming there is one list of firms, a systematic sample could be taken; the sampling interval would be 443,540/200 or 2,218. Use a random number table to pick a random start between 1 and 2,218 and then select every 2,218th firm after that. Alternatively, if the firms are numbered, 200 firms could be chosen at random using a random number table.

Exercise 7-6.

a. Students will choose five conflicts—wide variation here.
b. Students should be able to explain the rationale behind their choices in part a. The point of the exercise is to think about how researchers go about selecting cases for comparison.

Exercise 7-7.

a. The answers will vary by school and many different approaches are acceptable. The key is that the student is thinking about how to capture all of the different subgroups within the population.
b. A critically minded student should be able to identify parts of the plan in part a that would be more difficult to implement or identify groups in the population that may be harder to include.
c. If the sampling frame does not include the entire population of interest the sample that is drawn from the sampling frame will not allow inference about the population of interest.

Exercise 7-8.

a. The underlying NES sample included 290 black respondents. The authors added 310 more black respondents. By adding additional respondents from one strata, the sample is now disproportionately weighted.

b. The disproportionate sample is better because the 290 original black respondents would not be large enough to answer questions about black political behavior. By making the sample disproportionate the authors are able to use the data to answer these questions.

c. “It significantly enhances our ability to gauge the range, diversity, and determinants of African-American political opinion and vote choice. Furthermore, it facilitates informed comparisons to whites and Latinos at a time when such comparisons are especially useful.
to our conceptions of politics and representation.” This would not be possible without the data so the data are useful to society at large to better understand behavior.

**Exercise 7-9.**
Tell the students not worry about the unrealism of this situation. It is true that demographic data and maps of the sort envisioned here might not be available. Still, everyone can assume that they are. I have in mind a cluster sample as described in Chapter 7 of the textbook. The idea would be to enumerate primary sampling districts, draw a random sample of them (weighted according to population), list blocks or housing units in the chosen districts, and have interviewers draw a sample from the adults in the selected households.

### Chapter 8: Making Empirical Observations: Firsthand Observation

**Exercise 8-1.**
In some situations, provoking behavior may cause harm, for example, if one were to make a comment with racial or gender content that may be offensive to some just to see how others would react. However, as Bositis points out, in some situations provoking behavior is natural, such as an intern asking for clarifying information about institutional or organizational norms and procedures and observing the responses from supervisors and fellow workers. Stating a policy or political preference in a group may provoke responses. Students who have attended political rallies or public meetings and asked questions or made comments will have been able to observe the response and reactions of others. In other situations, the researcher is not providing the provocation or stimuli, but is simply observing how people respond to stimuli provided by other participants in the setting. Many students will have attended legislative sessions or official sessions of organizations such as the United Nations or the Supreme Court.

**Exercise 8-2.**
The author frequented several public places where she observed people being harassed and how they reacted. It is not clear that the individuals with whom she arranged interviews were ones whom she observed being harassed, but her observations allowed her to assess how people react to such comments. The fact that she was present in these locations and approached people to ask them if they would be willing to be interviewed increased her ability to include appropriate subjects in her study and very likely reduced the refusal rate over other methods of selecting subjects or respondents.

**Exercise 8-3.**
Participant observation is important if you want to study the strategizing of candidates and campaign support staff, directly see how the campaign is conducted, and gain contacts for lengthy interviews after the campaign.

Most important, the author talks about the importance of not overscheduling his time so that he was able to follow events as they were unfolding and to continue valuable conversations without having to keep to a tight schedule.
Students might be able to make observations in their classes or campus organizations about whether males or females talk the most, who takes charge of making decisions, and how people respond to comments. Is there a tendency to acknowledge or praise some people’s comments? Are good ideas that are raised by women taken over by men or ignored until they are repeated by a man? Is credit given to the men when they repeat ideas raised earlier by women? Do instructors tend to respond differently to students based on the ideological orientation of a student’s comment or a student’s personal attributes?

Ask students to keep a record of the extent to which and in what settings they observe other students talking politics.

Exercise 8-4.
Answers will vary depending on meeting attended and hypothesis. Students should demonstrate that they made observations relevant to their hypothesis.

Exercise 8-5.
a. Students should record observations of the interaction between political actors.
b. Students should list three hypotheses that are in related to the observations they made.

Exercise 8-6.
a. The description of the laboratory observation plan should indicate a decision participants must make, like who to vote for, which policy to support, etc. It should also indicate how information will be easy or hard to collect for different participants. Perhaps some get a series of URL addresses and other have to search for information—something like that.
b. Students should be able to offer a reasonable explanation for the choice of design.

Exercise 8-7.
a. Students should indicate when, where and for how long they made observations. They should also include observation notes about what they observed.
b. Students should make interpretations that are related to the notes they made.
c. There are many possible answers. Students might identify student groups they might speak with, events they might attend, other places on campus they might make observations or administrators or faculty they might wish to speak with.

Chapter 9: Document Analysis: Using the Written Record

Exercise 9-1.
a. An example hypothesis might be: written addresses will have more words while spoken addresses will have fewer words.
b. An example hypothesis might be: presidents invite more guests over time.
c. An example hypothesis might be: Democratic presidents invite a higher proportion of women to men than Republican presidents.
d. Names can be tricky. Some names are used by both men and women. Coding based on name alone might create bias in the data and conclusions drawn from the data would lead to biased conclusions. Inferences would therefore be compromised depending on the quality of the coding.

**Exercise 9-2.**
a. Each of the examples (words, sentences, paragraphs) are equally valid. Students should demonstrate that they have deliberated over the issue and come up with a logical explanation for their choice. For example, if paragraph is chosen how would the first paragraph be coded? It includes references to both positions and candidates.

b. Again, either choice is correct. Students should have a logical explanation for their choice.

c. Students might consider using a time period dictionary.

d. In a platform like this, all the words and phrases are positive in nature about the party’s own candidate so students should be looking for how to differentiate between effusive praise and simple acknowledgement, or something similar. For example, Theodore Frelighuysen’s service as Attorney General is described only by a count of the number of years of service. This would be an acknowledgement. But, “periled life in the contest for liberty,” is much more effusive.

**Exercise 9-3.**
a. List of five characteristics. They should be characteristics that are commonly found in ads like the examples given.

b. The explanations should be sufficient for applying the coding scheme.

c. Students should record their data.

d. Students should demonstrate deliberation and critical thought in their improvements.

**Exercise 9-4.**
a. Students should complete the table.

b. A suitable hypothesis is something like an increase in population causes an increase in CO2 emissions.

c. There are a great many answers here. Each of the four should be plausibly related to climate change.

**Exercise 9-5.**
a. When analyzing a single narrow topic, having access to all of the stories will allow the researcher to have a more complete and deeper understanding of the topic. If story content within a single issue area is largely similar or homogenous, then a random sample will work quite well. If, however, there is a great deal of difference between news articles on the subject, or heterogeneity, then a sample will not be as beneficial.

b. If one is interested in analyzing front page coverage, there would be a greater opportunity for nuclear proliferation to appear on the front page in years when there are
more stories per page. There would be a lesser opportunity for nuclear proliferation to appear on the front page in years when there are fewer stories per page.

c. More pages in the index means that more items were included in the sample and fewer pages means fewer items were sampled. The varying number of entries per year may influence the number of stories about each subject in a particular year introducing bias.

Exercise 9-6.
a. Consistency and reliability mean that when multiple coders code material, they code the in the same manner and end up with similar results. This is important because if multiple coders are coding content differently, the results from each coder may be substantially different and the data as a whole and the conclusions from it will be biased.

b. The student should make a logical argument. Pew is a well-respected organization and their decision to use 80% was not taken lightly.

c. The student may pick any of the three limitations and explain why it is the most important. There should be a logical explanation.

Chapter 10: Survey Research and Interviewing

Exercise 10-1.
a. Students should supply a closed-ended question with answer choices. Students should include just one party per choice, and include an “other” and “none” to make the choices exhaustive.

b. An open ended question without answer choices.

c. the question should highlight the virtues of the Republicans or ills of other parties to lead the respondents.

Exercise 10-2.
a. This question would be appropriate to ask in a closed-ended format, although response categories should include “don’t know” in addition to “yes” and “no.”

b. An open-ended question would reveal how much the respondent has thought about the issue and is aware of possible actions. Asking closed-ended questions might not obtain valid and reliable responses if the action is not well-understood by the respondent. Better responses might be obtained if the interviewer provided some background information.

c. A well-worded closed-ended question could elicit for what reasons or under what circumstances a respondent would be in favor of reestablishing section 4. Some amount of explanation might be needed on section 4. An open-ended question might just reveal a
yes or no response, unless an interviewer could follow up with probing questions about reasons for support or opposition.

d. A closed-ended question such as “Do you approve or disapprove of the way Congress is handling its job?” (This question came from a Quinnipiac University poll, Oct. 23–29, 2007.)

e. Open-ended would work, although some people might not be willing to admit they didn’t work. Ready response categories such as “less than 10 hours per week” or “temporarily unemployed/between jobs” might help this problem.

f. Opened-ended question.

g. A rating scale such as a 100-point feeling thermometer.

h. An open-ended question would allow respondents to indicate the level of support they have for restructuring Medicare. The answers might be tedious to code if respondents give qualified yes or no answers. It would probably be better to use a closed-ended question in which the respondent is given choices that reflect the most typical options or circumstances.

Exercise 10-3.

a. Although this question appears to be objective and straightforward, many “average” citizens may not be familiar with the subject. At a minimum provide an opportunity for the respondent to say or check “don’t know” or “haven’t heard too much about…” Otherwise, some unknown percentage of people may say “favor” or “oppose” in order to appear knowledgeable or to try to cooperate with the interviewer. It is possible to frame the question by providing a short definition. “There has been some discussion these days about the so-called flat tax, which is…” In short, this question cannot stand on its own.

b. This question seems clearly reactive because of the phrase “most of them.” Who wants to be for something most everyone else opposes?

c. This question seems very ambiguous—how much would be sold is not clear. Also, can one assume that most respondents will be familiar with these financial terms, and would the mention of children’s health care trigger a favorable response?

d. The statement about terrorist attacks is likely to prime the respondent for answering “yes,” because it raises what might be painful memories of a previous attack. Without the statement before the question, a researcher may find different results.

e. Direct questions like this one probably do not solicit much useful information. Discuss the randomized response technique. Students could review studies of substance abuse to see how researchers handle the problem.

Exercise 10-4.
The definitions of each concept will vary, as will the survey questions. More concrete concepts like party identification and level of participation will be easier to write questions for than the more abstract concepts like ideology. Students should pay special attention to the abstract terms. The key to each response is that the student provides a relatively clean definition—one that leaves very little room for debate about the meaning of the term. Once well-defined, the survey question should be easier to write. The explanation for content validity should focus on the student’s definition and measurement.

Exercise 10-5.

a. One cannot simply translate questions word for word. The meanings of the words and phrases may change in different languages. The key is maintaining validity. Is the question, in its various language versions, measuring the same concept? Will respondent’s using different languages understand the same question in the same way? If so, the question is useful comparing answers across cultures/countries. If not, the question is not useful.

b. Pew uses previous translations to maintain comparability. A different translation might lead to different answers and researchers could not compare the responses over time.

c. Local organizations work with local populations on a daily basis and better understand the uniqueness of the population. Cellphones may be more widely relied upon in some countries than in others. Different hours of the day may be more suitable for surveys in some areas than others. In some countries surveys must be done door to door and that may require special attention to the kinds of questions that are asked. Local organizations can make Pew aware of these issues and local customs or norms.

d. A shorter time period means major world events are less likely to occur. These events might change respondent’s answers so that answers given after the event are biased by the event creating bias in the results.

Exercise 10-6.

a. The results are different because of the additional “somewhat” categories. The somewhat categories are often interpreted as more moderate levels of support which might cause the weakest supporters to choose “somewhat disapprove” and be lumped in with those who more strongly disapprove when categories are collapsed.

b. By changing the question wording we can sometimes change the responses. Therefore, it is important to not only examine the results of a poll question but to also pay attention to the wording of the question to see what was asked. Or, if someone wanted to find more or less support for the president, they might change the wording of the question slightly to reach the desired result.

Exercise 10-7.
The answers will vary with the content of the survey questions but the rules broken are limited to the tips on writing good questions in the text. There should be plenty of double-barreled questions, double negatives, vague references to complex concepts, questions about the distant past or other information the respondent is not likely to know, closed ended questions with answer categories that are not exclusive or exhaustive, etc. The important part of the assignment is that the students can explain why they have broken the rules.

**Chapter 11: Making Sense of Data: First Steps**

*Note to instructors:* These problems demand that the student read the instructions carefully, take time finding the solutions, and be neat and organized. Applied statistics entails judgments. Many questions can be answered in different ways. Some questions explicitly call for discussion. What is important is justifying reasons with accurate calculations and knowledge of the concepts.

Depending on your class’s level and background you will of course want to pick and choose exercises.

**Exercise 11-1.**

a. Relative frequency: Percentage or proportion of total number of observations in a frequency distribution that have a particular value. Cumulative frequency adds the relative frequencies of each category—present plus previous—and so on. The difference is that relative frequency only includes one category while cumulative is additive.

b. Frequency is a count of observations in each category while proportion is a percentage of the total for the category. The key difference is count vs. percentage of total.

c. Central tendency measures the typical value while dispersion measures the distribution of values around the typical case. The key difference is that while both are classes of descriptive statistics, central tendency describes the center of a frequency distribution while dispersion describes the rest of distribution—in general.

d. The range measures the distance between the minimum and maximum values. The interquartile range measures the distance between the first and third quartiles. The key difference is that the interquartile range examines the middle of the frequency distribution instead of the whole distribution.

e. A frequency distribution plots the variable values on the x axis and the number of
observations for each value on the y axis. A normal distribution is one way
variables can be distributed with known properties. The difference is that the
normal distribution describes the properties of a variable while a frequency
distribution describes only the number of observations for each value of the
variable.

Exercise 11-2.
a. The mean would be inappropriate for a nominal level variable. Mode is the best choice.
b. The mode is the best choice for this nominal level variable.
c. The mean is the best choice for this ratio level variable. We would be wary of the
possibility of states with large senior populations like Florida skewing the data. In that
case, median might be better.
d. Mean could be misleading if there are extreme outliers with very high incomes.
Median might be better.
e. Median would probably be the best choice for the ordinal level variable. Mode might
be useful as well.
f. Mode would not be the best choice for this interval level rating scale. Mean or median
would be better. Likely too much variation in scores for mode to be meaningful.

Exercise 11-3
a. The mean, median and mode all equal 5.
b. The range for sample 1 is 0, and for sample 2 is 9. The standard deviation for sample 1
is 0 and the sample standard deviation for sample 2 is 3.24.

Sample 2 equation:

<table>
<thead>
<tr>
<th>x</th>
<th>$x_i - \bar{x}$</th>
<th>$(x_i - \bar{x})^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-4</td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>42</td>
</tr>
</tbody>
</table>

$$\hat{\sigma} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

$$\hat{\sigma} = \sqrt{\frac{42}{4}}$$
\[ \hat{\sigma} = \sqrt{10.5} \]

\[ \hat{\sigma} = 3.24 \]

c. It is important to calculate both because central tendency only describes the typical case in a variable. Dispersion describes how the rest of the observations are distributed around the typical case. In this example, dispersion differentiates between the two samples while central tendency did not.

**Exercise 11-4.**

a. | Response | Frequency | Percent | Valid Percent | Cumulative Percent |
---|---|---|---|---|
1 (decrease) | 82 | 7.03 | 8.18 | 8.18 |
2 | 82 | 7.03 | 8.18 | 16.36 |
3 | 128 | 10.97 | 12.76 | 29.12 |
4 | 280 | 24.00 | 27.92 | 57.04 |
5 | 210 | 17.99 | 20.94 | 77.98 |
6 | 128 | 10.97 | 12.76 | 90.74 |
7 (increase) | 93 | 7.97 | 9.27 | 100.01* |
Don’t know-Missing | 164 | 14.05 | — | |
Total | 1,167 | 100% | 100% | |

* More than 100 due to rounding.

b. The modal category is in the middle of the scale: number 4.

**Exercise 11-5.**

1) Select if SOUTH=0 (37 states remain)

   Select if ROMNEYVOTE\_PERCENTAGE\_2012 > 45 (20 states remain)

2) PA, OH, IN, AZ, WI; OK, NE, KS, ID, UT

3) HIGHTURNOUT should have five high turnout one’s and fifteen low turnout zero’s.

4&5) 5 high turnout; 15 otherwise
6 and 7) Means: Romney 55.66; Obama 42.14

Mean Vote for Romney and Obama in Select States, 2012

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent vote Romney 2012</td>
<td>20</td>
<td>55.66</td>
</tr>
<tr>
<td>Percent vote Obama 2012</td>
<td>20</td>
<td>42.14</td>
</tr>
<tr>
<td>Valid N (total)</td>
<td>20</td>
<td></td>
</tr>
</tbody>
</table>

8) Chart, Table and titles should be submitted with the workbook page.

Exercise 11-6.

<table>
<thead>
<tr>
<th>X</th>
<th>X_i-Xbar</th>
<th>(X_i-Xbar)^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1.5</td>
<td>2.25</td>
</tr>
<tr>
<td>3</td>
<td>0.5</td>
<td>0.25</td>
</tr>
<tr>
<td>0</td>
<td>-2.5</td>
<td>6.25</td>
</tr>
<tr>
<td>1</td>
<td>-1.5</td>
<td>2.25</td>
</tr>
<tr>
<td>7</td>
<td>4.5</td>
<td>20.25</td>
</tr>
<tr>
<td>3</td>
<td>0.5</td>
<td>0.25</td>
</tr>
<tr>
<td>1</td>
<td>-1.5</td>
<td>2.25</td>
</tr>
<tr>
<td>0</td>
<td>-2.5</td>
<td>6.25</td>
</tr>
<tr>
<td>5</td>
<td>2.5</td>
<td>6.25</td>
</tr>
<tr>
<td>1</td>
<td>-1.5</td>
<td>2.25</td>
</tr>
</tbody>
</table>

Xbar = 2.5

Σ(X_i-Xbar)^2 = 48.5

σ^2 = 4.85

s^2 = 5.39

σ = 2.20

s = 2.32

The population calculations include the sample size in the denominator while the sample calculations include the sample size minus one in the denominator. We subtract one in
the denominator in the sample version to account for our greater uncertainty when using sample data. Mathematically, the sample versions of the statistic will therefore always be larger because we are dividing by a smaller number. Theoretically, this makes sense because if we have less certainty with sample data, the values of the variance and standard deviation should be larger reflecting that uncertainty.

**Exercise 11-7.**
The purpose of this question is to encourage students to think about the nuance of central tendency and dispersion statistics. The US is a clear outlier and affects the mean and range.

a. Mean = 98.83

b. Median = 46.5

c. Trimmed mean = 55.8

d. The difference between the mean and median is enormous because of the US outlier. The mean is twice as large as the median. The trimmed mean or the median are probably the best choices to represent the typical case in this example. Students should have a logical explanation for their choice. If students are concerned about China as an additional outlier then the median is the best choice.

e. Maximum = 581; minimum = 23; range = 558

f. Q1 = 28, Q3 = 75.5, IQR = 47.5

g. The value of the interquartile range is that it is not affected by the US outlier and gives a clearer picture of dispersion.

**Exercise 11-8.**
a. Mean = 5

b. Population Variance = 14.86

c. Population Standard deviation = 3.85

d. There were an average of five executions in 2014 in states that executed prisoners with a standard deviation of 3.85 executions.

<table>
<thead>
<tr>
<th>x</th>
<th>x_i - \bar{x}</th>
<th>(x_i - \bar{x})^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-4</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>-3</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>1</td>
<td>-4</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>-2</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>25</td>
</tr>
</tbody>
</table>
Population standard deviation (and variance)

$$
\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{N}}
$$

$$
\sigma = \sqrt{\frac{104}{7}}
$$

$$
\sigma = \sqrt{14.86}
$$

$$
\sigma = 3.85
$$

**Exercise 11-9.**
a. Mean = $8.10
b. Sample Variance = 0.45
c. Sample Standard Deviation = $0.67
d. The average minimum wage is $8.10 with a standard deviation of $0.67.

<table>
<thead>
<tr>
<th>State</th>
<th>x</th>
<th>$x_i - \bar{x}$</th>
<th>$(x_i - \bar{x})^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alaska</td>
<td>8.75</td>
<td>0.65</td>
<td>0.42</td>
</tr>
<tr>
<td>Colorado</td>
<td>8.23</td>
<td>0.13</td>
<td>0.02</td>
</tr>
<tr>
<td>Idaho</td>
<td>7.25</td>
<td>-0.85</td>
<td>0.72</td>
</tr>
<tr>
<td>Illinois</td>
<td>8.25</td>
<td>0.15</td>
<td>0.02</td>
</tr>
<tr>
<td>Nebraska</td>
<td>8</td>
<td>-0.1</td>
<td>0.01</td>
</tr>
<tr>
<td>Rhode Island</td>
<td>9</td>
<td>0.9</td>
<td>0.81</td>
</tr>
<tr>
<td>Virginia</td>
<td>7.25</td>
<td>-0.85</td>
<td>0.72</td>
</tr>
</tbody>
</table>

Sample standard deviation (and variance)

$$
\hat{\sigma} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}
$$

$$
\hat{\sigma} = \sqrt{\frac{2.72}{6}}
$$

$$
\hat{\sigma} = \sqrt{.45}
$$

$$
\hat{\sigma} = .67
$$
Exercise 11-10.

a. B
b. A
c. They don’t; both distributions are skewed.
d. 100
e. 100

Exercise 11-11.
The authors don’t explain how they arrived at their numbers. Yet it seems clear there is some selective use of the mean in these arguments. The White House no doubt calculates the total tax cuts and divides by the number of people receiving them. If a few people were at the extreme end of the distribution while everyone else was in the lower part, the average (mean) would be higher than the median. The critics are probably looking at only those below a certain income level such as $75,000 and finding for this group that the cuts average only about $289. So both sides speak the truth, but they are talking about different things. It is not too hard for ordinary citizens to sort out this matter if they understand, even roughly, the properties of the arithmetic mean.

Exercise 11-12.

a. Standard deviation of the sample values is a measure of variation around the sample mean. It is calculated by summing the squared deviations around the mean and taking the square root of the total.

b. The 50th percentile; half the observations lie below and half lie above the median.

c. Most frequent observation.

d. The mean with a given percentage of observations deleted from each end of the distribution. Resistant to [not as affected by] deviant value.

e. The difference between the 75th and 25th percentiles. Half the observations lie between these two quartiles.

Exercise 11-13.

Min = 66; Max = 86; Median = 78
1st Quartile = 76; 3rd Quartile = 83; IQR = 7

1.5 * IQR = 10.5

Boxplot of Wage Gap in Select States, 2013

LW Q1 MD Q3 UW
Chapter 12: Statistical Inference

Exercise 12-1.
The students should stress in their examples that a type 1 error is the incorrect or mistaken rejection of a true null hypothesis. The consequence is that the researcher mistakenly believes that there is a relationship between two variables when in the true state of the world, there is not. This can lead to a great deal of wasted effort or resources going forward with this incorrect understanding. The $100 million is wasted on a program that does not increase turnout.

The consequence of a type II error is that a program that would have worked is not tried because the null hypothesis is incorrectly accepted. On the upside, the money is not wasted and perhaps the value of a civics class would be unearthed in another project.

Exercise 12-2.
Each student’s coin toss will be slightly different, but students should be able to demonstrate in their bar graphs an emerging normal distribution. The written answer should indicate that given an endless number of independent samples of size N from a fixed population that has a mean of μ and a standard deviation σ. Each time, you calculate the sample mean, $\bar{Y}$, and the standard deviation, $\hat{\sigma}$. In the end, you will have literally thousands of $\bar{Y}$’s. If you drew a graph of their distribution, the result would be what is called a sampling distribution. How do we know? Statistical theory demonstrates that the distribution of sample means is normal with these two characteristics: the mean of the sampling distribution would be μ, the population mean; and its standard deviation would be $\sigma/=\sqrt{N}$.

Exercise 12-3.
1. no
2. yes
3. no
4. no
5. yes
6. no
7. no
8. yes
9. no
10. yes

Exercise 12-4.
- a. $H_0: \mu = 8287$
- b. $H_A: \mu < 8287$
- c. Because $N$ is relatively small, a small-sample test of a mean or $t$ test would be used.
- d. The $t$ distribution

Exercise 12-5.
- a. The distribution should be labeled to indicate the shaded area (the right side of the distribution beyond 7) under the curve along with the mean and value of interest. It should also include the x and z scales.
- b. The mean is 6.33, the population standard deviation is 1.25.

Population standard deviation

\[
\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{N}}
\]

\[
\sigma = \sqrt{\frac{14.01}{9}}
\]

\[
\sigma = \sqrt{1.56}
\]

\[
\sigma = 1.25
\]

c. The Z score is .536

Z score

\[
Z = \frac{(x_i - \mu)}{\sigma}
\]

\[
Z = \frac{(7 - 6.33)}{1.25}
\]

\[
Z = \frac{(.67)}{1.25}
\]

\[
Z = .536
\]
d. \( p = .2946 \). There is a 29.46% chance of observing seven speeches or more.

**Exercise 12-6.**

A. .1151  
B. .0122  
C. .55  
D. 2.5  
E. 2.228  
F. 2.602  
G. The table represents half of the distribution. A full distribution equals a probability of 1, so half is .5.  
H. A larger \( z \)-score is further from the mean so there is a lower associated probability.  
I. The degrees of freedom allow us to account for the size of the sample. A larger sample provides more certainty so the \( t \)-score is lower. Fewer degrees of freedom represent less certainty so the \( t \)-score is bigger.  
J. The \( t \)-score is 1.96 and the probability is .025. This is an example of how the \( t \)-distribution equals the normal distribution with a sufficiently large sample. \( 2 \times .025 \) equals .05, which was the alpha level used to find the \( t \)-score.

**Exercise 12-7.**

<table>
<thead>
<tr>
<th>i</th>
<th>( x )</th>
<th>( x_i - \bar{x} )</th>
<th>( x_i - \bar{x}^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>-3</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

\( \bar{x} = 3 \)

\( \Sigma (x_i - \bar{x})^2 = 20 \)

\( s = 2 \)

\( t = 2.44 \)

The \( t \) value is statistically significant when using the one tailed test but not when using the two tailed test. This demonstrates that the two tailed test divides the critical range in half with half distributed on each end of the distribution while the one tail test has all of the critical range in one tail, leading to a lower hurdle for statistical significance.

**Exercise 12-8.**

\( X \quad x_i - \bar{x} \quad x_i - \bar{x}^2 \)
Xbar = 4.8

Σ(x_i-xbar)^2 = 71.6

s = 2.82

t= -.78

Accept the null hypothesis. The two tailed test should be used because a direction of is not specified.

Exercise 12-9.

<table>
<thead>
<tr>
<th>Country</th>
<th>X</th>
<th>x_i-xbar</th>
<th>x_i-xbar^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil</td>
<td>8</td>
<td>-1.2</td>
<td>1.44</td>
</tr>
<tr>
<td>Canada</td>
<td>16</td>
<td>6.8</td>
<td>46.24</td>
</tr>
<tr>
<td>Denmark</td>
<td>11</td>
<td>1.8</td>
<td>3.24</td>
</tr>
<tr>
<td>Ghana</td>
<td>5</td>
<td>-4.2</td>
<td>17.64</td>
</tr>
<tr>
<td>Kuwait</td>
<td>6</td>
<td>-3.2</td>
<td>10.24</td>
</tr>
<tr>
<td>India</td>
<td>9</td>
<td>-0.2</td>
<td>0.04</td>
</tr>
<tr>
<td>Malaysia</td>
<td>1</td>
<td>-8.2</td>
<td>67.24</td>
</tr>
<tr>
<td>South Africa</td>
<td>8</td>
<td>-1.2</td>
<td>1.44</td>
</tr>
<tr>
<td>South Korea</td>
<td>10</td>
<td>0.8</td>
<td>0.64</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>18</td>
<td>8.8</td>
<td>77.44</td>
</tr>
</tbody>
</table>

Xbar = 9.2

Σ(x_i-xbar)^2 = 225.6

s = 5.01

t= -2.51
There is a statistically significant difference between the sample and population mean so the student should reject the use of sample data. It is not representative of the population data.

**Exercise 12-10.**

A. A larger standard deviation indicates that there is more variation in the data and therefore we have less certainty. Thus, the larger standard deviation will cause the standard error to be larger giving us a lower likelihood of finding statistical significance.

B. A larger sample size indicates that we have more certainty that the data is representative of the population. Thus, the larger sample size will lead to smaller standard deviation and as well as a larger denominator in the standard error in the t or z test calculation. Both will lead to a smaller denominator in the t or z test calculation and will in turn produce a larger observed test value.

C. A larger difference of the means indicates that there is a bigger difference between the sample and population mean and the bigger the difference the more likely the difference is statistically significant.

D. The standard error of the mean is the standard deviation divided by the square root of the sample size. This value tells us about the certainty we have in our data. Both the standard deviation and sample size tell us about the data. A larger standard deviation means more variation and less certainty. A larger sample size means more certainty.

**Exercise 12-11.**

*Population Confidence Interval*

The base confidence interval is $4.82 < \mu < 5.18$. The standard error of the mean is .18 and the Z score is 1.96. The confidence limit is .18.

\[
CI = 5 \pm (Z_{\alpha/2})(\sigma/\sqrt{n})
\]

\[
CI = 5 \pm (1.96)(2/\sqrt{500})
\]

\[
CI = 5 \pm (1.96)(0.09)
\]

\[
CI = 5 \pm (0.18)
\]

\[
CI = (4.82 < \mu < 5.18), .05
\]

*Changing Standard Deviation*

Increasing the standard deviation doubles the standard error to .18 and increases the confidence limit to .35. This widens the confidence interval to $4.65 < \mu < 5.35$. A larger standard deviation means we have less certainty with the added variability—and the confidence interval must be wider to reflect that uncertainty.

\[
CI = 5 \pm (Z_{\alpha/2})(\sigma/\sqrt{n})
\]

\[
CI = 5 \pm (1.96)(4/\sqrt{500})
\]
Changing Sample Size

Increasing the sample size decreases the standard error to .06 and decreases the confidence limit to .06. This narrows the confidence interval to $4.88 < \mu < 5.12$. A larger sample size means we have more certainty—and the confidence interval is narrower to reflect that certainty.

Changing Confidence Level

Using a lower confidence level of .9 will cause the width of the confidence interval to shrink because we will have less confidence that the population mean lies within the interval. The $z$ score is now 1.64 which changes the interval to $4.85 < \mu < 5.15$.

Exercise 12.12.

<table>
<thead>
<tr>
<th>X</th>
<th>$x_i-x\bar{x}$</th>
<th>$x_i-x\bar{x}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>-4</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>-1</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>-2</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>-1</td>
</tr>
</tbody>
</table>
\[ \hat{\sigma} = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}} \]

\[ \hat{\sigma} = \frac{56}{7 - 1} \]

\[ \hat{\sigma} = \sqrt{9.33} \]

\[ \hat{\sigma} = 3.05 \]

\[ CI = 7 \pm (t)(\hat{\sigma}/\sqrt{n}) \]

\[ CI = 7 \pm (2.447)(3.05/\sqrt{7}) \]

\[ CI = 7 \pm (2.447)(1.15) \]

\[ CI = 7 \pm (2.81) \]

\[ CI = (4.19 < \mu < 9.81), .05 \]

The hypothesized value of 9 lies within the confidence interval so we can say that there is not a statistically significant difference between the observed value mean of 7 and the hypothesized value of 9 at the .05 level using a two tailed test. This means that we cannot reject 9 as a possibility.

**Exercise 12-13.**

a. The t observed = -2.201; the df = 186; the mean difference is -1.42; the test value is statistically significantly different from the mean of 70.58. We can reject the null hypothesis and accept the alternative hypothesis.

**Exercise 12-14.**

a. Proportion (mean) = .206; Sample standard deviation = .405

b.

\[ t_{obs} = \frac{\hat{p} - P}{\sigma} \]

\[ t_{obs} = \frac{.206 - .2}{.405} \]

\[ t_{obs} = .06 \]

\[ t_{obs} = .405 \]

\[ t_{obs} = .148 \]
c. The $t$ observed value of .148 is not greater than the $t$ critical of 1.96. This indicates that there is not a statistically significant difference between the hypothesized value of .2 and the sample proportion of .206. The student should accept the professor’s assertion as valid.

**Exercise 12-15.**

a. mean = 24.21, standard deviation = 6.87  
b. standard error of the mean = .97  
c. $(22.26 < µ < 26.16), .05$

---

**Chapter 13: Investigating Relationships between Two Variables**

**Exercise 13-1.**

a. Data that are recorded using categories.

b. Continuous data would create far too many rows or columns—potentially thousands or millions—for meaningful analysis. The categories in categorical data is much more manageable to put in a table.

c. The column percentages must equal 100 because it is the sum of the percentages of cases within each category. It is necessary to use the column totals as the bases (denominators) for the percentage calculations.

d. One would use Gamma to measure the strength and direction of a relationship between ordinal variables.

e. $1 =$ perfect positive relationship, $-1 =$ perfect negative relationship.

f. An absolute value of greater than .4 or .5 merits consideration as a strong relationship.

**Exercise 13-2.**

a. Strong negative monotonic correlation. A linear relationship would fit these data—the correlation coefficient would be $- .9$ or “higher”—but there is a discernible curve.

b. Strong positive monotonic or linear correlation. It is hard to tell. Moreover, the text stresses that statistics involve an element of judgment. But at least two terms should be part of the answer: “strong” and “positive.”

c. Moderate to weak positive correlation.

d. Strong curvilinear relationship.

**Exercise 13-3.**

a.  

<table>
<thead>
<tr>
<th>Participation</th>
<th>White</th>
<th>Non-white</th>
</tr>
</thead>
</table>

b. Modest relationship; 75% of whites vote versus 60% for non-whites.

c. Probabilities are .60 and .75, respectively.

d. \( \frac{150}{50} = 3 \) or 3 to 1; \( \frac{30}{20} = 1.5 \) or 1.5 to 1.

e. \( \frac{3}{1.5} = 2 \). The odds of whites voting are twice those of non-whites. There is thus an association. Note: Students may inadvertently reverse the question and obtain the odds ratio of not voting (50/150 = .33 and 20/60 = .67 for an odds ratio of .33/.67 = .5 [about]). The odds of whites not voting are about half those of non-whites, as expected. Other students may report the ratio of probabilities or the relative risk of voting [not voting], but this is not what the question asks.

**Exercise 13-4.**
a. Yes, Democratic senators tend to represent states in the Northeast and the Midwest, while Republican senators tend to come from the South and the West. Almost 70% of the senators from the Northeast and 54% of those from the Midwest are Democrats, while about 63% of the senators from the South are Republicans, as are about 62% of those from the West.

b. Republican, because there are more Republicans (50) than Democrats (49).

c. Lambda, because both variables are nominal.

d. Lambda = \( \frac{(49 - [7 + 9 + 12 + 10])}{49} = \frac{(49 - 38)}{49} = 11/49 = .2245 \).

**Exercise 13-5.**
a. The independent, or explanatory or outcome, variable is ideology and the dependent variable is frequency of praying. The more conservative a person is, the higher that person’s reported frequency of praying. As many as 41.7% of conservatives, 23.9% of moderates, and 18% of liberals report praying several times a day. In contrast, 30.1% of liberals, 24.2% of moderates, and 10.6% of conservatives report praying less than once a week.

This is a good example of the need to examine a table carefully and fully. Liberals and conservatives differ at the “extremes.” If, for instance, you divide the table into two parts (or more formally, partition the table and associated chi-square), it’s clear that the ideological groups are pretty much the same in the middle; there is little to no relationship. Looking at rows 1 and 5, by contrast, we see great differences. So we suspect that most of the association is due to those who are very religious and those who
are much less so. (We notice also that the sample is divided about in half on these two levels. It is very surprising [to us at least] that those who admitted “never” and “less than once a week” are almost as numerous as the “several times a day” people.) One possible interpretation is that the “religious center” really doesn’t divide on ideology. In short, the data illustrate the value of and the need for looking at data as a whole.

b. Gamma is an appropriate measure of association because both variables are ordinal. Gamma indicates a weak-to-moderate relationship. If we know the political orientation of two people, we can improve our prediction of which person prays more frequently over our prediction if we don’t know their political orientation. If rewritten, gamma can be given a proportional-reduction-in-error (PRE) interpretation. But as we define it, gamma doesn’t have the neat interpretation one might want. But neither does its cousin $r$. Gamma has a negative sign because frequency of praying is coded going from higher frequency to lower frequency. Political orientation is coded going from liberal to conservative. Therefore the cases are clustered along the diagonal going from bottom left to upper right.

**Exercise 13-6.**

a. Observed chi square is 9.74

$$\frac{[(20-25)^2}{25}+ [(30-25)^2}{25}+ [23-29)^2}{29}+ [35-29)^2}{29}+ [(57-46)^2}{46}+ [(35-46)^2}{46}]$$

$$\frac{([-5)^2}{25}+ \frac{[(5)^2}{25}+ [(-6)^2}{29}+ [(6)^2}{29}+ \frac{[(11)^2}{46}+ [(-11)^2}{46}]$$

$$\frac{[25/25]+ [25/25]+ [36/29]+ [36/29]+ [121/46]+ [121/46]}$$

$$1 + 1 + 1.24 + 1.24 + 2.63 + 2.63$$

9.74

b. There is 1 degree of freedom.

c. Critical chi square is 3.84

d. The variables are not statistically independent. Observed chi square is greater than critical chi square.

**Exercise 13-7.**
a. Comparison of means or analysis of variance.

b. The comparison of means: Is “type” of prison related to the number of inmate deaths.

b. Apparently yes, because the mean death rate was much higher in state institutions than in federal prisons. Both had higher rates than private facilities. At first glance, the private prisons seem safest.
d. The null hypothesis can be stated as the (population) means are equal. The alternative hypothesis, the one that seems to underlie the introduction, is that privately run institutions will have higher death rates than publicly managed prisons.

e. The $F$ test shows a statistically significant result. The observed $F$ is, of course, significant at the .0006 level. Two or more of the means probably differ in the “population.” That is all that can be concluded.

f. There are indeed differences among the types of facilities but not in the way suggested in the question setup. In the absence of other information, we can say that state prisons are considerably more “dangerous” than federal or private institutions. Private prisons, in fact, appear to be safe havens compared with public correctional facilities. Part of the reason for this finding may lie in the type of people being housed in each type. Perhaps states assign young, comparatively well-behaved inmates to private operators. State penitentiaries, on the other hand, may hold older and repeat offenders. The data by themselves offer no clues. But this might be an interesting topic for class discussion.

Exercise 13-8.

a.

PIAL.1e ...Have you encouraged anyone to vote for Barack Obama or Mitt Romney by -Posting on a social networking site such as Facebook or Twitter- in the last 30 days? * PARTY. In politics TODAY, do you consider yourself a Republican, Democrat, or Independent? Cross tabulation

<table>
<thead>
<tr>
<th>Count</th>
<th>PARTY. In politics TODAY, do you consider yourself a Republican, Democrat, or Independent?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Republican</td>
</tr>
<tr>
<td>PIAL.1e ...Have you encouraged anyone to vote for Barack Obama or Mitt Romney by -Posting on a social networking site such as Facebook or Twitter- in the last 30 days?</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Total</td>
<td>285</td>
</tr>
</tbody>
</table>
Chi-Square Tests

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>df</th>
<th>Asymptotic Significance (2-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Chi-Square</td>
<td>2.598a</td>
<td>2</td>
<td>.273</td>
</tr>
<tr>
<td>Likelihood Ratio</td>
<td>2.568</td>
<td>2</td>
<td>.277</td>
</tr>
<tr>
<td>Linear-by-Linear</td>
<td>.175</td>
<td>1</td>
<td>.675</td>
</tr>
<tr>
<td>N of Valid Cases</td>
<td>914</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 52.39.

b. 

\[
\left[\frac{(50-52.39)^2}{52.39}\right] + \left[\frac{(51-57.35)^2}{57.35}\right] + \left[\frac{(235-236.61)^2}{236.61}\right] + \left[\frac{(250-258.73)^2}{258.73}\right] + \left[\frac{(261-254.65)^2}{254.65}\right] \\
\left[\frac{(-2.39)^2}{52.39}\right] + \left[\frac{(8.73)^2}{58.27}\right] + \left[\frac{(-6.35)^2}{57.35}\right] + \left[\frac{(2.39)^2}{236.61}\right] + \left[\frac{(-8.73)^2}{258.73}\right] + \left[\frac{(6.35)^2}{254.65}\right] \\
\left[\frac{5.71}{52.39}\right] + \left[\frac{76.21}{58.27}\right] + \left[\frac{40.32}{57.35}\right] + \left[\frac{5.71}{236.61}\right] + \left[\frac{76.21}{258.73}\right] + \left[\frac{40.32}{254.65}\right] \\
2.59
\]

Exercise 13-9.

a. Party leaders’ (that is, delegates’) issue positions differ from those of “average party members.” (Actually, Herrera’s main concern was with the perceptual accuracy of “superdelegates” as opposed to those chosen in caucuses, conventions, and primaries.)

b. In general, the null hypothesis is that the difference of means in the population is 0; that is, \( \mu_{\text{voters}} - \mu_{\text{delegates}} = 0 \) or \( \mu_{\text{voters}} = \mu_{\text{delegates}} \), where \( \mu \) is a population mean attitude.

c. In view of Herrera’s notation, all of them could be rejected at the .01 level. We hope students will be able to express the idea that “significance” means that the chances of observing these differences or larger ones is pretty small if, in the U.S. population, leaders and followers really do not disagree.
d. You could use this question to bring together lots of issues in empirical research. The statistical analysis leads to a rejection of the null hypotheses; consequently, the research proposition is supported. There is a difference in the political “centers of gravity” of leaders and partisans. (Note that Herrera’s research is much more sophisticated and involves many other issues.)

Lots of analysts might leave the matter at that. But, you can think about, say, Chapter 10. Are the different kinds of respondents interpreting the seven-point questions the same way? Do they share a frame of reference? More important, what does a one-point difference of means say about American politics? You could ask students, for instance, if party leaders really are out of step with the rank and file. The data seem to indicate yes. But is this difference meaningful? Many average Americans might expect and accept such gaps in attitudes. Or it’s possible that public opinion is so imperfectly embodied in political leaders that something in American democracy is amiss. Whatever the case, all of this discussion points to the same lesson: data analysis is only an aid in understanding the world; it seldom provides final answers.

Exercise 13-10.
a. BDF=k-1=3-1=2; WDF=n-k=15-3=12; TDF=n-1=15-1=14

\[
\begin{array}{ccc}
\text{D} & \text{R} & \text{I} \\
0 & 0 & 0 \\
2(4) & 0 & 0 \\
3(9) & 4(16) & 1(1) \\
5(25) & 6(36) & 3(9) \\
7(49) & 9(81) & 4(16) \\
\end{array}
\]

\[
T = (\Sigma Y) = 17 \quad 19 \quad 8
\]

\[
T = 44
\]

\[
A = \Sigma T^2/n = 17^2/5 \quad 19^2/5 \quad 8^2/5
\]

\[
= 57.8 \quad 72.2 \quad 12.8
\]

\[
A = 142.8
\]

\[
B = (\Sigma Y)^2/n = 44^2/5
\]

\[
B = 129.07
\]

\[
C = \Sigma Y^2
\]

\[
C = 246
\]

b. BSS = C-B = 246-129.07 = 13.73

c. WSS = C-A = 246-142.8 = 103.2

d. TSS = C-B = 246-129.07 = 116.93
e. \( F = \frac{[\text{BSS}/(k-1)]}{[\text{WSS}/(n-k)]} = \frac{[13.73/2]}{[103.2/12]} = 6.87/8.6 = .8 \)

f. 2 df\(_1\); 12 df\(_2\); .05 level = 3.88

g. Not statistically significant. The observed F value is not greater than the critical value.

**Exercise 13-11.**

a. Null: There is no relationship between life expectancy and mean years of schooling.
   Alternative: An increase in life expectancy causes an increase in mean years of schooling.

b. Sample output:

<table>
<thead>
<tr>
<th>Model</th>
<th>Variables Entered</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Life expectancy at birth</td>
<td>Enter</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adj R Square</th>
<th>Std Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.252</td>
<td>.064</td>
<td>.059</td>
<td>2.418</td>
</tr>
</tbody>
</table>

   a. Predictors: (Constant), Life expectancy at birth (years)

<table>
<thead>
<tr>
<th>Model</th>
<th>Out of</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Regression</td>
<td>1</td>
<td>24.841</td>
<td>1</td>
<td>24.841</td>
</tr>
<tr>
<td></td>
<td>Predicted</td>
<td>105</td>
<td>4.411</td>
<td>1</td>
<td>24.841</td>
</tr>
<tr>
<td>Total</td>
<td>174.0</td>
<td>196</td>
<td>4.411</td>
<td>1</td>
<td>24.841</td>
</tr>
</tbody>
</table>

   a. Predictors: (Constant), Life expectancy at birth (years)

   c. The coefficient is .252. A one unit increase (one year) in life expectancy causes a .252 years increase in mean years of schooling.

d. The expected positive relationship is confirmed by the positive coefficient. Substantively, the alternative hypothesis is correct. Statistically, the alternative hypothesis is correct and the null is rejected. The t score of 14.47 is larger than the t critical of 1.96.

e. The student should explain how they used the direction of the coefficient and the t-score to establish statistical significance.

**Exercise 13-12.**

a. A regression line passes straight through a cloud of data points. We can measure the distance from the mean to the regression line and from the mean to a data point—whichever we encounter first. This distance is variation explained by the regression. We can also measure the distance beyond the regression line to a data point or between the data point and the regression line—this is unexplained variation—or the residual error. Summing the residual error would equal zero. But, by squaring the residual error, we get
a sum. Different regression lines would create different sums. The smallest sum represents the least amount of unexplained variation. There is only one combination of the values that minimizes the squared errors and provides the best fit—regression gives us the best fit line by summing the squared errors.

b. Observed values of Y are those we observe in the data—the observations of the dependent variable. Predicted values are the values we would expect given a specific value of X when using the regression equation. The difference between them is that unless there is no unexplained error, the predicted values will be different than the observed values because of unexplained variation.

c. In this case there would be no error. The regression line would perfectly explain the data points that would all be lined up on the regression line or the mean.

**Exercise 13-13.**
Xbar = 3.2
Sum of X-Xbar = 55.6
Y bar = 3.6
Sum of Y-Ybar = 68.4

\[
b = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}
\]

\[
b = \frac{48.8}{55.6}
b = 0.88
\]

\[
a = \bar{Y} - (b)\bar{X}
\]

\[a = 3.6 - (0.88)3.2\]

\[a = 3.6 - 2.82\]

\[a = -0.78\]

<table>
<thead>
<tr>
<th>X</th>
<th>Yhat</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>.98</td>
</tr>
<tr>
<td>1</td>
<td>.1</td>
</tr>
<tr>
<td>1</td>
<td>.1</td>
</tr>
<tr>
<td>5</td>
<td>3.62</td>
</tr>
<tr>
<td>4</td>
<td>2.74</td>
</tr>
<tr>
<td>7</td>
<td>5.38</td>
</tr>
<tr>
<td>7</td>
<td>5.38</td>
</tr>
<tr>
<td>2</td>
<td>.98</td>
</tr>
<tr>
<td>3</td>
<td>1.86</td>
</tr>
</tbody>
</table>
Exercise 13-14.

A. R²: R-Squared. R-Squared is useful in explaining a bivariate relationship because it indicates the percentage of the variation in the dependent variable explained by variation in the independent variable.

B. r: Pearson’s r: This correlation coefficient explains how correlated two variables are by indicating the magnitude and direction of the relationship.

C. Ŷ: Predicted Y: Y hat is calculated using the regression equation. Y hat can be used to draw a regression line by plotting at least two points and drawing the regression line through the points. When compared to observed values of Y, Predicted Y can also be used to calculate residuals which are useful in examining why certain cases are better or worse explained by the regression.

D. Yᵢ: Observed values of Y: We need observed values of Y in order to analyze the variable Y. We use observed Y in calculating a regression.

E. α: The intercept. The intercept is the point on a regression plot where the regression line crosses the Y axis. If X equals zero, the predicted Y value will equal the intercept.

F. b: Regression coefficient: The regression coefficient explains the magnitude and direction of the relationship between the independent and dependent variables. It is also the slope of the regression line.

G. ε: The error term. This represents the unexplained variation in a regression. It is not observed—hence why it is unexplained.
a. It is a strong positive relationship.
b. r is .903. This demonstrates a strong positive relationship. It means that there was a higher percentage of votes in a state for Romney when there was a higher percentage of conservatives.
c. predicted Y = -12.22 + (1.574*X)
d. The coefficient of 1.574 means that a one percentage point increase in conservative caused a 1.574 percentage point increase in Romney’s vote percentage in a state.
e. Yes it is statistically significant. A t observed value of 14.521 is much greater than the t critical value of 2.021.
**Exercise 13-16.**

a. \(79.31/12.36 = 6.417\).

b. \(-6.569/2.207 = -2.427\). (Insist on the minus sign, even if students use the absolute value.)

c. Since \(N = 34\), \(df = 34 - 2 = 32\).

d. Just about 2.042. We used 30 \(df\) as an approximation, as discussed in the text.

e. Yes, because the observed \(t\) is slightly larger in absolute value than the critical value at the .05 level.

f. Since \(R^2 = .155\), the proportion of explained variation is about 15.5%.

g. \(r = -.394\). The minus sign is important, as students see next.

h. No! The coefficient is significant, which means one could argue that there is a systematic connection between stability and growth—it even might be causal—but the direction (the nature) of the effect of \(X\) on \(Y\) is the exact opposite of the prediction. A 1% increase in GDP is associated with a 6.6% decrease in stability. Rapidly developing economies experience less stability, contrary to what the researcher expected. Perhaps growth brings disruptive changes in social relationships, established traditions, and the like, all of which place strains on government.

**Exercise 13-17.**

See the boxed formula for Small-Sample Difference of Means Test. First calculate the standard error:

\[
\sqrt{\frac{(29)(16) + (24)(9)}{53}}
\]

\[
\sqrt{\frac{680}{53}}
\]

\[
3.58
\]

\[
3.58\sqrt{(1/n_1)+(1/n_2))}
\]

\[
3.58\sqrt{(.03+.04))}
\]

\[
3.58\sqrt{(.07)}
\]

\[
3.58(.26)
\]

\[
.93
\]

Difference of means divided by standard error to find Observed \(t\) value

\[
(30-25)/.93
\]

\[
5/.93
\]
5.38

Calculate degrees of freedom

(30+25-2)=53

T critical for two tailed test, 95% confidence, 53 df = 2.021

Statistically significant difference. Mean was higher for older group. Accept hypothesis.

Chapter 14: Multivariate Analysis

Exercise 14-1.
a. The implication is that exposure to televised violence might predispose people to be violent, because such exposure (1) lowers inhibitions, (2) provides behavioral models and cues, or (3) creates anxiety and hyperactivity. An alternative explanation is that individuals who are prone to violence because of family and neighborhood socialization experiences tend to watch a lot more television than do people without such experiences and predispositions. This might be a good chance to talk about (causal) models in general. The first causal picture (a) could be expanded into a developmental sequence or more complicated structure (c).
b. Ethnicity is (causally?) related to school completion rates. It seems likely that if there were such a connection it would be explained by family income and similar factors.

c. Socioeconomic background and parental smoking habits come to mind. Peer pressure may also play a role. But this is a tough one, for the article also reports, “The link was not affected by social and economic factors, and there was no link between smoking and the total number of movies watched…”

d. There are at least two arguments here. First, “relaxation” of marriage laws (that is, legalization or recognition of civil unions) leads to lower marriage rates. Second, the decline in marriage rates contributes to increases in out-of-wedlock births. Both assertions seem questionable. There is of course the time order problem. Marriage rates have been declining for 30 or 40 years and certainly before gay marriage became an issue. More important, changes in social norms in general may be responsible for the variation in marriage rates and out-of-wedlock births. [These are seldom referred to as “illegitimate” births anymore.] Acceptance of gay rights, including civil unions and possibly marriage, might be an effect of these changed mores. An explanatory causal diagram might look something like this:
e. The causal assertions seem straightforward. One side claims that material conditions (e.g., poverty) cause crime, while the other side says that fatherless homes do. It would be important to encourage students to think in multivariate terms, to consider the possibility that the antecedents work together to produce effects. Instead of there being a spurious association between, say, fatherless homes and criminal behavior, there could be a “developmental sequence”:

**Exercise 14-2.**

a. The way the situation has been framed, logistic regression seems most appropriate.

b. A multivariate categorical data analysis with opinions being the dependent, or response, variable, region the explanatory factor, and socioeconomic characteristics the control variables. Students can presumably envisage constructing three-way tables and can be told about more advanced techniques such as log linear analysis and association models.

c. The project calls for multiple regression in which health spending is the dependent variable.

d. We hope that students will realize that they can use dummy variables in a multiple regression. For party bias they will create two dummy variables and use the remaining category as the reference category. Discuss with them how their choice of reference category is likely to affect the sign of the partial regression coefficients for the dummy variables. If they use “neutral” as the dummy category, the partial regression coefficient is likely to be positive for “pro-Democratic” and negative for “pro-Republican.” If they use “pro-Democratic” as the reference category, the signs of the coefficients for “neutral”
and “pro-Republican” are likely to be negative. For region they will have just one dummy variable.

**Exercise 14-3.**
a. A very slight relationship exists: about 7% more women than men voted for Kerry.

b. 3.56; \((2 - 1) \times (2 - 1) = 1\); the critical value for a chi-square with 1 degree of freedom is 3.84. The observed value is slightly less, so do not reject.

c. No, 90% of men and 92% of women voted for Kerry. The chi-square (.290) with 1 degree of freedom is not significant.

d. There is a relationship that is a bit stronger than in the original two-way table: 64% of women as opposed to 55% of men voted for Kerry. The association is not strong—a difference of 9 percentage points—but it suggests that Democrats may have a slight advantage among independent women. Also note that more than half of both sexes supported Kerry. However, the chi-square is not close to being significant. An astute student will notice that the total \(N\) in this table (41) is quite small, which partly explains the low value.

e. As in the two-way table for Democrats, there is no relationship between gender and vote among Republicans. A total of 92% of the men and 93% of the women voted for Bush. This is reflected also in the nonsignificant chi-square.

**Exercise 14-4.**
a. A simple cross-tabulation and chi-square should suffice.

<table>
<thead>
<tr>
<th>Death Penalty?</th>
<th>Defendant’s Race</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>White</td>
</tr>
<tr>
<td>Yes</td>
<td>11.9%</td>
</tr>
<tr>
<td></td>
<td>19</td>
</tr>
<tr>
<td>No</td>
<td>88.1%</td>
</tr>
<tr>
<td></td>
<td>141</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>(160)</td>
</tr>
</tbody>
</table>

Chi-square = .221, 1 \(df\).
There doesn’t appear to be any relationship at all between the defendant’s race and receiving the sentence.

b. Here is a three-way table:

<table>
<thead>
<tr>
<th>Death penalty? (Y)</th>
<th>Victim’s Race (Z)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>White</td>
</tr>
<tr>
<td>Defendant’s race (X)</td>
<td>(X)</td>
</tr>
<tr>
<td>Yes</td>
<td>12.6%</td>
</tr>
<tr>
<td></td>
<td>(19)</td>
</tr>
<tr>
<td>-------</td>
<td>----------</td>
</tr>
<tr>
<td>No</td>
<td>87.4 (132)</td>
</tr>
<tr>
<td>Total</td>
<td>100% (151)</td>
</tr>
</tbody>
</table>

Here is a case of interaction or specification. The original interpretation was that people (white or black) who kill whites stand a much greater chance of receiving the death penalty than if they murder blacks. That seemed to be the major point. But students will also notice that blacks are somewhat more likely (about 5%) to get the death penalty no matter what the victim’s race. They can explain the apparent contradiction by noticing that whites tend to kill whites (70% of white victims were killed by whites) and blacks murder blacks (the corresponding percentage is 92) and that the death of a white is more likely to bring a death penalty. Of course, statistical arguments about the fairness of capital punishment rage on and on.

**Exercise 14-5.**
Country 1 predicted membership

\[ a + (b_1 X_1) + (b_2 X_2) \]
\[ 100 + (0.5 \times 300) + (42 \times 1.8) \]
\[ 100 + 150 + 75.6 \]
\[ 325.6 \]

(round to 326 for people)

Country 2

\[ a + (b_1 X_1) + (b_2 X_2) \]
\[ 100 + (0.5 \times 85) + (42 \times 0.6) \]
\[ 100 + 42.5 + 25.2 \]
\[ 167.7 \]

(round to 168 for people). Hence, country 1 has a larger predicted assembly membership.

**Exercise 14-6.**

a. \[ \hat{Y} = 0.9533 - 1.4558X_1 + 0.00000798X_2 - 0.0145X_3. \]

b. A “one-unit increase in status decreases VA by about 1.456 units when the other variables have been held constant.”

c. Partial coefficients give a numerical indication of the direct effects of an independent variable. The variable’s indirect effects have been factored out. That is, status might affect voice and accountability not only directly but also through its influence on the other variables. But the estimated coefficient (−1.4558) represents only the direct effect and hence is called a partial coefficient.
Exercise 14-7.

a. An indicator variable using dummy coding or more simply a dummy variable.

b. \[ \hat{Y} = 11.99 - 1.57X_1 - .76X_2 + .03X_3 + .43X_4 - .03X_5 + .20X_6 + .58X_7 + .03X_8. \]

c. Income has a slightly positive relationship to opposition. The wealthier people are, the greater the opposition to government-funded welfare programs. More specifically, moving up one quintile increases the predicted opposition score by .03 units, assuming that all the other variables are held constant. Since income is undoubtedly related to other variables in the model, one has to keep them constant to measure the direct effects of \( X_3 \).

d. Men are slightly more likely to express opposition to social welfare policies than women. How much more? About \( \frac{3}{4} \) of a point more. This coefficient is significant, but you could discuss the question of its practical importance.

e. 11.99

f. This question provides an occasion to discuss the meaning of the intercept. The text does not make much of the concept since in some situations a few analysts consider it a nuisance parameter. But here it might help students understand the numerical value of the regression coefficients. Think of the intercept as the average opposition score and then interpret the coefficients as “adjustments” to it as, for example, “becoming male” “leads to” a .76 increase in opposition. See the question g.

g. The numerical effect is to decrease the opposition to government welfare provision by 1.57 units; the predicted score becomes 11.99 – 1.57(1) = 10.42.

h. 11.99 – 1.57(1) – .76(1) + .03(2.57) = 9.737.

i. 11.99 – 1.57(1) – .76(1) + .03(2) + .43(3) – .03(50) + .20(13.70) + .58(3.53) + .03(2.87) = 14.384.

j. There is a 3-unit decrease in party identification so the switch leads to a drop of \( 3 \times .43 = 1.29 \) in the opposition score. The predicted value is 14.384 – 1.29 = 13.094. Or substitute the new value in the previous equation: 11.99 – 1.57(1) – .76(1) + .03(2) + .43(0) – .03(50) + .20(13.70) + .58(3.53) + .03(2.87) = 13.094.

k. This test of the overall model strongly suggest that at least one of the partial regression coefficients is not zero. By itself it is not terribly informative since an investigator wants to know which variables have a significant relationship with \( Y \).

l. The \( z \) or standard normal distribution. A large-sample \( z \) test is appropriate.

m. \( z_{obs} = .76 / .30 = 2.5333 \); the critical value at the .05 level is 1.96; at the .01 level it is 2.53. The result is barely significant at the .01 level.
n. \( z_{obs} = \frac{.03}{.07} = .4286 \); not significant at either level.

o. Of course \( R^2 = .48 \) can be read as “about half—48% to be exact—of the variance in the opposition scores is explained by this set of independent variables.” We stress that in this context “explain” is a statistical term. There is still a lot of residual variation. It might be concluded that there is a modestly good fit, but on the other hand, by the standards of social science research the \( R^2 \) is large. However, it is possible that one or two of the variables are doing the lion’s share of the explaining.

**Exercise 14-8**

We of course have in mind using dummy variables as described in the multiple regression section of the chapter. There are \( K = 5 \) categories, so one could create four dummy variables of the form “\( X_1 = 1 \) if respondent is Protestant, 0 otherwise.” It might be helpful to talk about appropriate reference categories, which would depend on the researcher’s purpose. “None” might be appropriate, but others would be possible. You could ask that the students demonstrate their knowledge by interpreting the coefficients in general terms. (For example, “Compared with those not committed to a religion, Catholics are 10 points more favorable to....”) Analysis of variance can be used as well, and some instructors show the connections between the procedures with a problem of this sort.²

**Exercise 14-9**

a. Not necessarily. Put aside the statistical significance issue, which comes up later. As we stressed in Chapter 12, students need to remember the measurement scale of \( X_1 \). A dollar change in per-capita income is not much. If they multiply by 10 or 100, the connection becomes more apparent. At a minimum, there is a negative correlation: as income rises, crime decreases, although it is not necessarily a causal effect.

b. The observed \( t \) is \(-.0076/.041 = -.1853\); for the sake of completeness don’t forget the minus sign.

c. No, we wouldn’t, because at the .05 level (two-tailed test) the critical \( t \) (we used infinite degrees of freedom) is 1.96, which exceeds the absolute value of the observed \( t \).

d. It’s the same old drill. There is a positive relationship: the more money spent on public safety, the more crime. Obviously, this might be a good time to reinforce the difference between correlation and causation. Having more police and firefighters around doesn’t make people more violent, and perhaps crime should be treated as an independent variable. (There might be a dynamic relationship, as in reciprocal causation.) Also, both variables have an independent connection to \( Y \), although the first is not statistically significant.

e. Yes, since the observed \( t \) (4.126) exceeds the critical value (at the above significance level.) In fact, it could be rejected at the .001 level As always we are assuming a random sample.
f. Our calculations show that the standardized regression coefficient of $Y$ on $X_1$ is $-0.0766(2410.431/882.158) = -0.209$ and for $Y$ on $X_2$ it is $3.090(128.018/882.158) = 0.448$. You probably won’t get these results exactly because we used the standard deviations in the workbook table, whereas the computer program presumably calculates them after the missing values have been removed and before the regression analysis starts. SPSS gives $-0.208$ and $0.46$, respectively. In any case, we can say the absolute value of one standardized coefficient is roughly twice the other, but there is no reason to draw any substantive conclusions about relative importance in this context. Indeed, as we mentioned above, we just know that the variables are (statistically) linked, not too much else. After all, using standardization alone as a criterion, one would have to say public spending provides a better “explanation” of the variation in crime than does income, and that doesn’t make much sense.

### Exercise 14-10

<table>
<thead>
<tr>
<th></th>
<th>Estimated Coefficient</th>
<th>Standard Error</th>
<th>$t$ Statistic</th>
<th>Level of Significance (two-tailed)²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substantial</td>
<td>$-0.39$</td>
<td>$0.13$</td>
<td>$-3.00$</td>
<td>Yes (.005)</td>
</tr>
<tr>
<td>Limited</td>
<td>$-0.23$</td>
<td>$0.12$</td>
<td>$-1.92$</td>
<td>Yes (.05)</td>
</tr>
<tr>
<td>Circulation</td>
<td>$0.04$</td>
<td>$0.02$</td>
<td>$2.00$</td>
<td>Yes (.05)</td>
</tr>
<tr>
<td>Endorsement</td>
<td>$0.04$</td>
<td>$0.05$</td>
<td>$0.80$</td>
<td>No</td>
</tr>
<tr>
<td>Percent of revenue from TV</td>
<td>$-0.05$</td>
<td>$0.35$</td>
<td>$-0.14$</td>
<td>No</td>
</tr>
</tbody>
</table>

$R^2 = .35; \ N = 27$.  
²The researchers used a one-tailed test.

a. $\hat{Y} = -0.38\text{Substantial} - 0.23\text{Limited} + 0.04\text{Circulation} + 0.04\text{Endorsement} - 0.05\text{Revenue}$.

b. The independent variables explain about a third of the (statistical) variation in negative news coverage or, one might say, there is a reasonable fit to a linear model. Hence, there appears to be a relationship, but this indicator by itself tells us relatively little. Students might be reminded that the model can “fit” the data without any of the coefficients being especially large explanatory factors.

c. Students may get confused if they apply the literal introductory textbook definition of a “one-unit increase in . . . leads to a 39% decrease in negative coverage.” They can be encouraged to think about a hypothetical situation in which a paper changes from owning no TV stations to buying 9 or more or a limited number, 2 to 5. The result is a 39 or 23 percentage point reduction in negative coverage, if nothing else changes, that is, if circulation, partisanship, and revenue are held constant. This might be a good time to stress that sometimes “small” numbers carry big meanings.

d. The observed $t$ values are: $\text{Substantial} = -3.00; \text{Limited} = -1.92; \text{Circulation} = 2.00; \text{Endorsement} = 0.80; \text{Percent revenue from TV} = -0.14$. 
e. 2.074; Substantial = Yes (.005); Limited = Yes (.05); Circulation = Yes (.05); Endorsement = No; Percent revenue from TV = No.

**Exercise 14-11** (See below for alternative answer to Exercise 13-11 using SPSS Student Version.)

a. Here’s our version of the table. We used the weighted variable.

<table>
<thead>
<tr>
<th></th>
<th>Strongly Approve</th>
<th>Approve</th>
<th>Neither</th>
<th>Disapprove</th>
<th>Strongly Disapprove</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labour</td>
<td>60.4%</td>
<td>54.0%</td>
<td>57.9%</td>
<td>37.2%</td>
<td>34.8%</td>
</tr>
<tr>
<td>Liberal Democrat</td>
<td>29.7</td>
<td>23.3</td>
<td>17.7</td>
<td>16.5</td>
<td>15.5</td>
</tr>
<tr>
<td>Conservative</td>
<td>9.9</td>
<td>22.7</td>
<td>24.4</td>
<td>46.3</td>
<td>49.8</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

b. You might strongly encourage, if not require, students to generate the table on their own. But note our remarks about recoding the vote variable (var 04). There appears to be a moderate association between the two variables. First, the chi-square statistic is significant at the .001 level, which suggests rejection of the null hypothesis of statistical independence. Second, the two measures of ordinal correlation show a weak to modest relationship. Of course, without the full table, students can’t go much further, which is why we encourage them to try to cross-tabulate the variables themselves. If they do, they will find that those who support joining the EU were disproportionately inclined to vote for Labour. The respondents opposing membership, by contrast, tended to support the Conservative Party. (Roughly 60% to 55% of those who approved strongly or approved voting Labour, but about half of those disapproving or disapproving strongly backed the Conservatives.) In this instance, attitudes do seem to “explain” vote preferences.

c. We feel it helps if students examine the contingency tables themselves, although in this case the 3 × 4 × 3 multi-way table can be intimidating at first sight. The results here can be interpreted in many different ways. The original relationship in the 3 × 4 table appears to weaken, except for the last control category (Conservative). The chi-squares are not directly comparable (in this context), but they are suggestive of a reduced association. (You might tell advanced students about partitioning the likelihood ratio chi-square and other relatively new developments in “ransacking” cross-classifications.) One is not significant and one is significant at just below the .05 level. The measures of association in the first two categories are less than in the original table and the weighted averages of tau-b and gamma are .03 and .09, respectively. Hence, one might say that partisanship “explains” the relationship. Or, it could be said to “specify” the relationship between vote and opinion by giving the conditions under which it holds (Conservative identifiers) and under which it does not (Liberal Democrats and Labour). This is an example of “interaction.” Of course, it helps to use the tables themselves to elucidate the specification. (Most of the strong relationship in subtable three (Conservatives) seems to stem from the overwhelming [95% and above] backing that those opposed to EU membership give to their party.)
Exercise 14-11. (SPSS Student Version answer)

Table and statistics for a and b. These statistics are based on unweighted data. Note that these figures will not correspond exactly to the ones shown in the workbook because of the different sample sizes.

### Attitude Toward British Membership EU * party voted for recode Crosstabulation

<table>
<thead>
<tr>
<th>Attitude Toward British Membership EU</th>
<th>party voted for recode</th>
<th>labour</th>
<th>liberal democrat</th>
<th>conservative</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Count</td>
<td>26</td>
<td>9</td>
<td>7</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>% within party voted for recode</td>
<td>7.9%</td>
<td>8.5%</td>
<td>3.8%</td>
<td>6.8%</td>
</tr>
<tr>
<td>Strongly approve</td>
<td>Count</td>
<td>151</td>
<td>44</td>
<td>49</td>
<td>244</td>
</tr>
<tr>
<td></td>
<td>% within party voted for recode</td>
<td>46.0%</td>
<td>41.5%</td>
<td>26.9%</td>
<td>39.6%</td>
</tr>
<tr>
<td>Approve</td>
<td>Count</td>
<td>76</td>
<td>21</td>
<td>27</td>
<td>124</td>
</tr>
<tr>
<td></td>
<td>% within party voted for recode</td>
<td>23.2%</td>
<td>19.8%</td>
<td>14.8%</td>
<td>20.1%</td>
</tr>
<tr>
<td>Neither</td>
<td>Count</td>
<td>55</td>
<td>21</td>
<td>58</td>
<td>134</td>
</tr>
<tr>
<td></td>
<td>% within party voted for recode</td>
<td>16.8%</td>
<td>19.8%</td>
<td>31.9%</td>
<td>21.8%</td>
</tr>
<tr>
<td>Disapprove</td>
<td>Count</td>
<td>20</td>
<td>11</td>
<td>41</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>% within party voted for recode</td>
<td>6.1%</td>
<td>10.4%</td>
<td>22.5%</td>
<td>11.7%</td>
</tr>
<tr>
<td>Strongly disapprove</td>
<td>Count</td>
<td>328</td>
<td>106</td>
<td>182</td>
<td>616</td>
</tr>
<tr>
<td></td>
<td>% within party voted for recode</td>
<td>100.0%</td>
<td>100.0%</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Chi-square  df  Tau-c  Gamma  N
58.055  8  .223  .328  616

The statistics will not match unless the party identification variable is recoded, because the number of cells in the tables will be different. This will affect the degrees of freedom as well as the calculations for gamma and tau.

Respondents who voted for more conservative parties in the election were more likely to disapprove of British membership in the EU. For example, 22.9% of Labour voters either disapproved or strongly disapproved, compared with 30.2% of Liberal Democrats and 54.4% of Conservatives; 53.9% of Labour voters either approved or strongly approved, compared with 49% of Liberal Democrats and 30.7% of Conservatives. However, the percentage of voters who approved or disapproved strongly is relatively small compared with those who held less strong views or responded “neither.” Thus the relationship is weak, although it is statistically significant. The probability that the relationship observed in the table was not found in the population is less than .001. Gamma and tau-c are both positive, which means that in this case, given the way the variables are coded, the more conservative a respondent’s party affiliation, the more opposed to EU membership the respondent tends to be.
c.

Chi-Square Tests

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>df</th>
<th>Asymp. Sig. (2-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Chi-Square</td>
<td>58.055</td>
<td>8</td>
<td>.000</td>
</tr>
<tr>
<td>Likelihood Ratio</td>
<td>56.583</td>
<td>8</td>
<td>.000</td>
</tr>
<tr>
<td>Linear-by-Linear Association</td>
<td>46.732</td>
<td>1</td>
<td>.000</td>
</tr>
<tr>
<td>N of Valid Cases</td>
<td>616</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 7.23.

Symmetric Measures

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>Asymp. Std. Error</th>
<th>Approx. T</th>
<th>Approx. Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ordinal by Ordinal</td>
<td>Kendall's tau-b</td>
<td>.223</td>
<td>.034</td>
<td>6.478</td>
</tr>
<tr>
<td></td>
<td>Kendall's tau-c</td>
<td>.223</td>
<td>.034</td>
<td>6.478</td>
</tr>
<tr>
<td></td>
<td>Gamma</td>
<td>.328</td>
<td>.049</td>
<td>6.478</td>
</tr>
<tr>
<td>N of Valid Cases</td>
<td>616</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Exercise 14-12

a. 

\[ \hat{P} = \frac{e^\hat{z}}{1 + e^\hat{z}}, \text{where} \]

\[ \hat{z} = -4.4223 - .2005X_1 + .2652X_2 + .0337X_3 + .0295X_4 + .3824X_5 + 2.0460X_6 + .3195X_7 - .0053X_8 + .8183X_9 + .7279X_{10} + 1.783X_{11} + .1962X_{12} \]

b. 

\[ \hat{Logits} = -4.4223 - .2005X_1 + .2652X_2 + .0337X_3 + .0295X_4 + .3824X_5 + 2.0460X_6 + .3195X_7 - .0053X_8 + .8183X_9 + .7279X_{10} + 1.783X_{11} + .1962X_{12} \]

c. \( \hat{P} = .0119 \). This gives the estimated probability of voting for a male who did not comment on any ads, who has zero (that is, no) interest in newspaper news and campaigns, who is a pure independent, who is nonwhite, and who was never born.

d. Higher; coefficient is positive.
e. \( \hat{P} = .050 \). Aging increases the probability of voting. But at this point we can’t say a one-year increase leads to an \( X \) percent increase in the probability of voting. That comes next.

f. Estimated logit = \(-2.947\); estimated odds \(.052\) to 1. Odds are against his voting. Each one-year increase in age is associated with a \(.0295\) increase in the log odds.

g. The predicted probability is \(.905\) and the odds of voting are \(9.501\) to 1.

h. Can be read from the equation. According to the model the effect of “going from” male to female is to increase the log odds of voting. So the female is more likely to vote. For the man, \( \hat{P} = .872 \); odds are \(6.9026\) to one, both of which are lower.

i. Newspaper interest index, age, somewhat and very interested in campaign, gender, leaning, weak and strong partisans, and constant. Starting from the top of the table, the observed \( z \) scores are \(-1.119, .945, 2.982, 5.463, 2.287, 6.276, 2.084, -.233, 3.091, 2.862, 6.067, .954, \) and \(-10.474\).

j. Given their data and other assumptions and measures it appears so. When the various factors that might affect turnout are taken into account, the coefficient for “negative comments” does not appear to be large even if it has the predicted sign. The easiest solution is to let all the variables equal 0, in which case (as we found above) the estimated probability of voting is about \(.01185\) and the odds are \(.012\) to 1. Compare that value with the case where “negative mention” is 1. The probability is now \(.0097\) and the odds are \(.010\) to 1, hardly any decline at all. This finding contrasts with some of the cases above, where a change in a variable did seem to lead to a discernible difference. Still, one has to assume that Wattenberg and Brians’ measure of negative campaign exposure is comparable with those of other researchers. This question takes students back to the research design and measurement chapters.

**Exercises 14-13**

a. There is a positive relationship between beer and Obama. Each additional brewery adds \(5.4\)% to Obama’s vote percentage in a state. The beer variable is highly, statistically significant with a \( t \)-score of \(2.64\).

b. The relationship is statistically significant but substantive significance is lacking. There is no meaningful relationship between the number of breweries in a state and the percentage vote for Obama. The relationship that is probably being picked up is that larger states have more breweries and also happen to have a higher percentage vote for Obama.

c. In this regression conservative has an important effect. Increase the percentage of conservatives in a state by one percentage point decreases the percent vote for Obama by \(1.54\) percentage points. The conservative variable is statistically significant with a \( t \) score of over 12. Beer is not statistically significant when controlling for conservative. The
coefficient is near zero suggesting that the addition of a brewery has little effect on the dependent variable.

d. the conservative variable was an important predictor of Obama’s percent vote in 2012. This is a theoretically relevant relationship as voter ideology should affect vote percentages. Because beer was a spurious relationship all along it becomes irrelevant when controlling for an actually important variable like conservative.

<table>
<thead>
<tr>
<th>Variable Entered Removed</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NUMBER OF CRAFT BREWERIES IN A STATE</td>
<td>0.33</td>
<td>0.003</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CONSERVATIVE VOTER PERFORMANCE</td>
<td>0.33</td>
<td>0.003</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable Entered Removed</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CONSERVATIVE VOTER PERFORMANCE</td>
<td>0.33</td>
<td>0.003</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CONSERVATIVE VOTER PERFORMANCE</td>
<td>0.33</td>
<td>0.003</td>
</tr>
</tbody>
</table>
Exercises 14-14, and 14-15
Answers will vary.

2 Once more, this might make a good research paper topic.

Exercise 14-16.

a. A dummy variable has just two values: 0 for the presence (or absence) of a characteristic, group membership, condition and so on and 1 for its absence (or presence).

b. 1 and 0 are typically used because 0 indicates that there is no amount of the characteristic captured by the variable present. 1 simply means that the characteristic is present.

c. For example, students might choose race—a nominal variable that captures white, black, Hispanic, Asian and other. One dummy variable might be called Hispanic. The dummy variable will take on the values of 1 if the characteristic (Hispanic) is present and 0 otherwise. This variable will indicate Hispanic or not Hispanic. All other races fall in the 0 category.

d. Students should explain how they would multiply the categorical variable by the dummy variable. The student would then include the original two variables along with the new interaction term in the model.

Exercise 14-17.
Answers will vary with variables chosen.