

SAMPLING BIAS

A sampling method is 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

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

The most basic random method is simple

A 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 has an 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 is not significant.

SAMPLING VARIABILITY

is called **sampling variability**.

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

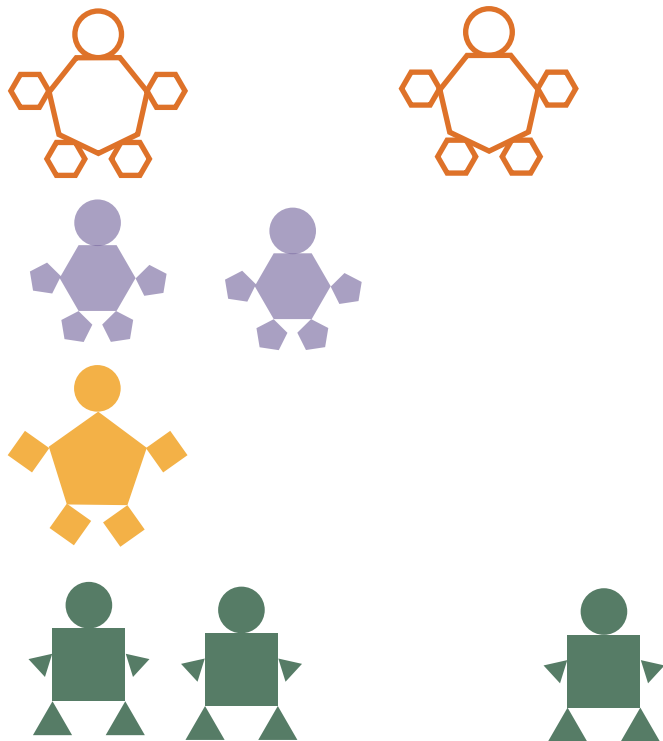
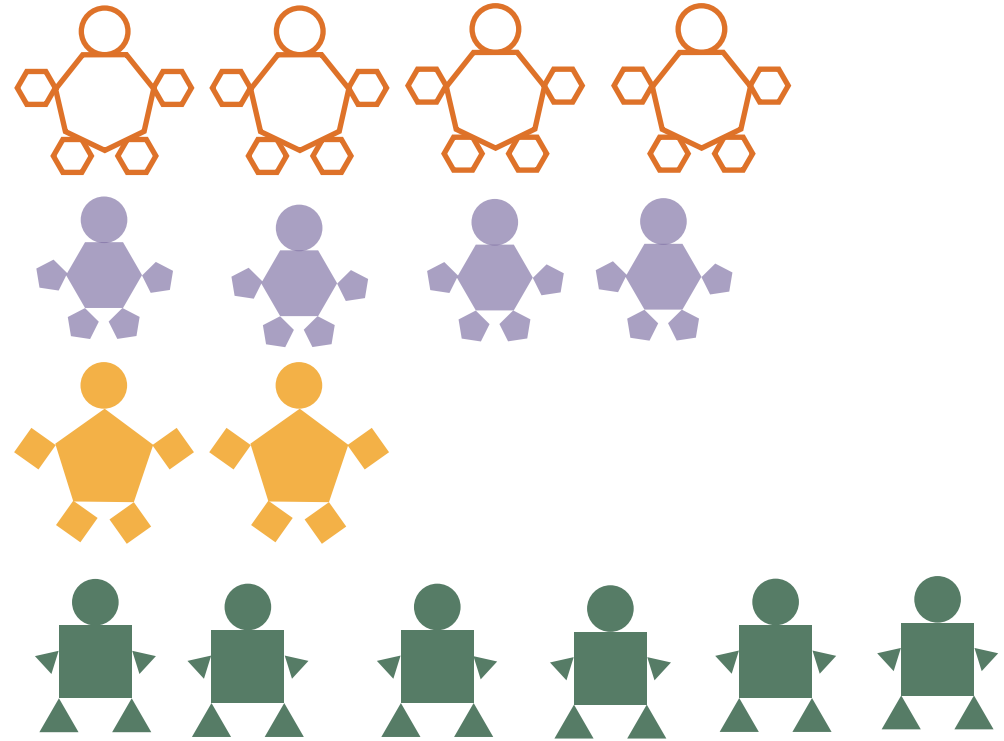
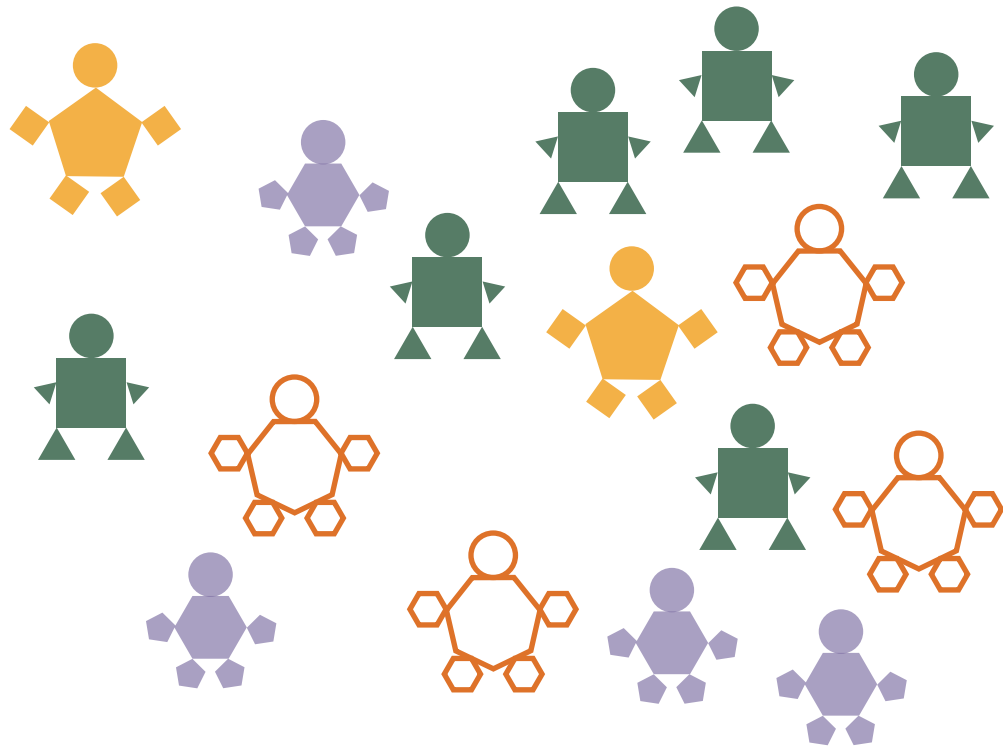
To help account for variability, pollsters might instead use a .

STRATIFIED SAMPLING

In stratified sampling, a population is

(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 500**), **50 employees** should be selected.

So, out of the 200 people sampled:

80 employees will come from Sales

70 from Engineering

50 from Marketing

EMESMEEESSMEESSSEMMEEMEESEMES-
 SEMMMMSEMSEMESSEEMSESSES-
 SEMEMEESMMSMESSEEESEESSMMM
 SEMESMSESSESMSSESSMSMSEMSEEMSS
 MSEEMMMESMEMSEMEMSSEEMSSEEEESS
 EEESSESMMSSSSEMESSEMMEEMSEEMEM
 SSMSSMSSMMSESSESESESSMSSSSSSSM
 SSEMEESMSEEMSSSMEMSMSESEESSEM
 ESMSEEESSMSSESSEMMMMEMEMSSEM
 MMSEMMEEMMEESMSSSSEEMEMSMEEES
 ESEMSMESMEEMEEEESSSSSESESSMSSE
 SMSSSESEEMSMESMESMEEESSESESE
 SEMESEMEMEEMSEESMSSEMMSEEEEMESS
 SSESMSMMEESEEMMSEMEMSMESE
 ESSSSMEMESSSEMESSMESMSSESSESESE
 SEMSSEEMMMSESESESMSSSMEMEEESS
 ESS

SSSSSSSSSSSSSSSSSSSSSSSSSSSSSS
 SSSSSSSSSSSSSSSSSSSSSSSSSSSSS
 SSSSSSSSSSSSSSSSSSSSSSSSSSSSS
 SSSSSSSSSSSSSSSSSSSSSSSSSSSSS
 SSSSSSSSSSSSSSSSSSSSSSSSSSSSS
 SSSSSSSSSSSSSSSSSSSSSSSSSSSSS
 SSSSSSSSSSSSSSSSSSSSSSSSSSSSS
 SSSSSSSSSSSSSSSSSSSSSSSSSSSSS
 SSSSSSSSSSSSSSSSSSSSSSSSSSSSS
 SSSSSSSSSSSSSSSSSSSSSSSSSSSSS
 SSSSSSSSSSSSSSSSSSSSSSSSSSSSS
 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 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

EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE
 EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE
 EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE
 EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE
 EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE
 EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE
 EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE
 EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE
 EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE

SSSSS SSS SS SSSSS S M M M M M
 SSSSSSS S SS M M MM MM
 SS S S SS S SS M M M M
 SS SSS SS SS S M M MM M
 SS SSS S S S M M M M M
 S S S S SSS SSSS SSS MM
 SSS S S SS SSSSS S M M MM M M
 S S SSS S S S S S M M M M M M
 SSSS S SSS SS SS S M M M M
 SSSSS S S SSSSSS M M M M M
 M M M

EE EEE EEE E E EEE EE
 E E E E EE EEE EE EE
 EEE EEE E EE E E EE E
 E EEEE E E E EEE E EE
 EEEE EE E E E EE E EE E
 EE E E EE EE EEE EE E
 EE EEEEE EE EE E EEEEE

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

CLUSTER SAMPLING

In cluster sampling, the population is divided into subgroups (clusters), and a



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.

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.

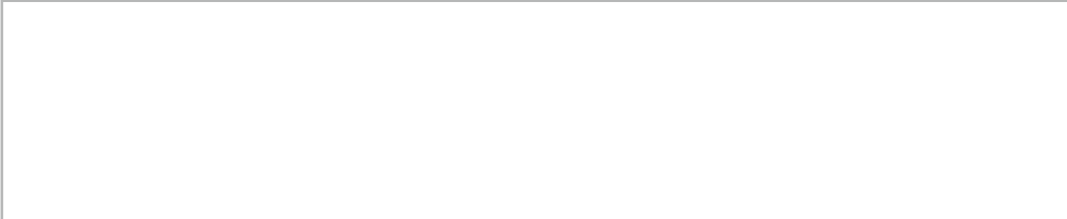
This approach ensures that each academic year is represented in the sample, but it may not be fully random since the selection is based on meeting specific quotas rather than being entirely random.

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,



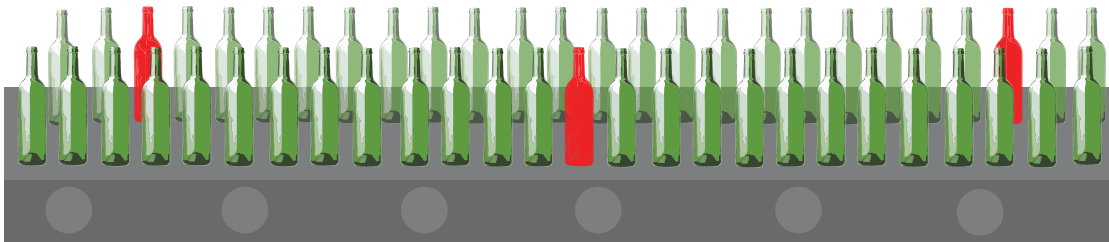
of the population is selected to be in the sample.

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 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

and

CONVENIENCE & VOLUNTARY SAMPLING

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

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

Which sampling method is represented by this scenario?

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

What sampling method was used?

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

Question

What sampling method was used?

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

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?

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?

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.

SOURCES OF BIAS

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

– when the sample is not representative of the population

Voluntary response bias – the sampling bias that often occurs when

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

Response bias – when the responder

– when the responder fears giving an honest answer might negatively affect them

Loaded questions – when the question wording

– 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.

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.

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.

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