

- (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)?
  - *What is your favorite thing about Cal?* (favorite\_cal)
  - *“How many body piercings do you have?”* (piercings)
  - *What is the chance that a fire alarm goes off in one of the dorms this week?* (fire\_alarm)
- (3) What values would you expect major to take? Do you anticipate any challenges when analyzing this data (and why)?
- (4) Are there any other variables in the data set which would present the same challenge you mentioned in Question 3? List them here.

The remaining questions can be answered using data visualization. To lay the groundwork for answering these questions using the glimpse of the survey data set that you have, 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. There is no right or wrong at this point, because you have not seen the full data set yet! You do not need to retroactively change your answers once you do see it.
- (5) Do students prefer to spend time at the beach or in the mountains?

(6) Is there an association between students' favorite season and terrain preference (beach or mountains)?

(7) How many students who preferred the mountains selected Winter as their favorite season?

(8) Consider the question:

How much coding experience do students have?

Imagine you were working with the “words” version of this variable as shown in the glimpse and wanted to make a one-variable bar chart using this data. Do you anticipate any possible challenges with your visualization properly representing the scale in words, and why?