

Problem 1

The Big Five Inventory is a commonly used personality assessment that measures the levels of Agreeableness, Conscientiousness, Extraversion, Neuroticism, and Openness for a particular individual. The assessment is a series of statements like, “I enjoy being with people” or “I enjoy participating in social gatherings”. Someone taking the assessment gives a response of 1, 2, 3, 4, or 5 indicating how much they agree with each statement. At the end, the responses are added together to give a total score for each trait.

In a 1987 study, 1,421 undergraduate students took this personality assessment and indicated whether they were willing to volunteer for future studies¹. The study found that females were slightly more likely to volunteer than males, and generally, extraverts (those who score high on the Extraversion scale) are more likely to volunteer than introverts (those who score low on the Extraversion scale).

- (a) Identify the main research question of the study. (1 pt)
- (b) Who are the subjects in this study, and how many are included? (1 pt)
- (c) What are the variables in the study? Identify each variable as numerical or categorical. If numerical, state whether the variable is discrete or continuous. If categorical, state whether the variable is ordinal. (2 pts)

Problem 2: OIS Exercise 1.18, page 30

Suppose you want to estimate the percentage of videos on YouTube that are cat videos. It is impossible for you to watch all videos on YouTube so you use a random video picker to select 1000 videos for you. You find that 2% of these videos are cat videos. Determine which of the following is an observation, a variable, a sample statistic (value calculated based on the observed sample), or a population parameter.

- (a) Percentage of all videos on YouTube that are cat videos. (1 pt)
- (b) 2%. (1 pt)
- (c) A video in your sample. (1 pt)
- (d) Whether or not a video is a cat video. (1 pt)

Problem 3

A study published in the *Journal of Public Economics* observed a group of 28,000 women who visited Colorado Title X clinics from 2008 to 2015². As part of the Colorado Family Planning Initiative, the women visiting these clinics were eligible for free long-acting reversible contraceptives (LARCs) such as intrauterine devices and subdermal implants.

The authors of the study analyzed trends in contraception choices, as well as trends in births by age, race, ethnicity, and zip code, before and after the implementation of the Colorado Family Planning Initiative. Overall, they found that births in the zip codes nearest to the Title X clinics tended to drop by 10%, but the drop decreases for zip codes farther and farther away from the clinics. This was true for all age groups, but especially so for women aged 15-19.

¹<https://bpspsychub.onlinelibrary.wiley.com/doi/abs/10.1111/j.2044-8309.1987.tb00769.x>

²[urlhttps://www.sciencedirect.com/science/article/pii/S0047272720301523](https://www.sciencedirect.com/science/article/pii/S0047272720301523)

- (a) What are the cases? (1 pt)
- (b) What is (are) the response variable(s) in this study? (1 pt)
- (c) What is (are) the explanatory variable(s) in this study? (1 pt)
- (d) Does the study employ random sampling? (1 pt)
- (e) Is this an observational study or an experiment? Explain your reasoning. (2 pts)

Problem 4: OIS Exercise 1.26, page 31

Suppose we want to estimate household size, where a “household” is defined as people living together in the same dwelling, and sharing living accommodations. If we select students at random at an elementary school and ask them what their family size is, will this be a good measure of household size? Or will our average be biased? If so, will it overestimate or underestimate the true value? (2 pts)

Problem 5: OIS Exercise 1.27, page 31

A statistics student who is curious about the relationship between the amount of time students spend on social networking sites and their performance at school decides to conduct a survey. Various research strategies for collecting data are described below. In each, name the sampling method proposed and any bias you might expect. (4 pts_)

- (a) He randomly samples 40 students from the study's population, gives them the survey, asks them to fill it out and bring it back the next day.
- (b) He gives out the survey only to his friends, making sure each one of them fills out the survey.
- (c) He posts a link to an online survey on Facebook and asks his friends to fill out the survey.
- (d) He randomly samples 5 classes and asks a random sample of students from those classes to fill out the survey.