Assignment 1

Assignment Description

In this assignment you are asked to manipulate data, estimate statistical relationships, and interpret the findings. The questions below guide you through the process of statistical estimation. You are given an R script that generates datasets for you to use. In places where it asks you to execute additional commands (like creating a variable, running a regression), refer to the R script we worked through in class. If anything is unclear, you can use the R help function by typing a question mark and then the name of the command. For example, if I wanted help with the mutate command, I would type "?mutate()". If you are still looking for help, usually googling the name of the command and the word "example" will provide many useful resources. I am also of course available to help.

I strongly suggest that you start this assignment early because it will not be possible (in my opinion) to do well if you start close to the due date. There are parts that you may find difficult; you will want to identify them and leave enough time to ask questions if necessary.

Assignment Instructions

Data analysis

In mylearningspace, you will find an R script called "EC655 Assign1.R" that contains some code you need to complete the assignment. All students must use this file to generate their data and write the code to answer the questions. Before using the R script, you need to do the following:

- Rename the file from "EC655 Assign1.R" to your family name followed by your student number (no spaces)
- In the two places where the command set.seed() appears, replace INSERT YOUR STUDENT NUMBER HERE with your full student number. Be sure to include the set.seed() command whenever you run the file.

Leave all other commands that currently exist in the R script untouched. Write your code to answer the question in the area indicated inside the comments.

Note that this file generates random data depending on your student number, so each student's data will be different, and therefore answers will also be different.

Submission

You are required to submit two documents according to the following instructions:

- A report containing your answers to all the questions. I outline below how I would like your report to look. The overall goal is that the answers to each question must be easily identifiable in a readable, professional-looking document. Hand in the report in Gradescope and the MyLearningspace Dropbox;
- R script. Hand in to the MyLearningspace Dropbox only

In the report described in (a) above, please answer all questions in the same order as they are stated on the question sheet. For each question and sub-question, include the relevant R code (if any) that you used, the output generated by that command if there was any, and an interpretation if you are asked to provide it. For example, if you were answering the following hypothetical question, it might look like this:

You could also format your own output tables rather than copying and pasting R output if you find it easier. The key is that as long as the questions are answered in order, and the R commands used for each question and associated output are clear, it will be fine.

A note on plagiarism: this is an independent assignment, which I expect you to complete on your own. It is plagiarism to copy someone else's work verbatim, which includes R code. Any work you submit should be yours only.