Assignment 1

This assignment is due on Canvas on Monday 9/20 before class, at 10:15 am.

General instructions for all assignments:

- Use this file as the template for your submission. You can delete unnecessary text (e.g. these instructions) but make sure to keep the nicely formatted "Problem 1", "Problem 2", "a.", "b.", etc
- Upload a your knitted PDF file to the Homework 01 submission section on gradescope. Name this file as: [PennID]_crim250_hw01.html (e.g. and "mcuellar_crim250_hw01.html")
- Each answer must be supported by written statements (unless otherwise specified)
- Include the name of anyone you collaborated with at the top of the assignment

Assignment 1

1. Install the datasets package on the console below using `install.packages("datasets")`. Now load the library.

Load the USArrests dataset and rename it dat. Note that this dataset comes with R, in the package datasets, so there's no need to load data from your computer. Why is it useful to rename the dataset?

2. Use this command to make the state names into a new variable called State. (This dataset has the state names as row names, so we just want to make them into a new variable. We also make them all lower case, because that will help us draw a map later - the map function requires the states to be lower case.)

List the variables in the dataset `USArrests`.

3. What type of variable (from the DVB chapter) is Murder?

Answer:

What R Type of variable is it?

4. What is this dataset showing? What do the numbers mean?

Answer:

5. Draw a histogram of Murder with proper labels and title.
6. Please summarize Murder quantitatively. What are its mean and median? What is the difference between mean and median? What is a quartile, and why do you think R gives you the 1st Qu. and 3rd Qu.?
7. Repeat the same steps you followed for Murder, for the variables Assault and Rape. Now plot all three histograms together by using the command `par(mfrow=c(3,1))` and then plotting each of the three.

What does the command par do, in your own words (you can look this up by asking R `?par`)?

What can you learn from plotting the histograms together?

Answer:

8. Delete the hashtags below and run this code, after installing the packages maps and ggplot2 in the console:

```
#library('maps')
#library('ggplot2')

#ggplot(dat, aes(map_id=state, fill=Murder)) +
#  geom_map(map=map_data("state")) +
#  expand_limits(x=map_data("state")$long, y=map_data("state")$lat)
```

What does this code do? Explain what each line is doing.

Answer: