

Project

You should identify a publicly available dataset that you are interested in working with. Your project will consist of three phases:

- Project proposal
- Data acquisition, scrubbing, and wrangling
- Story telling

There are 5 examples of this sort of activity in chapter 13 of our textbook. Feel free to read through this chapter to get some ideas. I am confident that your project will be better than any of these.

1. Your proposal should include a brief description of a dataset that you are interested in exploring, along with a link to where it can be accessed. The data set must be large enough and complex enough to support an interesting project.

You should also include a description of the sort of investigation that you intend to conduct. You can include questions you'd like to ask of the data, trends that you think may be of interest, categories of analysis that will allow you to slice the data in different ways, etc.

Deliverable: You will submit the proposal and receive approval or feedback from the instructor before you continue on to the other phases.

2. In the data preparation phase, you will write Python code that accesses, formats, filters, and otherwise manipulates the data set into a form that can be used in phase 3. The details of this will be highly dependent on the specifics of your project.

Your code should be organized, modular (i.e., methods), well-commented, and have a main method.

Deliverable: All of the code that you produce in phase 2 should be submitted as a single Python file.

3. In the final phase of your project you will add additional Python code that pulls out the specific data that answers the questions or produces the visualizations that you need to tell the story that you are after.

Deliverable: All of the additional code that you produce in phase 3 can be added to the same source code file you created in phase 2. Make sure you have a clean separation of

the code from the two phases. Also prepare either a written document (Word) or a presentation (PowerPoint) that uses the results of your work to tell the story.

4. Email your project as a zip attachment to prof. stucki.

Due: Friday, 4/12/2024 by 11:59pm