

Designing Samples:

Suppose we want to gather information about a group of people.

- If the group is small (for example, all students in this class) we can study each group member directly.
- If, however, the group is very large (for example, all students in the school), studying each member of the group may not be feasible.

As an alternative, we can select a smaller group of people who fairly represent the entire group.

The entire group of individuals (not necessarily people) that we want information about is called the population.

The part of the population in the study is called the sample.

The method we use to select the sample is called the sample design. The design of the sample is very important. If the design is poor, the sample will not accurately represent the population.

Types of Sample Designs:

Voluntary Response Sample

- People choose themselves to be in the sample by responding to a general appeal
- *Example:* We post an advertisement in *Big Red* asking GBHS students to respond
- *Problem:* People with strong opinions (often strong negative opinions) tend to reply, so they are overrepresented

Convenience Sample

- Individuals who are easiest to reach are chosen for the sample
- *Example:* We use students in this class as our sample
- *Problem:* This group may not be diverse enough to accurately represent all students at GBHS

Both Voluntary Response Samples and Convenience Samples result in a sample that is not representative of the population. These are biased samples because they favor certain outcomes.

Random selection eliminates bias from sample choice.

Simple Random Sample (SRS)

- Individuals are selected so that all possible combinations of individuals are equally likely to be in the sample
- *Example:* Generate a list of student ID numbers for all students at GBHS; then randomly select student ID numbers and choose those students for the sample

Systematic Random Sample

- The first individual is chosen at random; then a system or rule is used to choose all other individuals
- *Example:* Obtain an alphabetized list of all students at GBHS. Choose every 5th person on the list.

Stratified Random Sample

- Divide the population into groups of similar individuals; choose a SRS in each group to form the full sample
- *Example:* Divide all of the students at GBHS into four groups: freshmen, sophomores, juniors, and seniors; then choose a SRS from each grade level

Multistage Sample

- Select several groups; within each group, select a subgroup; within each subgroup select individuals for the sample.
- *Example:* Select several departments within the school (Math, English, Art). Within each of those departments, select several teachers. Choose several students within each class.

Cluster Sample

- Select several groups; within each group, select several subgroups; within each subgroup select ALL individuals for the sample.
- *Example:* Select several departments within the school (Math, English, Art). Within each of those departments, select several teachers. Choose ALL students in each class.

Although random selection eliminates bias from our choice of sample, it does not guarantee that our sample is representative of the population.

Potential problems include:

Undercoverage:

- some groups are left out of the process of choosing the sample
- *Example:* Students in vo-tech, early release, on suspension, or absent may be left out of the sample

Nonresponse:

- An individual chosen for the sample cannot be contacted or refuses to cooperate
- *Example:* A student chosen for the sample may refuse to divulge information or may be absent

Response Bias

- The behavior of the individual or interviewer may influence the accuracy of the response
- *Example:* Students may lie about drug or alcohol use

Wording of Questions

- Confusing or leading questions influence responses; poorly worded questions will not yield accurate responses.
- *Example 1:* "In a recent study, students in an Algebra I course were given a 25 question basic skills test. On average, students used a graphing calculator to answer 21 out of 25 questions. Do you think graphing calculators are overused?"

- *Example 2:* “By using a graphing calculator, students in an Algebra I course are able to make visual connection between equations and their graphs, reinforcing difficult concepts. Do you think graphing calculators are overused?”
- *Example 3:* “Do you like English or Math?”
- *Example 4:* “Do you like school?”

QUIZ: Question Bias

Name: _____

Date: _____ Pd: _____

Directions: The following questions will appear on a survey given to all AP Statistics students for the purpose of learning more about the students in the class. Choose 15 (if you choose more than 15, only the first 15 will be graded) of the following questions. Explain what is wrong with the way the question is worded. Then rewrite the question to reduce bias.

1. How many hours do you watch TV a night?
2. How many times have you skipped this class?
3. Are you involved in music?
4. What is your ideal GPA?
5. How long do you take doing homework?
6. Do you procrastinate?
7. In your opinion, what is the funniest YouTube video of all time?
8. Do you know what college you're going to?
9. What color are your eyes?

10. Have you ever snuck a boy/girl into your house?

11. Why did you choose to take Stats instead of Calc?

12. What is the lowest grade you have earned on a test so far?

13. Have you been bitten by a vampire?

14. Do you play any sports?

15. How many AP classes are you in?

16. How big is your head?

17. Were you satisfied with the presidential election?

18. What is your favorite color?

19. How many hours do you spend playing video games?

20. How many text messages would you say you send during class?

21. What happened to Quint's chin?

22. Do you play an instrument?

23. How helpful are the notes in this class?

24. Do the self-reflections help you with your weaknesses?

25. Should Mrs. Krummel have short or long hair?

Random Rectangles Activity

1. Guess the average area of the rectangles on the page: _____

2. Choose five representative rectangles, record their areas, and calculate the average area:

	Number	Area
1	_____	_____
2	_____	_____
3	_____	_____
4	_____	_____
5	_____	_____

TOTAL: _____ AVERAGE = _____

3. Use the random number table to choose five rectangles at random, record their areas, and compute their average:

	Number	Area
1	_____	_____
2	_____	_____
3	_____	_____
4	_____	_____
5	_____	_____

TOTAL: _____ AVERAGE = _____

