Teacher Notes for Introduction

The Introduction is designed to give students a glimpse of the statistical problem-solving process. Several specific research questions that students will answer during later investigations are presented in the opening paragraph. We then offer an example that shows how researchers attempted to get an answer to the question: Do most people wash their hands after using the bathroom? Results from a survey and two observational studies are presented. Following the example, we outline the four steps in conducting a statistical study:

- Defining a research question
- Collecting data
- Analyzing the data
- Interpreting results

Along the way, we explain how information from a carefully chosen *sample* can be used to make inferences about a larger *population*. In addition, we distinguish the three primary methods of producing data—*observational studies*, *surveys*, and *experiments*.

In Investigation #1: Did You Wash Your Hands?, students answer a series of questions that help them make sense of the hand-washing studies described in the Introduction.

Prerequisites

None

Learning Objectives

As a result of completing the Introduction, students should be able to:

- Distinguish between a population and a sample
- Briefly describe the three primary methods for producing data
- Identify the four steps in conducting a statistical study
- Critique information provided in an article or headline

Teaching Tips

Give students practice defining their own research questions. Statistics questions require some variability from individual to individual. For instance, "On average, how far from school do students in this class live?" is a statistics question. Students in the class live different distances from school, so individual-to-individual variation is present. Contrast this with the question, "How far does Tyler live from school?" This is not a statistics question, because there is no individual-to-individual variation present. "What percent of students own an iPod?" is a statistics question, while "Does Kayla own an iPod?" isn't.

For each acceptable research question that a student defines, ask the class which method of data collection—observational study, survey, or experiment—would be optimal.

Then, ask students to give some additional examples of research questions that would be best answered using each of the three methods of data production.

Ask students why researchers don't just get data from every member of the population—a census. Some issues that make a census impractical include the difficulty of contacting some individuals, time constraints, measurement issues, and budgetary limitations. If a manufacturer wanted to determine the average lifetime of its batteries, for instance, the company would definitely not be willing to test every battery produced!

Help students explore situations in which a *sample* would and would not provide good information about a *population*. For example, to estimate the average height of students in their school, students shouldn't use the school's basketball team as a sample. How about the students in a particular English class? Or, if students wanted to find out what percent of students in their school like the food served in the cafeteria, should they send an e-mail to all students in the school and use the first 30 students who respond as the sample?

Discuss the paragraph on "data ethics" with your students.

Are hot dogs unhealthy? What percent of people wear their seat belts when driving? Which works better—a low fat diet or a low carbohydrate diet? Would most teenagers keep an extra $10 they received in incorrect change at a store, or return it? Does listening to music hurt students' concentration and ability to study? How are peoples' heights and foot lengths related? These are just a few examples of the types of questions that statistics can help us answer. Getting clear answers to such questions requires data that have been produced according to a careful plan, as the following example illustrates.

Research question: *Do most people wash their hands after using the bathroom?*

Not according to a December 2005 newspaper article titled "Many Adults Report Not Washing Their Hands When They Should, and More People Claim to Wash Their Hands than Who Actually Do."[1] But before you believe such a headline, you should always ask, "Where did the data come from?"

The article mentioned in the previous paragraph was based on two studies that were done in August 2005. In the first study, 1,013 U.S. adults were asked questions about their hand-washing habits by telephone. This is an example of a **survey**. In the second study, observers watched and recorded the actual hand-washing behaviors of 6,336 adults in public restrooms in four major U.S. cities. This is an example of an **observational study**. Both studies were carried out by Harris Interactive, a company that specializes in these kinds of statistical research.

Now that you know how the data were produced, you might be interested in some results from the two studies.

> While 91% of surveyed adults *claimed* to always wash their hands after using the bathroom, only 83% of the adults in the observational study did so.
>
> In the survey, 94% of women claimed to always wash their hands after using the bathroom, compared with 88% of men. In the observational study, 90% of the women actually washed their hands, compared with 75% of men.
>
> A similar observational study done in 2003 revealed that 78% of the adults observed actually washed their hands after using the bathroom. In that study, 83% of the women and 74% of the men were observed washing their hands.

Based on these studies, what can we conclude? Can we conclude that 83% of *all* U.S. adults always wash their hands after using the bathroom? No, because researchers only observed a **sample** of 6,336 adults, not the entire **population** of U.S. adults. If another group of 6,336 adults was observed on a different day, the percent who washed would probably not be exactly 83%. Can we at least say that the actual percent of all U.S. adults who always wash their hands after using the bathroom is "close" to 83%? That depends on what you mean by "close."

1 Harris Interactive, December 14, 2005.

The process of carrying out a statistical study—like the survey or observational study in the previous example—begins with the clear statement of a **research question**. Basically, the research question describes what you want to know in simple terms. Most research questions relate to some population of interest—a group of people, animals, or things. Once a research question has been established, you need to collect some useful data. It's usually not practical to get data from every individual in the population (a **census**). Instead, we usually try to obtain data from a representative sample of individuals chosen from the population. So how do we get the data?

There are three preferred methods for producing data—**observational studies**, **surveys**, and **experiments**. In an experiment, we deliberately do something to one or more groups of individuals—such as giving a drug to people who are sick—and then measure their responses. Observational studies and surveys, on the other hand, attempt to gather data on individuals as they are. In an observational study, we record values of one or more variables—like gender or height—for a sample of individuals. We can obtain these values from direct observation, measurement, or existing data. In a survey, we select a sample of people and have them answer one or more questions. You have already seen examples of a survey and an observational study about people's hand washing habits. How might an experiment shed more light on this subject?

Some people might argue that having an observer present in the restroom—even if the observer isn't washing his or her hands—could influence an individual's hand-washing behaviors. To test this idea, we could design an experiment. Half of the time, we would station an observer at one of the sinks. The other half of the time, we would "hide" the observer in one of the bathroom stalls with a clear view of the sink area. Then, we could compare the percent of people who washed their hands under each of these experimental conditions, called **treatments**.

Each data production method comes with advantages and limitations that you need to understand before you can plan a study. The method used to produce the data also determines the kinds of conclusions that can be drawn. Choosing the best method for a given research question requires careful thought and a lot of practice.

Once we have our data in hand, we must try to figure out what they're telling us. We begin by making graphs and calculating numerical summaries. Then, we interpret the results of our analysis. Of course, our goal is to answer the original research question. Finally, we can communicate our findings to others who might be interested.

Here is a brief outline that summarizes the entire process.

Carrying Out a Statistical Study[2]

I. Formulate the research question

2 *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report*, The American Statistical Association, January 2007. *www.amstat.org/education/gaise*

Do some background research to understand the nature of the problem.

Think carefully about what you expect to find and why.

II. Collect data

Decide what to measure and how to obtain the measurements. Which method—survey, observational study, or experiment—would be best?

Think about how you will analyze the data.

Be sure to consider ethical issues.

Produce data according to your stated plan.

III. Analyze the data

Use graphical and numerical summaries to describe the data.

If appropriate, use inference methods to estimate population values or test claims about characteristics of the population.

IV. Interpret your results

Draw conclusions from your data analysis. Remember to answer the research question!

Address any limitations in your conclusions that result from the process of data collection and data analysis.

Communicate your findings.

In this module, you will learn how to analyze surveys, observational studies, and experiments that have been planned by others. Then, you'll get to design and carry out your own studies. As you go, keep this in mind: You can't draw sound conclusions from badly produced data.

Here's another important principle to remember: Statistical studies should be conducted in an ethical manner. Avoid the use of deception whenever possible and ensure that survey participants and experimental subjects are informed about the purpose of the study and any potential risks associated with their participation. If study subjects are people, they must provide their informed consent to participate after being made aware of any potential risks that may result from taking part in the study. For studies involving minors, parent/guardian permission is required. If the study uses animals as subjects, researchers should follow published guidelines for humane treatment of animals, such as those published by the American Psychological Association (see *www.apa.org/science/anguide.html*). Researchers should also ensure the anonymity and confidentiality of peoples' responses and behaviors unless participants have been informed that responses will not be confidential. For experiments, it is common to have a review board approve the experimental design in advance and monitor the results of the experiment as data are collected.

Teacher Notes for Investigation #1: Did You Wash Your Hands?

This investigation builds on the two hand-washing studies that were discussed in the Introduction. The questions posed here are designed to get students thinking about statistics in practice and to provoke discussion in the classroom. We try to alert students to several important issues, such as:

- The way in which data are produced affects the kinds of conclusions that can be drawn. Only well-designed experiments can be used to make cause-and-effect conclusions.
- An estimate we obtain from a sample could differ greatly from the truth about the population if:
  - Our sample doesn't represent the population well
  - The question we asked is unclear or misleading
  - People don't respond accurately or honestly
  - Some people refuse to respond
  - The observer influences the observed

Prerequisites

Students should be able to determine from a narrative description whether data were produced with a survey, an experiment, or an observational study.

Learning Objectives

As a result of completing this investigation, students should be able to:

- Decide which method of producing data—a survey, experiment, or observational study—is most appropriate for answering a given research question
- Describe how certain practical difficulties may affect the results of surveys, experiments, and observational studies
- Define a research question

Teaching Tips

Consider whether you want students to answer the questions in this first investigation individually or with a partner. Either way, be sure to allow time for class discussion of the questions.

Suggested Answers to Questions

Many of the questions do not have "right" or "wrong" answers. Students should be encouraged to defend their answers with specific evidence, much as an attorney would in a legal case.

1. Answers will vary. Students might focus on the unsanitary or disgusting nature of not washing hands after using the bathroom. There are also health-related implications. According to the Harris Interactive news release from the hand-washing study, "Infectious diseases, many caused by unclean hands, are the leading causes of death and disease

worldwide and the third leading cause of death in the United States. The Centers for Disease Control and Prevention (CDC) says that hand washing is the single most important means of preventing the spread of infection."

2. (a) Those who do not have telephones. Also, if calls were placed only to landlines, then those who only have cellphones would be left out.

(b) Answers will vary. Here's one possible answer. As adults who do not have telephones would tend to be poor, perhaps even homeless, they might not have access to proper facilities for washing their hands after using the bathroom. Without these people's behaviors represented in the survey, the 91% estimate would be too high.

(c) Answers will vary. Here's one possible answer. Some people would be embarrassed to admit that they don't usually wash their hands after using the bathroom, and so might refuse to answer. Other individuals might feel that they are too busy to participate in a survey.

(d) Since this survey asks about a potentially embarrassing issue, it seems likely that some people will give a socially acceptable "always" answer, even if this is not the truth. It is also possible that some individuals will have inaccurate recollections of their hand-washing habits.

3. (a) Answers will vary. Here's one possible answer. When others are present in the bathroom, some people might be more likely to wash their hands due to implicit "peer pressure." If that's the case, then a hidden camera would have revealed fewer than 83% who washed their hands.

(b) Answers will vary. Here's one possible answer. People may be generally more likely to follow societal expectations to wash their hands when they are in public than when they are at home. If so, then a hidden camera study would reveal less than 83% who washed their hands at home.

4. (a) Answers will vary. Here's one possible answer. From these two studies, it certainly appears that a higher percent of adults claim to wash their hands after using the bathroom than the percent who actually do so when observed. However, the survey and the observational study involved different groups of people. It is possible that the difference in the results of the two studies (91% who claimed they washed their hands versus 83% who actually did) is due to differences in hygiene habits between the people in these two groups, and not from people's tendency to overestimate their hand-washing tendencies. After all, if researchers had observed a different group of 6,336 adults in the same public restrooms on a different day, the percent who were seen washing their hands would probably not have been exactly 83%. Likewise, if researchers had surveyed another sample of 1,013 adults by telephone, it is unlikely that exactly 91% would say they always washed their hands after using the restroom. The difference in the results of these two studies may just be due to the natural variability that occurs from one sample to another.

(b) Answers will vary. Here's one possible answer. Conduct a hidden camera observational study of a sample of individuals in public restrooms. Then survey those same individuals about their hand-washing habits after the bathroom visit.

5. It's always important to know who sponsored (paid for) a statistical study. In this case, the sponsoring agency would likely be hoping to encourage people to do a better job of washing their hands after using the bathroom, thereby using more soap. A study that reveals a lower percent of people washing their hands would suit this agency's purpose.

6. (a) For example, "What percent of teenagers always wash their hands after using the bathroom?"

(b) A survey or an observational study would both be reasonable methods for producing data. An experiment wouldn't be appropriate, as the goal of the study is to record teenagers' normal hand-washing tendencies, not to try to do something to influence those tendencies.

(c) Answers will vary. Here's one possible answer. Teenagers tend to pay less attention to personal hygiene and health than adults do, so the percent of teenagers who always wash their hands after using the bathroom would probably be lower.

7. (a) A survey. Since it would be difficult to actually observe teenagers brushing their teeth, an observational study wouldn't be practical. Because the goal of the study is to record teenagers' normal tooth-brushing habits, not to try to do something to influence those habits, an experiment would not be appropriate. If possible, the teens who participate in the survey should be selected at random from the population of teens in the area. Note that it may be necessary to obtain parental permission before surveying teenagers.

(b) An experiment. If researchers deliberately give drug A to one group of individuals and drug B to another, then they can compare the differences in the percent in each group who experience nausea following their migraine headaches. If the decisions about which individuals get drug A and which drug B are made at random, then researchers can determine whether the difference in the percents who experience nausea is sizable enough to suggest a difference in the effectiveness of drug A and drug B. (By assigning the drugs at random to the migraine sufferers, researchers help ensure that the group of individuals taking drug A and the group taking drug B are fairly balanced in all ways that might affect their response to the drug treatments. If the two groups are similar to begin with, then any sizable differences that emerge between the two groups after the drugs are administered can be attributed to the effects of the drugs themselves.)

Note that a comparative observational study using two groups of people—one group who have used drug A and the other group who have used drug B—would not allow researchers to establish any kind of cause-and-effect relationship between the drug administered and people's tendency to have nausea later. Since people have chosen

whether to use drug A or drug B, it is possible that the two groups of individuals differ systematically in other ways that might affect their likelihood of becoming nauseated after having a migraine.

(c) Either an observational study or a survey. Asking a (random) sample of males and a (random) sample of females to report how many numbers are stored in their phones will allow for direct comparison of the average number of contacts for the two genders. However, some individuals may report inaccurate values. Actually observing the phone lists of the randomly selected people might result in more reliable data. An experiment would not be appropriate since we only want to observe what is true, not try to influence the state of affairs.

(d) Observational study. Watching the actual behavior of drivers at the stop sign in question would be more effective than asking them whether they stop. It would be best if observers could watch without being noticed by the drivers, since the presence of an observer may influence the behavior of the driver. An experiment would not be appropriate because we are simply trying to observe and record whether drivers stop completely, not to influence whether they stop.

(e) An experiment, with half of the customers receiving a bill with suggested tip amounts at the bottom and half of the customers receiving no suggested tip amounts. Ideally, the determination of which customers get the bills with suggested tip amounts should be made at random so that the two groups of customers will be as similar as possible in every respect that might influence the amount they decide to tip other than the intended "treatment"—suggested tip amounts on the bill versus no suggested tip amounts on the bill. Then, any substantial difference that emerges between the average tip amounts in the two groups could be attributed to whether suggested tip amounts were printed on the restaurant bill.

8. Answers will vary. The survey results from the two years were very similar—91% of respondents said they always washed their hands after using the bathroom in 2005, compared with 92% in 2007. However, the observational study results from the two years were quite different. In 2005, 83% of those observed washed their hands after using the bathroom. In the 2007 observational study, only 77% washed their hands after using the bathroom. One possible explanation for this decrease is the decline in the percent of men who washed their hands—from 75% in 2005 to 66% in 2007.

Possible Extensions

Ask students to develop a research question that would best be answered by (a) an observational study, (b) a survey, (c) an experiment.

Have students locate an article describing the results of a survey in printed or electronic media. Then ask them to identify the research question, the population, the sample, and any concerns they have about the results reported in the article.

Investigation #1: Did you wash your hands?

Corresponds to pp. 4-9
in Student Module

1. Why should we care whether people wash their hands after using the bathroom?

2. In the Harris Interactive survey, people were contacted by telephone. One of the questions the interviewers asked was, "How often do you wash your hands after using a public restroom?"

(a) Which U.S. adults were not included in this study?

(b) The survey estimated that 91% of all U.S. adults would claim that they always wash their hands after using the bathroom. Do you think this estimate is too high, too low, or about right given your answer to (a)? Explain.

(c) Several people refused to participate in the survey. Give a reason that this might happen.

(d) In any survey, it is possible that some people will not answer a question accurately or honestly. Thinking about the hand-washing survey, do you think this is likely to happen? Explain your answer.

3. The observational study of hand washing was conducted at a baseball field in Atlanta, a museum and an aquarium in Chicago, a bus and train terminal in New York, and a farmers' market in San Francisco.

(a) Observers in the public bathrooms combed their hair or put on make-up at one of the available sinks while they were watching individuals' hand-washing behaviors. If the observation had been done by hidden camera instead (with no observer present), do you think the percent who washed their hands would have been greater than, less than, or about the same as 83%? Justify your answer.

(b) Suppose the observational study had been conducted using hidden cameras in the homes of the same 6,336 adults. Do you think the percent of these individuals who washed their hands would have been greater than, less than, or about the same as 83%? Justify your answer.

4. (a) Comment on the conclusion reached in the newspaper headline: "More People Claim to Wash Their Hands than Who Actually Do."

(b) Describe a study design involving only one group of people that might help us better evaluate the validity of the quoted claim in part (a).

5. Both studies were paid for by the American Society for Microbiology and the Soap and Detergent Association. Should you take this information into account when interpreting the results of the studies? If so, how?

6. You have been asked to help design a study to investigate how often teenagers wash their hands after using the bathroom.

(a) Define a research question for your study.

(b) Would you recommend using a survey, an observational study, or an experiment to produce the data? Explain.

(c) Do you think the percent of teenagers who always wash their hands after using the bathroom is higher than, lower than, or about the same as the percent of adults who do so? Justify your answer.

7. For each of the following research questions, decide which method of data production—a survey, an experiment, or an observational study—would be most appropriate. Justify your choice of method.

(a) What percent of teenagers leave the water running while they brush their teeth?

(b) Which of two drugs is more effective at preventing nausea following the onset of a migraine headache?

(c) Do male teenagers or female teenagers tend to have more numbers stored in their cell phones?

(d) What percent of drivers come to a complete stop at a stop sign near a local elementary school?

(e) Does printing suggested tip amounts on the bottom of a restaurant bill increase the average amount that customers leave in tips?

8. A follow-up study conducted in 2007 by Harris Interactive revealed that while 92% of adults said that they always washed their hands after using the bathroom, only 77% of the adults observed in public restrooms actually did. According to Harris Interactive's Hand Washing Fact Sheet, "The overall decline in hand washing observations is largely due to males. The percentage of males observed washing their hands fell from 75% in 2005 to 66% in 2007. Overall, the percentage of females observed washing their hands is down slightly from 90% in 2005 to 88% in 2007."

Did people's hand washing habits improve or get worse from 2005 to 2007? Justify your answer with specific evidence from the reports describing the Harris Interactive studies.