#### Week 1 Practice Quiz

**TOTAL POINTS 5**

1.

Question 1

Which of the following classifications of variable types is **false**?

**1 / 1 point**



Whether a student has previously taken a statistics course → categorical



Customer satisfaction: very unsatisfied, unsatisfied, satisfied, very satisfied → ordinal categorical



Student height → continuous numerical



Population of each state in the US → continuous numerical

**Correct**

This question refers to the following learning objective(s):

Identify variables as numerical and categorical.

* If variable is numerical, further classify as continuous or discrete based on whether or not the variable can take on an infinite number of values or only non-negative whole numbers, respectively.
* If variable is categorical, determine if it is ordinal based on whether or not the levels have a natural ordering.

Counted data are discrete numerical variables since they can’t take on non-whole values.

2.

Question 2

**True or False:** If subjects are randomly assigned to treatments, conclusions can be generalized to the population.

**1 / 1 point**



True



False

**Correct**

This question refers to the following learning objective(s):

Classify a study as observational or experimental, and determine whether the study’s results can be generalized to the population and whether they suggest correlation or causation.

* If random sampling has been employed in data collection, the results should be generalizable to the target population.
* If random assignment has been employed in study design, the results suggest causality.

Random assignment allows us to make causal conclusions. For generalizability, we need random sampling.

3.

Question 3

As part of a statistics project, Andrea would like to collect data on household size in her city. To do so, she asks each person in her statistics class for the size of their household, and reports that her sample is a simple random sample. However, this is not a simple random sample. Which of the following is the **best** reasoning for why this is not a random sample that is appropriate for this research question?

**1 / 1 point**



Andrea did not block for any variables that might influence the response.



Andrea did not use any randomization; she took a convenience sample.



Andrea did not use a stratified sample.

**Correct**

This question refers to the following learning objective(s):

Distinguish between simple random, stratified, and cluster sampling, and recognize the benefits and drawbacks of choosing one sampling scheme over another.

4.

Question 4

Which of the following is not one of the four principles of experimental design?

**1 / 1 point**



randomize



replicate



stratify

**Correct**

The question refers to the following learning objective: Identify the four principles of experimental design and recognize their purposes: control any possible confounders, randomize into treatment and control groups, replicate by using a sufficiently large sample or repeating the experiment, and block any variables that might influence the response.

5.

Question 5

True or False: Stratified sampling allows for controlling for possible confounders in the sampling stage, while blocking allows for controlling for such variables during random assignment.

**1 / 1 point**



True



False

**Correct**

Stratifying and blocking both allow for controlling for potential confounders, but at different stages of the study design. We stratify when we sample (divide population into strata and sample from within each stratum), and block in the process of random assignment (divide sample into blocks and randomly assign from within each block to treatment groups).

## Suggested Readings and Practice

[OpenIntro Statistics, 3rd edition](https://www.openintro.org/stat/textbook.php?stat_book=os)

**Suggested reading:** Chapter 1, Sections 1.1 - 1.5

**Practice exercises:** End of chapter exercises in Chapter 1: 1.1, 1.3, 1.11, 1.13, 1.17, 1.19, 1.25, 1.27, 1.31

(Reminder: the solutions to the end of chapter exercises are at the end of the OpenIntro Statistics book)

#### Week 1 Quiz

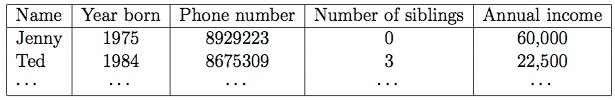
**LATEST SUBMISSION GRADE**

100%

1.

Question 1

Consider the table below describing a data set of individuals who have registered to volunteer at a public school.



Which of the variables are categorical and which are numerical?

**1 / 1 point**



categorical:Name, Phone number, Number of Siblings.

numerical: Year born, Annual income.



categorical: Name, Phone number

numerical: Year born, Number of siblings, Annual income.



All are categorical. None are numerical.



categorical: Name

numerical: Year born, Phone number, Number of siblings, Annual income.

**Correct**

This question refers to the following learning objective(s):

Identify variables as numerical and categorical.

* If variable is numerical, further classify as continuous or discrete based on whether or not the variable can take on an infinite number of values or only non-negative whole numbers, respectively.
* If variable is categorical, determine if it is ordinal based on whether or not the levels have a natural ordering.

2.

Question 2

The General Social Survey conducted annually in the United States asks how many friends people have and how they would rate their happiness level (very happy, pretty happy, not too happy). In order to evaluate the relationship between these two variables a researcher calculates the average number of friends for people who categorize themselves as very happy, pretty happy, and not too happy. Which of the following correctly identifies the variables used in the study as explanatory and response?

**1 / 1 point**



explanatory:very happy, pretty happy, not too happy

response: number of friends



explanatory:number of friends

response: very happy, pretty happy, not too happy



explanatory:number of friends

response: happiness level (categorical with 3 levels)



explanatory:happiness level (categorical with 3 levels)

response: number of friends

**Correct**

This question refers to the following learning objective(s):

Identify the explanatory variable in a pair of variables as the variable suspected of affecting the other, however note that labeling variables as explanatory and response does not guarantee that the relationship between the two is actually causal, even if there is an association identified between the two variables.

Having more friends might cause people to be happier or being happier might cause people to have more friends. So we can’t easily determine which variable is the explanatory and which the response based on which we might expect to affect which. However in this particular analysis the happiness level is the explanatory variable since we first divide the data into groups based on this variable, and then analyze summary statistics of number of friends of people who fall into these three categories. Therefore, number of friends is the response variable. Happiness level is a categorical variable that can take on three possible values (very happy, pretty happy, not too happy). These possible values are called levels.

3.

Question 3

In a study published in 2011 in The Proceedings of the National Academy of Sciences, researchers randomly assigned 120 elderly men and women who volunteered to be a part of this study (average age mid-60s) to one of two exercise groups. One group walked around a track three times a week; the other did a variety of less aerobic exercises, including yoga and resistance training with bands. After a year, brain scans showed that among the walkers, the hippocampus (part of the brain responsible for forming memories) had increased in volume by about 2% on average; in the others, it had declined by about 1.4%. Which of the following is **false**?

**1 / 1 point**



The results of this study can be generalized to all elderly.



A causal link between walking and expansion of the hippocampus can be inferred based on these results.



The explanatory variable is the type of exercise, and the response variable is the change in volume of the hippocampus.

**Correct**

This question refers to the following learning objective(s):

Classify a study as observational or experimental, and determine whether the study’s results can be generalized to the population and whether they suggest correlation or causation.

* If random sampling has been employed in data collection, the results should be generalizable to the target population.
* If random assignment has been employed in study design, the results suggest causality.

Results cannot be generalized to the population of all elderly since random sampling was not employed; the participants volunteered to be a part of the study. This introduces volunteer bias, as those who volunteer for this study may not necessarily constitute a representative sample from the population of all elderly. For example, they may be people who pay more attention to their health compared to average.

4.

Question 4

A school district is considering whether it will no longer allow students to park at school after two recent accidents where students were severely injured. As a first step, they survey parents of high school students by mail, asking them whether or not the parents would object to this policy change. Of 5,799 surveys that go out, 1,209 are returned. Of these 1,209 surveys that were completed, 926 agreed with the policy change and 283 disagreed. Which of the following statements is the **most** plausible?

**1 / 1 point**



The survey is unlikely to have any bias because all parents were mailed a survey.



The school district has strong support from parents to move forward with the policy approval.



It is possible that 80% of the parents of high school students disagree with the policy change.

**Correct**

This question refers to the following learning objective(s):

Question confounding variables and sources of bias in a given study.

It is possible that all who did not return surveys actually disagree with the policy change.

5.

Question 5

As part of a statistics project, Andrea would like to collect data on household size in her city. To do so, she asks each person in her statistics class for the size of their household, and reports that her sample is a simple random sample. However, this is not a simple random sample. Which of the following is the best reasoning for why this is not a random sample that is appropriate for this research question?

**1 / 1 point**



Andrea asked everybody in her class instead of asking her classmates to volunteer.



In this investigation of household size, each household represents a case. Andrea incorrectly sampled individuals instead of households.



Andrea did not use a random number table to randomize the order in which she collected the students’ responses, so the sample cannot be random.



Andrea did not use any randomization; she took a convenience sample.

**Correct**

This question refers to the following learning objective(s):

Distinguish between simple random, stratified, and cluster sampling, and recognize the benefits and drawbacks of choosing one sampling scheme over another.

6.

Question 6

In an experiment, what purpose does blocking serve?

**1 / 1 point**



Control for variables that might influence the response.



Prevent skewed results.



Increase sample size.



Obtain a random sample.

**Correct**

If there are variables that are known or suspected to affect the response variable, we first group subjects into blocks based on these variables, and then randomize cases within each block to treatment groups. This allows us to control for possible effects of these confounding variables by making sure they are equally represented in all treatment groups.

7.

Question 7

Which of the following is one of the four principles of experimental design?

**1 / 1 point**



randomize



stratify



cluster

**Correct**

Please refer to the section on "Supplementary R resources" under <https://www.coursera.org/learn/probability-intro/resources/crMc4> before choosing an option.