Teacher Notes for Section I: Observational Studies

In Section I of the module, students examine observational studies. There are five investigations in this section. The first three involve data that are produced with no random selection. In the last two investigations, the data production process involves random selection. As we discuss in the section overview, random selection helps ensure that a "representative" sample is obtained from some larger population of interest. When random selection is used in an observational study, we can generalize results from the sample to the larger population with confidence. Without random selection, our ability to generalize is limited.

The five investigations in this section are:

Investigation #2: Get Your Hot Dogs Here!

Students analyze nutritional data on different brands and types of hot dogs from an observational study carried out by Consumers Union.

Investigation #3: What's in a Name?

Students examine the popularity of the first names of students in their class.

Investigation #4: If the Shoe Fits …

Students play the role of statistical detective as they try to identify a culprit using data on the foot lengths and heights of a random sample of students at a school.

Investigation #5: Buckle Up

Students explore whether seat belt use among drivers is improving in the states.

Investigation #6: It's Golden (and It's Not Silence)

In this culminating investigation, students design, implement, analyze data from, and draw conclusions from an observational study involving people's preference for the golden ratio.

Prerequisites

Students should be able to:

- Distinguish an observational study from a survey or an experiment
- Describe the distribution of a categorical variable using counts, percents, and bar graphs
- Describe the shape, center, and spread of the distribution of a quantitative variable using a dotplot and numerical summaries (mean, median; range, interquartile range (IQR), standard deviation; five-number summary)
- Identify potential outliers
- Compare distributions of a quantitative variable using back-to-back stemplots, parallel boxplots, and numerical summaries (mean, median, mode; range, interquartile range (IQR), standard deviation)

Describe the relationship between two quantitative variables using a scatterplot

Examine the effect of adding a categorical variable to a scatterplot on the relationship between two quantitative variables

Use a scatterplot with or without a summary line to make predictions of y from x

Interpret the slope and y-intercept of a summary line in the context of a problem

Learning Objectives

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

Explain why random selection in an observational study allows sample results to be generalized to a larger population of interest

Describe the relative standing (percentile or z score) of one value within a distribution

Explain why an observational study might be preferable to a survey in describing individuals' behavior

Construct an appropriate graphical display (dotplot, stemplot, or histogram) of a quantitative variable

Describe how the method of data production affects their ability to draw conclusions from the data

Carry out a complete analysis of an observational study involving one or more categorical variables, using counts, percents, and bar graphs to support their narrative comments

Carry out a complete analysis of an observational study involving one or more quantitative variables, using dotplots, stemplots, boxplots, and numerical measures of center and spread to support their narrative comments

Design a practical sampling plan that incorporates random selection

Conduct an observational study in a way that should produce reliable data

Draw conclusions about a population based on graphical and numerical information from a representative (random) sample

Teaching Tips

Take time to go through each of the examples in the overview for this section. We discuss each one in detail below.

The first example addresses the research question: Who talks more, men or women? Note the individual-to-individual variation that's present in the number of words spoken per day. This interesting (and perhaps surprising) observational study appears to have debunked an earlier claim that women talk about three times as much as men. Be sure to point out one obvious limitation of this study: All the participants were college students from the United States and Mexico. These results may not apply to college students in other countries, or to people of other ages.

In the second example, we consider a quality control setting. This time, the research question is: Do the potato chips produced today meet specifications in terms of their salt content? Of course, we expect some variation in salt content from chip to chip. Based on a representative sample of chips from the day's production, we should be able to draw a conclusion about the salt content in the larger population of chips produced today.

How can we get a "representative" sample of individuals from some population of interest? Using random selection. In an ideal world, we would put a slip of paper representing each individual in the population in a hat, mix the slips thoroughly, and then draw out a sample of the desired size. Random selection entails letting chance decide which individuals from the population of interest end up in a sample.

Using the idealized hat method, every individual in the population has an equal chance of being selected for the sample. In addition, every subgroup of n individuals in the population is equally likely to be chosen as the sample (for any sample size n).

In the third example, a high-school student (Kayla) undertakes a study to answer the research question: What is the average number of contacts stored in seniors' cell phones?

We show the mechanics of using a table of random digits and a random number generator to mimic the hat method for Kayla's study. It is important to stress that before heading for the random digits table or random number generator, students should have already assigned a unique numeric label to each individual in the population.

It is usually not practical to obtain a true random sample. We ask students to think about some reasonable alternatives involving random selection for Kayla's cell phone study and for the potato chip quality control study.

Possible Extension

You might want to introduce students to other methods of sampling that involve random selection, such as stratified sampling, cluster sampling, systematic sampling, and multi-stage sampling. For more information about these sampling methods, consult any AP Statistics textbook.

Corresponds to pp. 10-13
in Student Module

You can learn a lot just by watching. That's the idea of an observational study. If you want to know how often people wash their hands after using the bathroom, don't ask them! Observe them. As we saw in the Introduction, what people say and what they actually do can be quite different. But be sure to keep in mind the old adage: "The observer influences the observed." Merely having an observer present in the restroom might affect the percent of people who wash their hands.

In her book, *The Female Brain*, Dr. Louann Brizendine claimed that women talk almost three times as much as men. Some researchers at the University of Arizona were skeptical, so they designed an observational study to examine this claim. About 400 male and female college students participated in the study. The students wore specially designed recording equipment that turned on automatically at pre-set intervals over several days without the students' knowledge. Researchers then counted words used by the male and female participants. Their findings? Both males and females tended to speak an average of about 16,000 words per day. Dr. Brizendine later admitted that her claim had little factual basis.

Let's consider one further example from industry. Suppose you are in charge of quality control at a factory that produces potato chips. Imagine a string of thousands of very similar looking chips moving one behind the other down a conveyor belt, hour after hour, day after day. At some point in the process, salt is added to each chip. How can you be sure that the chips your factory is producing today don't contain too much or too little salt? Do you have to measure the salt content of every potato chip made today? Of course not. It isn't practical to observe every chip. Even if it were, you wouldn't choose to do that, because measuring the amount of salt on a chip actually destroys the chip. If you examined the salt content of every chip produced that day, you'd have no potato chips left to sell! What should you do instead? Select a sample of chips from that day's production and measure the salt content of the chips in the sample.

The potato chip example reminds us of an issue that was discussed briefly in the Introduction. If we want to get information about some characteristic of a population, such as the salt content of the potato chips produced today, we often tend to measure that characteristic on a sample of individuals chosen from the population of interest. We'd like to draw conclusions about the population based on results from the sample. To generalize from sample to population in this way, we need to know that the sample is representative of the population as a whole.

Suppose you measured the salt content of the last 100 potato chips produced at the factory today and found that the chips were generally too salty. Should you conclude that the entire batch of chips produced today is too salty? Not necessarily. Something may have happened during the last hour of production that affected the saltiness of the chips made at the end of the day. The last 100 chips produced may not be a representative sample from the population of today's potato chips.

So how do we get a representative sample? If we choose the first 100 potato chips, or the last 100, or even 100 chips "willy-nilly" off the conveyor belt, we may obtain a sample in which the chips tend to be consistently saltier than or less salty than the entire batch of chips produced that day. The best way to avoid this problem is to let chance select the sample. For example, you might choose one time "at random" in each of the 10 hours of production and measure the salt content of the next 10 potato chips that pass a certain point on the conveyor belt at those times. This incorporates **random selection** into the way the sample is chosen.

Random selection involves using some sort of chance process—such as tossing a coin or rolling a die—to determine which individuals in a population are included in a sample. If the individuals are people, one simple method of random selection is to write people's names on identical slips of paper, put the slips of paper in a hat, mix them thoroughly, and then draw out one slip at a time until we have the number of individuals we want for our sample. An alternative would be to give each individual in the population a distinct number and use the "hat method" with this collection of numbers, instead of people's names. Notice that this variation would work just as well if the individuals in the population were animals or things instead of people.

The hat method works fine if the population isn't too large. If there are too many individuals in the population, however, we would need a very big hat and many small slips of paper. In such cases, it would be easier to "pretend" that we're using the hat method, but to choose the numbers in a more efficient (but equivalent) way.

Technology is the answer. Computers and many calculators have the ability to select numbers "at random" within a specified range, just like drawing the numbers out of a hat. These devices can generate many numbers at random in a short period of time.

Many statistics textbooks contain entire pages filled with rows of "random digits"—numbers from 0 to 9 generated at random using technology. Such tables of random digits were especially useful before the invention of graphing calculators. Here are four rows of random digits that might appear in such a table:

5 2 7 1 1	3 8 8 8 9	9 3 0 7 4	6 0 2 2 7
4 0 0 1 1	8 5 8 4 8	4 8 7 6 7	5 2 5 7 3
9 5 5 9 2	9 4 0 0 7	6 9 9 7 1	9 1 4 8 1
6 0 7 7 9	5 3 7 9 1	1 7 2 9 7	5 9 3 3 5

Now let's consider an example. Kayla wants to conduct an observational study investigating the average number of contacts stored in teenagers' cell phones. She decides to restrict her attention to seniors, most of whom have cell phones. There are 780 seniors in her high school. How might Kayla use random selection to choose a sample of 30 seniors to participate in the cell phone study?

It would be tedious to write 780 names on slips of paper, so Kayla decides to pretend that she's using the hat method. After getting an alphabetized list of the school's seniors from the office, Kayla numbers the students from 1 to 780 in alphabetical order. To choose 30 seniors at random, Kayla can then use either a random digits table or a random number generator.

Random digits table: To use a random digits table, Kayla could look at groups of three digits, which could range from `000` to `999`. If she lets `001` correspond to student 1 on the list, `002` correspond to student 2, and so forth, then `780` would correspond to student 780, the last senior on the list. Numbers `781, 782, ..., 000` would not correspond to any of the students on the list. By starting at the left-hand side of a row in the table and reading across three digits at a time, Kayla would continue until she had chosen 30 distinct numbers between `001` and `780`. The corresponding seniors would be the chosen sample.

Using the lines of random digits on the previous page, for example,

5 2 7 **1 1**	**3** 8 8 8 **9**	**9 3** 0 7 4	**6 0 2** 2 7
4 **0 0 1** 1	8 5 **8 4 8**	4 8 7 **6 7**	**5** 2 5 7 **3**
9 5 5 9 2	**9 4 0** 0 7	6 **9 9 7** 1	9 1 **4 8 1**
6 0 7 **7 9**	**5** 3 7 9 **1**	**1 7** 2 9 7	**5 9 3** 3 5

the senior numbered `527` would be chosen first, and the senior numbered `113` would be selected second. Kayla would skip the numbers `888` and `993` because they don't correspond to any seniors, and so on. Continuing likewise, the first 10 students in the sample would be the seniors numbered `527, 113, 074, 602, 274, 001, 185, 487, 675,` and `257`. The eleventh student selected would be the senior numbered `395`. Do you see why?

Random number generator: Kayla could also use her calculator or computer to generate a "random integer" from 1 to 780. She would repeat this until she got 30 distinct numbers from 1 to 780. The seniors on the alphabetized list with the corresponding numbers would be the chosen sample.

In this example, Kayla entered the command `randInt(1,780)` on a TI-84 calculator and pressed ENTER several times to repeat the command. The first ten resulting numbers were `718, 512, 653, 416, 190, 89, 689, 519, 470`, and `44`. So the seniors with these numbers would be included in her sample.

We used the "random integer generator" at *www.random.org* as an alternative and came up with the numbers here.

Random Integer Generator

Here are your random numbers:

741	72	355	297	755
559	398	629	47	310
536	304	752	397	483
388	405	149	634	699
739	152	721	516	640
293	589	714	771	566

If random selection is accomplished by using the hat method or mimicking it with random numbers, the resulting sample is called a **random sample**. To be classified as a random sample, the *n* selected individuals must have been chosen by a method that ensures:

(1) each individual in the population has an equal chance to be included in the sample

(2) each group of *n* individuals in the population is equally likely to be chosen as the sample

In the cell phone study example, Kayla did obtain a random sample. Once she selected the students for her observational study, it might have been quite difficult for Kayla to locate the 30 seniors who were chosen in a school with so many students, however. For practical reasons, Kayla might have used a method of random selection that didn't result in a truly random sample.

If, for example, the 780 seniors were assigned to 30 homerooms of 26 seniors each based on their last names, Kayla might have decided to select one student at random from each homeroom for her cell phone study. Notice that this alternative method of random selection does give each senior in Kayla's school an equal chance to be included in the sample, but it does not give every group of 30 seniors an equal chance to actually be chosen as the sample. In fact, with this method, the chance of getting a sample with two or more students from the same homeroom is zero!

Think back to the potato chip example for a minute. Can you imagine how difficult it would be to take a random sample from all of the potato chips produced in one day? Just picture someone numbering the individual potato chips for starters! It would be much more feasible to select, say, 10 consecutive potato chips from a particular spot on the conveyor belt by choosing a time at random during each hour of production.

Some observational studies do not use random selection to select the individuals who participate. In the hand-washing study from the Introduction, for example, observers simply watched whoever happened to be in public restrooms at the time. Perhaps the kinds of people who use public restrooms at sporting events, in museums or aquariums, and in train stations have different hand-washing habits than the population of adults as a whole.

The researchers from the University of Arizona used volunteer college students from the United States and Mexico in their observational study of talking patterns by gender. Because of the way in which their sample was chosen, their conclusion about male and female talking tendencies wouldn't necessarily apply to older adults or to college students from other countries. In fact, the results might not even extend to all college students, since some—perhaps those who talk a lot—might have refused to participate in the study. Lack of random selection limits our ability to generalize from the sample to a larger population of interest.

In the investigations that follow, you will learn more about designing and analyzing results from observational studies. You will see firsthand how the presence or absence of random selection affects our ability to generalize.

Teacher Notes for Investigation #2: Get Your Hot Dogs Here!

The first investigation in the Observational Studies section asks students to use graphical and numerical tools of data analysis, combined with a heavy dose of common sense, to examine data from a study that did not involve random selection. In this study, Consumers Union selected a convenience sample consisting of one package each of 54 brands of hot dogs. For each brand, they recorded information on several variables, including calorie content, protein-to-fat rating, sodium content, and an overall sensory rating. In this investigation, students are gradually taken through the process of analyzing the hot dog data collected by Consumers Union. As students answer the questions, they should begin to understand the connection between the way in which the data were produced and the kinds of conclusions that can and cannot be drawn from them.

Prerequisites

Students should be able to:

- Distinguish an observational study from a survey or an experiment
- Describe the distribution of a categorical variable using counts, percents, and bar graphs
- Describe the shape, center, and spread of the distribution of a quantitative variable from a dotplot
- Compare distributions of a quantitative variable using back-to-back stemplots, parallel boxplots, and numerical summaries (mean, median, mode; range, interquartile range (IQR), standard deviation)
- Describe the relationship between two quantitative variables using a scatterplot
- Examine the effect of adding a categorical variable to a scatterplot on the relationship between two quantitative variables
- Interpret the slope and y-intercept of a summary line in the context of a problem
- Use a scatterplot with a summary line to make predictions of y from x
- Explain why lack of random selection limits our ability to generalize sample results to a larger population of interest

Learning Objectives

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

- Describe how the method of data production affects their ability to draw conclusions from the data
- Carry out a complete analysis of an observational study involving one or more categorical variables, using counts, percents, and bar graphs to support their narrative comments
- Carry out a complete analysis of an observational study involving one or more quantitative variables, using dotplots, stemplots, boxplots, and numerical measures of center and spread to support their narrative comments

Teaching Tips

This investigation is designed to review the essential graphical and numerical tools for describing distributions of categorical and quantitative variables, and for describing relationships between two or more variables. Depending on your students' background with techniques of data analysis, you can choose to spend more time on questions involving methods that are less familiar to them. The chart below summarizes the exploratory data analysis tools that will be required in this investigation.

Setting	**Graphs**	**Numerical Summaries**
Categorical variables	Bar graphs, pie charts	Counts, percents, proportions
Quantitative variables	Dotplots, stemplots, histograms, boxplots	Center: Mean, median; Spread: range, standard deviation, interquartile range (IQR)
Relationships between categorical variables	Comparative bar graphs	Two-way tables; comparisons of counts, percents, proportions
Relationships between quantitative variables	Scatterplots	Means and standard deviations; correlation

Here is a question-by-question breakdown of the investigation:

Questions 1 through 6 focus on the data production process.

In questions 7 through 10, students use counts, percents, and different types of bar graphs to analyze the protein-to-fat rating and its relationship to the type of hot dog.

Questions 11 through 14 examine data on calories per frank. Students are asked to use graphical displays—dotplots, stemplots, and boxplots—and numerical summaries to describe calories per frank for the three types of hot dogs.

In questions 15 and 16, students examine the relationship between sodium content and calories per frank using scatterplots, correlation, and summary lines.

Question 17 asks students to acknowledge that lack of random selection limits their ability to generalize conclusions from the sample of hot dogs that was tested by Consumers Union.

The final two questions (18 and 19) allow students to demonstrate their understanding of data analysis techniques for categorical and quantitative variables by looking at sensory rating (includes taste) and sodium content. These two questions could be used for homework or as an alternative assessment.

Suggested Answers to Questions

1. Consumers Union carried out an observational study. They did not ask the hot dogs any questions and they did not deliberately do anything to the hot dogs in this study in order to measure a response.

2. We would expect some variation in the calorie content of individual Oscar Mayer beef frankfurters. Perhaps 148 calories is a typical or average amount of calories in hot dogs of this brand.

3. By following a standard preparation method, Consumers Union attempted to ensure that any differences in sensory qualities identified by the raters were due to the difference in brand of hot dog, and not to any difference in the way the hot dogs were cooked.

4. Because individual hot dogs vary, it would have been better to get an average sensory rating for each brand.

5. Probably not. There may be some reason that these packages were easy to reach in the store—because they were recently delivered, or perhaps because they are nearing their sell-by date. In either case, the conveniently chosen single package of hot dogs of a given brand may not accurately represent the characteristics of all packages of hot dogs of that brand at the store in question, much less in the population of all packages of hot dogs of that brand.

6. Since Consumers Union obtained its random sample from the packages of Armour beef hot dogs that were in this store at the time, they would really only be safe generalizing to the population of packages of Armour beef hot dogs in this store. It may be that all of these packages of hot dogs came on the same truck from the same warehouse or factory and the same batch of production, or they could have come from multiple shipments, factories, or batches.

7. (a) 5/15, or about 33% (b) 5/17, or about 29.4%

8. Most of the protein-to-fat rating values for the beef hot dogs are poor or below average.

9. (a) The graph on the left is a side-by-side bar graph comparing the counts of beef and meat hot dog brands falling into each of the protein-to-fat rating categories. The graph on the right is a segmented bar graph showing the distribution of the percent of each type of hot dog having poor, below average, and average protein-to-fat ratings. Because the number of hot dog brands of each type in the study are not equal (beef = 20; meat = 17), it would probably be better to use the segmented bar graph for comparing the protein-to-fat ratings. To see why, note that the same number of brands of beef and meat hot dogs earned poor protein-to-fat rating scores, but that about 10 percent more meat hot dog brands received poor scores.

(b) If we look only at the percent of brands with poor protein-to-fat ratings, then meat hot dogs are somewhat less healthy. If we compare the combined percents of hot dog brands with poor and below average ratios; however, beef hot dogs are worse by about 10 percent. Neither type of hot dog is very healthy when it comes to protein-to-fat rating.

10. (a) The graph on the left shows the distribution of type of hot dog for each protein-to-fat rating category. For example, we see that of hot dogs with poor protein-to-fat ratings, 50% are beef and 50% are meat. The graph on the right shows the distribution of protein-to-fat rating for each type of hot dog. For example, we see that for beef hot dogs, about 60% of brands have poor protein-to-fat ratings, about 37% have below average protein-to-fat ratings, and about 3% have average protein-to-fat ratings. The graph on the right allows for better comparison of the protein-to-fat ratings among the different types of hot dogs.

(b) Poultry hot dogs are healthiest in terms of protein-to-fat ratings. If we combine the poor and below average ratings, we find 97% of beef hot dog brands, 88% of meat hot dog brands, but only 30% of poultry hot dog brands. In addition, about 35% of the poultry hot dog brands have above average or excellent protein-to-fat ratings.

11. (a) You might have conjectured that the three distinct clusters correspond to the three different types of hot dogs—beef, meat, and poultry. That's really not the case, however. It is certainly true that more of the low-calorie hot dog brands are poultry.

(b) The marked point is for Best's kosher beef lower fat.

12. The dotplot shows two or three distinct clusters of calorie content for the beef hot dogs, with one unusually low value. A "typical" value for a beef hot dog appears to be around 150 calories. There is a lot of variability in the calorie content of beef hot dogs in the study—from about 110 to 190 calories per frank.

13. (a)

Meat		Beef
7	10	
	11	1
	12	
9 8 6 5	13	1 2 5 9
7 6 0	14	1 8 9 9
3	15	2 3 7 8
	16	
9 5 3 2	17	5 6
2	18	1 4 6
5 1 0	19	0 0

Key: 13| 1 means this brand advertises 131 calories per frank

(b) Both distributions appear to have two distinct clusters of calorie content values—one in the 170s–190s and the other in the 130s–150s—and one brand with noticeably lower calorie content. A "typical" value for a meat hot dog is about 153 calories (the median) and a "typical" value for a beef hot dog is about 152.5 calories (the median). The spread of calorie content values is slightly higher for meat hot dogs—the range is 88 for the meat hot dogs and 79 for the beef hot dogs; the interquartile range (IQR) is

42 for the meat hot dogs and 38.5 for the beef hot dogs. Eat Slim Veal brand hot dogs have much lower calorie content (107) than the other brands of meat hot dogs. Best's kosher beef lower fat brand hot dogs have the lowest calorie content (111) among beef hot dogs.

14. (a) The dotplot shows the calorie content for each brand of hot dog, while the boxplot only shows summary values—minimum, first quartile (Q1), median, third quartile (Q3), and maximum.

(b) The boxplot makes it visually easier to compare the center and spread of calorie content for the three types of hot dogs.

(c) From the comparative boxplot, we see that beef and meat hot dogs have similar calorie content, while poultry hot dogs have fewer calories, on average. In fact, 75% of the poultry hot dog brands have fewer calories than the median calorie content (around 153) of beef and meat hot dogs. The spread (variability) in calorie content for each of the three types is fairly similar, as seen by the nearly equal ranges (between 85 and 90) and interquartile ranges (between 40 and 43). From the comparative dotplot, we see that for each type of hot dog, there seem to be two distinct clusters of brands with respect to calorie content, along with one unusual value. For the poultry hot dogs, the unusual value is the maximum (Foster Farms Jumbo Chicken, which has 170 calories). For the beef and meat hot dogs, the unusual value is the minimum.

15. (a) There appears to be a weak, positive relationship between the amount of sodium and the calorie content of hot dogs. That is, hot dogs with more sodium tend to have more calories, and hot dogs with less sodium tend to have fewer calories.

(b) The highlighted point corresponds to a brand of hot dog with a relatively large amount of sodium (more than 500 mg) per frank, but a relatively low amount of calories (around 100) per frank. (This point corresponds to Kroger Turkey brand hot dogs.)

(c) There appears to be a much stronger positive relationship between sodium per frank and calories per frank for meat and beef hot dogs. Most of the points for poultry hot dogs are below the points for beef and meat hot dogs, which reaffirms that poultry hot dogs tend to have fewer calories than the other two types. It also appears that none of the poultry hot dog brands fall among the varieties of hot dog with the lowest amount of sodium.

16. (a) Considering all the brands tested by Consumers Union, the average calorie content per frank is about 147, and the average sodium content per frank is about 425 mg. Overall, the correlation of 0.52 confirms our visual impression of a moderate, positive, somewhat linear relationship between sodium and calories for these brands of hot dogs.

(b) The slope, 0.196, tells us that for each additional mg of sodium a brand of beef hot dog has, our summary line predicts an average increase in calorie content of about 0.2 calories per frank. The *y*-intercept, 78, means the summary line would predict 78 calories per frank for a brand of beef hot dog with 0 mg of sodium per frank. As all the brands of beef hot dogs tested by Consumers Union had around 300 mg of sodium per frank or more, it is not reasonable to trust such a prediction made with the summary line.

(c) For beef hot dogs: 78 + 0.196(300) = 136.8 or about 137 calories

For meat hot dogs: 62 + 0.232(300) = 131.6 or about 132 calories

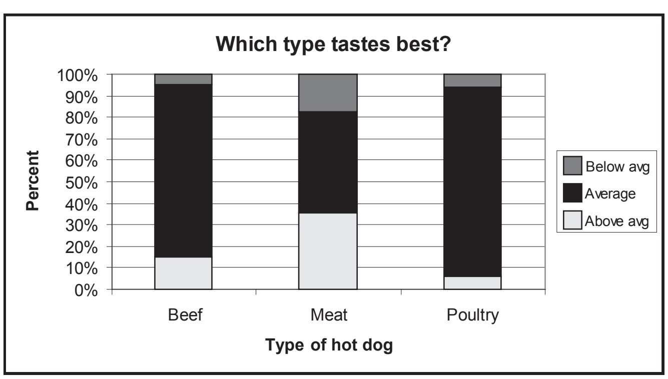
For poultry hot dogs: 24 + 0.214(300) = 88.2 or about 88 calories

(d) We would expect our prediction for poultry hot dogs to be least accurate, as those points vary the most around the corresponding summary line. We would expect our prediction for beef hot dogs to be most accurate, as those points vary the least around the corresponding summary line.

17. No. Consumers Union did not use random selection to choose the brands of beef, meat, and poultry hot dogs they tested. As a result, the chosen brands of each type of hot dog may not represent the population of all brands of that type.

18. Because the number of brands tested differs for the three types of hot dogs, it is more appropriate to compare percents or proportions, rather than counts. In the table below, we have converted the original count data to show the percent of brands in each category with above average, average, and below average sensory ratings.

		Sensory Rating		
Type of Hot Dog		Above Avg.	Average	Below Avg.
	Beef	15.0%	80.0%	5.0%
	Meat	35.3%	47.1%	17.6%
	Poultry	5.9%	88.2%	5.9%



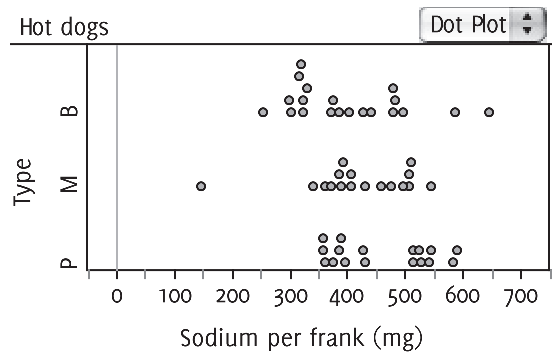
These distributions of sensory ratings are displayed in the bar graph to the left.

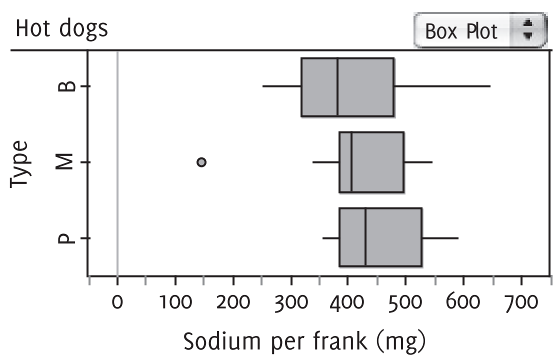
A higher percent of meat hot dog brands (35.3%) received above average sensory ratings than for either of the other two types of hot dogs. On the other hand, meat hot dog brands received more below average sensory ratings than did brands of beef and poultry hot dogs. Beef hot dog brands received the smallest percent (5%) of below average ratings, and a respectable percent (15%) of above average ratings. Poultry hot dog brands were generally rated as average in terms of their sensory qualities, including taste.

For the brands tested, if you want to get the best possible sensory experience, and you're willing to take some risk of having an unpleasant sensory experience, then meat brand hot dogs may be the way to go. If you want to have at least an average sensory experience, go with a brand of beef hot dogs. As Consumers Union did not use random selection to choose packages of hot dogs for testing, we should be hesitant to generalize these results to the larger populations of beef, meat, and poultry hot dogs.

19. Comparative graphs and numerical summaries of the sodium content for the three types of hot dogs are shown to the right and below.

From the comparative dotplot, we see that the distribution of sodium content for the tested brands of beef hot dogs has two distinct peaks—one at around 300 mg of sodium per frank and the other at around 475 mg of sodium per frank. For the tested brands of meat hot dogs, the dotplot also shows two peaks—one at around 400 mg of sodium per frank and one at around 500 mg per frank. One brand of meat hot dog, Eat Slim Veal, had unusually low sodium content (144 mg per frank). We can see this potential outlier clearly on the comparative boxplots. There appear to be two distinct clusters of poultry hot dog brands—those with sodium content between 350 and 450 mg per frank and those with sodium content of between 500 and 600 mg per frank.





Descriptive Statistics: Sodium per Frank (mg)

Variable	Type	N	Mean	SE Mean	StDev	Minimum	Q1
Sodium	B	20	401.2	22.9	102.4	253.0	319.8
	M	17	418.5	22.8	93.9	144.0	379.0
	P	17	459.0	20.6	84.7	357.0	379.0

Variable	Type	Median	Q3	Maximum
Sodium	B	380.5	478.5	645.0
	M	405.0	501.0	545.0
	P	430.0	535.0	588.0

We see from the boxplots and the numerical summaries (previous page) that poultry hot dogs have the highest median and mean sodium content (430 mg and 459 mg, respectively). Since the first quartile of the poultry and meat brand boxplots is at about the median for the beef brand boxplot, we see that about 75% of the poultry and meat brand hot dogs have higher sodium content than the median (380.5 mg) for beef hot dog brands. Notice that the potential outlier pulls the mean calorie content of the meat hot dog brands well below the mean for the poultry brand hot dogs, even though their two medians are much closer. Beef hot dog brands showed the most variability in sodium content, as we can see from the width of the boxes themselves (interquartile range), and the larger range (around 400 mg). Poultry brand hot dogs have more variability in the middle 50% of the distribution than meat brand hot dogs. The standard deviation of calorie content is higher for meat brand hot dogs due to the large distance of the unusually low value from the mean.

Because Consumers Union did not use random selection to choose packages of hot dogs for testing, we should be hesitant to generalize these results to the larger populations of beef, meat, and poultry hot dogs.

Possible Extensions

How healthy is fast food? Fast food companies often make nutritional data on the products they serve available online or in print. Students could use this available data to compare calories, fat, sodium, and other variables across different companies or different food categories (burgers, chicken sandwiches, etc.). As a starting point, we were able to access McDonald's nutrition facts online at *www.mcdonalds.com/usa/eat/nutrition_info.html*. For Burger King, start at *www.bk.com*. For Wendy's, try *www.wendys.com*.

Investigation #2: Get Your Hot Dogs Here!

Corresponds to pp. 14-28
in Student Module

If baseball is America's game, then hot dogs are America's food. Whether you are at a sporting event, a backyard barbecue, or even a local convenience store, you are bound to see folks wolfing down frankfurters. Why do so many people like to eat hot dogs? For the yummy taste, of course! But what makes hot dogs taste so good? Unfortunately for health-conscious eaters, it's probably the fat and sodium they contain. Not all hot dogs are created equal, however. Some are made from beef, others from poultry, and still others from a combination of meats. With so many varieties available, can hot dog lovers find a healthy option that still tastes great?

Several years ago, Consumers Union, an independent nonprofit organization, tested 54 brands of beef, meat, and poultry hot dogs. For each brand tested, they recorded calories, sodium, cost per ounce, a protein-to-fat rating, and an overall sensory rating that included taste, texture, and appearance. The table below and those on the following two pages summarize some of their findings, which were published in *Consumer Reports*.[1] Note that the hot dogs are categorized by type—meat, beef, and poultry.

Meat Hot Dogs

Brand	Protein-to-Fat	Calories per Frank	Sodium per Frank (mg)	Overall Sensory Rating
Armour Hot Dogs	Poor	146	387	Average
Ball Park	Poor	182	473	Above Avg.
Bryan Juicy Jumbos	Poor	175	507	Average
Eat Slim Veal	Average	107	144	Average
Eckrich Jumbo	Poor	179	405	Average
Eckrich Lean Supreme Jumbo	Average	136	393	Average
Farmer John Wieners	Below Avg.	139	386	Average
Hormel 8 Big	Below Avg.	173	458	Above Avg.
Hygrade's Hot Dogs	Poor	195	511	Average
John Morrell	Poor	153	372	Average
Kahn's Jumbo	Poor	191	506	Above Avg.
Kroger Jumbo Dinner	Poor	190	545	Above Avg.
Oscar Mayer Wieners	Poor	147	360	Above Avg.
Safeway Our Premium	Below Avg.	172	496	Above Avg.
Scotch Buy with Chicken & Beef	Poor	135	405	Below Avg.
Smok-A-Roma Natural Smoke	Poor	138	339	Below Avg.
Wilson	Poor	140	428	Below Avg.

1 "Hot dogs: There's not much good about them except the way they taste," *Consumer Reports*, June 1986.

Beef Hot Dogs

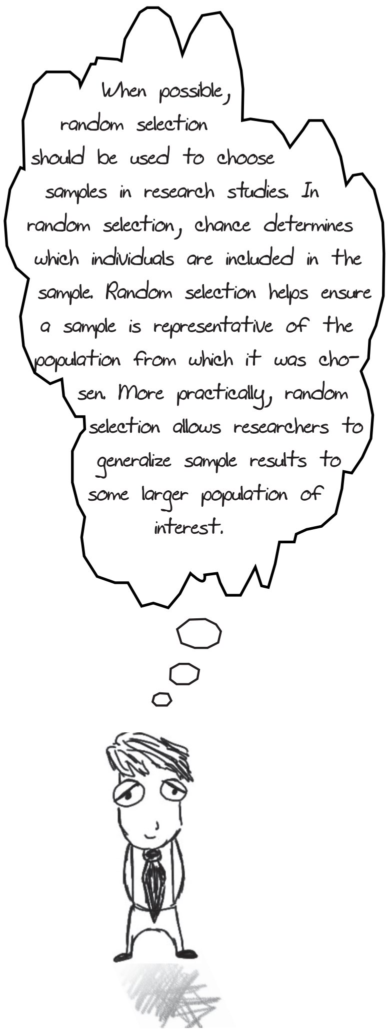
Brand	Protein-to-Fat	Calories per Frank	Sodium per Frank (mg)	Overall Sensory Rating
A & P Skinless Beef	Poor	157	440	Average
Armour Beef Hot Dogs	Poor	149	319	Average
Best's Kosher Beef	Below Avg.	131	317	Average
Best's Kosher Beef Lower Fat	Average	111	300	Average
Eckrich Beef	Poor	149	322	Average
Hebrew National Kosher Beef	Poor	152	330	Average
Hygrade's Beef	Poor	190	645	Average
John Morrell Jumbo Beef	Poor	184	482	Average
Kahn's Jumbo Beef	Poor	175	479	Average
Kroger Jumbo Dinner Beef	Poor	190	587	Average
Mogen David Kosher Skinless Beef	Below Avg.	139	322	Average
Nathan's Famous Skinless Beef	Below Avg.	181	477	Above Avg.
Oscar Mayer Beef	Poor	148	375	Average
Safeway Our Premium Beef	Poor	176	425	Above Avg.
Shofar Kosher Beef	Below Avg.	158	370	Average
Sinai 48 Kosher Beef	Below Avg.	132	253	Below Avg.
Smok-A-Roma Natural Smoke	Below Avg.	141	386	Average
Thorn Apple Valley Brand	Poor	186	495	Above Avg.
Vienna Beef	Below Avg.	135	298	Average
Wilson Beef	Poor	153	401	Average

Poultry Hot Dogs				
Brand	**Protein-to-Fat**	**Calories per Frank**	**Sodium per Frank (mg)**	**Overall Sensory Rating**
Foster Farms Jumbo Chicken	Below Avg.	170	528	Average
Gwaltney's Great Dogs Chicken	Below Avg.	152	588	Average
Holly Farms 8 Chicken	Below Avg.	146	522	Average
Hygrade's Grillmaster Chicken	Average	142	513	Average
Kroger Turkey	Excellent	102	542	Average
Longacre Family Chicken	Above Avg.	135	426	Average
Longacre Family Turkey	Above Avg.	94	387	Average
Louis Rich Turkey	Average	106	383	Average
Manor House Chicken (Safeway)	Average	86	358	Average
Manor House Turkey (Safeway)	Excellent	113	513	Average
Mr. Turkey	Average	102	396	Average
Perdue Chicken	Average	143	581	Average
Shenandoah Turkey Lower Fat	Above Avg.	99	357	Average
Shorgood Chicken	Below Avg.	132	375	Average
Tyson Butcher's Best Chicken	Average	144	545	Below Avg.
Weaver Chicken	Below Avg.	129	430	Above Avg.
Weight Watchers Turkey	Excellent	87	359	Average

The *Consumer Reports* article did not provide many details about how the hot dog data were produced. Our best guess is that Consumers Union first obtained one package of each of the 54 brands of hot dogs they intended to test. For each brand, they could then determine the protein-to-fat rating and the calories and sodium per frank from information provided on the package. To prepare the hot dogs for taste testing, Consumers Union cooked each frankfurter in boiling water.

1. Did Consumers Union produce these data using a survey, an experiment, or an observational study? Justify your answer.

2. According to the data table, Oscar Mayer beef hot dogs have 148 calories per frank. Does this mean that *every* Oscar Mayer beef hot dog has exactly 148 calories, or is there some variability in calorie count from frank to frank? Explain.



3. Why didn't Consumers Union cook some hot dogs in the microwave, others on a grill, and the rest in boiling water?

4. For the taste testing, would it have been better to rate one hot dog of each brand, or to get an average sensory rating for several hot dogs of each brand? Why?

5. It is possible that someone from Consumers Union went to one grocery store in a particular city and picked up one easy-to-reach packet of each brand of hot dogs. Would this **convenience sampling** method result in a representative sample of each brand of hot dogs? Why or why not?

6. Suppose Consumers Union had used random selection to choose a package of Armour beef hot dogs from a single grocery store for testing. If they obtained an average sensory rating for all the hot dogs in the selected package, to what population could they generalize their results—all Armour beef hot dogs ever produced, all Armour beef hot dogs that have ever been sent to this store, or all Armour beef hot dogs in this store at the time the sample was chosen? Justify your answer.

In this study, Consumers Union recorded several variables for each brand of hot dog, including type of hot dog, protein-to-fat rating, calories, sodium, and sensory rating. Two of these are **quantitative variables**—calories and sodium. Type of hot dog, protein-to-fat rating, and sensory rating are **categorical variables**. When we analyze data, the types of graphs and numerical summaries we should use are determined by the type of data we are analyzing. We begin by examining two of the categorical variables: type of hot dog and protein-to-fat rating.

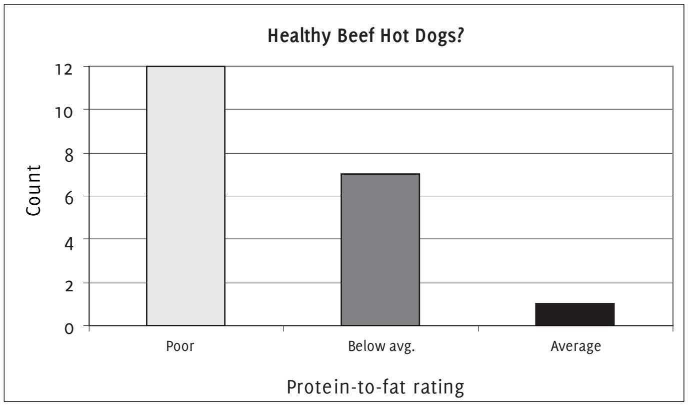
7. Here is a **two-way table** that summarizes the protein-to-fat ratings by type of hot dog.

		Type of Hot Dog		
Protein-to-Fat Rating		Beef	Meat	Poultry
	Poor	12	12	0
	Below Avg.	7	3	5
	Average	1	2	6
	Above Avg.	0	0	3
	Excellent	0	0	3

(a) What percent of hot dogs with a below average protein-to-fat rating were made from poultry?

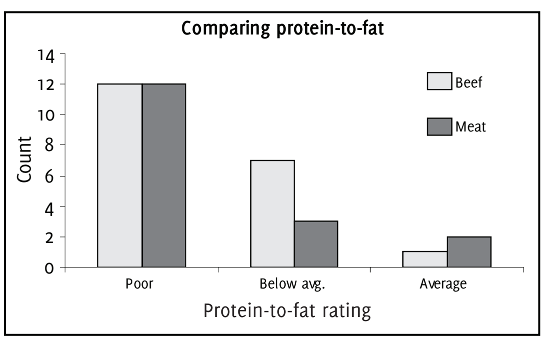
(b) What percent of poultry hot dogs had below average protein-to-fat ratings?

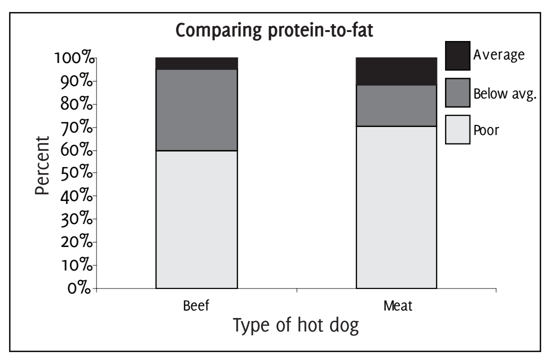
8. Here is an Excel bar graph of the protein-to-fat rating data for the beef hot dogs.



Describe what the graph tells you about protein-to-fat ratings in beef hot dogs.

9. Two Excel bar graphs that could be used for comparing the protein-to-fat ratings for beef and meat hot dogs are displayed below.

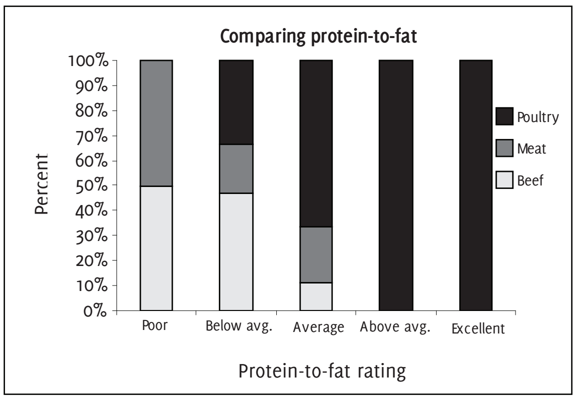


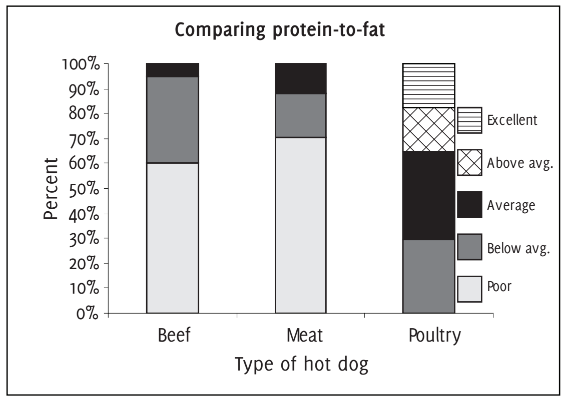


(a) Which graph is more appropriate for making this comparison? Explain.

(b) Write a few sentences comparing protein-to-fat ratings for beef and meat hot dogs.

10. Two different bar graphs that could be used for comparing the protein-to-fat ratings for all three types of hot dogs are displayed below.



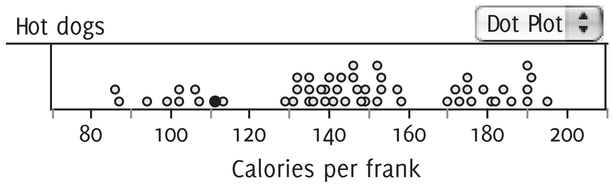


(a) Which graph is more appropriate for making this comparison? Explain.

(b) In terms of protein-to-fat ratings, which type of hot dogs is healthiest? Justify your answer with appropriate graphical and numerical evidence.

Now let's look at the calorie content for different brands of hot dogs.

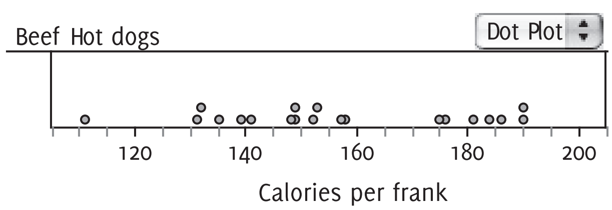
11. A dotplot of the calorie data for all 54 brands of hot dogs is shown below.



(a) Why do you think this distribution has three distinct clusters? Check whether your hunch is accurate.

(b) Identify the brand and type of hot dog for the highlighted point.

12. A dotplot of the calorie content for the 20 brands of beef hot dogs is shown below. Describe the interesting features of this distribution.



13. How does the calorie content of beef and meat hot dogs compare? A partially completed back-to-back stemplot of the calorie data for these two types of hot dogs is shown below.

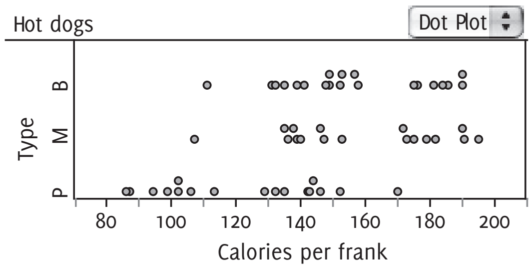
```
Meat                Beef

          |10|
          |11| 1
          |12|
          |13| 1 2 5 9
          |14| 1 8 9 9
          |15| 2 3 7 8
          |16|
          |17| 5 6
          |18| 1 4 6
          |19| 0 0
```

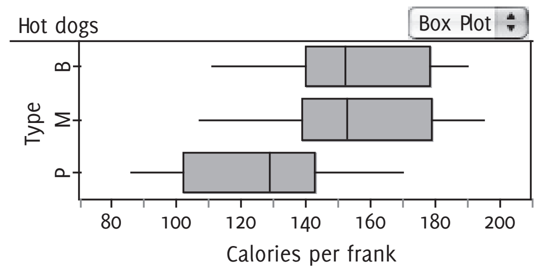
Key: `13| 1` means this brand advertises 131 calories per frank

(a) Add the calorie data for the meat hot dogs to the stemplot. Note that in a back-to-back stemplot, the "leaves" increase in value as you move away from the "stem" in the center of the graph.

(b) Comment on any similarities and differences in the distributions of calories per frank for these two types of hot dogs. Be sure to address center, shape, and spread, as well as any unusual values.

14. To compare calories per frank for all three types of hot dogs, we used computer software to construct graphs and numerical summaries.





Descriptive Statistics: Calories per Frank by Type

Variable	Type	N	Mean	Median	TrMean	StDev
Calories	B	20	156.85	152.50	157.56	22.64
	M	17	158.71	153.00	159.73	25.24
	P	17	122.47	129.00	121.73	25.48
Variable	**Type**	**SE Mean**	**Minimum**	**Maximum**	**Q1**	**Q3**
Calories	B	5.06	111.00	190.00	139.50	179.75
	M	6.12	107.00	195.00	138.50	180.50
	P	6.18	86.00	170.00	100.50	143.50

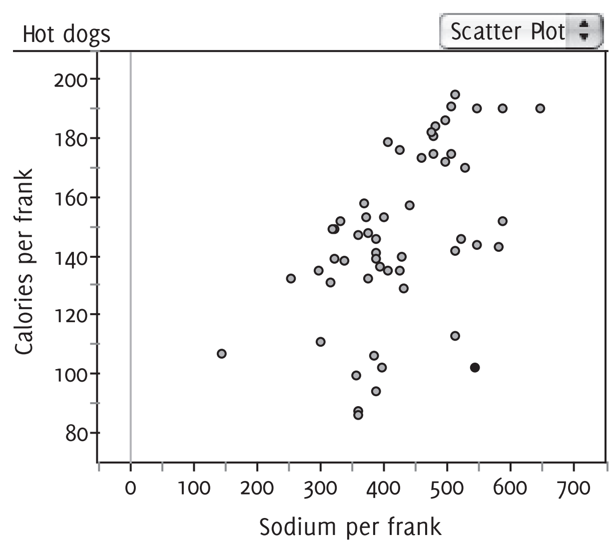
(a) Describe one advantage of using the dotplot instead of the boxplot to display these data.

(b) Describe one advantage of using the boxplot instead of the dotplot to display these data.

(c) How do beef, meat, and poultry hot dogs compare in terms of calorie content? Justify your answer using appropriate graphical and numerical information.

Research Question: Is there a relationship between the calorie content and the amount of sodium per frank in these brands of hot dogs?

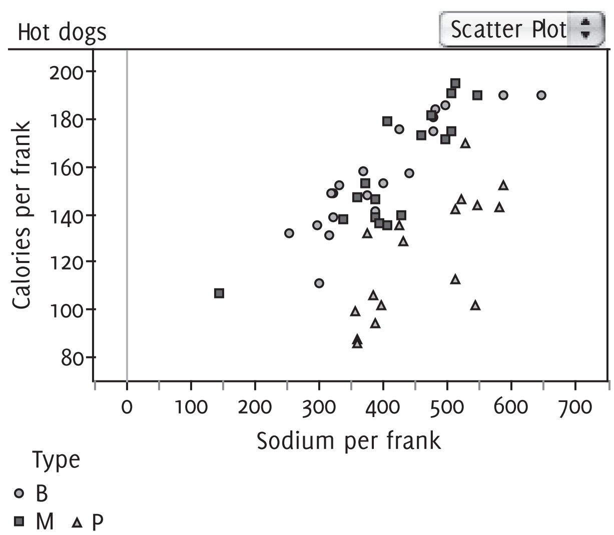
15. The scatterplot below summarizes the sodium and calorie data for the 54 brands of hot dogs in the Consumers Union study.



(a) Describe any interesting features of the scatterplot in the context of this study.

(b) What is unusual about the highlighted point in the scatterplot on the previous page?

Here is another scatterplot of the sodium and calorie data with the type of hot dog identified.



(c) What more can you say about the relationship between sodium and calories per frank when type of hot dog is considered?

16. The next two displays show some numerical summaries of the calorie and sodium data.

Hot dogs

	Calories per frank	Sodium per frank (mg)
	146.611 29.0773	424.833 95.8564

S1 = mean ( )
S2 = stdDev ( )

Hot dogs

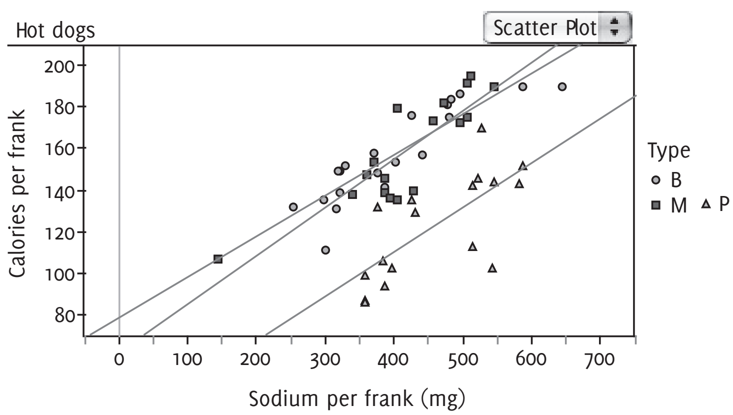
	Calories per frank
Sodium per frank (mg)	**0.516054**

S1 = correlation ( )

(a) What additional information about the relationship between sodium and calorie content of hot dogs do these numerical summaries provide?

The graph below includes three summary lines—one describing the relationship for each type of hot dog.

(b) Interpret the slope and the *y*-intercept of the summary line for beef hot dogs.



— ○ Calories per frank = 78 + 0.196Sodium per frank (mg); $r^2 = 0.79$
— ■ Calories per frank = 62 + 0.232Sodium per frank (mg); $r^2 = 0.75$
— △ Calories per frank = 24 + 0.214Sodium per frank (mg); $r^2 = 0.51$

(c) Suppose Consumers Union had chosen another brand of meat hot dog, beef hot dog, and poultry hot dog, each having 300 milligrams of sodium per frank. What would you predict for the calories per frank in each case? Explain how you made your prediction.

(d) Based on the graph on the previous page, which of the predictions in the previous question do you think would be most accurate? Explain.

17. In the Consumers Union study, beef hot dogs had a mean calorie content of 156.85 calories per frank, compared to 158.71 calories per frank for meat hot dogs and 122.47 calories per frank for poultry hot dogs. Would you feel comfortable generalizing this result about calorie content to the *population* of all brands of beef, meat, and poultry hot dogs? Why or why not?

18. What about the taste? Consumers Union gave an overall sensory rating, which included texture, taste, and appearance. The following table summarizes the ratings by type of hot dog.

		Sensory Rating		
Type of Hot Dog		Above Avg.	Average	Below Avg.
	Beef	3	16	1
	Meat	6	8	3
	Poultry	1	15	1

Which type of hot dog had the best overall sensory ratings? Prepare a brief report that includes graphical and numerical evidence to support your answer.

19. How salty are they? Which have more sodium per frank—beef, meat, or poultry hot dogs? Carry out an analysis that includes graphs and numerical summaries to help answer this question. Write a brief report that summarizes your analysis on a separate piece of paper.

Teacher Notes for Investigation #3: What's in a Name?

This second investigation in the Observational Studies section presents students with another example of a study in which no random selection is used. In the previous investigation, students used data that had been produced by someone else. This time, they will collect their own data on the popularity of the first names of students in their class. Then, students will analyze the data with appropriate graphs and numerical summaries. Finally, students are asked to reflect on how the method of data production affects their ability to generalize results.

Prerequisites

Students should be able to:

- Use proportions to answer questions involving categorical variables
- Describe the shape, center, and spread of the distribution of a quantitative variable
- Distinguish an observational study from an experiment
- Explain why lack of random selection limits our ability to generalize sample results to a larger population of interest

Learning Objectives

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

- Construct an appropriate graphical display (dotplot, stemplot, or histogram) of a quantitative variable
- Describe how the method of data production affects their ability to draw conclusions from the data

Teaching Tips

This is a fairly straightforward observational study based on research showing that people's first names can affect their lives in unexpected ways. For example, two economics professors conducted a study titled "Are Emily and Brendan More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination."[1] These researchers constructed sets of resumes for fictitious individuals, some of whom were highly qualified and others of whom were less qualified. For each resume, they gave the fictitious applicant either a "white sounding" or "black sounding" name. Then, they sent about 5,000 resumes out in response to more than 1,000 jobs that were advertised in Chicago and Boston. What did they find? Candidates with "white sounding" names were 50% more likely to be called back for an interview than candidates with "black sounding" names. These results held for jobs of all types, and for candidates who were extremely qualified and those who were less qualified. It's no wonder many parents spend so much time choosing their children's names!

---

1 "Are Emily and Brendan More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," by Dr. Marianne Bertrand and Dr. Sendhil Mullainathan, NBER Working Paper No. 9873, July, 2003.

Suggested Answers to Questions

For questions **1** through **16**, answers will vary. Note that on question 12, the best choice of graph—dotplot, stemplot, histogram, or boxplot—will depend on the number of values and the spread of those values. A dotplot or stemplot will display all of the actual data values, but they may be awkward to construct if there are too many data values, or if the values are too spread out. In such cases, a histogram or a boxplot would be easier to construct.

17. This study is not an experiment, because nothing was deliberately done to the students in your class to measure their responses. Instead, students' names, genders, and birth years were recorded, and the corresponding decade ranks data were accessed online.

18. No. The data were collected from students in our class. Since random selection was not used to select the students who took part in the study, we can't be sure that our class represents the larger school population well. We should therefore be hesitant to generalize the results about our class' names to the population of students in our school.

19. If the goal was to generalize to all students at our school, we should use random selection to choose a sample of students from the school population, and then record the names of those students. This could be accomplished using a variation of the hat method with an alphabetical roster of students enrolled at our school. In order to generalize to all high-school students in the district, we would need to use random selection to choose a sample of students from the population of high-school students in the district, and then record the names of the selected students. It might be difficult to obtain a complete list of all high-school students in the district. If so, we could use random selection to choose, say, five students from each high school, and then record the names of those students. To generalize to all high-school students in our state, we would need to use random selection to choose a sample of students from the population of all high-school students in the state. It probably wouldn't be practical to compile a list of all high-school students in the state, or to use random selection to choose a few students from every high school in the state. Instead, we could use random selection to choose, say, 10 high schools from our state, and then use random selection to choose five students from each of those schools.

Possible Extensions

Students might enjoy reading an excerpt from Chapter 6 in *Freakonomics*, by Steven D. Levitt and Stephen J. Dubner, William Morrow/HarperCollins publishers, 2005. The chapter's title is catchy: "Perfect Parenting, Part II; or: Would a Roshanda by Any Other Name Smell as Sweet?"

Investigation #3: What's in a Name?

Corresponds to pp. 29-32 in Student Module

According to the *Seattle Times* (Oct. 5, 2003), there will be a lot of Jacobs and Emilys in the high-school graduating class of 2020—those were the most popular baby names in the United States in 2002 according to Social Security card applications.

It's nice to be popular, and great to be "cool." The authors of the book *Cool Names for Babies* (Satran, Pamela & Rosenkrantz, Linda, Harper Collins Publishers, 2004) say that it is the unusual names that are most cool.

In this activity, you will carry out an observational study to assess the popularity and coolness of your class based on the names of the students in class.

Getting Started

To complete this activity, you will need to use the Social Security Administration's Popular Baby Name web site. It can be found at *www.ssa.gov/OACT/babynames*.

On this site, you will be able to find lists of the 10 most popular baby names for boys and girls in each year starting in 1880. These lists were compiled using a random sample consisting of 1% of all babies born in a particular year who subsequently applied for a social security card. You will also find a list of the top 1,000 names for each decade from the 1900s to the 2000s.

Spend a few minutes familiarizing yourself with the information available on this web site. Then, start answering the questions that follow.

1. Let's start with an easy question! What is your first name?

2. Are you male or female?

3. In what year were you born?

4. Is your name one of the 10 most popular names for the year in which you were born?

5. Is your name one of the 10 most popular names for the most recent year for which data are available?

6. Is your name one of the most popular 1,000 for the decade in which you were born? If so, record your name's rank. If your name is not in the top 1,000, just record that your name is "cool"!

7. After each student in your class has answered questions 1–6, enter the data from the entire class into the following table.

First Name	Gender	Year Born	In Top 10 for Year Born? (Yes or No)	In Top 10 for Most Recent Year? (Yes or No)	Rank for Decade Born (1-1,000 or cool)

8. Is there a most common name for the class? If so, what is the most common name?

9. What is the most common year of birth for the class?

10. In the year that was the most common birth year for the class, what is the most popular name for boys according to the popular baby names web site? For girls? Does anyone in the class have these most popular names?

11. What proportion of the class has “cool” names?

12. Omitting the cool names from the data set, construct a graphical display that shows the distribution of the decade ranks data. How would you describe this distribution? (Comment on shape, center, spread, and any unusual values.)

13. What proportion of the class has names that were in the top 10 names for the year in which they were born?

14. Based on your answers to questions 11 and 13, is your class more popular or more "cool?"

15. What proportion of the class has names that are listed in the top 10 for the most recent year for which data are available?

16. Is the proportion from question 15 lower than, about the same as, or higher than the proportion from question 13? How does this suggest that the popularity of the class' names has changed over time?

17. What makes this study an observational study, rather than an experiment?

18. Was there random selection in the data collection for this study? How does this affect your ability to generalize from the study?

19. How might you modify this study if your goal was to generalize to all students at your school? To all high-school students in your school district? To all high-school students in your state?

Teacher Notes for Investigation #4: If the Shoe Fits ...

This investigation puts students in the role of data detectives as they attempt to use a footprint left behind at the scene of the crime to help school administrators identify the perpetrator. First, students examine shoe print length data for a random sample of male and female students. Next, students explore the relationship between height and shoe print length for both male and female students. With their preliminary analysis complete, students must then decide whether the shoe print left at the scene belonged to a male or a female, and predict the height of the culprit. Unlike the two previous investigations, the random selection of students in this observational study allows your students to generalize their findings to the population of all students at the high school.

Prerequisites

Students should be able to:

- Describe the shape, center, and spread of the distribution of a quantitative variable using a dotplot and numerical summaries (mean, median; range, interquartile range (IQR), standard deviation)
- Identify potential outliers
- Distinguish an observational study from an experiment
- Describe the relationship between two quantitative variables using a scatterplot
- Use a scatterplot to make predictions of *y* from *x*

Learning Objectives

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

- Explain why random selection in an observational study allows sample results to be generalized to a larger population of interest
- Examine the effect of adding a categorical variable to a scatterplot on the relationship between two quantitative variables
- Draw conclusions about a population based on graphical and numerical information from a random sample
- Make decisions in the face of uncertainty using graphical and numerical information

Teaching Tips

CSI stands for Crime Scene Investigation.

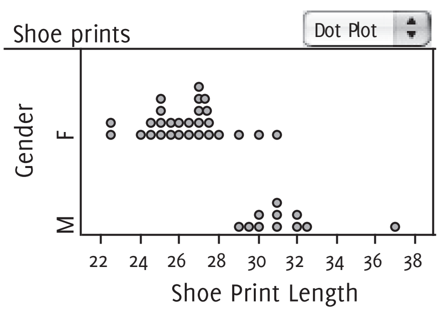
Note that the shoe print lengths were measured in centimeters, but the heights were measured in inches. This was done deliberately to force students to think carefully about units of measurement.

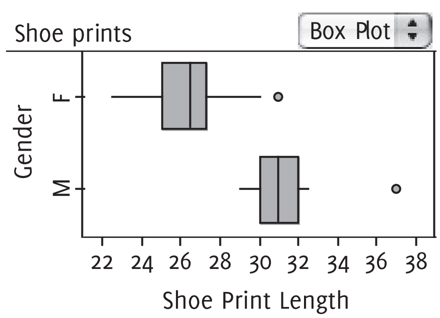
Be sure to emphasize an essential difference between this investigation and the previous two investigations: the use of random selection to produce the data. Random selection is our best attempt to ensure that the sample we choose is representative of the population of interest. Random selection allows us to generalize the results of a sample to the population at large.

Include discussion of using a line of best fit (Q-Q line, median-median line, least-squares line) to make height predictions if your students have the necessary background.

Suggested Answers to Questions

1. Here are comparative dotplots and comparative boxplots of the shoe print lengths from Fathom software.





2. We can see from the plots that the males in this sample tended to have longer shoe prints than did the females. A "typical" male in the sample had a shoe print length of about 31 cm, while a "typical" female in the sample had a shoe print length of about 27 cm. There was one male in the sample with an unusually long shoe print—37 cm, or about 14.5 inches! This individual is an outlier according to the 1.5*IQR* rule:

$$Q_3 + 1.5IQR = 32 + 1.5(2) = 35 \text{ is the upper cutoff.}$$

You might be surprised to discover that the female with the longest shoe print is also identified as an outlier by the 1.5*IQR* rule:

$$Q_3 + 1.5IQR = 27.325 + 1.5(2.325) = 30.8125 \text{ is the upper cutoff.}$$

There is slightly more variability in the female shoe print length distribution (range = 8.5 cm; *IQR* = 2.325 cm) than in the male shoe print length distribution (range = 8 cm; *IQR* = 2 cm). The shape of the male shoe print length distribution is roughly symmetric with one extremely high outlier. The female shoe print length distribution appears somewhat bimodal.

3. Since nothing was deliberately done to the individuals in this study to measure their responses, this was not an experiment. Instead, the individuals were simply observed and their shoe print lengths, heights, and genders were recorded. That makes this an observational study.

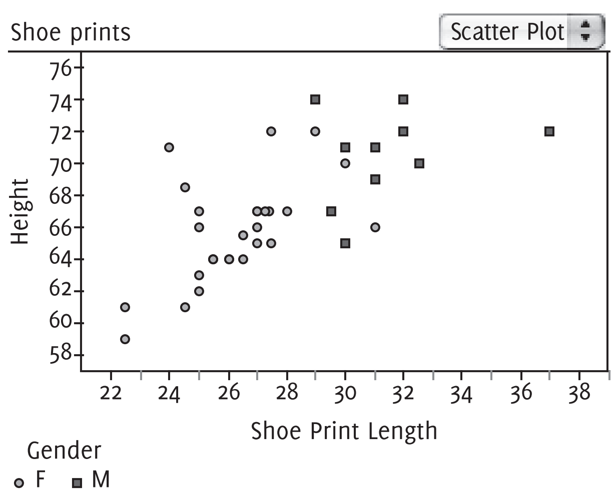
4. The administrators probably wanted to use information from this sample of students to draw conclusions about the population of all students at the school. Their best method for attempting to obtain a representative sample was to choose the students for this study using random selection.

5. A male. There were no females in the random sample with a shoe print length as long as 32 cm, but there were four males with shoe print lengths of 32 cm or more. Although it is possible that a 32 cm shoe print length could be from a female student at the school (not one of the ones in this sample), a more plausible explanation is that the shoe print came from a male student.

6. If the suspect's shoe print length were 27 cm, we would suspect that the suspect was a female. None of the males in the random sample had shoe print lengths less than 29 cm, while about half of the females in the sample had shoe print lengths of 27 cm or less. Although it is possible that a 27 cm shoe print length could be from a male student at the school (not one of the ones in this sample), a more plausible explanation is that the shoe print came from a female student.

If the suspect's shoe print length were 29 cm, we would have a much more difficult time deciding whether the culprit was a male or a female student. A shoe print length of 29 cm falls toward the top end of the distribution of shoe print lengths for the females in the random sample. Three of the 28 females in the sample had shoe print lengths of 29 cm or higher. However, one of the 11 male students in the random sample had a shoe print length of 29 cm. Based on the sample data, it is plausible that a 29 cm shoe print could have come from either a male or female student.

7. The Fathom screen shot below shows the height versus shoe print length data for the students in this random sample. There appears to be a moderately strong linear relationship between shoe print length and height for these students.



8. No. It appears that two different lines would be needed to summarize the relationship between shoe print length and height—one for females with a steeper slope and lower *y*-intercept, and one for males with a less steep slope but higher *y*-intercept.

9. As students don't know whether the intruder was male or female, they will have to base their predictions on the entire set of sample data. For a shoe print length of 30 cm, the scatterplot suggests a height of about 69 inches. This is close to the average height of the three students in the random sample that had shoe print lengths of 30 cm. In addition, the least-squares regression line for the entire sample of values is:

$$\text{Height} = 44 + 0.832(\text{Shoe print length}).$$

If we substitute 30 for the shoe print length, we get a value of 68.96 inches.

10. From the original doplot, we see that a shoe print length of 31 cm was at the center of the distribution for males in the random sample, but equal to the maximum shoe print length for the 28 females in the sample. As a result, we would infer that the suspect was probably a male. Using the scatterplot of height versus shoe print length, for a male with a 31 cm long shoe print, we would predict a height of around 70 cm. The least-squares line using only the male data is:

$$\text{Height} = 59 + 0.362(\text{Shoe print length}).$$

If we substitute 31 for the shoe print length, we get a value of 70.22 inches. Alternatively, we could have used the average height for the three males in the random sample who had 31 cm long shoe prints: 69.67 inches.

Possible Extensions

1. You might want to have students collect and use data from a random sample of students at your school to draw a conclusion about the perpetrator of the crime. To make shoe prints, have each student step in water and then step on a piece of newspaper. Be sure to measure the length of the shoe print at its longest point for consistency.

2. As a variation on this investigation, you could create a scenario in which a hand print was found at the scene of the "crime," instead of a foot print.

Investigation #4: If the Shoe Fits . . .

Welcome to CSI at School. Over the weekend, a student entered the school grounds without permission. Even though it appears the culprit was just looking for a quiet place to study undisturbed by friends, school administrators are anxious to identify the offender and have asked for your help. The only available evidence is a suspicious footprint outside the library door.

Corresponds to pp. 33-36 in Student Module

In this activity, you will use data on shoe print length, height, and gender to help develop a tentative description of the person who entered the school.

After the incident, school administrators arranged for the data in the table below to be obtained from a random sample of this high school's students. The table shows the shoe print length (in cm), height (in inches), and gender for each individual in the sample.

Shoe Print Length	Height	Gender	Shoe Print Length	Height	Gender
24	71	F	24.5	68.5	F
32	74	M	22.5	59	F
27	65	F	29	74	M
26	64	F	24.5	61	F
25.5	64	F	25	66	F
30	65	M	37	72	M
31	71	M	27	67	F
29.5	67	M	32.5	70	M
29	72	F	27	66	F
25	63	F	27.5	65	F
27.5	72	F	25	62	F
25.5	64	F	31	69	M
27	67	F	32	72	M
31	69	M	27.4	67	F
26	64	F	30	71	M
27	67	F	25	67	F
28	67	F	26.5	65.5	F
26.5	64	F	27.25	67	F
22.5	61	F	30	70	F
			31	66	F

Use the data provided to answer the questions that follow.

1. Construct an appropriate graph for comparing the shoe print lengths for males and females.

2. Describe the similarities and differences in the shoe print length distributions for the males and females in this sample.

3. Explain why this study was an observational study and not an experiment.

4. Why do you think the school's administrators chose to collect data on a random sample of students from the school? What benefit might a random sample offer?

5. If the length of a student's shoe print was 32 cm, would you think the print was made by a male or a female? How sure are you that you are correct? Explain your reasoning.

6. How would you answer question 5 if the suspect's shoe print length was 27 cm? 29 cm?

7. Construct a scatterplot of height versus shoe print length using different colors or different plotting symbols to represent the data for males and females. Does it look like there is a linear relationship between height and shoe print length?

8. Does it look like the same straight line could be used to summarize the relationship between shoe print length and height for both males and females? Explain.

9. Based on the scatterplot, if a student's shoe print length was 30 cm, approximately what height would you predict for the person who made the shoe print? Explain how you arrived at your prediction.

10. The shoe print found outside the library actually had a length of 31 cm. Based on the given data and the analysis of questions 1–9, write a description of the person who you think may have left the print. Explain the reasoning that led to your description and give some indication of how confident you are that your description is correct.

Teacher Notes for Investigation #5: Buckle Up

This investigation asks students to examine the change in seat belt use in the states from 2004 to 2005. Data were collected by observers at a random sample of road locations in each state. As in the previous investigation, the random selection should allow students to generalize the sample results to the population of road locations in each state. Students are first asked to consider whether the seat belt use data from the two years should be compared as two separate lists of values, or whether they should analyze the difference in seat belt use from 2004 to 2005 for the states. Then, they must use graphs and numerical summaries to describe the overall change in seat belt use. Students are later asked about the position of their state within the distribution of change in seat belt use. After reviewing details of how the data were produced, students must produce their own summary analysis of the change in seat belt use by drivers.

Prerequisites

Students should be able to:

- Describe the shape, center, and spread of the distribution of a quantitative variable using a dotplot and numerical summaries (mean, median, mode; range, interquartile range (IQR), standard deviation)
- Identify potential outliers
- Distinguish an observational study from a survey or an experiment
- Describe the relationship between two quantitative variables using a scatterplot

Learning Objectives

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

- Explain why random selection in an observational study allows sample results to be generalized to a larger population of interest
- Decide whether to use a comparative dotplot or a dotplot of differences to analyze data from two samples
- Describe the relative standing of one value within a distribution
- Explain why an observational study might be preferable to a survey in describing individuals' behavior
- Draw conclusions about a population based on graphical and numerical information from a random sample

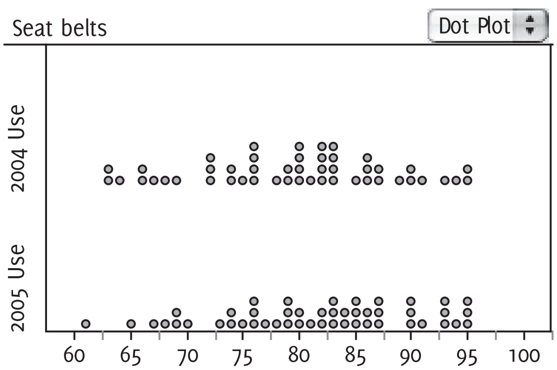
Teaching Tips

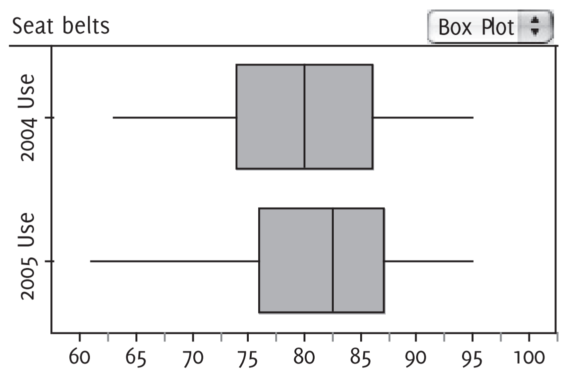
At the time we were writing, the NHTSA's 2007 seat belt use survey was completed, but the data had not yet been released for the individual states. You may want to use more current data from the NHTSA's web site, *www.nhtsa.dot.gov*, for this investigation.

In NHTSA's seat belt usage study, a random sample of road locations within each state was chosen, and then drivers' seat belt use was recorded at those locations. The same method of producing data was used in both 2004 and 2005. Random selection of road locations within each state should allow the seat belt use results to be generalized to all road locations in the state. One value—the overall percent of drivers in the state who were observed wearing seat belts—was recorded for each state in 2004 and again in 2005. These two lists of 48 values are not two independent sets of values. We would expect similar seat belt use within a state from one year to the next. Consequently, we should analyze the differences in seat belt use from 2004 to 2005, and not the two sets of values separately. Through this investigation, we're trying to help students learn an important statistical lesson: The way in which the data were produced determines the proper way of analyzing the data.

Suggested Answers to Questions

1. The Fathom screen shots below show comparative dotplots and comparative boxplots of seat belt use in 2004 and 2005.



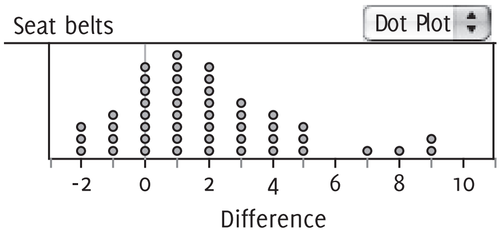


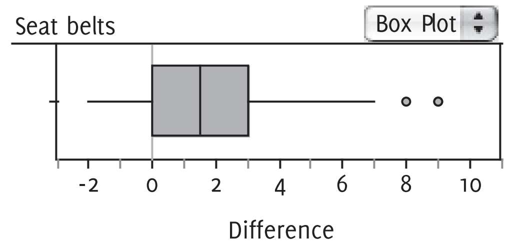
In 2004, between 63% and 95% of the samples of drivers observed in each state were wearing their seat belts. A typical (median) value for seat belt use was 80%. The distribution is roughly symmetric. There are no potential outliers according to the 1.5*IQR* rule.

In 2005, between 61% and 95% of the samples of drivers observed in each state were wearing their seat belts. A typical (median) value for seat belt use was 82.5%. The distribution is skewed to the left. There are no potential outliers according to the 1.5*IQR* rule.

The center of the 2005 seat belt use distribution (median = 82.5%) is slightly larger than for the 2004 distribution (median = 80%). Variability in seat belt use among states for the two years is pretty similar (2005 range = 34%; 2004 range = 32%). There appear to be more states toward the higher end of the seat belt use distribution in 2005 than in 2004.

2. A Fathom dotplot and boxplot of the change in seat belt use are shown below. This distribution is skewed to the right. The middle 50% of states showed increases in seat belt use (based on the samples) of between 0% and 3%. Thirty-three of the 48 states showed increases in seat belt use (based on the samples) from 2004 to 2005. The median increase in seat belt use (based on the samples) was 1.5%.





3. The comparative dotplot and boxplot look at the data as two separate lists of values, which hides the paired nature of the data. When we examine the differences in seat belt use via the graphs in question 2, we get a much clearer picture of how seat belt use has changed from 2004 to 2005.

4. Based on the sample results, yes. Thirty-three of the 48 differences are positive, which indicates an increase in seat belt use for those 33 states from 2004 to 2005.

5. Mean = 1.96%; Median = 1.5%

6. The long tail to the right, which includes the three high potential outliers, pulled the mean in that direction. The median is resistant to these extreme values.

7. Use the median, so that the few extremely high values don't make the improvement in seat belt use look better than perhaps it was.

8. There are four potential outliers—7% (Texas), 8% (Nevada), 9% (North Dakota), and 9% (West Virginia). According to the 1.5*IQR* rule:

$$Q_3 + 1.5IQR = 3 + 1.5(3) = 7.5 \text{ is the upper cutoff.}$$

Only Nevada, North Dakota, and West Virginia are identified as outliers. These states showed dramatic increases in seat belt use (based on the samples) from 2004 to 2005.

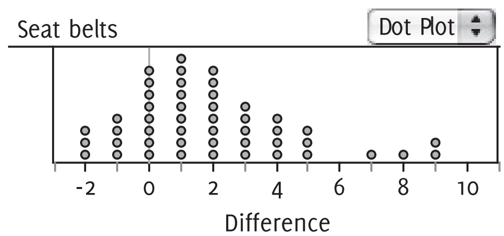
9. Answers will vary. In New Jersey, for example, seat belt use increased by 4% (based on the samples). This result puts New Jersey in the top 25% of the distribution. So, New Jersey's change in seat belt use is not that typical.

10. Drivers were observed in their vehicles to see if they were wearing seat belts. Drivers were not asked whether they wore seat belts, so this wasn't a survey. Researchers did not deliberately do anything to vehicle drivers to measure their responses, so this wasn't an experiment.

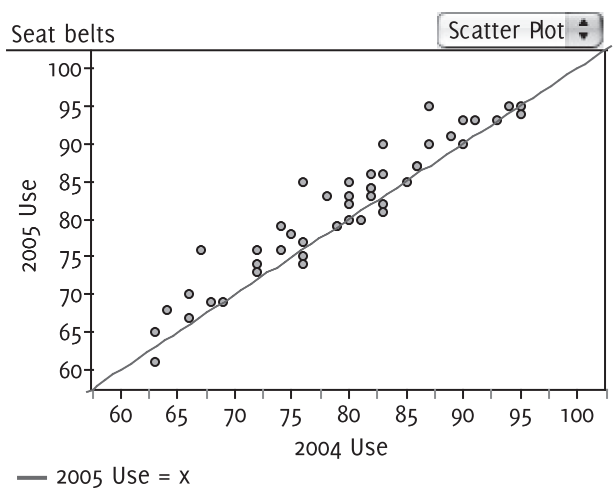
11. Observations of drivers' behavior should produce more accurate information than drivers' self-reported behavior. Some people might claim they wear seat belts when they don't.

12. Because the data were based on observations at a random sample of roadway sites in a given state, we should be able to generalize the results to the population of roadway sites in that state. Note, however, that because the seat belt use value in each state was based on a sample of drivers, the actual seat belt use by drivers in that state might differ a little. A different random sample of roadway sites would probably have led to slightly different seat belt use statistics.

13. Drivers' seat belt use in the states appears to have improved between 2004 and 2005. In that two-year period, 33 of 48 states showed increases in their seat belt use statistics. A typical increase in seat belt use was about 1.5% (the median). The dotplot to the right shows the distribution of changes in seat belt use by state from 2004 to 2005. Note the long right tail, which includes four states with unusually high increases in seat belt use: Texas (7%), Nevada (8%), North Dakota (9%), and West Virginia (9%).



The scatterplot to the right shows a strong, positive, linear relationship between observed seat belt use in 2004 and in 2005. We have added the line $y = x$ for reference. Points above the line represent states with higher observed seat belt use in 2005 than in 2004. Points below the line represent states with lower observed seat belt use in 2005 than in 2004. As most of the points are above the line, we see that observed seat belt use improved in most states during this two-year period.



Possible Extensions

1. You might want to have students conduct their own observational study of seat belt use at your school, or in a nearby location. Be sure to get permission from appropriate officials before allowing students to observe drivers' behavior. Also, ensure that students will be safe while making their observations.

2. Students could investigate whether drivers lock their car doors when they park at home versus when they park elsewhere.

Investigation #5: Buckle Up

Corresponds to pp. 37-40
in Student Module

Do you wear your seat belt when driving? Do most people? Is seat belt use changing over time? To answer questions such as these (well, at least the last two questions—only you know the answer to the first question, but we sure hope the answer is yes!), the National Center for Statistics and Analysis published data on seat belt use for 48 states. No data were available for New Hampshire or Wyoming.

The data shown in the table at the top of the next page are from a large-scale study conducted annually by the National Highway Traffic Safety Administration.[1] The study involves actual observation of drivers' seat belt use at a random selection of roadway sites in each state.

The table gives the percentage of drivers observed who used seat belts in 2004 and in 2005. The table also shows the change in seat belt use percentage from 2004 to 2005 (computed as 2005 use percentage – 2004 use percentage).

Use the data in the table to answer the following questions.

1. Would comparative dotplots or comparative boxplots be better for comparing the seat belt use rates for 2004 and 2005? Make the graph that you pick. Then write a sentence or two describing the similarities and differences in the seat belt use rate distributions in 2004 and 2005.

2. Construct an appropriate graph that shows the change in seat belt use by state from 2004 to 2005. Comment on any interesting features of the distribution.

1 "Seat Belt Use in 2006—Use Rates in the States and Territories," Traffic Safety Facts, National Highway Traffic Safety Administration, January 2007.

State	2004 Use	2005 Use	Difference	State	2004 Use	2005 Use	Difference
Alabama	80	82	2	Missouri	76	77	1
Alaska	78	83	5	Montana	81	80	-1
Arizona	95	94	-1	Nebraska	79	79	0
Arkansas	64	68	4	Nevada	87	95	8
California	90	93	3	New Jersey	82	86	4
Colorado	79	79	0	New Mexico	90	90	0
Connecticut	83	82	-1	New York	85	85	0
Delaware	82	84	2	No. Carolina	86	87	1
Florida	76	74	-2	North Dakota	67	76	9
Georgia	87	90	3	Ohio	74	79	5
Hawaii	95	95	0	Oklahoma	80	83	3
Idaho	74	76	2	Oregon	93	93	0
Illinois	83	86	3	Pennsylvania	82	83	1
Indiana	83	81	-2	Rhode Island	76	75	-1
Iowa	86	87	1	So. Carolina	66	70	4
Kansas	68	69	1	South Dakota	69	69	0
Kentucky	66	67	1	Tennessee	72	74	2
Louisiana	75	78	3	Texas	83	90	7
Maine	72	76	4	Utah	86	87	1
Maryland	89	91	2	Vermont	80	85	5
Massachusetts	63	65	2	Virginia	80	80	0
Michigan	91	93	2	Washington	94	95	1
Minnesota	82	84	2	West Virginia	76	85	9
Mississippi	63	61	-2	Wisconsin	72	73	1

3. In what way is the graph in question 2 more informative than the graph in question 1?

4. Did most states increase seat belt use from 2004 to 2005? What aspect of the graph you made in question 2 could be used to justify your answer?

5. Compute the mean and median change in seat belt use.

6. What aspect of the graph you made in question 2 explains the large difference between the mean and the median?

7. Would you recommend using the mean or the median to describe the seat belt use change data? Why?

8. Are there any states that stand out as unusual in this data set? If so, which states and what makes them unusual?

9. How did seat belt use in your state change from 2004 to 2005? Would you describe your state as typical with respect to seat belt use change? Explain. (If your state is one of the two states for which no data are given, choose a neighboring state and answer this question for that state.)

10. What makes this seat belt use study observational, rather than an experiment?

11. Why do you think the study was based on actual observation of drivers, rather than a survey of drivers asking if they use a seat belt when driving?

12. Based on the sampling method used in this study, do you think it would be reasonable to generalize the seat belt use results to drivers at all locations in a given state? Explain.

13. Write a brief summary report describing how seat belt use changed from 2004 to 2005. Include graphs and numerical summaries as appropriate.

Teacher Notes for Investigation #6: It's Golden (and It's Not Silence)

In this investigation, students will design and carry out an observational study to determine whether students prefer golden rectangles to nongolden rectangles. Students must first decide on a sampling plan that will allow them to generalize their sample results to all students at the school. After tweaking the plan for practical considerations, they will implement the plan to collect data. With the data in hand, students will perform a preliminary graphical and numerical analysis. Finally, students are asked to prepare a report that summarizes their findings.

Prerequisites

Students should be able to:

- Analyze a categorical variable using bar graphs, counts, and percents
- Explain how to obtain a true random sample from a population of interest

Learning Objectives

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

- Design a practical sampling plan that incorporates random selection
- Conduct an observational study in a way that should produce reliable data
- Consider whether the outcome of an observational study could simply be due to chance, rather than an actual preference among individuals
- Draw conclusions about a population based on graphical and numerical information from a representative sample

Teaching Tips

For the three rectangles shown on the student handout, the ratio of the longest side to the shortest side is Rectangle #1: 5.33, Rectangle #2: 2.00, Rectangle # 3: 1.60.

Students will need access to a list of students at your school if they are going to select individual students at random. Otherwise, they will need a list that shows other possible sampling units—homerooms, grade levels, math classes, etc. You may need some lead time to acquire the necessary information.

A major purpose of this investigation is to reinforce the benefit of using random selection in choosing a sample—the ability to generalize to a larger population of interest. If it is feasible for students to carry out this observational study using a true random sample from the population of students at your school, they should do so. If true random sampling isn't possible, any practical adjustments your students propose should incorporate random selection. Students often confuse "haphazard" selection with random selection. Remind students that random selection requires a chance mechanism (such as the hat method) be used to choose the individuals in the sample.

There is no prescribed method for determining how many rectangles to use or what dimensions the nongolden rectangles should have in questions 4 and 5. Likewise, we did not specify the number of students who should take part in the observational study. Too few rectangles will make it difficult to distinguish whether students actually have a preference for the golden rectangle or are simply choosing a rectangle at random. Too many rectangles might result in small numbers of students choosing each of the rectangles, which would again make it difficult to detect a preference for the golden rectangle.

Here's an example to help you think through the decisions discussed in the previous tip. Suppose your students opt to use three rectangles, one of which is golden. Suppose further that students decide to choose 30 individuals at random to participate in the study. If individuals actually have no preference among the rectangles in terms of their "goldenness," we would expect about 10 individuals to choose each of the three rectangles. What if 15 students actually say they prefer the golden rectangle? Is it plausible that students are simply picking rectangles at random, and that, just by chance, 15 picked the golden rectangle? Consider rolling a die to simulate such a chance process. Let outcomes 1 and 2 represent students who pick the golden rectangle and outcomes 3, 4, 5, and 6 represent students who pick one of the other two rectangles. Roll the die 30 times, once for each student in the observational study. How often does such a simulation result in as many as 15 people picking the golden rectangle?

(The theoretical probability of 15 or more students picking the golden rectangle if all are choosing at random is about 0.043, using a binomial distribution with $n = 30$ and $p = 1/3$.)

You might want to show students Jim Loy's Most Pleasing Rectangle Poll web page, *www.jimloy.com/poll/poll.htm*, for interesting results about how the orientation of the rectangles might affect people's preferences.

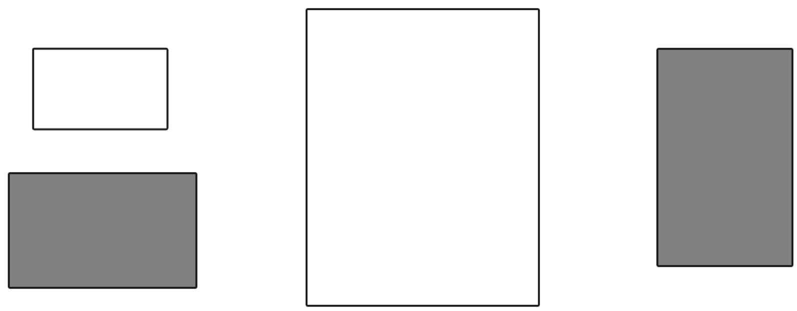
Suggested Answers to Questions

1. To get a true random sample, you would need to obtain a complete listing of all students in the school, and then use random digits or a random number generator to carry out some variation of the hat method. Note that the question asks about obtaining a random sample, not a census of all students at your school.

2. If there are a large number of students in your school, then it may not be practical to collect data from a true random sample of students. If there are a small number of students in your school, then it may be possible to collect data from a random sample.

3. If a true random sample isn't practical, you can still incorporate random selection into your sampling method. For example, if your school has grade-level meetings once per week, you might be able to take separate random samples from each grade to participate in the study. Or, if your school is organized in mixed-grade homerooms (say by alphabetical order of last names), you might be able to take a random sample of homerooms and let every student in the selected homeroom take part in the study. Whatever method you propose should include random selection and offer a reasonable chance of getting a representative cross-section of students from your school.

4. Here are two golden and two nongolden rectangles. The golden rectangles are shaded.



5. Answers will vary. We suggest giving students at least three rectangles from which to choose. If you only provided two rectangles and students really had no preference for one rectangle over the other, they would still pick the golden rectangle half the time, just by chance. This would be equivalent to flipping a coin to choose the preferred rectangle each time. With three rectangles, the chance a student with no particular preference among the rectangles would pick the golden rectangle just by chance is reduced to 1/3. On the other hand, offering the students too many rectangles from which to choose could make it harder to distinguish whether students clearly prefer the golden rectangle.

6. Answers will vary.

7. Answers will vary. Students should summarize individuals' choices in tabular form, showing the number who preferred each of the rectangles, as well as the percent who favored each. Since the variable being measured—preferred rectangle—is categorical, students should present their results graphically in a bar graph.

8. Answers will vary. In evaluating the quality of students' responses, you may want to consider both the accuracy and clarity of communication in:

Tabular presentation of the data

Graphical presentation of the results

Analysis of whether students showed a preference for the golden rectangle

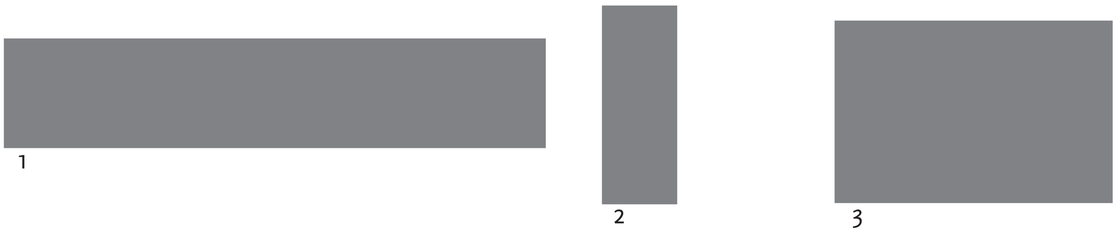
Discussion of the generalizability of results based on their sampling method

Possible Extensions

This observational study can be modified to have students determine whether people are more likely to pick the number "3" when asked to pick one of the numbers 1, 2, 3, or 4.

Investigation #6: It's Golden (and It's Not Silence)

Which of the three rectangles shown here do you find the most pleasing?



Corresponds to pp. 41-43
in Student Module

If you picked the third one, you selected the "golden" rectangle. Because they are generally thought to be the most pleasing, golden rectangles are common in art, architecture, and even in the boxes designed for packaging products that are sold in grocery stores.

A rectangle is "golden" if the ratio of its longest side to its shortest side is approximately 1.618.

In this activity, you will design and carry out an observational study to determine if students at your school do, in fact, find golden rectangles more pleasing than other, less-golden ones.

Since the goal is to be able to generalize the study findings to all students at your school, the first thing to think about is how you will select the students who will participate in your study.

1. Describe a way to select study participants that would result in a random sample of students from your school. Don't worry at this point if your plan cannot be easily implemented—instead, focus on what it would take to get a true random sample of students at your school.

2. Do you think it would be possible to actually implement the plan you described in the previous question? Explain.

3. If it would not be possible to carry out the selection plan described in question 1, describe another sampling method that you think would result in a "representative" sample, but not a truly random sample, from your school. Explain why you think a sample selected in the way you propose here could be considered representative of the students at your school.

Now let's think about how you will collect data from the selected students in a way that will enable you to determine if students really do find golden rectangles more pleasing than nongolden rectangles.

4. In the space below, draw a few rectangles that are golden and several nongolden rectangles.

5. In this study, you will be showing the selected students some rectangles and asking which of the rectangles is most pleasing. How many rectangles will you have the selected students choose between? Why did you select this number?

6. Prepare a separate page containing the rectangles to be shown to your study participants.

After your teacher has approved the data collection plan and your page of rectangles, you can proceed to collect the data for your study.

7. Summarize your data in table form and construct an appropriate graphical display of the data.

8. Write a brief report on separate paper that addresses the question "Do students at your school find golden rectangles to be the most pleasing?" Use tables and graphs to support your conclusions.