



Watch on  YouTube

### Definitions:

Population: the whole group

Sample: part of the group

Parameter: prediction about the *population*.

Statistic: numerical summary of a *sample*.

Categorical (Qualitative) Data: category or classification.

Quantitative Data: number that you can add or subtract and it makes sense.

*Hint*: Categorical is usually words and Quantitative is usually numbers with a few exceptions such as: zip codes cannot be added or subtracted. They are a classification of where you live.

### Level of Measurement:

1. Categorical options:

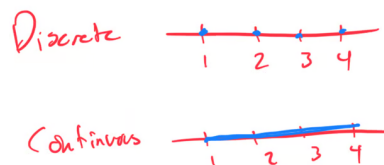
1. Nominal: has no order.
2. Ordinal: can be put in order.

2. Quantitative options:

1. Interval: the value 0 does exist in context. *Example*: 0 degrees is a real temperature.
2. Ratio: the value 0 does not exist in context. *Example*: baby weight of 0 lbs means there is no baby.

Discrete: you can *count* and finite

Continuous: you can *measure* and infinite



### Types of Sampling:

- Simple Random Sample: equal chance of being selected
- Stratified Random Sample: split into groups and randomly select from each group
- Cluster Random Sample: split into groups and randomly pick the entire group(s)
- Systematic Random Sample: example: pick every 3rd person
- Convenience Sampling: easily accessible, not reliable.

Note: can sample with replacement or without replacement

## ALEKS Topics:

### Topic 1: Differentiating between parameters and statistics

Problem 1: In a study of a town's population, the average age of all residents is 40 years, but a sample of 100 residents has an average age of 42 years. Identify which value is a parameter and which is a statistic, and explain why.

*Answer: The average age of all residents (40 years) is a parameter because it describes the entire population. The average age of the sample (42 years) is a statistic because it describes a subset of the population. A parameter is a fixed value for the whole population, while a statistic is calculated from a sample.*

Problem 2: A company claims the mean salary of all employees is \$50,000, while a survey of 200 employees yields a mean salary of \$48,500. Determine which is a parameter and which is a statistic, and justify your answer.

*Answer: The mean salary of all employees (\$50,000) is a parameter because it represents the entire employee population. The mean salary of the 200 employees (\$48,500) is a statistic because it is derived from a sample. Parameters apply to populations, while statistics apply to samples.*

### Topic 2: Classifying samples

(a) A botanist at a nursery wants to inspect the health of the plants at the nursery. Which of the following best describes a <b>stratified sample</b> of plants?
<input checked="" type="radio"/> The botanist takes a list of the plants and selects every 4 <sup>th</sup> plant until 64 plants are selected.
<input type="radio"/> All of the plants that are in one particular room at the nursery are easily accessible. So, the botanist selects the 64 plants in this particular room.
<input type="radio"/> The botanist forms 4 groups of plants based on the ages of the plants (in months). Then, he selects 16 plants at random from each group.
(b) A chemist at a pharmaceutical company wants to test the quality of a new batch of microscopes. Which of the following best describes a <b>convenience sample</b> of microscopes?
<input checked="" type="radio"/> The microscopes in the first shipment that was received are <b>easily accessible</b> . So, he selects all 54 of the microscopes in this shipment.
<input type="radio"/> The chemist forms groups of 9 microscopes based on the laboratories that the microscopes are in. Then, he randomly chooses 6 groups and selects all of the microscopes in these groups.
<input type="radio"/> The chemist assigns each microscope a different number. Using a random number table, he draws 54 of those numbers at random. Then, he selects the microscopes assigned to the drawn numbers. Every set of 54 microscopes is equally likely to be drawn using the random number table.
(c) A facilities supervisor at a sports stadium wants to rate the condition of the seats at the stadium. Which of the following best describes a <b>cluster sample</b> of seats?
<input type="radio"/> The supervisor forms 5 groups of seats based on the dates the seats were last replaced. Then, she selects 18 seats at random from each group.
<input checked="" type="radio"/> The supervisor forms groups of 30 seats based on the sections the seats are in. Then, she selects <b>all of the seats in 3 randomly chosen groups</b> .
<input type="radio"/> The supervisor takes a list of the seats and selects every 5 <sup>th</sup> seat until 90 seats are selected.

Problem 1: A researcher surveys 50 students from each grade level in a school to study study habits. Classify this sample as random, stratified, systematic, or convenience, and explain your reasoning.

*Answer: This is a stratified sample. The researcher divides the population into strata (grade levels) and selects a sample from each stratum (50 students per grade). This ensures representation from each group, unlike a random sample (equal chance for all) or convenience sample (easiest to access).*

Problem 2: A pollster interviews every 10th person entering a mall to gather opinions on a product. Classify this sample as random, stratified, systematic, or convenience, and provide justification.

*Answer: This is a systematic sample. The pollster selects every 10th person, following a fixed interval, which distinguishes it from a random sample (no pattern) or convenience sample (non-systematic selection). It is not stratified since no subgroups are defined.*

### Topic 3: Classification of variables

Problem 1: Classify the variable "number of cars owned by a household" as qualitative or quantitative, and explain why.

*Answer: The variable is quantitative because it represents a numerical count (e.g., 0, 1, 2 cars) that can be measured or compared mathematically, unlike qualitative variables, which describe categories or qualities (e.g., color).*

Problem 2: Determine if the variable "favorite music genre" is qualitative or quantitative, and justify your classification.

*Answer: The variable is qualitative because it describes a category or preference (e.g., rock, jazz) that cannot be quantified numerically, unlike quantitative variables, which involve measurable numerical values.*

#### Topic 4: Discrete versus continuous variables

Problem 1: Is the variable "number of students in a classroom" discrete or continuous? Explain why it fits the chosen category.

*Answer: The variable is discrete because it takes on specific, countable values (e.g., 20, 25 students) with no intermediate values possible. Continuous variables, like height, can take any value within a range.*

Problem 2: Classify the variable "height of a tree" as discrete or continuous, and provide reasoning for your answer.

*Answer: The variable is continuous because tree height can take any value within a range (e.g., 5.2 m, 5.25 m), including fractions, unlike discrete variables, which are limited to distinct, countable values.*

#### Topic 5: Choosing units of measurement and an appropriate method to gather data

Problem 1: To measure the length of fish in a lake, choose an appropriate unit (e.g., centimeters or meters) and suggest a method (e.g., observation, survey, or experiment) to collect the data. Explain your choices.

*Answer: Unit: Centimeters, as fish lengths are typically small and centimeters provide precise measurements. Method: Observation, by catching and measuring fish directly. This is appropriate because it provides accurate data for length, unlike surveys (subjective) or experiments (manipulating variables).*

Problem 2: For studying the weight of apples in an orchard, select a suitable unit (e.g., grams or kilograms) and a data collection method. Justify your selections.

*Answer: Unit: Grams, as apples have relatively small weights, and grams allow for precise measurements. Method: Observation, by weighing apples directly with a scale. This ensures accurate weight data, unlike surveys (less reliable) or experiments (unnecessary manipulation).*

#### Topic 6: Choosing an appropriate method to conduct a survey and making an estimation

Problem 1: To estimate the percentage of city residents who support a new park, choose a survey method (e.g., online, phone, or in-person) and estimate the sample size needed for accuracy. Explain your choices.

*Answer: Method: Online survey, as it reaches a broad audience efficiently and is cost-effective. Sample size: Approximately 400, based on standard statistical practice for a 95% confidence level with a 5% margin of error for a moderate population size. This ensures a representative sample without excessive cost.*

Problem 2: For estimating the average time students spend on homework, select a survey method and propose a reasonable sample size. Provide reasoning for your decisions.

*Answer: Method: In-person survey, as it allows direct interaction with students for accurate responses. Sample size: About 200 students, sufficient for a 95% confidence level with a reasonable margin of error for a school-sized population. This balances accuracy and practicality.*

### Topic 7: Identifying and reducing statistical bias

Problem 1: A survey on exercise habits is conducted only at a gym. Identify the type of bias (e.g., selection bias) and suggest a way to reduce it, such as expanding the sample to non-gym-goers.

*Answer: Bias: Selection bias, as the sample only includes gym-goers, who are likely more active, skewing results. Reduction: Expand the survey to include non-gym-goers, such as through online or community surveys, to represent a broader population with varied exercise habits.*

Problem 2: A poll on a new policy is conducted only during weekday mornings at a business district. Identify the bias and propose a method to reduce it, such as including evening or weekend surveys.

*Answer: Bias: Selection bias, as the poll only captures business district workers during weekday mornings, excluding others like evening workers or residents. Reduction: Conduct the poll at varied times (evenings, weekends) and locations (e.g., residential areas) to include a more diverse sample.*

### Activity:

Do you like football?

SRS  
 $\frac{3}{5} = 60\%$

cluster  
 $\frac{2}{5} = 40\%$

Stratified  
1-12  $\frac{4}{5} = 80\%$   
13-23  $\frac{3}{5} = 60\%$  } 70%

Avg: 62.5%

Systematic  
 $\frac{4}{5} = 80\%$

Census - ask entire population  
57%



[Tags](#) [Archive](#) [RSS feed](#) [Youtube](#) [QR Code](#) email [akennon@fscj.edu](mailto:akennon@fscj.edu) with any issues on this website

Made with [Montaigne](#) and [bigmission](#) 