

1. What is an example of statistics in the real world? Name a job or process that uses statistics and why it is important – this can be one we talked about in class or one you can think of on your own.

Statistics are used in baseball by general managers to make teams better. A good example of this is seen in moneyball where they try to maximize runs scored with a low payroll.

2. Last semester, there were 1,150 Statistics 201 students. We wanted to approximate the average height of a Statistics 201 student so we looked at 40 students and measured their height. The experiment showed that the average height of the 40 students was 165cm.

The experimenter found that the mandatory physicals, that students get when they start school, include each student's height and that this information was stored in the registrar's database. The experimenter requested this data to see how close the average of the 40 students was to the average of 1,150 students. The average of all of the students was 172 cm, which is only 7cm higher.

In this example identify the following:

**Population:** All 1,150 students in Stat 201

**Sample:** 40 Stat 201 students measured

**Variable:** Height (cm)

**Statistic (give numerical value):** 165 cm

**Parameter (give numerical value):** 172 cm

3. **Types of studies:**

- a. Time Magazine published an article about a study that recorded certain attributes of citizens over fifty and their communities. It was seen that the prevalence of fast-food restaurants and the amount of violence increase the heart-attack risk and the social cohesion, or how well neighbors get along, actually provides cardiovascular benefits. What type of study is this and why?

This is an observational study because there was no attempt to influence the cardiovascular health of the study's participants.

- b. A study, designed after the study in part a, placed volunteers for the study in a free 50+ living environment. The scientist placed actors and actresses in adjacent apartments and had them either act nicely or harshly and tracked the cardiovascular activity of the volunteers living next to them. What type of Study is this and why?

This is a designed study because the scientists make decision to see if they influence the cardiovascular health of participant.

4. In 2014, the hottest summer recorded in parts of Oregon (KVAL News,) ice cream sales and accidental drowning deaths were abnormally high in Salem, Oregon. If I said that eating ice cream caused accidental drowning deaths I would have created a good example of **confounding**, what would my lurking variable be?

The lurking variable would be the summer heat. It is unlikely that eating ice cream causes drowning but reasonable that both would increase in hot summers.

5. Politico.com published an article that talked about the US Census and how nearly 10 million Americans reported a different race or ethnicity than they did in the early 2000's. Respondents either switched, moved from choosing a single race or ethnicity to choosing multiple or vice versa. Answer all that apply below for the scenario described above:

a. Sampling Method: *Census*

b. Population: *Americans*

c. Sample: *N/A (Americans)*

d. Variable: *Race / Ethnicity*

6. Ron Burgundy, a fake news anchor, wants to tell his viewers the American reaction to Florida accidentally banning all computers and smart phones from internet cafes, as outlined in a real news article posted on Gizmodo.com. Ron called a software developer, a telecommunications worker, an internet café owner and two customers from their store to get quotes and reel for his nightly news show. Answer all that apply below for the scenario described above:

- a. Sampling Method: *Judgement*
- b. Population: *Americans*
- c. Sample: *The five interviewees*
- d. Variable: *Reaction to the Florida law*

7. In an attempt to estimate the percentage of USC undergraduate students that had sex in the last 24 hours I sent an anonymous survey to the students through Pearson as a homework assignment and used the percentage for the class as my estimate. Answer all that apply below for the scenario described above:

- a. Sampling Method: *Convenience*
- b. Population: *USC undergrads*
- c. Sample: *The class of students I sent the survey to*
- d. Variable: *Whether or not they had sex in the last 24 hours.*

8. Walking down the hallways of the luxurious LeConte building I noticed a flyer seeking students for a study on how alcohol effects REM sleep. The study stated that students should contact a provided email if they would like to participate in the study that provides them free drinks at a bar in Five Points and a place to sleep if they let the scientist attach brainwave readers to their head while they sleep – yes, that means you can't bring anyone back from the bar until the study is over. Answer all that apply below for the scenario described above:

- a. Sampling Method: *Volunteer*
- b. Population: *All Students*
- c. Sample: *The students that choose to participate*
- d. Variable: *effects on REM sleep*

9. In an experiment to see how being nerdy affects someone's 'selfie-ability' an at least normally attractive person wore a nerdy disguise and asked every fiftieth stranger on the street in the downtown Park City, Utah if they would pretend to be his girlfriend and take a selfie with them. (<https://www.youtube.com/watch?v=c4jJSUeRSss>)

Answer all that apply below for the scenario described above:

- a. Sampling Method: *Systematic*
- b. Population: *People in downtown Park City, Utah*
- c. Sample: *Every fiftieth Stranger*
- d. Variable: *Whether or not they will take a selfie with him*



10. In an attempt to estimate the percentage of USC undergraduate students that consumed any illegal drugs in the last weekend I obtained a sample by selecting three Stat 201 courses and polling all of the students in those class by cooperating with lecturers in my department. Answer all that apply below for the scenario described above:

- a. Sampling Method: Cluster (Convenience is okay but not best)
- b. Population: USC undergrads
- c. Sample: The 3 selected Stat 201 courses
- d. Variable: Whether or not they consumed illegal drugs in the last weekend.

11. In an attempt to estimate the percentage of USC undergraduate students that consumed any illegal drugs in the last weekend I obtained a sample by selecting ten Stat 201 courses and polling a random sample of the students in each class by cooperating with lecturers in my department. Answer all that apply below for the scenario described above:

- a. Sampling Method: Stratified (Convenience is okay but not best)
- b. Population: USC undergrads
- c. Sample: The random samples from the 10 selected STAT 201 courses
- d. Variable: Whether or not they consumed illegal drugs in the last weekend.

12. A sample, obtained through randomization, of 1,025 adults by the Washington Post showed that 42% of all respondents, 21% of Republicans, 67% of Democrats, and 37% of Independents approved of President Obama's handling of the situation in Iraq post Airstrikes against the militant group ISIS. ([http://www.washingtonpost.com/politics/polling/overall-barack-obama-iraq/2014/08/20/257e62a6-2859-11e4-8b10-7db129976abb\\_page.html](http://www.washingtonpost.com/politics/polling/overall-barack-obama-iraq/2014/08/20/257e62a6-2859-11e4-8b10-7db129976abb_page.html)) Answer all that apply below for the scenario described above:

- a. Sampling Method: *Simple Random Sample*
- b. Population: *American adults*
- c. Sample: *1,025 surveyed by the Washington Post*
- d. Variable: *Whether or not they approve of President Obama's handling of Iraq*

13. A referee of a basketball tournament used sample of twenty five middle school students, ages 11 to 14 year olds, that were participating in the tournament in Columbia, South Carolina to determine the average height of middle school students in Columbia, South Carolina. What is the bias here and what type of bias is it?

*The bias is that kids interested in playing basketball are taller on average. Since this sample favored taller kids, we have sampling bias.*

14. A sample was taken by a polling institute in 2012 to find the proportion of Americans that agreed with the June 28<sup>th</sup> ruling by the Supreme Court that upheld the Patient Protection and Affordable Care Act, also known as 'Obamacare.' The Chair of the institute told their employees to make their sample 50% Democrats, 35% Republican and 15% Independent. A quick Google search shows a Gallup.com article that shows that, in 2012, about 42% of Americans are at least Republican leaning, about 47% are at least democratic leaning and about 11% are independent. What do we call this mismatch of populations?

*When our sample doesn't follow the same proportions as the population we say that we have under coverage.*

15. Public Policy Polling reported the following:

[http://www.publicpolicypolling.com/pdf/2011/PPP\\_Release\\_Natl\\_010813\\_.pdf](http://www.publicpolicypolling.com/pdf/2011/PPP_Release_Natl_010813_.pdf)

*"When asked if they have a higher opinion of either Congress or a series of unpleasant or disliked things, voters said they had a higher opinion of root canals (32 for Congress and 56 for the dental procedure), NFL replacement refs (29-56), head lice (19-67), the rock band Nickelback (32-39), colonoscopies (31-58), Washington DC political pundits (34-37), carnies (31-39), traffic jams (34-56), cockroaches (43-45), Donald Trump (42-44), France (37-46), Genghis Khan (37-41), used-car salesmen (32-57), and Brussels sprouts (23-69) than Congress.*

*Congress did manage to beat out telemarketers (45-35), John Edwards (45-29), the Kardashians (49-36), lobbyists (48-30), North Korea (61-26), the Ebola virus (53-25), Lindsay Lohan (45-41), Fidel Castro (54-32), playground bullies (43-38), meth labs (60-21), communism (57-23), and gonorrhea (53-28)."*

Let's consider that the firm wanted to have the correct proportion of each political party so it threw out some of the responses so that the proportions of their samples matched the estimated true proportions. Most of the discarded responses were from independents who happen to be more likely to respond to surveys. What type of bias did they avoid by doing this?

Non-response bias, this is because Independents answer differently than Democrats and Republicans and we want the correct proportion of party affiliations so we can accurately judge approval.

16. In an attempt to estimate the percentage of USC undergraduate students that consumed any illegal drugs in the last weekend I randomly walked up to fifty students in front of the Russell House and asked them. Zero of the fifty said they did. I reported this statistic to a colleague and they suggested that couldn't be correct because of information posted on Casa Columbia's website. They suggested there may have been some bias in my sample – what type of bias are they suggesting? <http://www.casacolumbia.org/newsroom/op-eds/wasting-best-and-brightest-alcohol-and-drug-abuse-college-campuses>

Response bias, this is because students might lie to a stranger asking them about drugs because of shame or a fear of getting in trouble.



17. University of Missouri produced a research paper called 'Alcohol Effects on Performance monitoring and Adjustment: Affect Modulation and Impairment of evaluative Cognitive Control.' Researchers, hiding which type of drink was given to each group, gave a random sample of volunteer participants (Group A) non-alcoholic beverages and the rest (Group B) got identical, yet alcoholic beverages. The participants completed a task and then were asked about errors during the completion of the task. The conclusion of the study was that group B participants knew what they were doing as much as group A participants but the difference was that group A didn't care about the errors.

I. Answer all that apply below for the scenario described above:

- a. Sampling Type: *Volunteer Sample*
- b. Experimental Units: *Volunteers*
- c. Experimental Design: *Completely randomized design*
- d. Control Group: *Group A (non-alcoholic)*
- e. Placebo: *non-alcoholic drinks*
- f. Treatment: *alcoholic drinks*

II. Was blinding used in this experiment? If so, what type of blinding?

*Yes, this was a single blind study because only the recipient of the 'treatment' doesn't know what they're getting.*

III. Consider a scenario where we are interested in the difference between men and women and decide to put the men into one group, the women into another group and complete this experiment the same way as I've described above. What technique would this use and what would each group, men and women, be called?

*This technique is called blocking and each group is called a block*