DSCI 510: Principles of Programming for Data Science

Fall 2022

Homework 4

Preliminary Analyses

Due: 12/2/2022 23:59 PM PT

Task:
- Please submit a zip package of:
  - A PDF file answering the following 1 question. 5 points.
  - Data analysis / visualization code. 5 points.
  - At least 2 figures of your preliminary analysis / visualization result. 5 points.

1. What kind of analyses or visualizations did you do? (Similar to Homework 2 Q3, but now you should answer based on your progress, rather than just your plan). [5 points]

Rubric:

1. The analyses and visualizations should:
   a. Feasible. It should be an analysis that could be done with your data source. 2 points.
   b. Meaningful. It should be something that is not obvious, non-trivial, not in common sense. 2 points.
   c. Impactful. It should be something having an impact and is interesting to the audience. 1 point.
2. The analysis / visualization code should:
   a. Have a readme file specify how to run your code. 1 point.
   b. Could be run based on the instructions. 2 points.

     c. Have a link to your github repository. 1 point.

     d. Specify dependency using requirements.txt or environment.yml. 1 point.

3. The sample figures should:

     a. Relevant to your analyses / visualizations task. 2 points.

     b. Well made, easy to read. 2 points.

     c. Includes a short text description of the figure and your findings. 1point.