second project

Tue 10/10, 8:43 PM

Please follow the following for the second project:

1) Select a dataset of your own choice (simple, complicated, whatever your

level of comfort is)

2) Chop the dataset using pandas to fit whatever visualization you have in

mind.

3) Clean the dataset and drop missing values - if needed

4) Create two different types of visualization - minimum. Your choice, just

make sure they are different kinds - (for example a bar chart and a pie, but

not two pies).

5) You can have more visualizations if you want

6) Make sure your readme file is perfect!

7) Submit on github and kaggle (if using kaggle datasets), just github if you

are using data from other sources like drivedata.org or other census data.

8) Get in touch with me about anything that is worrisome - we can do skype or

speak on the phone. Any concern you may have, do not hesitate.

9) Also, if you have ideas on how to improve your data science program, please

write/speak with me. I am going to write to the department chair with a list of

things I believe need to change to make your experience better - it is

anonymous and I already have many things in mind (i.e., adding a statistics

course, having Github account mandatory from day one, have more career and

professional resources for you to get a job when you are done, etc).

All the best to you.