

## Psych 205 - Lab 2 - Prof. Key

The goal of this lab is to get practice describing data and working in Quarto.

Download this file as a .qmd from the course page.

### In Lecture.

1. One advantage of Quarto (and R Markdown) is that you can run code in a document. You do this using a “code block”. In the space below, insert an R code block, type out a math equation that used to give you difficulty as a kid into the code block below, and run the code to see the result. Below the code block, add text for humans to determine whether R got the math problem correct or not. Then render the document as a .pdf and .html file. Did this work?
2. Load the “grad onboarding” survey into the code block below, and answer the following questions. *Note : to successfully render code in the document, you must a) explicitly load the dataset into your Quarto document and b) make sure you have no errors in your code. :)*
  - Graph the variables `d$self.skills` and `d$class.skills` side by side using the `par()` function. Change the formatting of the graph to make it look ready for presentation. Add vertical lines to each graph to illustrate the mean (solid line) and standard deviation (dashed lines).
  - Below each graph, report the mean and standard deviation of both variables, and interpret what these statistics tell you about the individuals in our class. (Who cares? What do these statistics tell us?)
3. Split your graphics window into a 2x5 grid, and graph each of the 10 “can.\*” variables in the dataset (e.g., “can.import”, “can.clean”, etc.). Make sure each graph contains the name of the variable and that the graph looks good / is intelligible. *Note : you can and should use a for-loop to do this!* Then, report the frequencies of these variables - what are some things you observe about the data? Does this make sense given what you know about the participants?

### In Section

4. Choose one of the datasets from the class folder (avoid the one(s) labeled [repeated measures] as we will get to these later. I’ll be focusing on the “Perceptions of the Wealthy” dataset if you want to follow along with my key when it’s posted). Look over the accompanying article for a guide to the variables, and identify two numeric variables from the dataset that seem interesting to you. Graph these variables, report the relevant

descriptive statistics for each variable. Then explain what these statistics and graphs tell you about the individuals in the dataset, and what other questions you might ask about these individuals (e.g., what do the data NOT tell you?)

## On Your Own

5. Psychologists often want to combine information from multiple questions (that measure the same construct) into one variable. There are lots of different ways to do this, but the first that we will talk about is just the humble average. For example, the Rosenberg (1965) self-esteem scale has 10 questions all related to self-esteem; rather than work with 10 different variables, it might be nice to just average these 10 and report one average number. These are called “likert” scales. It’s technically pronounced ‘lick-ert’, but no one says that because it sounds kind of gross.

[Here’s an overview \(from my 101 class notes\) of how to work with such likert scales conceptually and computationally \(in R\).](#) You’ll need to load the psych library into R; to do this, run the following code in your console.

```
install.packages("psych") # installs the "psych" package. you only need to do this once.
```

```
library(psych) # loads the library. you need to do this every R session.
```

Use these notes to answer the following questions. Let me know if anything is unclear!

**Use the self-esteem dataset (from the class datasets folder). You should find a codebook that describes these data in the same folder.**

Check to make sure the data loaded correctly. (Note that 0s in this dataset mean the person was missing data.) Report the sample size of the dataset.

Create a self-esteem scale from the 10-items. Make sure to reverse-score the negatively-keyed items so that for every question, higher numbers measure higher self-esteem. Graph this variable as a histogram, and make the graph look nice (ready for publication).

Report the alpha reliability, mean, and standard deviation of this variable. Below your graph, describe what these statistics tell you about the self-esteem of the participants. What other questions do you have about this variable, or the output?

Then, graph the variable gender as a categorical factor, and report the number of people who identified as “female”, “male”, and “other”. You will need to do some data cleaning here.

**Challenge Problem (Optional, but Encouraged!)** Report the mean and standard deviation of self-esteem for people who are identified as female, male, and “other” in the dataset.

Note : there are MANY ways to do this in R; you can use the subset function or indexing to divide the dataset into three groups - “females”, “males”, and people who reported “other”.

Or other fancier methods we will talk about later. See how many different ways you can do it.