We'll start class today

by reading about Data Ethics

because one of the goals today is:

**Evaluate if a statisical 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.
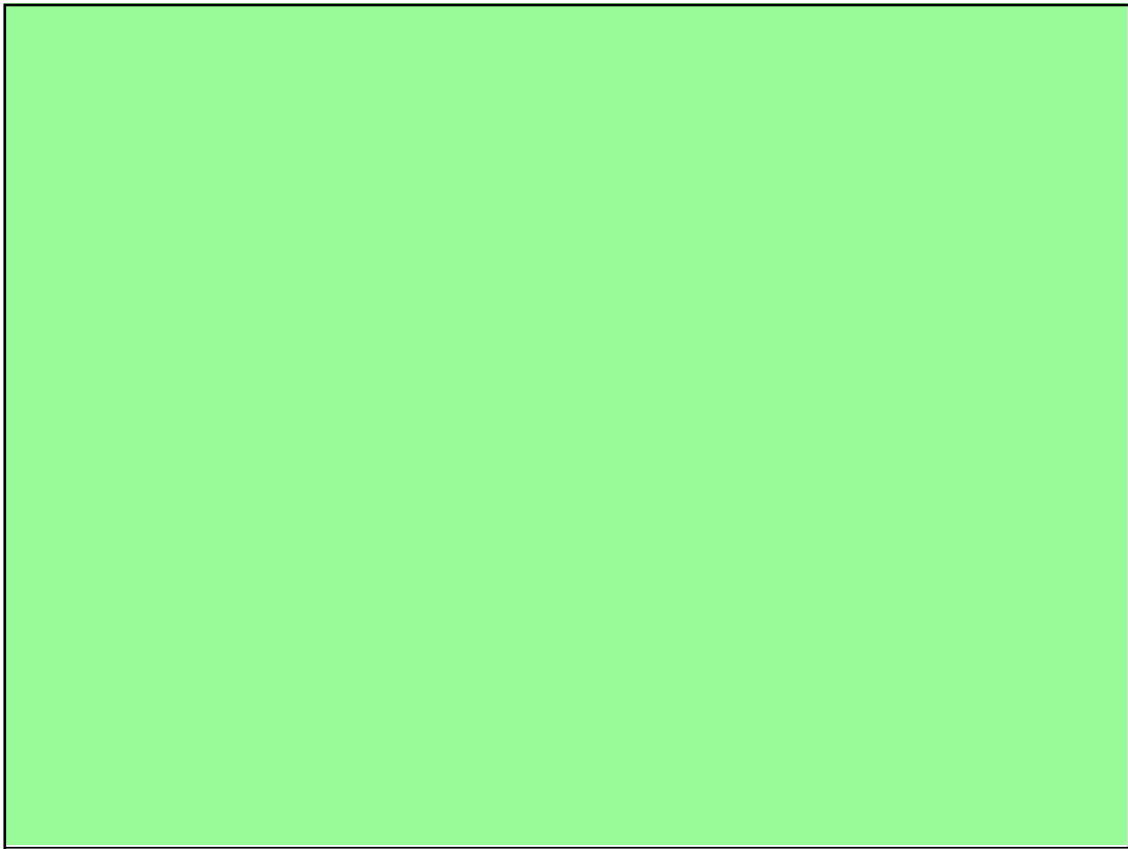
The Challenges of Establishing Causation

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

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.

① The association must be strong.

① The association must be strong.

② The association is <u>consistent</u>.

- many studies
- many countries

---

③ Larger values of the explanatory variable are associated with stronger respones.

people who smoke more cigarettes per day or who smoke over a longer period get lung cancer more often. people who stop smoking reduce their risk.

② The association is <u>consistent</u>.

③ Larger values of the explanatory variable are associated with stronger respones.

④ The alleged "cause" precedes the effect in time. •

Lung cancer develops over years of smoking. The number of men dying of lung cancer rose as smoking became more common, with a lag of 30 years

---

① The association must be strong. •

② The association is <u>consistent</u>.

③ Larger values of the explanatory variable are associated with stronger respones.

④ The alleged "cause" precedes the effect in time. •

⑤ the alleged "cause" is plausible

cant try to prove eating cheese will cure blindness
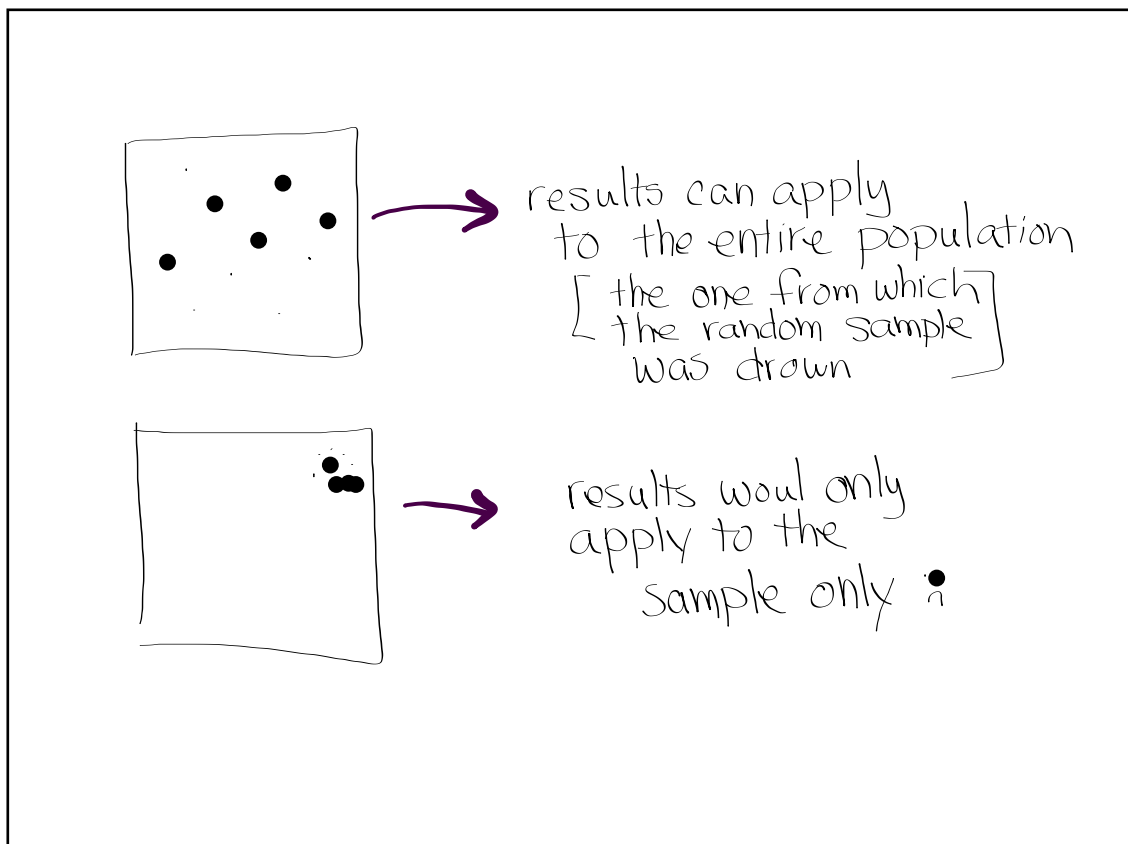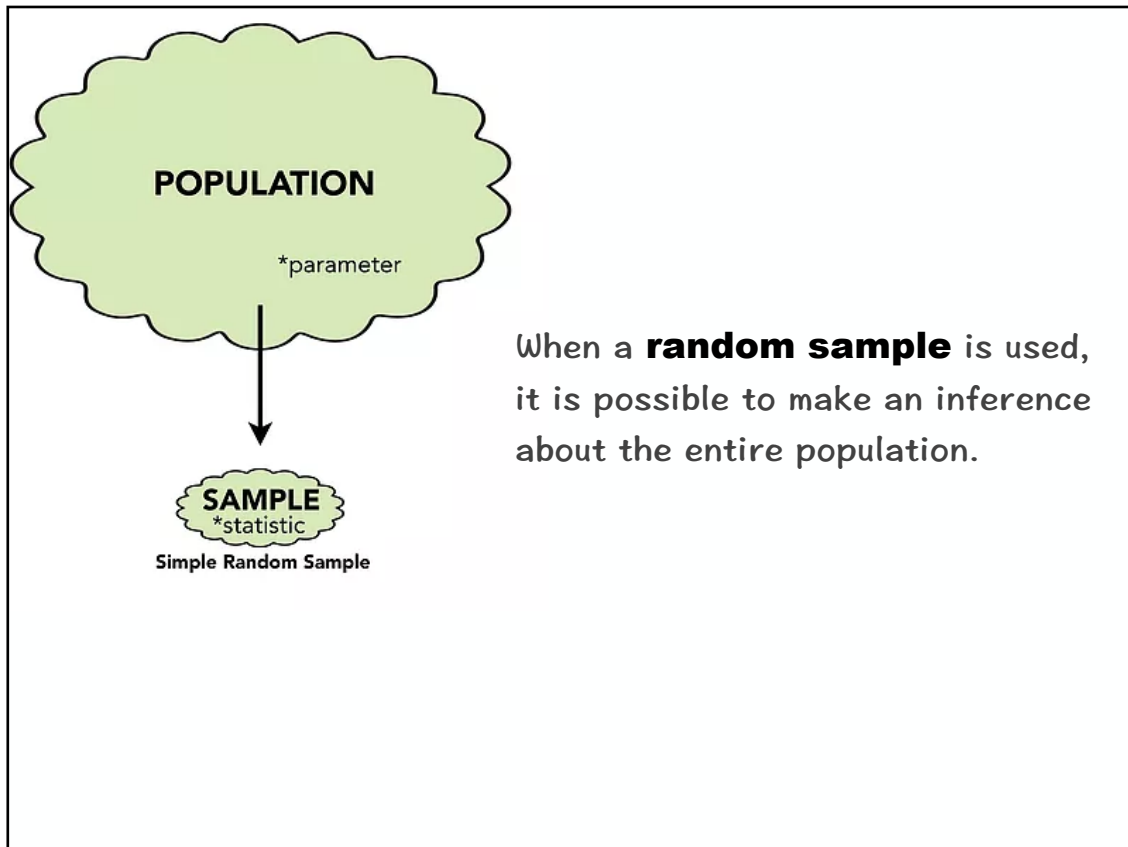
The Scope of Inference
## Putting it All Together
(pages 275–277)

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

POPULATION

*parameter

SAMPLE
*statistic

Simple Random Sample

When a **random sample** is used, it is possible to make an inference about the entire population.

---

results can apply to the entire population [the one from which the random sample was drawn]

results woul only apply to the sample only :

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 <u>treatment caused the change</u>.
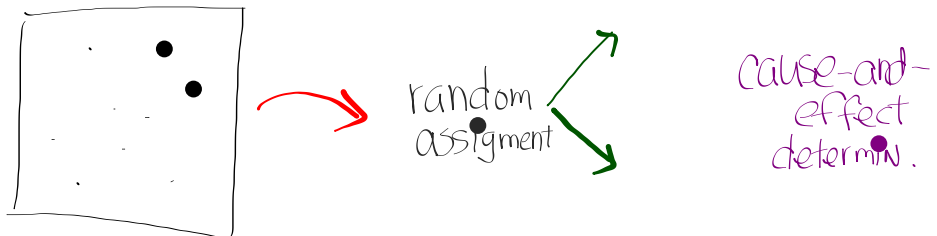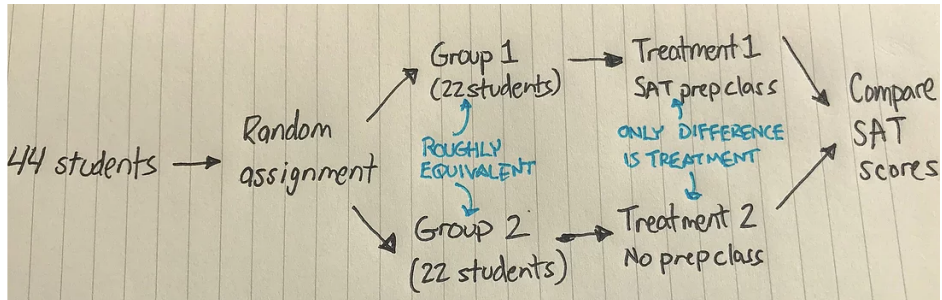
Pick Up
The
Handout

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

  _____.

		Were individuals randomly assigned to groups?	
		Yes	No
Were individuals randomly selected from a population?	Yes	Inferences about the population: _Yes_ Inferences about cause and effect: _Yes_ *(Rare in the real world)*	Inferences about the population: _Y_ Inferences about cause and effect: _N_ *(Some observational studies)*
	No	Inferences about the population: _N_ Inferences about cause and effect: _Yes_ *(Most experiments)*	Inferences about the population: _N_ Inferences about cause and effect: _N_ *(Some observational studies)*

		Were individuals randomly assigned to groups?	
		Yes	No
Were individuals randomly selected from a population?	Yes	Inferences about the population: _Yes_ Inferences about cause and effect: _Yes_ *(Rare in the real world)*	Inferences about the population: _Yes_ Inferences about cause and effect: _NO_ *(Some observational studies)*
	No	Inferences about the population: _NO_ Inferences about cause and effect: _Yes_ *(Most experiments)*	Inferences about the population: _NO_ Inferences about cause and effect: _NO_ *(Some observational studies)*

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?|_____   Random assignment?_____

   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?_____   Random assignment?_____

   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

No inference about popul. or causation

Conclusion: 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:

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? _____   Random assignment? _____

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? _____   Random assignment? _____

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__

Conclusion:

Inference about CAUSATION

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__

Conclusion:

Inference about Pop. & Causation

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

The Scope of Inference

Big Ideas:
INFerence

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

Data Ethics

Big Ideas:
INFERENCE

Random Sample →
Allows generalizing of
whole Population

Random Assignment →
Allows us to make
Cause and Effect
conclusions

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.

   NOT Rand. Sample

   NO Random Assig.

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

   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 <u>amount</u> 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**

and p.269....91

(Optional..... 109, 111, 113, 115....Optional because Data

Ethnics is not a required topic for AP Statistics.)

(finish PPC's (Personal Project check)