

Key

**Observational Study or Experiment?**

For each situation, determine whether the research conducted is an observational study or an experiment. Explain your reasoning.

**Obs** 1. The muscles of men aged 40 - 50 were 40% to 50% stronger after they participated in a 10 week, high-intensity, resistance training program twice a week.

**Obs** 2. Among a group of women aged 65 and older who were tracked for several years, those who had a vitamin B<sub>12</sub> deficiency were twice as likely to suffer severe depression as those who did not.

**EXP** 3. Forty volunteers suffering from insomnia were divided into two groups. The first group was assigned to a special no-desserts diet while the other continued desserts as usual. Half of the people in these groups were randomly assigned to an exercise program, while the others did not exercise. Those who ate no desserts and engaged in exercise showed the most improvement.

**Obs** 4. Some gardens prefer to use non-chemical methods to control insect pests in their gardens. Researchers have designed two kinds of traps and want to know which design will be more effective. They randomly choose 10 locations in a large garden and place one of each kind of trap at each location. After a week, they count the number of bugs in each trap.

**Obs** 5. In 2001, a report in the *Journal of the American Cancer Institute* indicated that women who work nights have a 60% greater risk of developing breast cancer. Researchers based these findings on the work histories of 763 women with breast cancer and 741 women without the disease.

**EXP** 6. Scientists at a major pharmaceutical firm investigated the effectiveness of an herbal compound to treat the common cold. They exposed each subject to a cold virus, and then gave him or her either the herbal compound or a sugar solution known to have no effect. Several days later, they assessed the patient's condition, using a cold severity scale of 0 to 5.

**EXP** 7. To research the effects of dietary patterns on blood pressure in 459 subjects, subjects were randomly assigned to three groups and had their meals prepared by dietitians. Those who were fed a diet low in fat and cholesterol lowered their systolic blood pressure by an average of 6.7 points when compared with subjects fed a control diet.

**EXP** 8. Some people who race greyhounds give the dogs large doses of vitamin C in the belief that the dogs will run faster. Investigators at the University of Florida tried three different diets in random order on each of five racing greyhounds. They were surprised to find that when the dogs ate high amounts of vitamin C, they ran more slowly.

9. An educational software company wants to compare the effectiveness of its computer animation for teaching biology with that of a textbook presentation. The company gives a biology pretest to each of a group of high school juniors, and then divides them into two groups. One group uses the animation, and the other studies the text. The company retests all students and compares the increase in biology test scores in the two groups.  
a. Is this an observational study or an experiment? Justify your answer.  
b. If the group using the computer animation has a much higher average increase in test scores than the group using the textbook, what conclusions, if any, could the company draw?  
10. What is the best way to answer each of the questions below: a survey, an experiment, or an observational study? Explain your choices. For each, write a few sentences about how such a study might be carried out.  
a. Are people generally satisfied with how things are going in the country right now?

Survey

Depends on how chosen - population or volunteers

b. Do college students learn basic accounting better in a classroom or using an online course? **EXP**  
c. How long do your teachers wait on average after they ask the class a question? **Observation**  
11. Can special study courses actually help raise SAT scores? One organization says that the 30 students they tutored achieved an average gain of 60 points when they retook the test.  
a. Explain why this does not necessarily prove that the special course caused the scores to go up. *May have done better for other reasons*  
b. Propose a design for an experiment that could test the effectiveness of the tutorial course.

**PARAMETER OR STATISTIC?**

1. For the studies described, identify the population, sample, population parameters, and sample statistics:

a) The Gallup Organization conducted a poll of 1003 Americans in its household panel to determine what percentage of people plan to cancel their summer vacation because of the increase in gasoline prices. *Pop: Americans Sample: 1003 American panel*

b) Harris Interactive surveyed 2435 U.S. adults nationwide and asked them to rate quality of American public schools. *Pop: all US Adults Sample 2435 polled*

c) The American Institute of Education conducts an annual study of attitudes of incoming college students by surveying approximately 261,000 first-year students at 462 colleges and universities. There are approximately 1.6 million first-year college students in this country. *Pop: 1.6 million 1st yr students Sample: 261,000 students*

d) In a USA Today Internet poll, readers responded voluntarily to the question "Do you consume at least one caffeinated beverage every day?" *Pop: All people Sample: readers who voluntarily resp.*

e) Astronomers typically determine the distance to galaxy by measuring the distances to just a few stars within it and taking the mean (average) of these distance measurements. *Pop: dist. to all stars Sample dist. to few stars*

2. Determine whether the numerical value is a parameter or a statistic (and explain):

a) A survey of 1103 students was taken from the university with 19,500 students. *Stat*

b) The 2006 team payroll of the New York Mets was \$101,084,963. *Par.*

c) In a recent study of physics majors at the university, 15 students were double majoring in math. *Par.*

d) A recent survey by the alumni of a major university indicated that the average salary of 10,000 of its 300,000 graduates was \$125,000. *Stat.*

e) The average salary of all assembly-line employees at a certain car manufacturer is \$33,000. *Par.*

f) The average late fee for 360 credit card holders was found to be \$56.75. *Stat.*

**Population Parameters and Sample Statistics Practice**

- For each statement, identify whether the numbers underlined are statistics or parameters.
  - Of all U.S. kindergarten teachers, 32% say that knowing the alphabet is an essential skill. *parameter*
  - Of the 800 U.S. kindergarten teachers polled, 34% say that knowing the alphabet is an essential skill. *stat.*
- Of the U.S. adult population, 36% has an allergy. A sample of 1200 randomly selected adults resulted in 33.2% reporting an allergy.
  - Describe the population. *US Adults*
  - What is the sample? *1200 random sel. adults*
  - Describe the variable. *having allergy*
  - Identify the statistic and give its value. *33.2%*
  - Identify the parameter and give its value. *36%*
- In your own words, explain why the parameter is fixed and the statistic varies.
 

*par = entire pop so stat. depends on sample*
- The admissions office wants to estimate the cost of textbooks at USC. Let the variable  $x$  be the total cost of all textbooks purchased by a student this semester. The plan is to randomly identify 100 students and obtain their total textbook costs. The average cost for the 100 students will be used to estimate the average cost for all students.
  - Describe the parameter the admissions office wishes to estimate. *avg cost*
  - Describe the population. *all student @ USC*
  - Describe the variable involved. *total cost of textbooks this sem*
  - Describe the sample. *100 students*
  - Describe the statistic and how you would use the 100 data values collected to calculate the statistic. *tot. textbook cost; get average of 100 sample*
- Select 10 students currently enrolled at USC and collect data for these three variables: 1) Number of courses enrolled in, 2) Total cost of textbooks and supplies, 3) Method of payment used for textbooks and supplies.
  - What is the population? *all students enrolled @ USC*
  - What is the sample? *10 students*
- Suppose a 12 year old asked you to explain the difference between a sample and a population, how would you explain it to him/her? How might you explain why you would want to take a sample, rather than surveying every member of the population? *sample = part of population*
- In your own words, explain the difference between a statistic and a parameter. *# about pop. about sample*
- Television station QUE wants to know the proportion of TV owners in Virginia who watch the station's new program at least once a week. The station asked a group of 1000 TV owners in Virginia if they watch the program at least once a week.
  - Identify the individuals of the study and the variable. *TV owners in VA - watching new*
  - Do the data comprise a sample? *yes - assuming more taken 1000 down the line*
  - If so, what is the underlying population? *all TV owners in VA*
  - Is the proportion of viewers in the sample who watch the new program at least once a week a statistic or a parameter? *statistic*
- A study reveals that there are exactly 100 Senators in the 109th Congress of the United States, and 55% of them are Republicans.
  - Identify the individuals of the study and the variable. *Senators of 109th Congress; Party*
  - Do the data comprise a sample or a population? *Pop. b/c all Senators*
  - Does the study represent a statistic or a parameter? *Parameter*
  - For a set population, does a parameter ever change? *NO*

- Identify the population and the sample. *Pop - all American households; sample = 1353*
- A survey of 1353 American households found that 18% of the households own a computer.
  - A recent survey of 2625 elementary school children found that 28% of the children could be classified obese. *Pop - all Elem school children; sample = 2625 elem*
  - The average weight of every sixth person entering the mall within 3 hour period was 146 lb. *Pop - all adults; sample - every 6th person in 3hr period*
- Determine whether the numerical value is a parameter or a statistic (and explain):
  - A recent survey by the alumni of a major university indicated that the average salary of 10,000 of its 300,000 graduates was 125,000. *statistic - from sample*
  - The average salary of all assembly-line employees at a certain car manufacturer is \$33,000. *par - entire pop.*
  - The average late fee for 360 credit card holders was found to be \$56.75. *stat*
- Consider the students in our class as the sample and all of the students at our high school as the population. Identify the following as a parameter or a statistic.
  - The proportion of students in our class who use instant messaging or text messaging on a daily basis. *stat.*
  - The proportion of students at our school who use instant messaging or text messaging on a daily basis. *par.*
  - The average number of hours that students at our school spent watching television last week. *par.*
  - The average number of hours that students in our class slept last night. *stat.*
- Identify each of the following as a parameter or a statistic. If you need to make an assumption about who or what the population is, explain your assumption.
  - The proportion of voters who voted for President Bush in the 2004 election. *par, stat.*
  - The proportion of voters surveyed by CNN who voted for John Kerry in the 2004 election. *par.*
  - The proportion of voters among our school's faculty who voted for Ralph Nader in the 2004 election. *par.*
  - The average number of points scored in a Super Bowl game. *par.*

**Sampling Techniques Worksheet**

- Multiple Choice**
- Identify the choice that best completes the statement or answers the question.
- What is the best sampling technique to use for determining the average speed of the cars on a section of highway?
    - simple random sample
    - systematic sample
    - convenience sample
    - a) or b)
  - Which method is most likely to produce a random sample of the members of your class?
    - listing the first six students that come to mind
    - choosing the five oldest students in the class
    - writing the name of each student on a separate piece of paper and then drawing these slips from a hat
    - selecting the first six students to arrive at class
  - A large corporation wants to find out which benefits plan its employees would prefer. Which procedure would be most likely to obtain a statistically unbiased sample?
    - surveying a random sample of employees from a list of all employees.
    - inviting all employees to indicate their choices by e-mail
    - placing suggestion boxes at random locations in the company's plant and offices
    - assembling a group with one member from each department and recording the preferences of these employees

B 4. A college president wants to find out which courses students consider to be of the most benefit to them. Which procedure would be most likely to produce a statistically unbiased sample?

- a. asking students to mail in a questionnaire
- b. surveying a random sample of students taken from the list of all students
- c. surveying the first hundred students on an alphabetical list
- d. having students complete a questionnaire on the college web site

B 5. A pollster wants to find out if citizens are satisfied with the city council. Which procedure would be most appropriate for obtaining a statistically unbiased sample?

- a. interviewing people at a popular local shopping mall
- b. surveying people whose names have been randomly chosen from the telephone book
- c. placing an advertisement in the local newspaper asking for mail-in responses
- d. mailing a questionnaire to people whose names have been chosen randomly from a list of customers of the municipal utility company

**Complete each statement.**

6. The group of individuals who actually are selected for a survey is called the sample.

7. The set of all individuals who belong to the group being studied by a survey is called the population.

nonvenience samp

8. A television reporter interviewed travelers stranded at an airport during a snowstorm about the efficiency of air travel in Canada. Name the sampling technique used.

Voluntary response

9. A soap company distributed free samples of a new laundry detergent to all households in several randomly selected neighborhoods. The company requested the recipients to return a postage-paid card indicating whether they thought the sample was better than their usual detergent. What sampling technique was the company using?

Systematic sample

10. A psychologist is studying the sleep patterns of the 3960 students at her university. She decides to start by asking a random sample of 30 students how many hours of sleep they get weekday nights. Identify the type of sample in each of the following survey methods.

Simple random sample

a) The psychologist assigns each student a number from 1 to 3960. She selects the sample by randomly choosing one of the first 132 numbers and every 132nd number thereafter.

Cluster sample

Stratified sample

b) The psychologist assigns each student a number from 0001 to 3960 and uses a computer to randomly generate a list of 30 numbers to select the students for the sample.

c) Students are listed by the neighborhood they live in. The psychologist randomly selects six neighborhoods and then randomly selects five students from each one.

d) An equal proportion of students are randomly selected from each discipline.

11. A particular school has 550 female students and 590 male students. A random sample of 30 students was surveyed for suggestions about social activities for the following school year.

- a) Is it possible that the sample included only male students? yes b/c random
- b) Would a sample consisting entirely of male students be representative of the school population? Explain your reasoning. No

**Margin of Error**

A survey of a sample population gathers information from a few people and then the results are used to reflect the opinions of a larger population. The reason that researchers and pollsters use sample population is that it is cheaper and easier to poll a few people rather than everybody. One key to successful surveys of sample populations is finding the appropriate size for the sample that will give accurate results without spending too much time or money. Determining a margin of error depends on whether you're working with a proportion or a mean.

**Proportions:** Suppose that 900 American teens were surveyed about their favorite ski category of the 2002 Winter Olympics in Park City, Utah. Ski jumping was the favorite for 20% of those surveyed. What is the true population proportion of teens who enjoy ski jumping?

First, find the standard deviation:  $s = \sqrt{\frac{p(1-p)}{n}}$

For our example,  $p = 0.2$ , so  $1 - p = 0.8$ .  $s = \sqrt{\frac{0.2(0.8)}{900}} = 0.013$

Since about 95% of the data will fall within almost two standard deviations, we will use the formula

$$1.96 \sqrt{\frac{p(1-p)}{n}}$$

The margin of error will be 1.96(0.013), or 0.02548. Let's round that to 0.025.

Because our survey did not ask every single teenager in America, we are basically making a guess here, so our margin of error provides a cushion around our guess. We believe that the true proportion will fall inside the interval created when we add and subtract the margin of error from our sample proportion:

$$0.20 \pm 0.025$$

Our interval is 0.175 to 0.225. We believe the true proportion will lie inside that interval.

**Find the margin of error for each of the following and create an interval for the true population proportion.**

1. A sample of 550 people leaving a shopping mall showed that 64% of shoppers claim to have spent over \$25.

$MOE = .04$      $CI = [.6 \pm 0.068]$      $[.6, .68]$

2. In a random sample of machine parts, 18 out of 225 were found to have been damaged in shipment.

$MOE = .035$      $CI = [.045, .115]$

3. A telephone survey of 1000 adults was taken shortly after the U.S. began bombing Iraq found that 832 adults voiced their support for this action.

$$MOE = .023 \quad CI [ .809, .855 ]$$

4. An assembly line does a quality check by sampling 50 of its products. It finds that 16% of the parts are defective.  $MOE = .102$  [ .058, .262 ]

Now let's look at how the sample size will affect the margin of error. Use the ski jumping example for the following:

1. Find the margin of error for a survey of 90 American teens.

$$.083$$

2. Find the margin of error for a survey of 9,000 teens.

$$.0083$$

3. Find the margin of error for a survey of 90,000 teens.

$$.0013$$

4. Draw a conclusion about the margin of error based on the size of the sample. Why do you think this is so?

As sample size ↑, MOE ↓

Mult by 4 ⇒ n is under √

5. If you want to cut your margin of error in half, what would you have to do to the sample size? Why?

Margins of error can also be used to estimate population means. Let's see how this will work. A company that produces white bread is concerned about the distribution of the amount of sodium in its bread. The company takes a simple random sample of 100 slices of bread and computes the sample mean to be 103 milligrams of sodium per slice and the sample standard deviation is 12 milligrams.

The margin of error formula for means is  $1.96 \frac{s}{\sqrt{n}}$ , so in this situation, our margin of error will be:

$$1.96 \frac{12}{\sqrt{100}} \text{ or } 2.352$$

That means that we will expect our true population mean to fall between  $103 \pm 2.352$ , or  $100.648 - 105.352$ .

What if the bread label states that the sodium content of the bread is 100 milligrams? Should the company be concerned? Why or why not?

Find the margin of error for the following and an interval that could contain the true mean:

1. You want to rent an unfurnished one-bedroom apartment for next semester. The mean monthly rent for a random sample of 10 apartments advertised in the local newspaper is \$540 with a standard deviation of \$80.

$MAE = 49.58$   $CI [ 490.42, 589.58 ]$

2. Your company sells exercise clothing and equipment on the Internet. To design the clothing, you collect data on the physical characteristics of your different types of customers. Here are the weights (in kilograms) for a sample of 24 male runners: 67.8 61.9 63.0 53.1 62.3 59.7 55.4 58.9 60.9 69.2 63.7 68.3 64.7 65.6 56.0 57.8 66.0 62.9 53.6 65.0 55.8 60.4 69.3 61.7

$$MOE = 1.92 \quad CI [ 59.87, 63.71 ]$$

3. A hardware manufacturer produces bolts used to assemble various machines. Suppose the average diameter of a simple random sample of 50 bolts is 5.11 mm and the standard deviation is 0.1 mm.

$$MOE = .0226 \quad CI = [ 5.082, 5.138 ]$$

4. We have IQ test scores of 31 seventh-grade girls in a Midwest school district. We have calculated that sample mean is 105.84 and the standard deviation is 14.27.

$$MOE = 5.023 \quad CI [ 100.817, 110.863 ]$$

5. Let's look at problem #4 again. How would the margin of error change if there were 90 girls instead of the 31?

$$MOE = 2.948$$

6. What if there were 250 girls?

$$MOE = 1.769$$

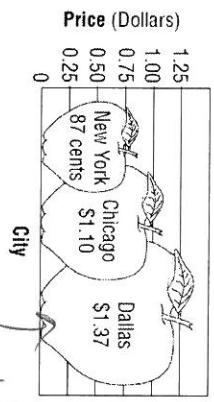
7. How does the sample size change the margin of error?

Sample ↑ MOE ↓

Analyzing Misleading Graphs

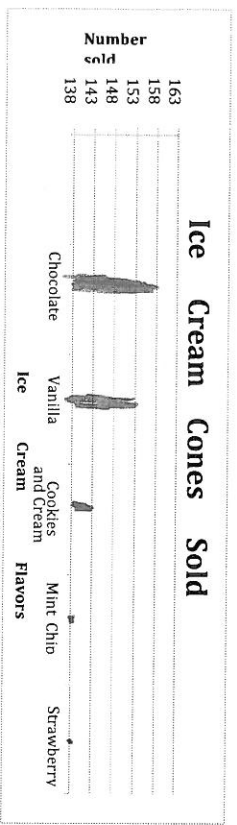
Look at each graph or set of data carefully with your partner. Discuss how each graph is misleading and what you could do to correct it.

1. The Price of a Pound of Apples in Selected Cities in September 2000



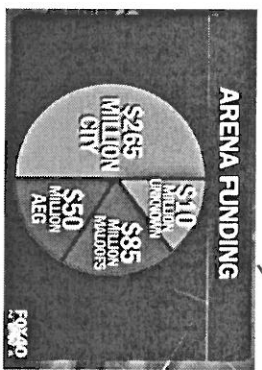
Should be higher on scale

2.



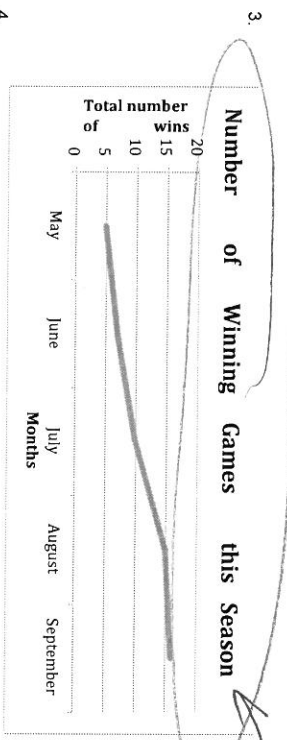
Doesn't start @ 0, so look like bigger diff. in values

5.



wedges not to scale  
\$10 mil. same as \$50 mil.

6.



Doesn't match data

4.

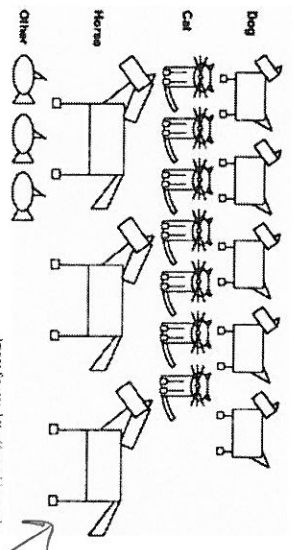


Image Source: <http://www.bbc.co.uk>

No title  
horse 50  
big looks like most from just  
even just

5.

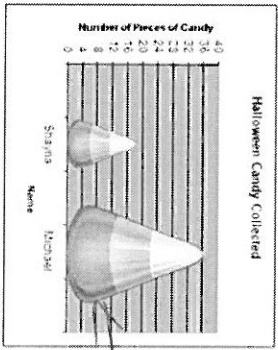


Image Source: <http://www.yale.edu>

Picture makes look like a lot more than Shayna really only 2x bigger.