

# Study Design

## Learning Outcomes

- Explain the purpose of random sampling and its effect on scope of inference
- Explain the purpose of random assignment and its effect on scope of inference
- Identify whether a study is observational or an experiment
- Identify confounding variables in observational studies and explain why they are confounding
- Identify the types of bias present in a study

## Terminology Review

Statistics is the study of how best to collect, analyze, and draw conclusions from data. Statistical inference will allow us to make a statement about a population parameter based on a sample statistic.

Some terms covered in this activity are...

- Population
- Sample
- Parameter
- Statistic
- Selection Bias
- Response Bias
- Non-response Bias
- Scope of Inference
- Explanatory Variable
- Response Variable
- Confounding Variable
- Experiments
- Observational Study

To review these concepts see Section 1.3 to 1.6 in the textbook.

## Types of Bias

There are two parts to study design: how the sample was selected and how the study was conducted. First we will look at sampling and types of bias.

In these next questions, identify the target population, the sample, the variable, and the type of bias present.

1. To determine if the proportion of out of state undergraduate students at Montana State University has increased in the last 10 years, a statistics instructor sent an email survey to 500 randomly selected current undergraduate students. One of the questions on the survey asked whether they had in-state or out-of-state residency. She only received 378 responses.

Target population:

Sample:

Variable:

Type of Bias:

2. PEW Research surveys US adults about many different topics. Recently a survey was conducted to assess current presidential approval. A random sample of 6395 US adults was taken. Of those surveyed, 42% say they agree with President Trump on many or nearly all of the top issues facing the country today.

Target population:

Sample:

Variable:

Type of Bias:

3. A television station is interested in predicting whether or not voters in its listening area are opposed to legalizing marijuana for adult use. It asks its viewers to phone in and indicate whether they are in favor of this or opposed to this. Of the 2241 viewers who phoned in, forty-five percent were opposed to legalizing marijuana.

Target population:

Sample:

Variable:

Type of Bias:

4. To gauge the interest in a new swimming pool, a local organization stood outside of the Bogart Pool during open hours. One of the questions they asked was, "Since the Bogart Pool is in such bad repair, don't you agree that the city should fund a new pool?"

Target population:

Sample:

Variable:

Type of Bias:

5. The Bozeman school district is interested in surveying parents of students about their opinions on returning to school this fall following the COVID-19 pandemic. They divided the school district into 10 divisions based on location and randomly surveyed 20 households within each division.

Target population:

Sample:

Variable:


Type of Bias:

## Study Design


The two main study designs we will cover are observational studies and experiments. Both the sampling method and the study design will help to determine the **scope of inference** for a study. Remember that only in a randomized experiment can we conclude a **causal** (cause and effect) relationship between the explanatory and response variable.

**Scope of Inference:** If evidence of an association is found in our sample, what can be concluded?

	Study Type		
Selection of cases	Randomized experiment	Observational study	
Random sample (and no other sampling bias)	Causal relationship, and can generalize results to population.	Cannot conclude causal relationship, but can generalize results to population.	→ Inferences to population can be made
No random sample (or other sampling bias)	Causal relationship, but cannot generalize results to a population.	Cannot conclude causal relationship, and cannot generalize results to a population.	→ Can only generalize to those similar to the sample due to potential sampling bias



Can draw cause-and-effect conclusions



Can only discuss association  
due to potential confounding  
variables

For the next exercises, identify the explanatory variable, the response variable, a potential confounding variable, and the study design.

- The pharmaceutical company, Moderna Therapeutics is working in conjunction with the National Institute of Health towards a vaccine for COVID-19 and has recently begun Phase 3 clinical trials. US Clinical research sites will enroll 30,000 volunteers without COVID-19 to participate. Participants will be randomly assigned to receive either the candidate vaccine or a saline placebo. They will then be followed to assess vaccine related symptoms and development of COVID-19. The trial is blinded, so the investigators and the participants will not know who is assigned to which group.

Explanatory Variable:

Response Variable:

Confounding Variable:

Study design:

What is the scope of inference for this study?

7. In another study, a local health department randomly selected 1000 US adults without COVID-19 to participate in a health survey. Each participant was assessed at the beginning of the study and then followed for 1 year. They were interested to see which participants elected to receive a vaccination for COVID-19 and whether any participants developed COVID-19.

Explanatory Variable:

Response Variable:

Confounding Variable:

Study design:

What is the scope of inference for this study?

8. What is a potential confounding variable for the study in question 7? Explain how this meets the definition of a confounding variable.