

Start class by Reading:  
p. 278 about Data Ethics

We'll start class today  
by reading about Data Ethics

*because one of the goals today is:*  
**Evaluate if a statistical study has been carried out in  
ethical manner.**

page 278

- All planned studies must be reviewed in advance by an *institutional review board* charged with protecting the safety and well-being of the subjects.
- All individuals who are subjects in a study must give their *informed consent* before data are collected.
- All individual data must be kept *confidential*. Only statistical summaries for groups of subjects may be made public.

## The Challenges of Establishing Causation

---

There are several criteria for establishing causation when we can't do an experiment:

After all, you can't randomly assign people to smoke.



There are several criteria for establishing causation when we can't do an experiment:

- **The association is strong.**



There are several criteria for establishing causation when we can't do an experiment:

- **The association is strong.**
- **The association is consistent.**

- Many studies
- many countries



There are several criteria for establishing causation when we can't do an experiment:

- **The association is strong.**
- **The association is consistent.**
- **Larger values of the explanatory variable are associated with stronger responses.**

People who smoke more cigarettes per day get lung cancer more often



There are several criteria for establishing causation when we can't do an experiment:

- **The association is strong.**
- **The association is consistent.**
- **Larger values of the explanatory variable are associated with stronger responses.**
- **The alleged cause precedes the effect in time.**

Lung cancer develops after the smoking.



There are several criteria for establishing causation when we can't do an experiment:

- **The association is strong.**
- **The association is consistent.**
- **Larger values of the explanatory variable are associated with stronger responses.**
- **The alleged cause precedes the effect in time.**
- **The alleged cause is plausible.**



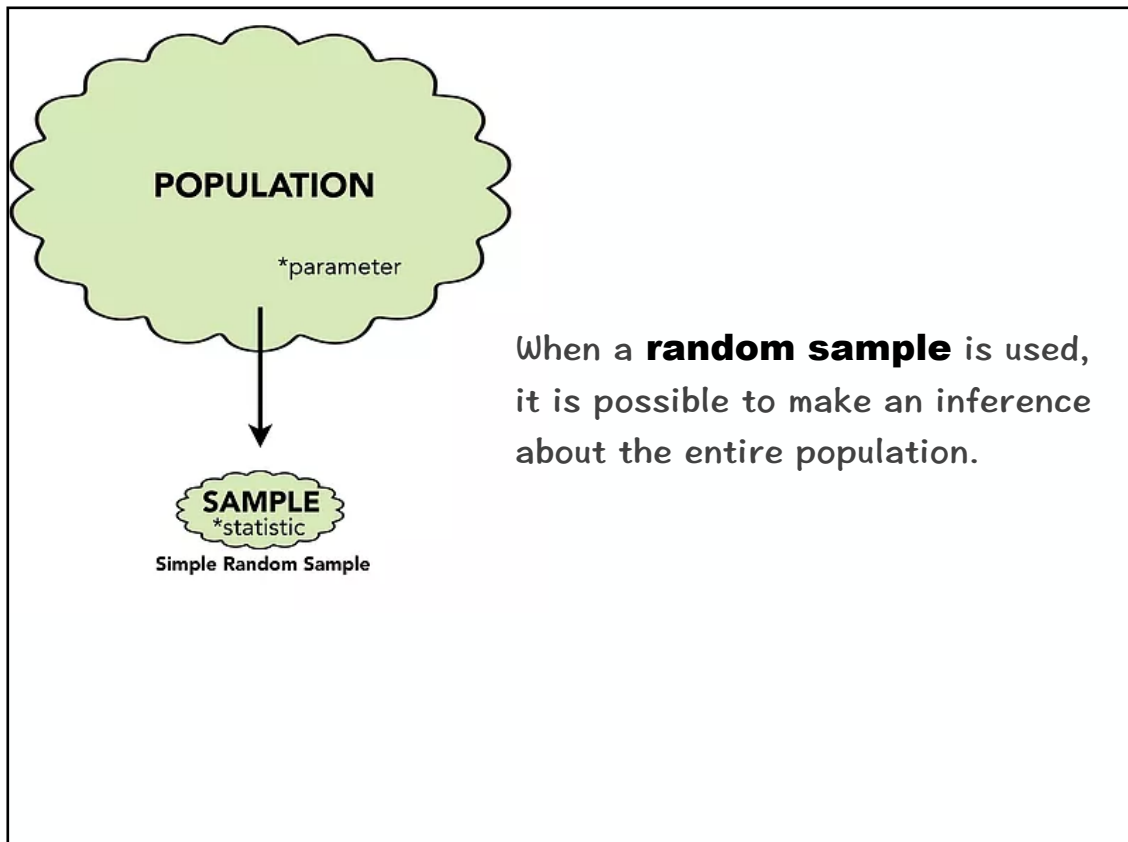
## The Scope of Inference Putting it All Together (pages 275–277)

*Last learning target*

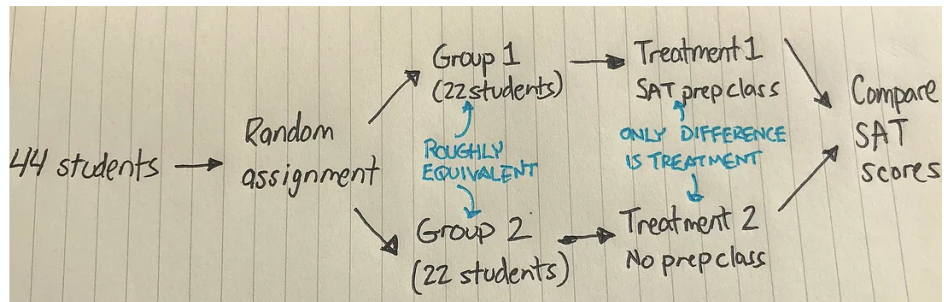
**When is it OK to make an inference about a larger population?**

**When is it OK to make an inference about cause and effect?**

# First some reminders



We use **random assignment** in an experiment to create two groups that are roughly equivalent, so that if there is a difference in the response variable at the end of the experiment, we can say the treatment caused the change.



Pick Up  
the  
Handout

**Scope of inference**

The way in which data are produced determines the types of conclusions we can make.

- Random sampling allows us to make an inference about  
\_\_\_\_\_.
- Random assignment allows us to make an inference about  
\_\_\_\_\_.

**Scope of inference**

The way in which data are produced determines the types of conclusions we can make.

- Random sampling allows us to make an inference about  
the entire population.
- Random assignment allows us to make an inference about  
Cause & Effect.



		<b>Were individuals randomly assigned to groups?</b>	
		Yes	No
<b>Were individuals randomly selected from a population?</b>	Yes	Inferences about the population: <u>Yes</u> Inferences about cause and effect: <u>Yes</u> <i>(Rare in the real world)</i>	Inferences about the population: <u>Yes</u> Inferences about cause and effect: <u>NO</u> <i>(Some observational studies)</i>
	No	Inferences about the population: <u>NO</u> Inferences about cause and effect: <u>Yes</u> <i>(Most experiments)</i>	Inferences about the population: <u>NO</u> Inferences about cause and effect: <u>NO</u> <i>(Some observational studies)</i>

		<b>Were individuals randomly assigned to groups?</b>	
		Yes	No
<b>Were individuals randomly selected from a population?</b>	Yes	Inferences about the population: <u>Yes</u> Inferences about cause and effect: <u>Yes</u> <i>(Rare in the real world)</i>	Inferences about the population: <u>Yes</u> Inferences about cause and effect: <u>NO</u> <i>(Some observational studies)</i>
	No	Inferences about the population: <u>NO</u> Inferences about cause and effect: <u>Yes</u> <i>(Most experiments)</i>	Inferences about the population: <u>NO</u> Inferences about cause and effect: <u>NO</u> <i>(Some observational studies)</i>

## Does listening to music improve GPA?



Some students at your school claim that listening to music while studying will help improve their GPA. Design a study to help discover if this claim is true.

Here are four proposed studies for investigating the question of the day. Suppose we found that the mean GPA of students who listen to music is significantly lower than the mean GPA of students who didn't listen to music. What conclusions could we make? Can we generalize and can we determine causation?

- ① Start by answering all 8 of the yes/no questions
- ② Then come back and write conclusions.

1. Get all the students in your statistics class to participate in a study. Ask them whether or not they study with music on and divide them into two groups based on their answer to this question.

Random sample? No      Random assignment? No

Conclusion:

2. Select a random sample of students from your school to participate in a study. Ask them whether or not they study with music on and divide them into two groups based on their answer to this question.

Random sample? Yes      Random assignment? No

Conclusion:

1. Get all the students in your statistics class to participate in a study. Ask them whether or not they study with music on and divide them into two groups based on their answer to this question.

Random sample? NO Random assignment? NO

Conclusion:

2. Select a random sample of students from your school to participate in a study. Ask them whether or not they study with music on and divide them into two groups based on their answer to this question.

Random sample? Yes Random assignment? NO

Conclusion:

1. Get all the students in your statistics class to participate in a study. Ask them whether or not they study with music on and divide them into two groups based on their answer to this question.

Random sample? NO Random assignment? NO

Conclusion:  
No inference  
about popul.  
or causation

For the students in this class, there is a correlation between listening to music while studying and lower GPA.

2. Select a random sample of students from your school to participate in a study. Ask them whether or not they study with music on and divide them into two groups based on their answer to this question.

Random sample? Yes Random assignment? NO

Conclusion:  
Inference  
about  
Population  
only

For all students from your school there is a correlation between listening to music while studying and lower GPA.

3. Get all the students in your statistics class to participate in a study. Randomly assign half of the students to listen to music while studying for the entire semester and have the remaining half abstain from listening to music while studying.

Random sample? NO      Random assignment? Yes

Conclusion:

4. Select a random sample of students from your school to participate in a study. Randomly assign half of the students to listen to music while studying for the entire semester and have the remaining half abstain from listening to music while studying.

Random sample? Yes      Random assignment? Yes

Conclusion:

3. Get all the students in your statistics class to participate in a study. Randomly assign half of the students to listen to music while studying for the entire semester and have the remaining half abstain from listening to music while studying.

Random sample? NO      Random assignment? Yes

Conclusion:

4. Select a random sample of students from your school to participate in a study. Randomly assign half of the students to listen to music while studying for the entire semester and have the remaining half abstain from listening to music while studying.

Random sample? Yes      Random assignment? Yes

Conclusion:

3. Get all the students in your statistics class to participate in a study. Randomly assign half of the students to listen to music while studying for the entire semester and have the remaining half abstain from listening to music while studying.

Random sample? NO Random assignment? Yes

Inference about CAUSATION

Conclusion: For students in the stat class, listening to music while studying causes lower GPA.

4. Select a random sample of students from your school to participate in a study. Randomly assign half of the students to listen to music while studying for the entire semester and have the remaining half abstain from listening to music while studying.

Random sample? Yes Random assignment? Yes

Inference about Pop. & causation

Conclusion: For all students in the school, listening to music while studying caused lower GPA.

The Scope of Inference

Big Ideas:

**INference**

Random Sample →  
Allows generalizing of whole population

Random Assignment →  
Allows us to make Cause and Effect conclusions

## The Scope of Inference

Big Ideas:

## INference

Random Sample →  
Allows generalizing of  
whole population

Random Assignment →  
Allows us to make  
Cause and Effect  
conclusions

## Data Ethics

- ① Institutional Review
- ② Informed Consent
- ③ Confidentiality

Check your  
Understanding

1. When an athlete suffers a sports-related concussion, does it help to remove the athlete from play immediately? Researchers recruited 95 athletes seeking care for a sports-related concussion at a medical clinic and followed their progress during recovery. Researchers found statistically significant evidence that athletes who were removed from play immediately recovered more quickly, on average, than athletes who continued to play.

(a) What conclusion can we draw from this study? Explain your answer.

(b) Would it be ethical to conduct an experiment to answer this question? Explain.

1. When an athlete suffers a sports-related concussion, does it help to remove the athlete from play immediately? Researchers recruited 95 athletes seeking care for a sports-related concussion at a medical clinic and followed their progress during recovery. Researchers found statistically significant evidence that athletes who were removed from play immediately recovered more quickly, on average, than athletes who continued to play.

(a) What conclusion can we draw from this study? Explain your answer.

NO Random Assig. NOT Rand. Sample  
 We can't make inference about the pop. or cause and effect. All we can say is that athletes in this study recovered more quickly if removed from play.

(b) Would it be ethical to conduct an experiment to answer this question? Explain.

No. You can't assign athletes to continue playing after they're injured.

2. Can eating nuts during pregnancy help children avoid nut allergies? Researchers studied over 8000 children who were born in the early 1990s to mothers who were part of the Nurses' Health Study II. Children whose mothers ate the most nuts during pregnancy (at least five times per week) were significantly less likely to develop nut allergies than children whose mothers ate the least amount of nuts during pregnancy (less than once per month).

(a) Does this study show that eating nuts during pregnancy causes a reduced risk of nut allergies in children?

(b) Would it be ethical to conduct an experiment to answer this question? Explain.

2. Can eating nuts during pregnancy help children avoid nut allergies? Researchers studied over 8000 children who were born in the early 1990s to mothers who were part of the Nurses' Health Study II. Children whose mothers ate the most nuts during pregnancy (at least five times per week) were significantly less likely to develop nut allergies than children whose mothers ate the least amount of nuts during pregnancy (less than once per month).

(a) Does this study show that eating nuts during pregnancy causes a reduced risk of nut allergies in children? No, there was no random assignment.

(b) Would it be ethical to conduct an experiment to answer this question? Explain.

No. the babies are receiving treatment and they can't give informed consent.



## 4.3....103, 105, 107, 117-118

(Optional..... 109, 111, 113, 115....Optional because Data  
Ethnics is not a required topic for AP Statistics.)

### Thoughts

- You may want to start doing some Review for Tuesday's Test,
- Flashcards on the Student site.
- Chapter Review Problems.