

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.”<sup>1</sup> 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<sup>1</sup>

### I. Formulate the research question

---

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





## Investigation #1: Did you wash your hands?



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.