Project 0 – Notes for TAs

When a student first opens a Jupyter Notebook link from Canvas, they might be asked to log in. Tell them to use their CSULB account information when logging in.

About the Dataset and A Snippet of the Data

* Says there are 301 rows in the dataset. After displaying some of the columns, Python says there are 300 rows. Why the difference? (Wait for responses)
  + Ans: The raw data includes a row with the variable names.
* Notice that in the Snippet the rows are numbered starting at 0, not 1.
* What is raw data? What does raw data consist of?
  + Ans: Raw data consists of variable names and the values of those variables for the observations/cases/individuals. An **observation** (or a case) is an individual entity in a study.
  + A **variable** is a characteristic that may differ among individuals.
  + Usually, data is stored and presented in a dataset where each row represents one case, and each column represents one variable.

1. Categorize Variables:
   1. Have students enter their answers.
2. Frequency Table:
   1. Explain that ‘CATEGORY’ will be analyzed, as evidenced by the “mc =…”  
       line.
   2. Run the code. Notice the random order of the categories.
3. Relative Frequency Table
   1. Note that the variable being analyze “df” was defined when making the frequency table.
   2. Ask students: How do we calculate relative frequency? Ans: category frequency / number of individuals. Walk through 2 of the calculations to show that the relative frequencies are correct.
4. Frequency Distribution Bar Chart
   1. Run the code
5. Describe the Distribution
   1. Write a minimum of 2 sentences to describe what the bar graph is showing us about the frequency distribution of blood pressure category.

Have students work on their own, and with their neighbors, to complete 6-10. Walk around and check their work and help those who need it. Students should be able to complete this and turn it in before they leave class. Let them know that they have until Monday night to complete it and turn it in if they don’t finish it in time.