

HW4- CMDA

Submit your .py script to Dropbox by Monday, October 20, 5pm. Name your script ("YourName_HW4.py"). Use the textbook, class notes and scripts, as well as Python-related internet resources. Ask questions in the Forum for your colleagues to answer. **Work in teams, but submit your individual files with your work.**

All tasks should be performed in Python

- 1) Create a new ipython notebook .
- 2) Discuss with your team and decide if you would like to use the same data set for your final project. You can stay with it or you can find a new one. Data can come from a csv file, a text file, or a JSON object.
- 3) Import your data set to pandas. Re-create any reshaping operations you did in R for homework 2. If your dataset is new, provide a short data manual as a separate pdf.
- 4) Get numerical summaries for all your variables.
- 5) Comment on and perform any needed treatment for missing values.
- 6) Comment on and perform any needed data transformations.
- 7) Get three visualizations that you consider meaningful.
- 8) Save your data as a pickle.
- 9) Update your project repository on Git with the pickled file. Include a comment of what you did in the description of the commit.
- 10) Download your HW4 notebook as .py, and upload it to Dropbox and GitHub.