# Sampling Principles and Strategies

Colby Community College

A **population** is the complete collection of all measurements or data that are being considered.

A **population** is the complete collection of all measurements or data that are being considered.

# Example 1

A researcher asks:

What is the average mercury content in swordfish in the Atlantic Ocean?

What is the target population?

A **population** is the complete collection of all measurements or data that are being considered.

# Example 1

A researcher asks:

What is the average mercury content in swordfish in the Atlantic Ocean?

What is the target population?

The target population is all the swordfish in the Atlantic ocean.

A **population** is the complete collection of all measurements or data that are being considered.

# Example 1

A researcher asks:

What is the average mercury content in swordfish in the Atlantic Ocean?

What is the target population?

The target population is all the swordfish in the Atlantic ocean.

### Definition

A **census** is the collection of data from every member of a population.

A **population** is the complete collection of all measurements or data that are being considered.

# Example 1

A researcher asks:

What is the average mercury content in swordfish in the Atlantic Ocean?

What is the target population?

The target population is all the swordfish in the Atlantic ocean.

### Definition

A **census** is the collection of data from every member of a population.

## Note

It is often prohibitively expensive to collect data from every member of a population.

A **sample** is a subset of members selected from a population.

A **sample** is a subset of members selected from a population.

# Example 2

In a journal article it was stated that there are 38 million carbon monoxide detectors installed in the United States. When 30 of them were randomly selected and tested, it was found that 12 of them failed an alarm in hazardous carbon monoxide conditions.

What is the population?

Ryan, T. J., & Arnold, K. J. (2011). Residential carbon monoxide detector failure rates in the United States. American journal of public health. 101(10), e15–e17. https://doi.org/10.2105/AJPH.2011.300274

A **sample** is a subset of members selected from a population.

# Example 2

In a journal article it was stated that there are 38 million carbon monoxide detectors installed in the United States. When 30 of them were randomly selected and tested, it was found that 12 of them failed an alarm in hazardous carbon monoxide conditions.

What is the population?

All 38 million carbon monoxide detectors in the United states.

Ryan, T. J., & Arnold, K. J. (2011). Residential carbon monoxide detector failure rates in the United States. American journal of public health. 101(10), e15–e17. https://doi.org/10.2105/AJPH.2011.300274

A **sample** is a subset of members selected from a population.

# Example 2

In a journal article it was stated that there are 38 million carbon monoxide detectors installed in the United States. When 30 of them were randomly selected and tested, it was found that 12 of them failed an alarm in hazardous carbon monoxide conditions.

What is the population?

All 38 million carbon monoxide detectors in the United states.

What is the sample?

Ryan, T. J., & Arnold, K. J. (2011). Residential carbon monoxide detector failure rates in the United States. American journal of public health, 101(10), e15–e17. https://doi.org/10.2105/AJPH.2011.300274

A **sample** is a subset of members selected from a population.

# Example 2

In a journal article it was stated that there are 38 million carbon monoxide detectors installed in the United States. When 30 of them were randomly selected and tested, it was found that 12 of them failed an alarm in hazardous carbon monoxide conditions.

What is the population?

All 38 million carbon monoxide detectors in the United states.

What is the sample?

The 30 carbon monoxide detectors that were selected and tested.

Ryan, T. J., & Arnold, K. J. (2011). Residential carbon monoxide detector failure rates in the United States. American journal of public health. 101(10), e15–e17. https://doi.org/10.2105/AJPH.2011.300274

**Anecdotal evidence** is a factual claim relying only on personal observation, collected in a casual or non-systematic manner.

**Anecdotal evidence** is a factual claim relying only on personal observation, collected in a casual or non-systematic manner.

# Example 3

A man on the news got mercury poisoning from eating swordfish, so the average mercury concentration in a swordfish must be dangerously high.

**Anecdotal evidence** is a factual claim relying only on personal observation, collected in a casual or non-systematic manner.

# Example 3

A man on the news got mercury poisoning from eating swordfish, so the average mercury concentration in a swordfish must be dangerously high.

# Example 4

My friend's dad had a heart attack and died after they game him a new heart disease drug, so the drug must not work.

**Anecdotal evidence** is a factual claim relying only on personal observation, collected in a casual or non-systematic manner.

# Example 3

A man on the news got mercury poisoning from eating swordfish, so the average mercury concentration in a swordfish must be dangerously high.

# Example 4

My friend's dad had a heart attack and died after they game him a new heart disease drug, so the drug must not work.

### Note

Such evidence may be true and verifiable, but it may only represent extraordinary cases.

Suppose we as a student who happens to be majoring in nutrition to select several graduates for a study.

What kind of students do you think she might collect?

Suppose we as a student who happens to be majoring in nutrition to select several graduates for a study.

What kind of students do you think she might collect?

Do you think her sample would be representative of all students?

Suppose we as a student who happens to be majoring in nutrition to select several graduates for a study.

What kind of students do you think she might collect?

Do you think her sample would be representative of all students?

## **Definition**

When data is gathered in a way that is not truly random, the sample is considered **biased**.

Suppose we as a student who happens to be majoring in nutrition to select several graduates for a study.

What kind of students do you think she might collect?

Do you think her sample would be representative of all students?

## **Definition**

When data is gathered in a way that is not truly random, the sample is considered **biased**.

### Note

Any time samples are selected by hand, there is a risk of picking a biased sample, even if the bias is not intentional.

Suppose we as a student who happens to be majoring in nutrition to select several graduates for a study.

What kind of students do you think she might collect?

Do you think her sample would be representative of all students?

### Definition

When data is gathered in a way that is not truly random, the sample is considered **biased**.

### Note

Any time samples are selected by hand, there is a risk of picking a biased sample, even if the bias is not intentional.

#### Note

Sample selection is often done by computer.

A **voluntary response sample** (or **self-selected sample**) is one which the respondents themselves decide whether to be included.

A **voluntary response sample** (or **self-selected sample**) is one which the respondents themselves decide whether to be included.

# Example 6

Internet polls, such as on Twitter, where people decided whether or not to participate.

A **voluntary response sample** (or **self-selected sample**) is one which the respondents themselves decide whether to be included.

# Example 6

Internet polls, such as on Twitter, where people decided whether or not to participate.

## Example 7

Call-in polls, where people are asked to call a special number to register an opinion.

A **loaded question** is one that is intentionally worded to elicit a desired response.

A **loaded question** is one that is intentionally worded to elicit a desired response.

# Example 8

When asked

"Should the President have the line item veto to eliminate waste?"

97% of respondents said yes.

A **loaded question** is one that is intentionally worded to elicit a desired response.

# Example 8

When asked

"Should the President have the line item veto to eliminate waste?"

97% of respondents said yes.

# Example 9

When asked

"Should the President have the line item veto?"

57% of respondents said yes.

A **non-response** occurs when someone refuses to respond or is not available.

A **non-response** occurs when someone refuses to respond or is not available.

# Example 10

Many telemarketers disguise their sales pitch as an opinion poll, causing the non-response problems to increase in recent decades.

A **non-response** occurs when someone refuses to respond or is not available.

# Example 10

Many telemarketers disguise their sales pitch as an opinion poll, causing the non-response problems to increase in recent decades.

# Example 11

A **prospective study** identifies individuals and collects information as events unfold.

A **prospective study** identifies individuals and collects information as events unfold.

# Example 12

Harvard, in 1938, began studying a group of 268 sophomores, and soon an additional study added 456 inner-city Bostonians. They are surveyed every two years and physical examinations given every five. The study continues to this day.

A **prospective study** identifies individuals and collects information as events unfold.

# Example 12

Harvard, in 1938, began studying a group of 268 sophomores, and soon an additional study added 456 inner-city Bostonians. They are surveyed every two years and physical examinations given every five. The study continues to this day.

### Definition

A retrospective study collects data after events have taken place.

A **prospective study** identifies individuals and collects information as events unfold.

# Example 12

Harvard, in 1938, began studying a group of 268 sophomores, and soon an additional study added 456 inner-city Bostonians. They are surveyed every two years and physical examinations given every five. The study continues to this day.

### Definition

A retrospective study collects data after events have taken place.

# Example 13