Lab 3: Class Survey

STAT 20 Spring 2022

Part I: Understanding the Context of the Data

1. What is the unit of observation in the survey of students in Stat 20?
2. Where in the Taxonomy of Data are the variables that correspond to each of the following survey questions (i.e., what is their type)?

- Question 6: What is your favorite thing about Cal?
- Question 13: Where would you rather be?
- Question 20: What is the chance that we remain in remote learning for the entire semester?

3. What values would you expect `major` to take? Do you anticipate any challenges when analyzing this data?

Questions 4 - 7 can be answered using data visualization. To lay the groundwork for answering these questions using the survey data set, plan out your plots and set your expectations for what the data might look like. Please:

- sketch (with paper and pencil) a plot that captures the distribution or relationship between the variables,
- label the axis and add axis tick marks with plausible values,
- depict a shape that reflects your expectation of the phenomena. Also:
- indicate the *aesthetic mappings* (the variables in use and which visual cues / aeshetic each one is mapped to) and
- indicate the *geometry* of your plot.

4. Do students prefer to spend time at the beach or in the mountains?
5. How much coding experience do students have?
6. What are students' perceptions of the chance that we remain in remote learning for the entire semester?
7. What is the relationship between students' optimism for cryptocurrency and their skepticism of the effect of technology on interpersonal relationships?

Part II: Computing on the Data

Using the real survey data, construct the plots that you proposed in Part I and calculate for the first three a measure of a typical observation (mean, median, mode). Use these summaries to answer the original question in a sentence or two, noting, where appropriate, the shape of the distribution and what a "typical" observation is.

8. Do students prefer to spend time at the beach or in the mountains?
9. How much coding experience do students have?
10. What are students' perceptions of the chance that we remain in remote learning for the entire semester?
11. What is the relationship between students' optimism for cryptocurrency and their skepticism of the effect of technology on interpersonal relationships?

Part III: Extensions

Three variables appear in the survey data frame that were derived from the original `major` question: `health`, `tech`, and `business`. These variables indicate whether or not a student is in a health-related, technology-related, or business-related field, respectively.

For the following questions, please answer with a plot and a sentence or two written response given the structure in the plot.

12. Is there an association between students majoring in a technology field and their optimism for cryptocurrency?
13. Is there an association between students majoring in a business field and their optimism for cryptocurrency?
14. Propose your own question involving one or two variables and answer it using a plot with a written interpretation.
15. Propose your own question involving two or three variables and answer it using a plot with a written interpretation.