POPULATION

The **population** of a study is the group the collected data is intended to describe.

POPULATION

The **population** of a study is the group the collected data is intended to describe.

Sometimes the intended population is called the **target population**, since in badly designed studies, the collected data may not be representative of the intended population.

EXAMPLE

A university sends out a survey to all students in the cafeteria asking about their satisfaction with campus facilities.

What is the population?

EXAMPLE

A university sends out a survey to all students in the cafeteria asking about their satisfaction with campus facilities.

What is the population?

While the intended population may have been all students, the real population of the survey is students who go to the cafeteria.

EXAMPLE

A local fitness center posts a survey on their website asking visitors how often they exercise each week.

What is the population?

EXAMPLE

A local fitness center posts a survey on their website asking visitors how often they exercise each week.

What is the population?

While the intended population may have been all members of the fitness center, the real population of the survey are people who on the website of the fitness center.

PARAMETER

A **parameter** is a value (average, percentage, etc.) calculated using all the data from a population

EXAMPLES

If you calculate the average income of every person in a city, that average is a parameter.

If you determine the percentage of all voters in a country who prefer a particular candidate, that percentage is a parameter.

CENSUS

A survey of an entire population is called a **census**.

EXAMPLE

National Census

Every ten years, the United States conducts a national census, where the government attempts to count every person living in the country. The data collected includes information such as age, gender, race, household size, and housing conditions. This information is used for things like determining the number of seats each state has in the U.S. House of Representatives.

EXAMPLE

School Enrollment Census

A school district may conduct a census at the beginning of each academic year to account for every student enrolled in its schools. The census gathers data on student demographics, such as grade level, age, and special education needs, to help allocate resources and plan educational programs.

Since surveying an entire population is often impractical, we usually select a sample to study.

SAMPLE

A **sample** is a smaller subset of the entire population, ideally one that is fairly representative of the whole population.

EXAMPLE

If you're interested in the average height of adults in a city, you might measure the height of 1,000 adults randomly selected from the city's population. This group of 1,000 people is your sample, and the average height calculated from this group is meant to represent the average height of all adults in the city.

EXAMPLE

In a political poll, a sample of voters (say, 1,500 people) is surveyed to estimate how the entire population of voters will vote in an upcoming election.

STATISTIC

A **statistic** is a value (average, percentage, etc.) calculated using the data from a sample.

EXAMPLE

If you conduct a survey with 500 people in a city and determine that 60% of them prefer a certain brand of coffee, that 60% is a statistic.

EXAMPLE

Medical Study

Researchers conduct a study with a sample of 500 patients to determine the average reduction in blood pressure after taking a new medication. They find that the average reduction in blood pressure among the sample is 10 mmHg. This 10 mmHg is a **statistic** because it is derived from the sample of patients.

EXAMPLE

Public Health Survey

A health department surveys 800 residents to estimate the prevalence of smoking in a community. They find that 10% of the respondents smoke regularly. This 10% smoking rate is a statistic because it is based on the data from the sample of residents.

QUESTION

A public health researcher wants to assess how residents of New Jersey feel about a new community initiative. To collect opinions, she goes to the New Jersey Farmers Market and randomly surveys 300 visitors. She finds that 70% of those surveyed support the initiative.

What is the sample and population?

Is the 70% value a parameter or a statistic?

QUESTION

A public health researcher wants to assess how residents of New Jersey feel about a new community initiative. To collect opinions, she goes to the New Jersey Farmers Market and randomly surveys 300 visitors. She finds that 70% of those surveyed support the initiative.

What is the sample and population?
Is the 70% value a parameter or a statistic?
The sample is the 300 visitors at the New Jersey Farmers Market who were surveyed.

QUESTION

An economist wants to find out how employees in San Francisco feel about a proposed change in city minimum wage laws. She visits the San Francisco Ferry Building and randomly interviews 350 workers there. The results show that 55% of those interviewed support the wage increase.

What is the sample and population?

Is the 55% value a parameter or a statistic?

QUESTION

An economist wants to find out how employees in San Francisco feel about a proposed change in city minimum wage laws. She visits the San Francisco Ferry Building and randomly interviews 350 workers there. The results show that 55% of those interviewed support the wage increase.

What is the sample and population?
Is the 55% value a parameter or a statistic?

The sample is the 350 workers interviewed at the San Francisco Ferry Building.

The population is all employees in San Francisco.

The 55% value is a statistic because it is calculated from the sample of workers, not from the entire population of San Francisco employees.

QUESTION

To estimate the average weight of apples in an orchard, a researcher picks 50 apples from different trees throughout the orchard and weighs them.

What is the sample and population in this study?

QUESTION

To estimate the average weight of apples in an orchard, a researcher picks 50 apples from different trees throughout the orchard and weighs them.

What is the sample and population in this study?

The sample is the 50 apples that were picked and weighed.

The population is all the apples in the orchard.

QUESTION

A hotel manager wants to estimate the average length of stay for guests at the hotel. She reviews the records of 500 randomly selected guests from the past month and calculates their average length of stay.

What is the sample and population in this study?

QUESTION

A hotel manager wants to estimate the average length of stay for guests at the hotel. She reviews the records of 500 randomly selected guests from the past month and calculates their average length of stay.

What is the sample and population in this study?

The sample is the 500 guests whose length of stay was reviewed.
The population is all guests who stayed at the hotel during the past month.

QUESTION

A company announces that the average annual salary of all its employees is $65,000.

Is this a statistic or a parameter?

QUESTION

A company announces that the average annual salary of all its employees is $65,000.

Is this a statistic or a parameter?

This average annual salary of $65,000 is a parameter because it describes the average salary of the entire population of employees within the company.

QUESTION

A survey conducted by a market research firm interviews 200 randomly selected employees from various companies in a city. The survey finds that the average commute time for these employees is 30 minutes.

Is this a statistic or a parameter?

QUESTION

A survey conducted by a market research firm interviews 200 randomly selected employees from various companies in a city. The survey finds that the average commute time for these employees is 30 minutes.

Is this a statistic or a parameter?

This average commute time of 30 minutes is a statistic because it is calculated from the sample of 200 employees.

CENSUS

Parameter

SURVEY

Statistic

CATEGORICAL DATA

Categorical (qualitative) data are pieces of information that allow us to classify the objects under investigation into various categories

QUANTITATIVE DATA

Quantitative data are responses that are numerical in nature and with which we can perform meaningful arithmetic calculations.

EXAMPLE

A researcher is conducting a study on the favorite colors of people in a community. She collects responses from 100 people, and the colors reported are red, blue, green, and yellow.

The data is **qualitative** because it describes categories of colors rather than numeric measurements.

EXAMPLE

An online bookstore tracks the number of books sold each month. In January, they sold 150 books, in February 200 books, and in March 175 books.

The data is **quantitative** because it consists of numerical values representing the number of books sold.

EXAMPLE

A company surveys its employees about their job satisfaction. The responses include categories such as "Very Satisfied," "Satisfied," "Neutral," "Dissatisfied," and "Very Dissatisfied."

The data is **qualitative** because it describes different levels of satisfaction using categorical responses.

EXAMPLE

A hospital records the number of patients admitted each day. On Monday, there were 30 patients, on Tuesday, 45 patients, and on Wednesday, 40 patients.

The data is **quantitative** because it involves numerical values representing the number of patients admitted.

QUESTION

A tourist attraction logs the different countries from which visitors come, such as France, Japan, Brazil, and Canada.

Is the data qualitative or quantitative?

QUESTION

A tourist attraction logs the different countries from which visitors come, such as France, Japan, Brazil, and Canada.

Is the data qualitative or quantitative?

The data is **qualitative** because it categorises visitors based on their countries of origin, without using numerical values.

QUESTION

A company records the number of new hires each month. The counts for the past three months are 8, 12, and 10 new hires.

Is the data qualitative or quantitative?

QUESTION

A company records the number of new hires each month. The counts for the past three months are 8, 12, and 10 new hires.

Is the data qualitative or quantitative?

The data is quantitative because it consists of numerical values representing the number of new hires each month.

QUESTION

A research study measures the daily water intake of participants in liters. The recorded intakes are 1.5 liters, 2.0 liters, and 2.5 liters.

Is the data qualitative or quantitative?

QUESTION

A research study measures the daily water intake of participants in liters. The recorded intakes are 1.5 liters, 2.0 liters, and 2.5 liters.

Is the data qualitative or quantitative?

The data is quantitative because it consists of numerical values representing the amount of water consumed.

QUESTION

A librarian categorizes books in the library by genre, such as fiction, non-fiction, mystery, and science fiction.

Is the data qualitative or quantitative?

QUESTION

A librarian categorizes books in the library by genre, such as fiction, non-fiction, mystery, and science fiction.

Is the data qualitative or quantitative?

The data is qualitative because it classifies books into different genres without using numerical values.

QUESTION

A company assigns employee ID numbers to its staff for identification purposes. One employee has the ID number 12345.

Is the data qualitative or quantitative?

QUESTION

A company assigns employee ID numbers to its staff for identification purposes. One employee has the ID number 12345.
Is the data qualitative or quantitative?

The data is qualitative because the employee ID number is used as a categorical identifier for individuals rather than for numerical analysis or measurement. Although it consists of numbers, it serves to distinguish between employees rather than to represent a measurable quantity.

QUESTION

A city uses zip codes to organize mail delivery. The zip code 30301 is assigned to a particular neighborhood.

Is the data qualitative or quantitative?

QUESTION

A city uses zip codes to organize mail delivery. The zip code 30301 is assigned to a particular neighborhood.

Is the data qualitative or quantitative?

The data is qualitative because the zip code 30301 is used as a categorical identifier for a specific geographic area. Although it consists of numbers, it functions as a label to categorise locations rather than a numerical value for calculations or measurements.

SAMPLING BIAS

A sampling method is **biased** if every member of the population doesn't have equal likelihood of being in the sample.

There are many ways to sample a population, but there is one goal we need to keep in mind: we would like the sample to be **representative of the population**.

There are many ways to sample a population, but there is one goal we need to keep in mind: we would like the sample to be **representative of the population**.

One way to ensure that the sample has a reasonable chance of mirroring the population is to employ **randomness**.

The most basic random method is simple **random sampling**.

RANDOM SAMPLE

A **random sample** is one in which each member of the population has an equal probability of being chosen.
A simple random sample is one in which every member of the population and any group of members has an
equal probability of being chosen.

EXAMPLES OF RANDOM SAMPLING

We could write peoples' names on separate slips of paper. All slips are placed in a box and names are drawn from the box.

We can use assign a number to each person and use a random number generator to select people.

Even a random sample might end up not being totally representative of the population.

If we repeatedly take samples of 1000 people from a population, some of these samples might tend to have a slightly higher percentage of older people and some samples might include more younger people; some samples may have a larger percentage of women than the general population. In most cases, this **sampling variability** is not significant.

SAMPLING VARIABIALITY

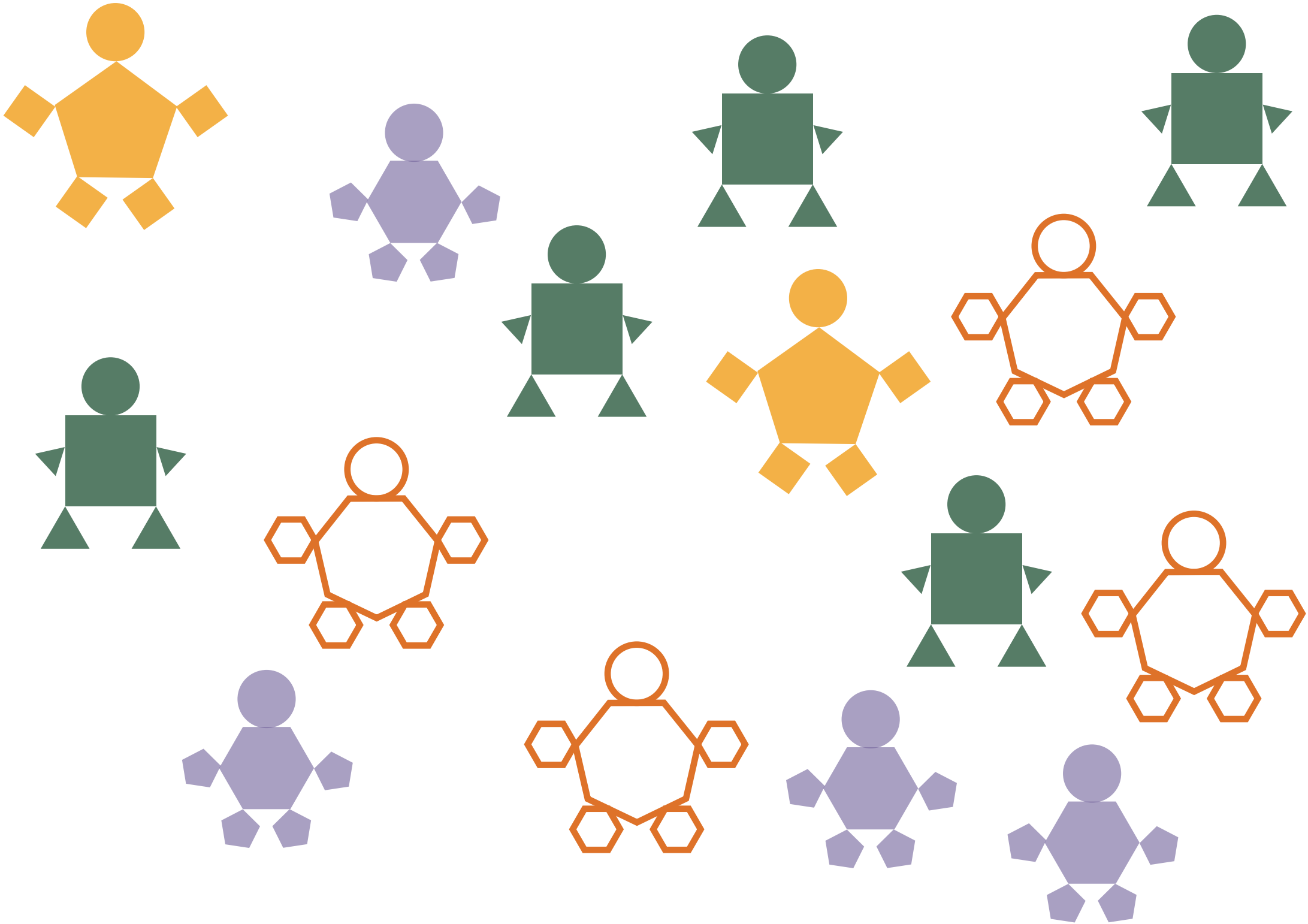
The natural variation of samples is called **sampling variability**.

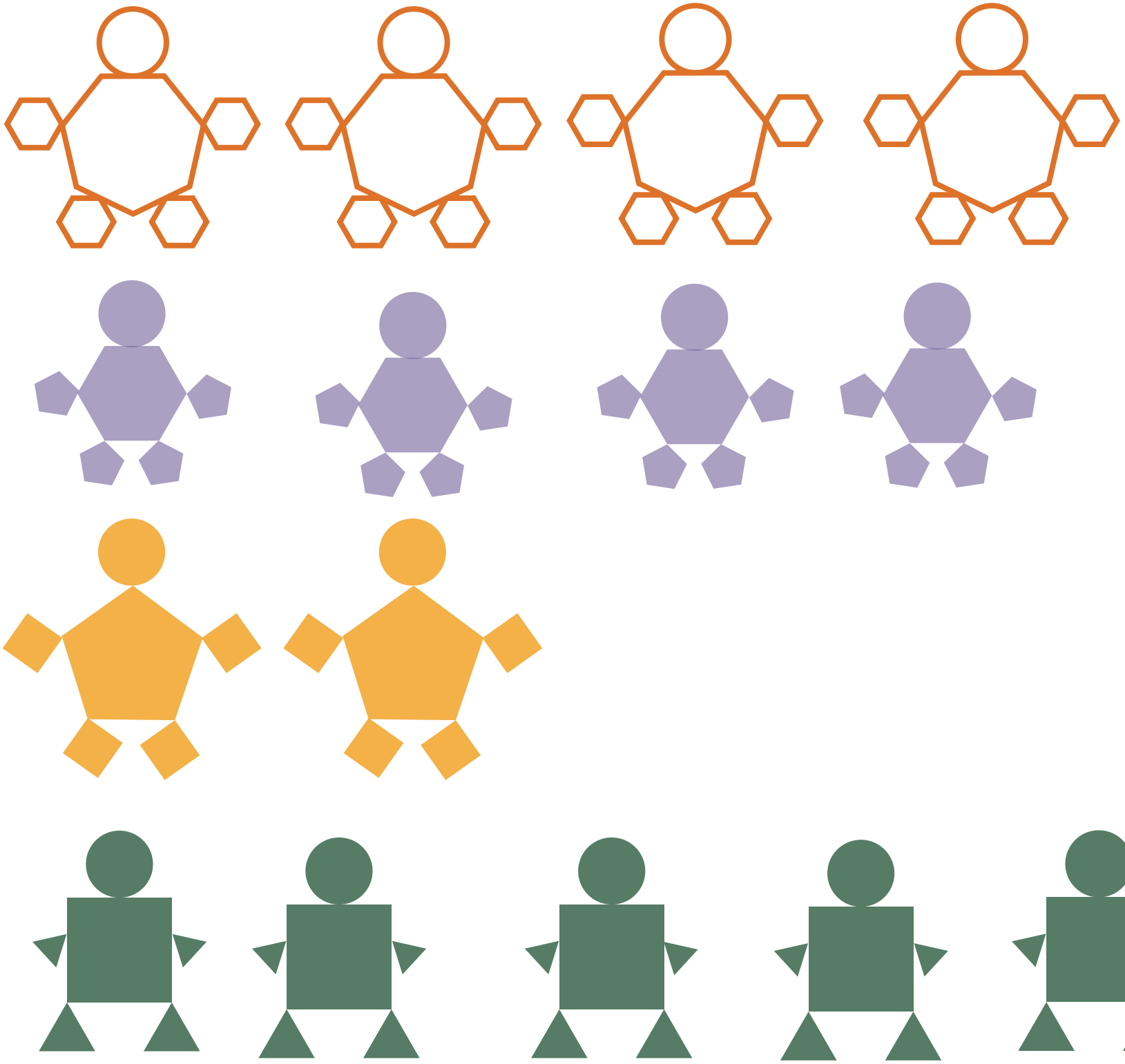
This is unavoidable and expected in random sampling, and in most cases is not an issue.

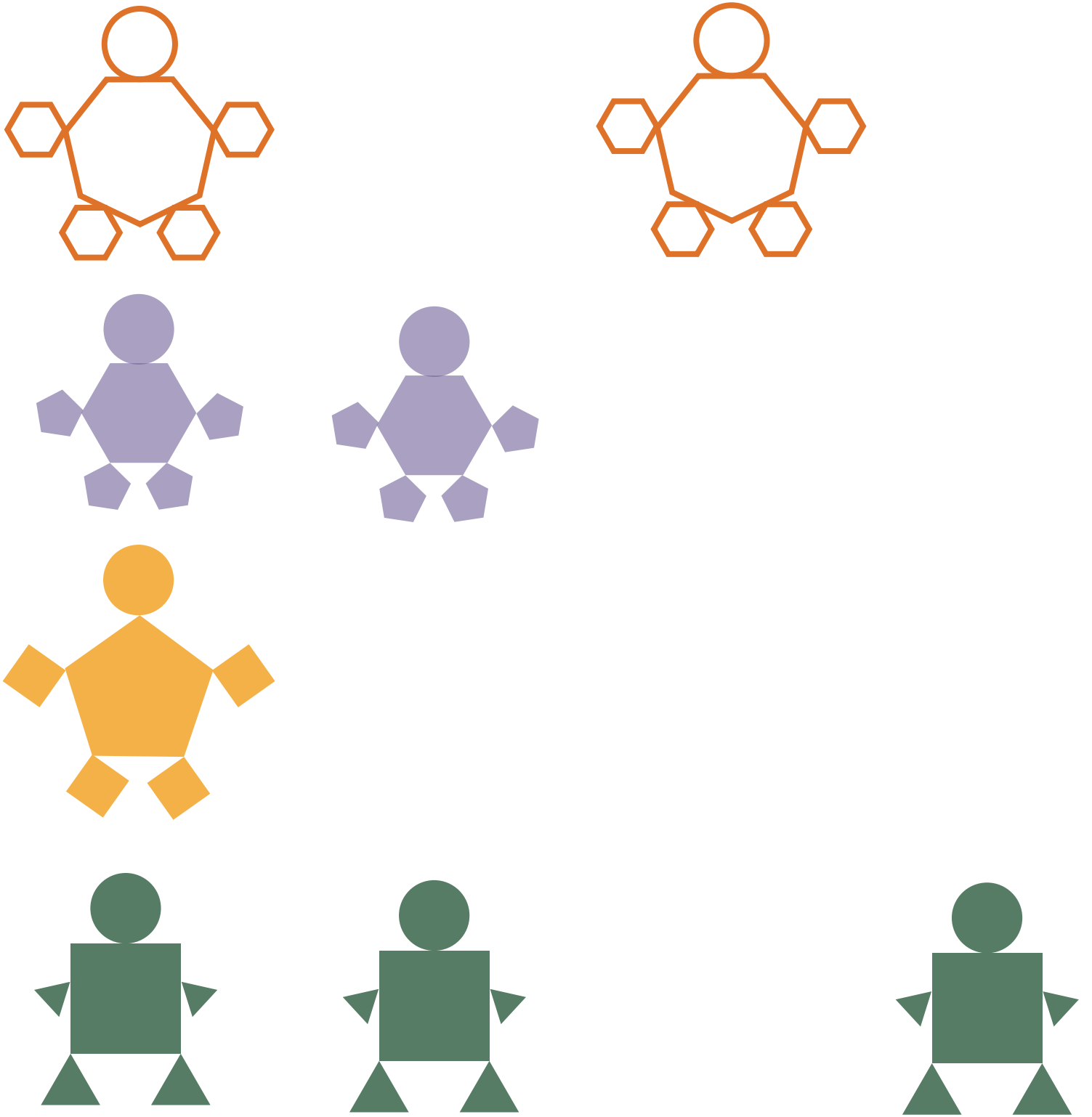
To help account for variability, pollsters might instead use a **stratified sample**.

STRATIFIED SAMPLING

In **stratified sampling**, a population is divided into a number of subgroups (or strata). Random samples are then taken from each subgroup with sample sizes proportional to the size of the subgroup in the population.







EXAMPLES OF STRATIFIED SAMPLING

A company has 500 employees. The company has three departments:

40% of the employees work in the **Sales**
35% work **engineering**
25% work in **marketing**.
The company wants to conduct a survey by sampling **200 employees**. To ensure that the survey accurately reflects the composition of the company, they could use stratified sampling.
From the Sales department **40% of 200=80 employees** should be selected.
From the Engineering department **35% of 200=70 employees** should be selected.
From the Marketing department (**25% of 200**), **50 employees** should be selected.
So, out of the 200 people sampled:
80 employees will come from Sales
70 from Engineering
50 from Marketing

EMESMEEESSSMEESSSEMMEMEEESEMES-
SEMMMMSEMSMSEMEMESSSEEMSESSES-
SEMEMEESESMMSMESSSEEESEESSMMMM
SEMESMSSESSESMSESSMSMSEMSEEMMSS
MSEEMMMESMEMSEMEMSSEEMSSEEEESS
EEESSSESMMSSSSEMESSEMMEMSEEMEEM
SSMSSMSSSMMSESSEESESSESMSSSSSSSSM
SSEMEESMSEEEMSSSMEMSMSESEESSSEM
ESMSEEESMSSSESSSEMMMMEMEMSSEMM
MMSEMMESMMEESMSSSSEEMEMSMEEES
ESEMSMESMEEMEEEESSSSSEESESSSMSSSE
SMSSSESEEMSMESMESMESMEEESSEESESE
SEMESEMSEMEEMSEESMSESMMSEEEMESS
SSESMSSMSMMEESEEMMSEMEMSMESMSE
ESSSSMEMESSSEMESSMESMSSESSEESEESE
SEMSSEEEMMMSEESEESMSSSSMEMEEESS
ESS


S S S S S S S S S S S S S S S S S S
S S S S S S S S S S S S S S S S S S
S S S S S S S S S S S S S S S S S S
S S S S S S S S S S S S S S S S S S
S S S S S S S S S S S S S S S S S S
S S S S S S S S S S S S S S S S S S
S S S S S S S S S S S S S S S S S S
S S S S S S S S S S S S S S S S S S
S S S S S S S S S S S S S S S S S S
S S S S S S S S S S S S S S S S S S
M M M M M M M M M M M M M M
M M M M M M M M M M M M M M
M M M M M M M M M M M M M M
M M M M M M M M M M M M M M
M M M M M M M M M M M M M M
M M M M M M M M M M M M M M
M M M M M M M M M M M M M M
M M M M M M M M M M M M M M
M M M M M M M M M M M M M M
M M M M M
E E E E E E E E E E E E E E E E E E E E E E
E E E E E E E E E E E E E E E E E E E E E E
E E E E E E E E E E E E E E E E E E E E E E
E E E E E E E E E E E E E E E E E E E E E E
E E E E E E E E E E E E E E E E E E E E E E
E E E E E E E E E E E E E E E E E E E E E E
E E E E E E E E E E E E E E E E E E E E E E

S S S S S S S S S S S S S S S S
S S S S S S S S S S
S S S S S S S S S
S S S S S S S S S S
S S S S S S S S S S
S S S S S S S S S S S S S S S
S S S S S S S S S S S S S
S S S S S S S S S S
S S S S S S S S S S S S S
S S S S S S S S S S S S S

M M M M M
M M M M M M
M M M M
M M M M M
M M M M
M M
M M M M M M
M M M M M M
M M M M
M M M M M
M M M

E E E E E E E E E E E E E E
E E E E E E E E E E E E E E
E E E E E E E E E E E E E E
E E E E E E E E E E E E E E
E E E E E E E E E E E E E E E
E E E E E E E E E E E E E E E
E E E E E E E E E E E E E E E E E E

QUOTA SAMPLING

Quota sampling is a variation on stratified sampling, wherein samples are collected in each subgroup until the desired quota is met.

Example

A university wants to conduct a survey on campus facilities satisfaction among its students. They decide to use quota sampling to ensure that their sample includes students from different academic years. They set quotas to include 100 students from each of the following categories:

Freshmen (1st year)
Sophomores (2nd year)
Juniors (3rd year)
Seniors (4th year)

Once they have filled the quota for a particular year (e.g., 100 Freshmen), they no longer include students from that year in the sample. They continue to gather responses from students in the remaining years until all quotas are met.

CLUSTER SAMPLING

In **cluster sampling**, the population is divided into subgroups (clusters), and a set of subgroups are selected to be in the sample.

Example

A city wants to survey public library usage. They randomly select 5 out of 20 library branches. They then survey all patrons who visit those 5 branches during a specific week.

Example

A national park wants to assess visitor experiences. They randomly select 5 out of 30 park campgrounds. They then survey all campers staying at those 5 campgrounds during a particular weekend.

SYSTEMATIC SAMPLING

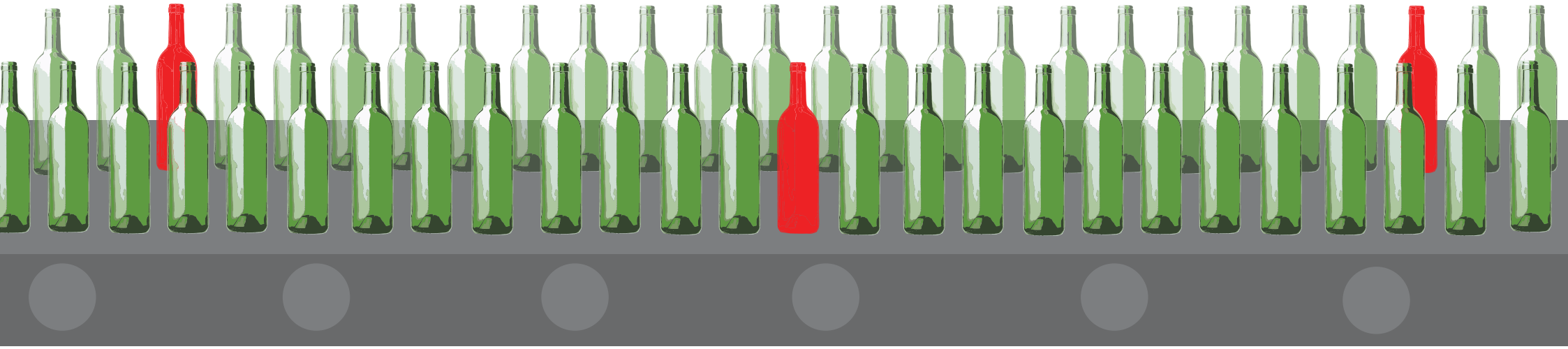
In **systematic sampling**, every nth member of the population is selected to be in the sample.

Example

A factory wants to inspect the quality of products on the assembly line. They use systematic sampling by checking every 50th item produced during a shift for quality control.

Example

A factory wants to inspect the quality of products on the assembly line. They use systematic sampling by checking every 20th item produced during a shift for quality control.



Example

A supermarket wants to assess the freshness of its produce. They use systematic sampling by inspecting every 40th item of fruit delivered each day.

The Worst Way to Sample

Perhaps the worst types of sampling methods are convenience samples and voluntary response samples.

CONVENIENCE & VOLUNTARY SAMPLING

Convenience sampling is the practice of samples chosen by selecting whoever is convenient.

Voluntary response sampling is allowing the sample to volunteer.

Question

A journalist interviews the first 20 people who enter a community center for a public event.

Which sampling method is represented by this scenario?

Question

A journalist interviews the first 20 people who enter a community center for a public event.

Which sampling method is represented by this scenario?

Convenience sampling

Question

A company surveys employees who are present at the office during the lunch break.

Which sampling method is represented by this scenario?

Question

A company surveys employees who are present at the office during the lunch break.

Which sampling method is represented by this scenario?

Convenience sampling

Question

A local community center posts a suggestion box for residents to express their opinions on new community programs.

Which sampling method is represented?

Question

A local community center posts a suggestion box for residents to express their opinions on new community programs.

Which sampling method is represented?

This is a self-selected sample, or voluntary response sample, where residents choose to participate by submitting their suggestions.

Question

What sampling method was used?

Every 6th customer entering a store was selected for a feedback survey.
survey.

Question

What sampling method was used?

Every 6th customer entering a store was selected for a feedback survey.
survey.

Systematic sampling

Question

What sampling method was used?

A sample was chosen to include 40 teenagers and 60 adults.

Question

What sampling method was used?

A sample was chosen to include 40 teenagers and 60 adults.

Stratified or Quota

Question

What sampling method was used?

Subscribers to a magazine are invited to share their opinions on a new feature through an email survey.

Question

What sampling method was used?

Subscribers to a magazine are invited to share their opinions on a new feature through an email survey.

Voluntary response

Question

What sampling method was used?

A company uses a random number generator to select 100 employees from their entire staff list for a satisfaction survey.

Question

What sampling method was used?

A company uses a random number generator to select 100 employees from their entire staff list for a satisfaction survey.

Simple random

Question

What sampling method was used?

To study public opinion on a new park, a research firm randomly selects 5 out of 20 neighborhoods and surveys all residents in those neighborhoods.

Question

What sampling method was used?

To study public opinion on a new park, a research firm randomly selects 5 out of 20 neighborhoods and surveys all residents in those neighborhoods.

Cluster

SOURCES OF BIAS

There are number of ways that a study can be ruined before you even start collecting data.

Sampling bias – when the sample is not representative of the population

Voluntary response bias – the sampling bias that often occurs when the sample is volunteers

Self-interest study – bias that can occur when the researchers have an interest in the outcome

Response bias – when the responder gives inaccurate responses for any reason

Perceived lack of anonymity – when the responder fears giving an honest answer might negatively affect them

Loaded questions – when the question wording influences the responses

Non-response bias – when people refusing to participate in the study can influence the validity of the outcome

Question

Identify the type of bias.

Consider a recent study that found eating a certain brand of cereal improves athletic performance in high school students. This study was funded by the cereal company itself. Identify the type of sampling bias found in this example.

Question

Identify the type of bias.

Consider a recent study that found eating a certain brand of cereal improves athletic performance in high school students. This study was funded by the cereal company itself. Identify the type of sampling bias found in this example.

Solution: This is an example of a **self-interest study**; the researchers have a vested interest in a positive outcome. This suggests that the study should be examined with caution for potential bias.

Question

Identify the type of bias.

A survey asks participants, “How often do you exercise each week?” What type of sampling bias might this lead to?

Question

Identify the type of bias.

A survey asks participants, “How often do you exercise each week?” What type of sampling bias might this lead to?

Solution: This might suffer from **response bias**, as participants might overestimate their exercise frequency to appear more health-conscious.

Question

Identify the type of bias.

An organization conducts a survey asking employees if they feel they receive adequate support for their work. Which sampling bias may occur in this scenario?

Question

Identify the type of bias.

An organization conducts a survey asking employees if they feel they receive adequate support for their work. Which sampling bias may occur in this scenario?

Solution: This survey may be affected by **perceived lack of anonymity**, as employees might fear that their responses could impact their job security or relationships with management, leading to less honest answers.

Question

Identify the type of bias.

A health survey asks, “How often do you visit your doctor for routine check-ups?” and 35% of people refuse to participate. Which sampling bias is represented?

Question

Identify the type of bias.

A health survey asks, “How often do you visit your doctor for routine check-ups?” and 35% of people refuse to participate. Which sampling bias is represented?

Solution: This is an example of **non-response bias**, as the high refusal rate may skew the results, making them less representative of the overall population.

Question

Identify the type of bias.

A university sends an email survey to alumni asking for feedback on their education experience. What type of sampling method is used, and what bias might occur?

Question

Identify the type of bias.

A university sends an email survey to alumni asking for feedback on their education experience. What type of sampling method is used, and what bias might occur?

Solution: This is a self-selected sample. It might introduce **voluntary response bias**, as alumni who choose to respond might have stronger opinions about their education, either positive or negative, which may not represent the entire alumni population.

Question

Identify the type of bias.

A political campaign conducts a survey by calling people who have previously donated to their party. What type of sampling bias might this lead to?

Question

Identify the type of bias.

A political campaign conducts a survey by calling people who have previously donated to their party. What type of sampling bias might this lead to?

Solution: This introduces sampling bias, as it only includes individuals who have previously supported the party financially, which may not represent the views of the general population.

Observational studies and experiments

An **observational study** is a study based on observations or measurements

Observational studies and experiments

An **observational study** is a study based on observations or measurements

An **experiment** is a study in which the effects of a treatment are measured.

Examples of experiments:

Agricultural Study: A farmer tests a new type of fertilizer by applying it to half of their crops while using traditional fertilizer on the other half. The new fertilizer is the treatment.

Environmental Science Experiment: A researcher tests the effect of a new water filtration system on river water quality by installing it at one site and comparing it to an unfiltered site. The filtration system is the treatment.

Mental Health Intervention: A therapist tests the effectiveness of a new meditation technique by introducing it to a group of patients and comparing their stress levels to a group that does not use the technique. The meditation technique is the treatment.

Question

Determine whether the following scenario describes an observational study or an experiment:

Blood pressure levels are recorded for 50 people at rest and after walking up a flight of stairs.

Question

Determine whether the following scenario describes an observational study or an experiment:

Blood pressure levels are recorded for 50 people at rest and after walking up a flight of stairs.

Experiment

Question

Determine whether the following scenario describes an observational study or an experiment:

Blood pressure levels are recorded for 50 people at rest and after walking up a flight of stairs.

Experiment

Question

Determine whether the following scenario describes an observational study or an experiment:

A researcher records the dietary habits of 100 individuals and tracks their health outcomes over five years.

Question

Determine whether the following scenario describes an observational study or an experiment:

A researcher records the dietary habits of 100 individuals and tracks their health outcomes over five years.

Observational study

Question

Determine whether the following scenario describes an observational study or an experiment:

Two groups of athletes are randomly assigned to different training programs, and their performance is measured after six months.

Question

Determine whether the following scenario describes an observational study or an experiment:

Two groups of athletes are randomly assigned to different training programs, and their performance is measured after six months.

Experiment

Question

Determine whether the following scenario describes an observational study or an experiment:

Researchers survey 1,000 adults about their sleep habits and analyze the relationship between sleep duration and reported mental health.

Question

Determine whether the following scenario describes an observational study or an experiment:

Researchers survey 1,000 adults about their sleep habits and analyze the relationship between sleep duration and reported mental health.

Observational study

A city wants to reduce traffic accidents at a dangerous intersection. To test a new traffic signal system, they install it at the intersection and compare accident rates before and after the change. After six months, the number of accidents decreases significantly.

However, at the same time, the city also began a public safety campaign about driving cautiously with greater police presence. It's unclear whether the drop in accidents is due to the new traffic signal, the increased police presence, or the public safety campaign.

A city wants to reduce traffic accidents at a dangerous intersection. To test a new traffic signal system, they install it at the intersection and compare accident rates before and after the change. After six months, the number of accidents decreases significantly.

However, at the same time, the city also began a public safety campaign about driving cautiously, and the police increased patrols in the area. It's unclear whether the drop in accidents is due to the new traffic signal, the increased police presence, or the public safety campaign.

This is called **confounding** – when it is not clear which factor or factors caused the observed effect. Confounding is the downfall of many experiments, though sometimes it is hidden.

Confounding

Confounding occurs when there are two potential variables that could have caused the outcome and it is not possible to determine which actually caused the result.

Example

A marine biologist tests a new type of fish food to see if it enhances the growth rate of juvenile fish. They feed one group of fish the new food and another group the standard food. Over a few months, the fish fed the new food show faster growth.

However, during this period, the biologist also installs new water filtration systems and increases the frequency of water changes. It's unclear whether the improved growth is due to the new fish food, the upgraded filtration system, or the more frequent water changes. These additional changes introduce confounding variables, making it difficult to determine the specific cause of the enhanced growth.

Example

A chemist tests a new method for purifying water by using a novel filtration material. They apply the method to one batch of contaminated water and compare its purity to another batch treated with the traditional filtration method.

At the same time, the chemist also changes the flow rate of the water through the filter and adjusts the temperature of the water. It's unclear whether the improvement in water purity is due to the new filtration material, the altered flow rate, or the temperature adjustments. These additional changes introduce confounding variables, making it difficult to determine which factor most significantly affected the water purity.

Question

To evaluate the effectiveness of a new sleep aid on improving sleep quality, a researcher conducts the following experiment: 80 participants (50 men and 30 women) who report having trouble sleeping are recruited. Their sleep quality is measured using a sleep diary and wearable devices. The researcher gives the men a new sleep aid pill and the women a placebo pill, but only the researcher is aware of this distinction.

Question

To evaluate the effectiveness of a new sleep aid on improving sleep quality, a researcher conducts the following experiment: 80 participants (50 men and 30 women) who report having trouble sleeping are recruited. Their sleep quality is measured using a sleep diary and wearable devices. The researcher gives the men a new sleep aid pill and the women a placebo pill, but only the researcher is aware of this distinction.

The results of the experiment are likely to be invalid mostly because:

A: The subjects did not know if they were receiving the real treatment.
B: The treatment group and control group were not the same size.
C: The subjects were volunteers.
D: The gender of the participant is a confounding variable in this experiment.
E: None of the above

Question

To evaluate the effectiveness of a new sleep aid on improving sleep quality, a researcher conducts the following experiment: 80 participants (50 men and 30 women) who report having trouble sleeping are recruited. Their sleep quality is measured using a sleep diary and wearable devices. The researcher gives the men a new sleep aid pill and the women a placebo pill, but only the researcher is aware of this distinction.

The results of the experiment are likely to be invalid mostly because:

A: The subjects did not know if they were receiving the real treatment.
B: The treatment group and control group were not the same size.
C: The subjects were volunteers.
D: The gender of the participant is a confounding variable in this experiment.
E: None of the above

There are a number of measures that can be introduced to help reduce the likelihood of confounding.
The primary measure is to use a **control group**.

Control Group

When using a control group, the participants are divided into two or more groups, typically a **control group** and a **treatment group**.

The treatment group receives the treatment being tested; the control group does not receive the treatment.

Example

A pharmaceutical company tests a new drug by administering it to one group of patients while another group receives the standard treatment. Both groups are monitored in the same clinic with similar conditions and protocols. The health outcomes are compared to control for differences in clinical environment and care.

Example

An environmental scientist tests a new method of cleaning oil spills by applying it to one section of a contaminated beach and using a standard method on another section. Both sections are exposed to similar environmental conditions. The effectiveness of the new cleaning method is compared to assess its impact on oil removal and beach recovery.

Sometimes not giving the control group anything does not completely control for confounding variables. For example, suppose a medicine study is testing a new headache pill by giving the treatment group the pill and the control group nothing. If the treatment group showed improvement, we would not know whether it was due to the medicine in the pill, or a response to have taken any pill. This is called a **placebo effect**.

Placebo effect

The **placebo effect** is when the effectiveness of a treatment is influenced by the patient's perception of how effective they think the treatment will be, so a result might be seen even if the treatment is ineffectual.

A 2013 study from the U.K. found that 97% of physicians acknowledged in a survey having used some form of placebo during their career. This might be as simple as expressing a strong belief in the likelihood that a patient will feel better from whatever treatment the doctor prescribes, even if the treatment itself is not chemically powerful.

Howick, Jeremy, Felicity L. Bishop, Carl J. Heneghan, Jane L Wolstenholme, Sarah L Stevens, FD Richard Hobbs and George Lewith. "Placebo Use in the United Kingdom: Results from a National Survey of Primary Care Practitioners." PLoS ONE 8 (2013): n. pag.

To control for the placebo effect, a placebo, or *dummy treatment*, is often given to the control group. This way, both groups are truly identical except for the specific treatment given.

Placebo and Placebo controlled experiments

A **placebo** is a dummy treatment given to control for the placebo effect.

An experiment that gives the control group a placebo is called a
placebo controlled experiment.

Example

In a clinical trial testing a new anti-nausea medication, participants in the control group receive a pill that looks identical to the medication but contains only a harmless substance. This allows researchers to determine whether the new medication is more effective than the placebo at reducing nausea.

Example

In a trial for a new antihistamine, one group of participants receives the actual drug while another group gets a placebo tablet made of an inert substance. Both groups are asked to track their allergy symptoms, allowing researchers to see if the new antihistamine provides better relief compared to the placebo.

In some cases, it is more appropriate to compare to a conventional treatment than a placebo. For example, in a cancer research study, it would not be ethical to deny any treatment to the control group or to give a placebo treatment. In this case, the currently acceptable medicine would be given to the second group, called a comparison group in this case.

When using a placebo, it would defeat the purpose if the participant knew they were receiving the placebo.

Blind studies

A **blind study** is one in which the participant does not know whether or not they are receiving the treatment or a placebo.

Blind studies

A **blind study** is one in which the participant does not know whether or not they are receiving the treatment or a placebo.

A **double-blind study** is one in which those interacting with the participants don't know who is in the treatment group and who is in the control group.

Example

In a study about anti-depression medicine, you would not want the psychological evaluator to know whether the patient is in the treatment or control group either, as it might influence their evaluation, so the experiment should be conducted as a double-blind study.

It should be noted that not every experiment needs a control group.

An example of an experiment without a control group is if a researcher is interested in the effect of a new medication on blood pressure, they could measure blood pressure in the same group of participants both before and after taking the medication.

Question

To evaluate a new dietary supplement,
researchers divide participants into two groups: one group receives the supplement, and the other group receives a placebo that looks identical to the supplement. The researchers administering the supplement do not know which participants are receiving the real supplement and which are receiving the placebo.

Does this experiment have a control group?

Is it blind, double-blind, or neither?

Question

To evaluate a new dietary supplement, researchers divide participants into two groups: one group receives the supplement, and the other group receives a placebo that looks identical to the supplement. The researchers administering the supplement do not know which participants are receiving the real supplement and which are receiving the placebo.

Does this experiment have a control group?

Yes, the experiment has a control group. The control group is the group receiving the placebo, which is used to compare against the experimental group receiving the actual dietary supplement.

Is it blind, double-blind, or neither?

The experiment is double-blind because neither the participants nor the researchers administering the supplement know which group the participants are in.

Question

A study is conducted to evaluate the effectiveness of a new cognitive training program designed to improve memory. Participants are randomly assigned to either the new training program or a traditional training program. The evaluators who assess the participants' memory improvement do not know which program each participant underwent.

Does this experiment have a control group?

Is it blind, double-blind, or neither?

Question

A study is conducted to evaluate the effectiveness of a new cognitive training program designed to improve memory. Participants are randomly assigned to either the new training program or a traditional training program. The evaluators who assess the participants' memory improvement do not know which program each participant underwent.

Does this experiment have a control group?

Yes, the experiment has a control group. The control group is the one receiving the traditional training program, which serves as a baseline to compare the effects of the new cognitive training program.

Is it blind, double-blind, or neither?

The experiment is single-blind because the evaluators assessing the memory improvement do not know which training program each participant received. However, the participants are aware of which training program they are undergoing, so it is not double-blind.

Question

A research team wants to assess the impact of a new classroom teaching method on students' math test scores. They conduct a study with 500 students from several schools. The students are randomly assigned to one of two groups. The first group of 250 students is taught using the new teaching method, while the second group of 250 students is taught using the standard teaching method. The students' math test scores are compared at the end of the school year. All students know which teaching method they are receiving.

Question

A research team wants to assess the impact of a new classroom teaching method on students' math test scores. They conduct a study with 500 students from several schools. The students are randomly assigned to one of two groups. The first group of 250 students is taught using the new teaching method, while the second group of 250 students is taught using the standard teaching method. The students' math test scores are compared at the end of the school year. All students know which teaching method they are receiving.

In this study, which is the control group?

A: The group of 250 students using the new teaching method
B: The 500 students in total
C: The group of 250 students using the standard teaching method
D: There is no control group.

Question

A research team wants to assess the impact of a new classroom teaching method on students' math test scores. They conduct a study with 500 students from several schools. The students are randomly assigned to one of two groups. The first group of 250 students is taught using the new teaching method, while the second group of 250 students is taught using the standard teaching method. The students' math test scores are compared at the end of the school year. All students know which teaching method they are receiving.

In this study, which is the control group?

A: The group of 250 students using the new teaching method
B: The 500 students in total
C: The group of 250 students using the standard teaching method
D: There is no control group.

Question

A research team wants to assess the impact of a new classroom teaching method on students' math test scores. They conduct a study with 500 students from several schools. The students are randomly assigned to one of two groups. The first group of 250 students is taught using the new teaching method, while the second group of 250 students is taught using the standard teaching method. The students' math test scores are compared at the end of the school year. All students know which teaching method they are receiving.

This study is:

A: Not an experiment
B: Not blind.
C: Double-blind
D: Blind, but not double-blind
E: None of the above

Question

A research team wants to assess the impact of a new classroom teaching method on students' math test scores. They conduct a study with 500 students from several schools. The students are randomly assigned to one of two groups. The first group of 250 students is taught using the new teaching method, while the second group of 250 students is taught using the standard teaching method. The students' math test scores are compared at the end of the school year. All students know which teaching method they are receiving.

This study is:

A: Not an experiment
B: Not blind.
C: Double-blind
D: Blind, but not double-blind
E: None of the above

Question

A researcher tests whether a new type of energy bar improves athletic performance. The study involves 100 participants who are divided into two groups. The first group of 50 participants eats the new energy bar before exercising, while the second group of 50 participants eats a regular energy bar. The participants' exercise performance is measured and compared. All participants know which type of energy bar they are consuming.

Question

A researcher tests whether a new type of energy bar improves athletic performance. The study involves 100 participants who are divided into two groups. The first group of 50 participants eats the new energy bar before exercising, while the second group of 50 participants eats a regular energy bar. The participants' exercise performance is measured and compared. All participants know which type of energy bar they are consuming.

In this study, which is the control group?

A The group of 50 participants eating the new energy bar
B The 100 participants in total
C The group of 50 participants eating the regular energy bar
D There is no control group.

Question

A researcher tests whether a new type of energy bar improves athletic performance. The study involves 100 participants who are divided into two groups. The first group of 50 participants eats the new energy bar before exercising, while the second group of 50 participants eats a regular energy bar. The participants' exercise performance is measured and compared. All participants know which type of energy bar they are consuming.

In this study, which is the control group?

A The group of 50 participants eating the new energy bar
B The 100 participants in total
C The group of 50 participants eating the regular energy bar
D There is no control group.

Question

A researcher tests whether a new type of energy bar improves athletic performance. The study involves 100 participants who are divided into two groups. The first group of 50 participants eats the new energy bar before exercising, while the second group of 50 participants eats a regular energy bar. The participants' exercise performance is measured and compared. All participants know which type of energy bar they are consuming.

This study is:

A Not an experiment
B Not blind.
C Double-blind
D Blind, but not double-blind
E None of the above

Question

A researcher tests whether a new type of energy bar improves athletic performance. The study involves 100 participants who are divided into two groups. The first group of 50 participants eats the new energy bar before exercising, while the second group of 50 participants eats a regular energy bar. The participants' exercise performance is measured and compared. All participants know which type of energy bar they are consuming.

This study is:

A Not an experiment
B Not blind.
C Double-blind
D Blind, but not double-blind
E None of the above

Question

A research study is conducted to test the effectiveness of a new anti-anxiety medication. Participants are randomly assigned to receive either the new medication or a placebo pill that looks identical but contains no active ingredients.

Participants: 100 participants are divided into two groups: one group receives the new medication, while the other group receives the placebo.
Healthcare Providers: The therapists who interact with the participants and assess their anxiety levels do not know which treatment each participant is receiving.

Question

A research study is conducted to test the effectiveness of a new anti-anxiety medication. Participants are randomly assigned to receive either the new medication or a placebo pill that looks identical but contains no active ingredients.

Participants: 100 participants are divided into two groups: one group receives the new medication, while the other group receives the placebo.
Healthcare Providers: The therapists who interact with the participants and assess their anxiety levels do not know which treatment each participant is receiving.

Q:Which is the control group?

Question

A research study is conducted to test the effectiveness of a new anti-anxiety medication. Participants are randomly assigned to receive either the new medication or a placebo pill that looks identical but contains no active ingredients.

Participants: 100 participants are divided into two groups: one group receives the new medication, while the other group receives the placebo.
Healthcare Providers: The therapists who interact with the participants and assess their anxiety levels do not know which treatment each participant is receiving.

Q:Which is the control group?
The group receiving the placebo pill.

Question

A research study is conducted to test the effectiveness of a new anti-anxiety medication. Participants are randomly assigned to receive either the new medication or a placebo pill that looks identical but contains no active ingredients.

Participants: 100 participants are divided into two groups: one group receives the new medication, while the other group receives the placebo.
Healthcare Providers: The therapists who interact with the participants and assess their anxiety levels do not know which treatment each participant is receiving.

Q. What type of study is this?

Question

A research study is conducted to test the effectiveness of a new anti-anxiety medication. Participants are randomly assigned to receive either the new medication or a placebo pill that looks identical but contains no active ingredients.

Participants: 100 participants are divided into two groups: one group receives the new medication, while the other group receives the placebo.
Healthcare Providers: The therapists who interact with the participants and assess their anxiety levels do not know which treatment each participant is receiving.

Q. What type of study is this?

Double-blind. Both the participants and the healthcare providers assessing the participants' anxiety levels are unaware of which treatment is being administered. This design helps to minimize biases and ensures that the observed effects are due to the medication itself rather than expectations or prejudices.