CSCI 4580/5580 - Data Science – Fall 2020

Lab 2: Exploratory Data Analysis

Total Points: 100 Points

Launch Jupyter Notebook and open the provided notebook for Lab 2 in Jupyter Notebook. To be reminded of how you can open Jupyter Notebook, see Lab 0. Also, you can find a tutorial for Jupyter Notebook here (ignore the installation part):

- https://www.youtube.com/watch?v=HW29067qVWk&ab_channel=CoreySchafer

Installing Packages

If you see error messages like this:

```
ModuleNotFoundError: No module named 'numpy'
```

in the notebook, it means that you need to install the missing package(s) as explained in Lab 0.
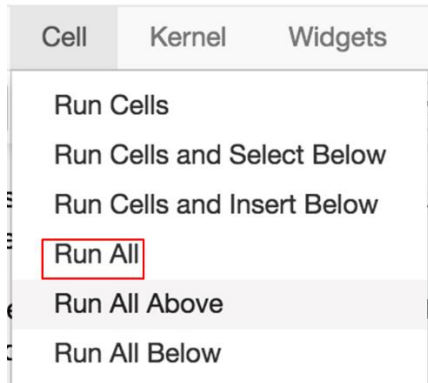
Supplementary Materials

We have added pointers to the tutorials for packages that are used in the labs on Canvas under Modules → Supplementary Materials. This module will be gradually updated throughout the semester.

Submission

You need to submit a single .ipynb file on Canvas, named *your-lastname_your-first-name.ipynb*. For example, if your name is John Smith, you should name the file *smith_john.ipynb*.

Answer Q1-Q11 in the designated cells in the lab notebook. Please do not add or remove any cells (you can add cells while doing the homework but remove them before submission).

Please download your submission file after submission and make sure it is not corrupted. Use the Run All option under the Cell menu and make sure all cells run without any issues.

We will not be responsible for corrupted submissions and will not be able to take a resubmission after the deadline.

Need Help?

In case you need help with this assignment, please ask your questions on the designated channel for Lab 2 on Microsoft Teams. This way, you may receive assistance from the TA's or your classmates (DO NOT share your solution) [Recommended].

If you need to include your solution in your question, please create a new chat with all the TA's (in the new chat, add all the TA's rather than sending your question to the TA's separately):