Remember that I will post these slides on Moodle later.

CH. 1 – Introduction

PY 221 Statistics & Research Methods I

Dr. Valenti

Overview for Today

1. Questions about the syllabus or anything about the course.
2. Romantic relationships survey data: *Were your hypotheses supported*?
3. Recap: statistics, research process
4. Continue content for Ch. 1

Outline for Ch. 1

1. What are statistics?
   - descriptive
   - inferential
2. The research process
3. Populations and samples
4. Types of designs
5. Types of variables

"Homework" for today

1. Please read the entire syllabus.
2. Look at your schedule for the semester and identify 4 hours per week to devote to this course, outside of class time.
3. Jot down Qs you have @ syllabus or anything from today.
4. Ask friends to complete a quick survey by scanning the QR code/ clicking link on p. 13 of syllabus. Survey is anonymous & they may skip Qs.
5. Explore our Moodle site.
6. Textbook – link to PDF is in syllabus. Read Chapter 1, and note p. 13 of syllabus where I suggest sections to leave out.
7. Obtain all the other required resources (and possibly some or all of the recommended resources).

Overview for Today

1. Questions about the syllabus or anything about the course.
2. Romantic relationships survey data: *Were your hypotheses supported*?
3. Recap: statistics, research process
4. Continue content for Ch. 1

Outline for Ch. 1

1. What are statistics?
   - descriptive
   - inferential
2. The research process
3. Populations and samples
4. Types of designs
5. Types of variables

Survey questions

- It is appropriate to go to your date's room or apartment on the first date.
  (answered on a *disagree* to *agree* scale)
- Are you in a serious romantic relationship right now?
- At what age, if ever, would you want to get married?
- How many total children, if any, would you want to have?
- What is the length of the longest romantic relationship you have ever been in?
- With which gender identity do you most identify?
- How religious are you?
- What range does your current overall GPA fall into?
- How career-motivated are you?

Hypotheses about Relationships, Careers, Etc. Among College Students

N = ~139 students in our sample

1. As religiosity increases, the age at which a person wants to get married decreases.
2. Higher career motivation is associated with a desire for fewer children.
3. Higher career motivation is associated with a higher age at which people want to get married.
4. People with higher GPAs also tend to have higher career motivation.
5. People who are more religious will agree less with the idea that it's appropriate to go to your date's room/apartment on the first date.

First, let's look at an example of a descriptive statistic....

I ran a **FREQUENCY** on the gender identity variable for our sample.

Frequencies *summarize* the data, telling us how many people chose each response.

Therefore, running a frequency table is considered computing a **descriptive** statistic.

gender

		Frequency	Percent
Valid	1 Male	21	14.7
	2 Female	109	76.2
	3 Transgender male	3	2.1
	5 Not listed (please specify if you wish)	2	1.4
	6 Non-binary	4	2.8
	Total	139	97.2
Missing	System	4	2.8
Total		143	100.0

Next I'll test the hypotheses by computing some inferential statistics…

#1 - As religiosity increases, the age at which a person wants to get married decreases.

#5- People who are more religious will agree less with the idea that it's appropriate to go to your date's room/apartment on the first date.

Correlations

		relig	agemarr	roomapt
relig	Pearson Correlation	1	-.431**	-.304**
	Sig. (2-tailed)		.000	.000
	N	139	139	138
agemarr	Pearson Correlation	-.431**	1	.092
	Sig. (2-tailed)	.000		.283
	N	139	139	138
roomapt	Pearson Correlation	-.304**	.092	1
	Sig. (2-tailed)	.000	.283	
	N	138	138	138

**. Correlation is significant at the 0.01 level (2-tailed).

To see if the data support our hypotheses, I ran a CORRELATION, which is an *inferential* statistical test.
The results show that both hypotheses were supported by the data.

Hypotheses:

2. Higher career motivation is associated with a desire for fewer children.
3. Higher career motivation is associated with a higher age at which people want to get married.

I ran another correlation, this time using the variables related to career motivation, number of desired children, and predicted age of marriage.

Correlations

		career	agemarr	numchild
career	Pearson Correlation	1	-.027	.095
	Sig. (2-tailed)		.752	.265
	N	139	139	139
agemarr	Pearson Correlation	-.027	1	-.338**
	Sig. (2-tailed)	.752		.000
	N	139	139	139
numchild	Pearson Correlation	.095	-.338**	1
	Sig. (2-tailed)	.265	.000	
	N	139	139	139

**. Correlation is significant at the 0.01 level (2-tailed).

Here, neither hypothesis was supported by the data.

Hypothesis:

#4 - People with higher GPAs also tend to have higher career motivation.

I ran an ANALYSIS OF VARIANCE (aka, an ANOVA), predicting career motivation from GPA.

ANOVA

career

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	6.391	2	3.195	4.880	.009
Within Groups	87.741	134	.655		
Total	94.131	136			

Descriptives

career

	N	Mean	Std. Deviation	Std. Error
1 Lower than 2.70 (C+ or lower)	8	3.75	.886	.313
2 Between 2.70 & 3.33 (between B- and B+)	55	3.64	.825	.111
3 Higher than 3.33 (higher than B+)	74	4.08	.790	.092
Total	137	3.88	.832	.071

Here, the results show there are some significant differences across the three GPA levels, in terms of students' career motivation.
We'd need to run a follow-up analysis to know exactly which groups differ significantly on career motivation, & if our hypothesis was supported.

Overview for Today

1. Questions about the syllabus or anything about the course.
2. Romantic relationships survey data: *Were your hypotheses supported*?
3. Recap: statistics, research process
4. Continue content for Ch. 1

Outline for Ch. 1

1. What are statistics?
   - descriptive
   - inferential
2. The research process
3. Populations and samples
4. Types of designs
5. Types of variables

PY 221 Pro Tip: Even though most of your assessments are open note, open book, *memorizing* terminology will help you process the lecture content more easily.

i.e., will prevent your brain from exploding during class (or, perhaps more likely, prevent you from tuning out)

REVIEW

1. What are the two types of “statistics”? How would you describe each type of statistics?

2. What are some reasons that psychology majors and those with careers in psychology learn statistics?

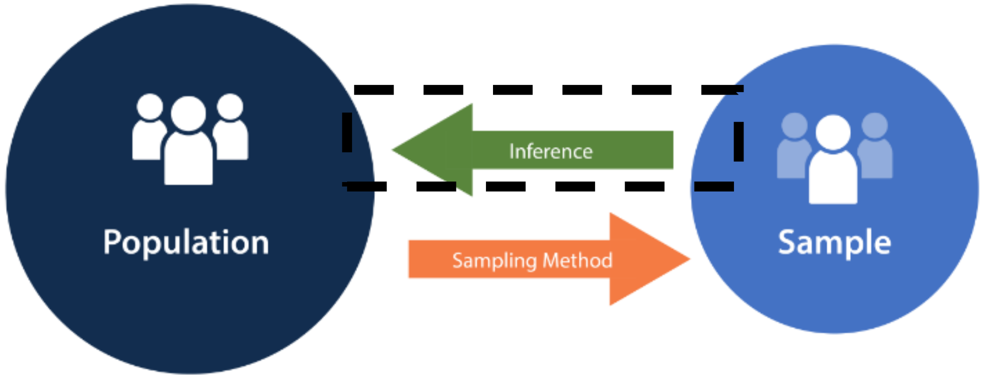
1. What is "statistics"?

descriptive statistics

1. ***Numbers*** created to summarize data, like a mean, sum, or standard deviation
   - E.g., The average PY 221 class at BSC has **18** students in it.

inferential statistics

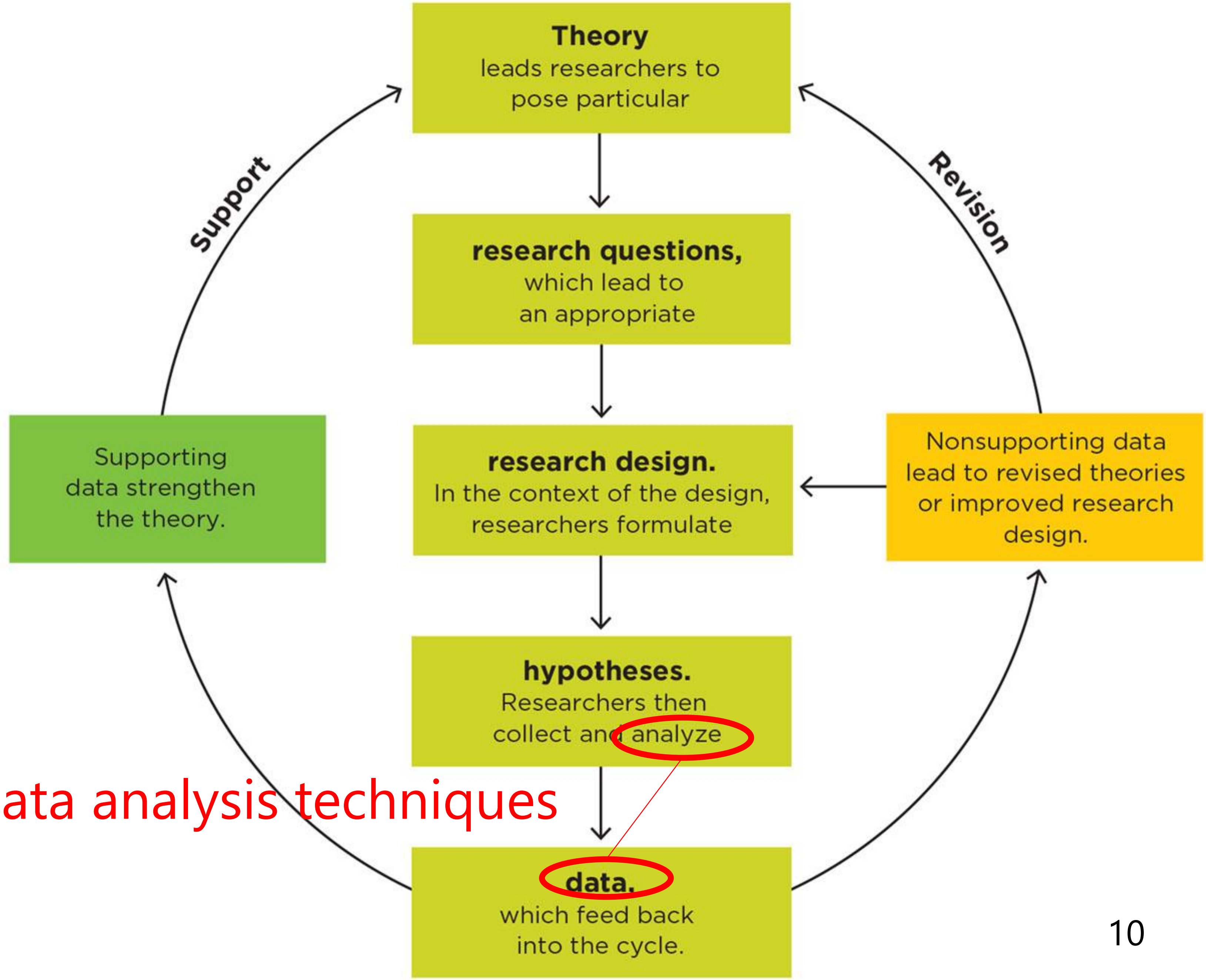
2. ***Techniques*** and ***procedures*** for drawing conclusions about a population from a sample
   - E.g., In general, do people with lower GPAs have less career motivation?



2. Why should psychology majors learn statistics?

- To **analyze** own **data** collected as an undergrad (grad student, professor)
- To **understand published work** relevant to what they do (e.g., therapists reviewing research on treatments)
- For **business** or **non-profit** world – e.g., marketing research
- For everyday life **– claims hear on tv, read about** – know what questions to ask to evaluate that research.

the research process



The focus of PY 221 is data analysis techniques

Overview for Today

1. Questions about the syllabus or anything about the course.
2. Romantic relationships survey data: *Were your hypotheses supported*?
3. Recap: statistics, research process
4. Continue content for Ch. 1

Outline for Ch. 1

1. What are statistics?
   - descriptive
   - inferential
2. The research process
3. Populations and samples
4. Types of designs
5. Types of variables

Outline for Ch. 1

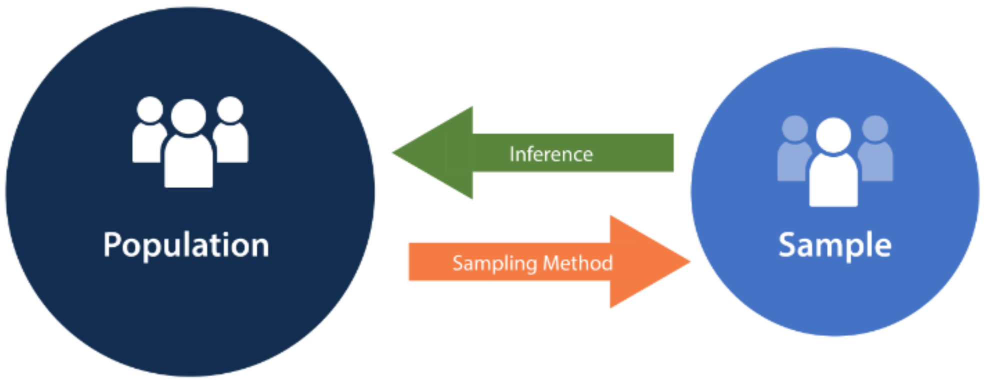
1. What are statistics?
   - descriptive
   - inferential
2. The research process
3. **Populations and samples**
4. Types of designs
5. Types of variables

Populations and Samples – definitions

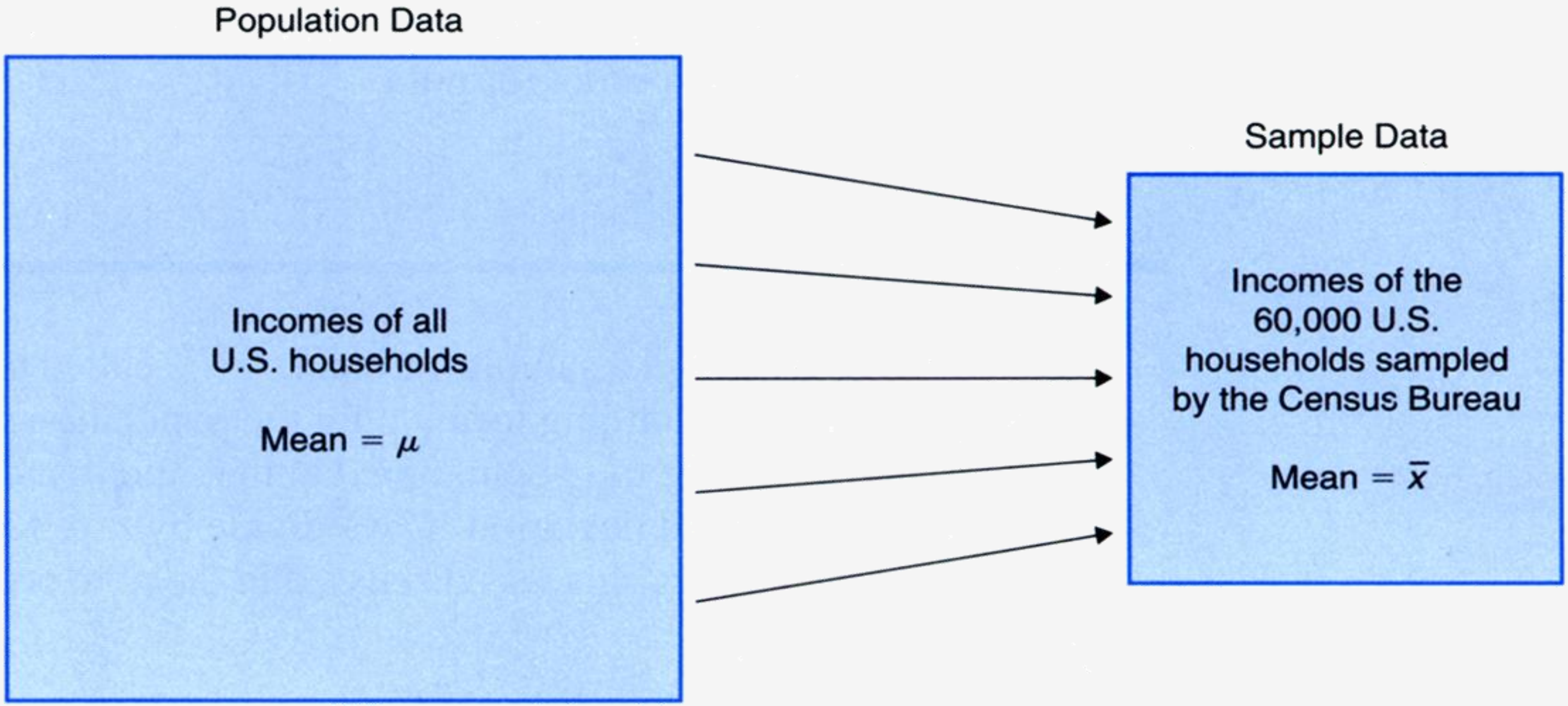
- Population
  - The entire collection of units to which we want to generalize a set of findings.

In psychology, these "units" are often *people.*

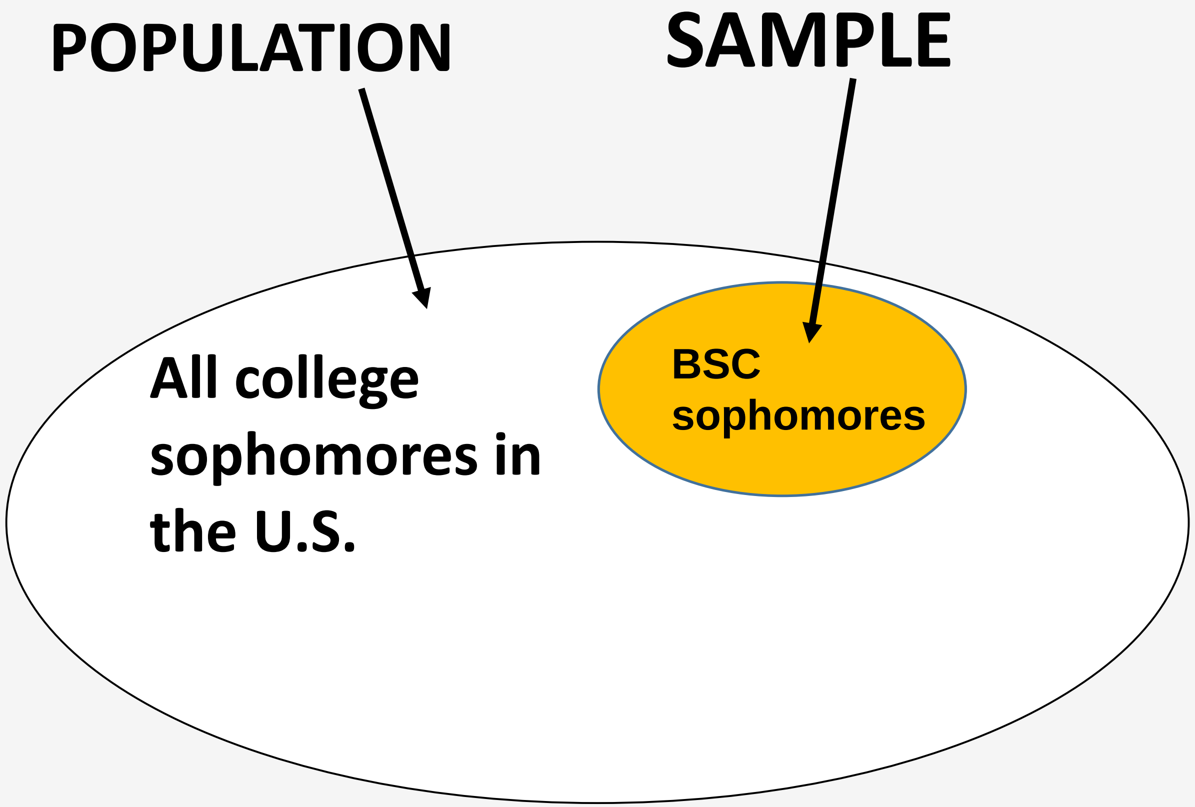
- Sample
  - A smaller collection of units, drawn from a population, used to determine truths about that population.



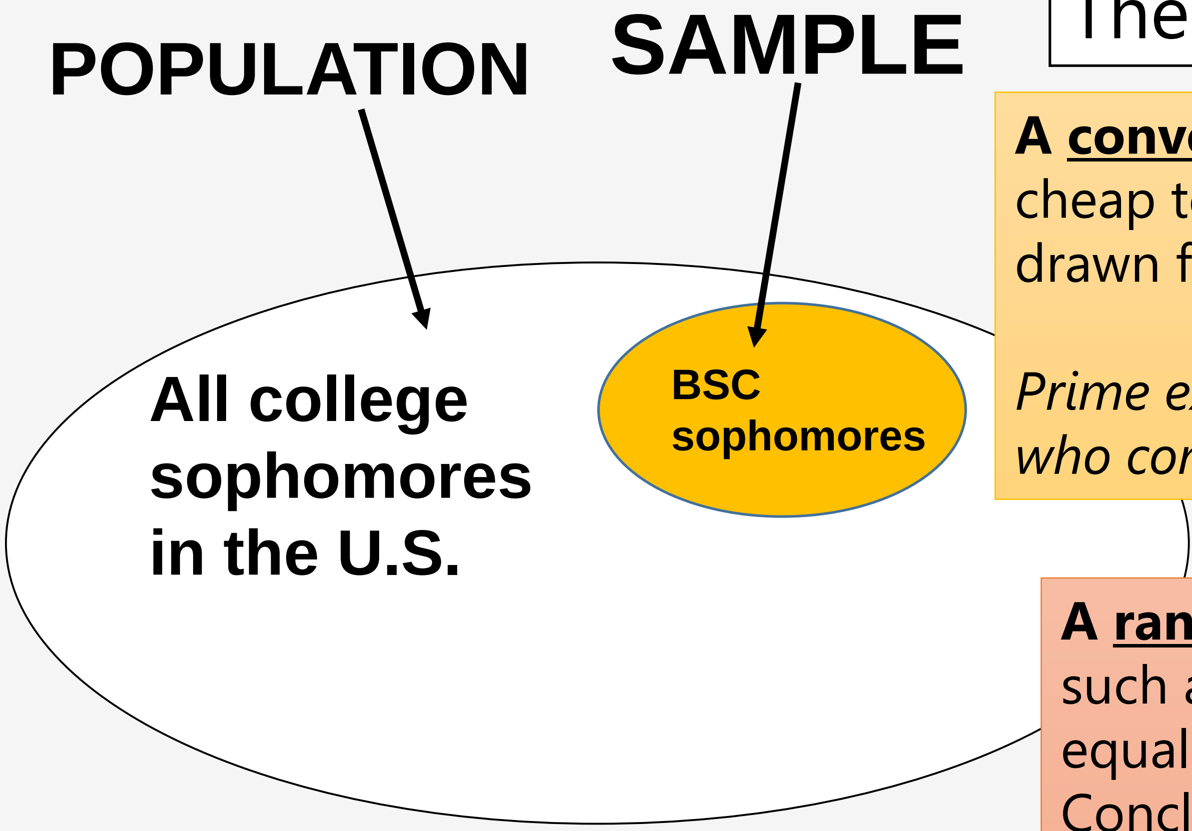
Two examples



Two examples



Two examples



There are different methods for sampling.

A <u>convenience</u> sample: a sample that is easy, quick, and cheap to recruit; Ps (i.e., Participants) are *not* randomly drawn from the broader population.

Prime example of convenience sample: BSC PY 101 students who complete ERO credit.

A <u>random</u> sample: a sample drawn from a population in such a way that all members of the population have an equal chance of being selected for the sample. Conclusions from the data from these samples are more likely to be *representative* of the population.

<u>(Possibly Surprising) Note</u>: Psychologists almost never use random samples!

Populations and Samples – important facts

- **Even with a large, random sample, our sample statistics (e.g., the sample *mean*) will always differ, at least slightly, from our population statistics (e.g., the population mean).**

- **Even if we pull several large, random samples from the same population, each sample will have slightly different sample statistics (e.g., slightly different sample means).**

Parameters and Sample Statistics

- Population: The entire collection of units to which we want to generalize a set of findings.
  - EX: All college sophomores
- Sample: A smaller collection of units, drawn from a population, used to determine truths about that population.
  - EX: BSC sophomores

- **Parameters**: any descriptive statistics calculated using *population* data.
- **Sample statistics**: any descriptive statistics calculated using *sample* data.

μ and σ are examples of **parameters.**

population mean = μ ("mu")
population standard deviation = σ ("sigma")

sample mean = $\overline{x}$
sample standard deviation = *s*

$\overline{x}$ and *s* are examples of **sample statistics.**

(aka, parameter estimates)

Review: ***Descriptive statistics*** are ***numbers*** created to summarize data, like a mean or a standard deviation.

Populations and Samples, revisited

μ and σ are **parameters**

population mean = μ ("mu")
population standard deviation = σ ("sigma")

- **Even with a large, random sample, our sample statistics (i.e., the parameter estimates) will always differ, at least slightly, from the actual values of the parameters.**

- **Even if we pull several large, random samples from the same population, each sample will have slightly different sample statistics (i.e., parameter estimates).**

sample mean = $\bar{x}$
sample standard deviation = s

$\bar{x}$ and s are **sample statistics**

(aka, parameter estimates)

Check your understanding

Suppose that a person sampled 50,000 U.S. households and asked whether any adult in the household had ever been diagnosed with an anxiety disorder. 20% of households said yes.

"20%" is a descriptive statistic, and also an example of a:

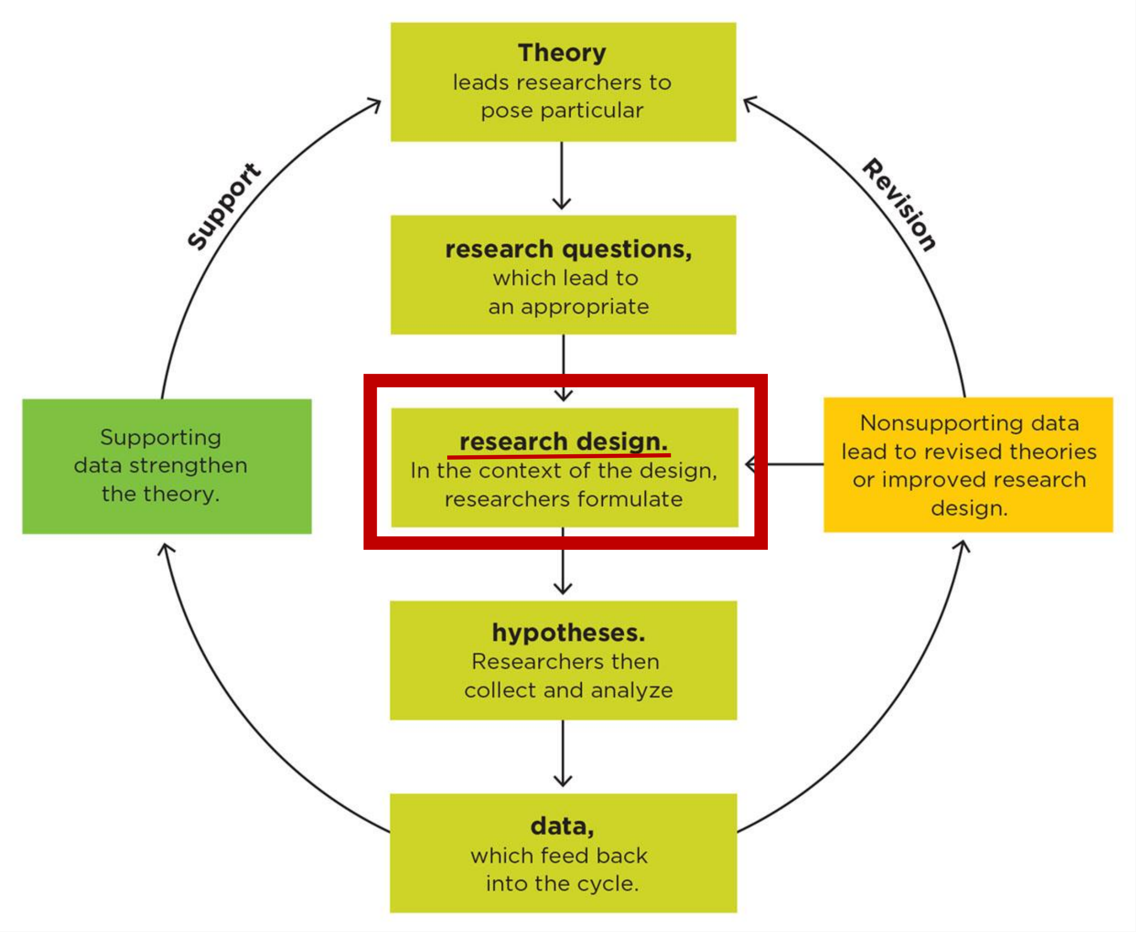
A. Parameter estimate
B. Parameter
C. Sample statistic
D. A and C
E. B and C

Answer: D. Because the 20% came from data from a sample, not from data from the entire population, that value is a *sample statistic*, aka a *parameter estimate*.

Outline for Ch. 1

1. What are statistics?
   - descriptive
   - inferential
2. The research process
3. Populations and samples
4. **Types of designs**
5. Types of variables

What do you remember about *research designs*?



- What are the two most common research designs in psychology?

- What do you know about these designs?

EX – You're interested in 2 variables: *exercise* & *life expectancy*

- HYPOTHESIS #1: People who exercise more hours per week also tend to live to a higher age.

Use a **correlational** design for your study

- HYPOTHESIS #2: Exercising more hours per week *causes* people to live to a higher age.

Use an **experimental** design for your study because you are interested in whether one variable *causes* a change in the other.

Practice distinguishing conclusions from correlational studies vs. experimental studies.

- Which of these conclusions came from a **correlational** study, and which came from an **experiment**?

cannot draw causal conclusions

causal relationship is explicitly stated

A study showed that . . .

1. having breast implants (vs. not) is associated with a greater likelihood of suicide.
2. suicides are more common among those who receive breast implants than those who don't.
3. getting breast implants makes a person more likely to commit suicide.
4. having breast implants and committing suicide are positively related; as one goes up, the other goes up.
5. having breast implants causes people to commit suicide.

Sorry for the weird example. It's from a textbook I used to use.

Practice with a weird example

- Which of these findings came from a **correlational** study, and which came from an **experiment**?

A study showed that . . .

1. having breast implants (vs. not) is **associated with** a greater likelihood of suicide.
2. suicides **are more common among** those who receive breast implants than those who don't.
3. getting breast implants makes a person more likely to commit suicide.
4. having breast implants and committing suicide are positively **related**; as one goes up, the other goes up.
5. having breast implants causes people to commit suicide.

Practice with a weird example

Look for words like "causes," "affects," "leads to," "influences," "makes [something] more/less likely".

- Which of these findings came from a **correlational** study, and which came from an **experiment**?

A study showed that . . .

1. having breast implants (vs. not) is associated with a greater likelihood of suicide.
2. suicides are more common among those who receive breast implants than those who don't.
3. getting breast implants **makes** a person more likely to commit suicide.
4. having breast implants and committing suicide are positively related; as one goes up, the other goes up.
5. having breast implants **causes** people to commit suicide.

The "self-graded HW" on Moodle provides additional practice with these concepts.

Practice distinguishing conclusions from correlational studies vs. experimental studies.

- Which of these conclusions came from a **correlational** study, and which came from an **experiment**?

cannot draw causal conclusions

causal relationship is explicitly stated

A study showed that . . .

1. having breast implants (vs. not) is associated with a greater likelihood of suicide.
2. suicides are more common among those who receive breast implants than those who don't.
3. getting breast implants makes a person more likely to commit suicide.
4. having breast implants and committing suicide are positively related; as one goes up, the other goes up.
5. having breast implants causes people to commit suicide.

Sorry for the weird example. It's from a textbook I used to use.

How do we conduct studies that use each design?

1. Correlational research involves…
   - observing how two or more variables naturally relate to one another, without directly interfering with those variables.
     - you ***measure*** all of your variables
     - You might find that your variables are completely uncorrelated (unrelated), or that they are positively correlated (moving in the same direction), or negatively correlated (moving in opposite directions)
2. Experimental research involves …
   - systematically manipulating one or more variable to see its/their effect on an outcome variable.
     - you ***manipulate*** at least one variable (you do this 1st)
     - you ***measure*** at least one variable (you do this 2nd)

Remember that I will post these slides on Moodle later.

Types of research designs

2. Experimental research involves …

- systematically manipulating one or more variable to see its/their effect on an outcome variable.

This variable is called the **independent variable**

- you ***manipulate*** at least one variable

Researcher directly changes it, directly alters it

- you ***measure*** at least one variable

Researcher simply records, assesses, observes it *or*
Participant may be asked to report it

This variable is called the **dependent** variable