Week 1 Reading Guide

## Chapter 1: Data, Types of Variables & Study Design

### Section 1.1

What is a summary statistic?

### Section 1.2.1: Observations, Variables & Data Matrices

What is a case / observational unit?

What is a variable?

What is data?

What is a data frame?

How is a data frame “tidy”?

### Section 1.2.2: Types of Variables

What are examples of numerical / quantitative variables?

What are examples of discrete variables?

What are examples of continuous variables?

What is an example of a variable recorded as a number which **is not** a numerical variable?

What are examples of categorical / qualitative variables?

What are the “levels” of a categorical variable?

What is an ordinal categorical variable?

Give an example of an ordinal variable using the county data:

What is a nominal categorical variable?

Give an example of a nominal variables using the county data:

**Note: Ordinal and nominal variables will be treated the same in this course. I recommend taking more statistics courses in the future to learn better methods of analysis for ordinal variables!**

### Section 1.2.3: Relationships Between Variables

What does it mean when we say there is an association between two variables?

What does it mean when we say two variables are independent?

### Section 1.2.4: Explanatory & Response Variables

What does it mean for a variable to be an “explanatory” variable?

What does it mean for a variable to be an “response” variable?

### Section 1.2.5: Observational Studies & Experiments

What is an experiment?

What is a randomized experiment?

What is a placebo and how is it used in a randomized experiment?

What is an observational study?

How is an observational study different from a randomized experiment?

What is a cohort and how is it used in an observational study?

#### Notes

In a data frame, rows correspond to:

In a data frame, columns correspond to:

True or False: A pair of variables can be both associated AND independent.

True or False: Given a pair of variables, one will always be the explanatory variable and one the response variable.

True or False: If a study does have an explanatory and a response variable, that means changes in the explanatory variable must **cause** changes in the response variable.

True or False: Observational studies can show a naturally occurring association between variables.

## Chapter 2: Sampling & Experimental Design

### Section 2.1.1: Populations & Samples

What is a population?

If a census was performed, how were observations sampled?

How is a sample different from a population?

### Section 2.1.2: Parameters & Statistics

What is a statistic?

What is a parameter?

What is the relationship between a statistic and a parameter?

### Section 2.1.3: Anecdotal Evidence

What is anecdotal evidence?

### Section 2.1.4: Sampling from a Population

What does it mean for something to be biased?

How can a sample be biased?

How are observations sampled when a researcher uses a simple random sample?

What does it mean for a sample to be “representative” of the population?

What does it mean for a study to have non-response bias?

What does it mean for a study to have response bias?

How are observations sampled when a researcher uses a convenience sample?

#### Notes

* Someone must first be *chosen* to be in a study and refuse to participate in order to have **non-response bias**.
* There must be a valid reason for someone to lie or be untruthful to justify saying **response bias** is present. Yes, anyone could lie at any time to any question. Response bias is when those lies are **predictable** and **systematic** based on outside influences. An example would be surveying Cal Poly students and asking them if they have ever driven drunk.

True or False: Convenience sampling tends to result in non-response bias.

True or False: Volunteer sampling tends to result in response bias.

### Section 2.1.5: Four Sampling Methods

How are observations sampled when a researcher uses a stratified sampling method?

What is a strata and how is it used in stratified sampling?

How are observations sampled when a researcher uses a cluster sampling method?

How are clusters made in cluster sampling?

How are clusters similar to and different from strata?

How are observations sampled when a researcher uses a multistage sampling method?

#### Notes

Ideally, how should we sample cases from our target population? Using what sampling method?

True or False: Random sampling helps to resolve selection bias, but has no impact on non-response or response bias.

### Section 2.2.1: Principles of Experimental Design

What does it me for an experiment to control for a specific variable?

What is a control and how is it used in an experiment?

What is randomization and how is it used in an experiment?

What is a confounding variable?

What does it mean for an experiment to have replication?

What is a replicate?

What is blocking and how is it used in an experiment?

What are blocks?

### Section 2.2.2: Reducing Bias in Human Experiments

In an experiment, what is a treatment group?

In an experiment, what is a control group?

What does it mean for an experiment to be blind?

What does it mean for an experiment to be double blind?

What is a placebo effect?

### Section 2.3: Observational Studies

What is a prospective study?

What is a retrospective study?

#### Notes

What are the four principles of a well-designed randomized experiment?

True or False: Observational studies can show an association between two variables, but cannot determine a causal relationship.

True or False: In order for an experiment to be valid, a placebo must be used.

True or False: If random sampling of the target population is used, and no other types of bias is suspected, results from the sample can be generalized to the entire target population.

True or False: If random sampling of the target population is used, and no other types of bias are suspected, results from the sample can be inferred as a causal relationship between the explanatory and response variables.