BSDS 100: Intro to Data Science with `R`
Assignment 1
Due 9/1 at 11:59pm

Before you begin, make sure that you have R and RStudio properly installed.

Directions: Write a single R markdown (.Rmd) file that answers each of these questions and produces a knitted .pdf which holds your responses. Make sure that all code can be successfully run on any computer. Put your name and date at the top of the document. For all questions that require written responses, write the answer (numbered appropriately) in markdown, not as a comment in the code. Turn in the .Rmd file and .pdf file on Canvas. Late assignments are not accepted.

1. Tell me a little about yourself. For instance,

    (a) Why are you interested in taking this course?

    (b) Have you ever used `R` before this course? If so, to what extent?

    (c) Have you ever used another programming language before? How would you rate your skill level?

    (d) What major are you? What do you plan on doing when you graduate from USF?

    (e) Tell me at least one thing that you hope to do with `R` this semester.

    (f) Feel free to tell me anything else interesting about yourself that you'd like to share.

2. Read Chapter 1 of Doing Data Science by Cathy O'Neil and Rachel Schutt entitled "What is Data Science?" here: https://www.safaribooksonline.com/library/view/doing-data-science/9781449363871/ch01.html.

    (a) From reading the article, what in your opinion is the biggest issue facing the field of "data science," and what do you think can be done to help fix this issue?

    (b) What, in your opinion, is the biggest difference between traditional statistics and data science? How does technology play a role in this difference?

    (c) What, in your opinion, is the biggest difference between data scientists in academia and industry?

    (d) What kind of data scientist do you want to be? What skills are you most interested in developing?

    (e) Based on this article, a major obstacle in understanding data science in the first place is understanding what data science actually *is*. From what you've read, and what you understand so far, how would **you** define the term "data science"? Note: I'm interested in hearing **your** definition, not one from the internet or an article or someone else.

3. Perform the following tasks:

   (a) Assign x=3 and y=4.
   (b) Calculate $\ln(x + y)$.
   (c) Calculate $\log_{10}(\frac{xy}{2})$.
   (d) Calculate $2x^{1/3} + y^{1/4}$.
   (e) Calculate $10^{x-y} + e^{xy}$.

4. While in RStudio, press Alt + Shift + K. What happens? How can you get to the same place using the menus?

5. The term Big Data is used frequently in data science, but a formal definition doesn't really exist. We talked about one commonly used definition in lecture. Discuss in what contexts you've run across the term and what it means. If you have not witnessed the term before, find an article or website that uses it and discuss its use in that article.