

DSCI 510: Principles of Programming for Data Science

Fall 2022

Homework 4

Preliminary Analyses and Bonus

Due: 12/9/2022 23:59 PM PT

Task:

- Please submit a zip package of:
 - A PDF, Markdown, or Jupyter notebook file answering the following 1 question. 5 points.
 - If you did the bonus, also answer the # 2 question.
 - Subfolder **code**.
 - Data analysis / visualization code. 5 points.
 - If you did the bonus, also include its code here.
 - Subfolder **result**.
 - At least 2 figures of your preliminary analysis / visualization result. 5 points.
 - If you did the bonus, also include its results here. If it's an animated/interactive visualization, please record a video of it.
-
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]
 2. [Bonus] Did you do any advanced visualizations? Briefly describe how you do it, and why it's a great visualization.

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. Show your progress on analysis. You don't have to finish all the work by then. 4 points.
 - a. Have a link to your github repository. If it's a private repository, you need to add mail@gleb.com and yuzhongh@usc.edu as your collaborator in the repo settings. 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. 1 point.
4. [Bonus]
- a. Homework 4 takes 5 points in your final score. Bonus could add up to 5 points in your final score.
 - b. You could re-use the advanced visualization in bonus assignment as part of Homework 4.
 - c. It should be non-trivial, ideally interactive. It can **NOT** be a simple plot like line plot, scatter plot, bar plot, histogram, etc. They are considered the basic requirements in Homework 4.
 - d. As a bonus assignment, we won't give out 5 points easily, but will give 0-5 points based on the quality of visualization. We will give 5 points only if it's intriguing. Examples that considered 5 points:
 - i. Interactive line growth chart. [Link](#)

- ii. Zoomable sunburst. [Link](#)
 - iii. Choropleth Map. [Link](#)
- e. You could reference existing visualization examples, but you **MUST** modify and visualize the data used in the final project, instead of the data that has been used in the visualization examples.
- f. You may reference Week14 slides, page 16-17 to see advanced visualization methods. For example:
 - i. Advanced charts in [seaborn](#), [plotly](#), [bokeh](#).
 - ii. Interactive, web based visualization using [d3.js](#), [echarts](#). You may need to learn and write some javascript code.