# Technical Requirements

The technical requirements for Project 1 are as follows.

\* [ ] Use Pandas to clean and format your data set(s)

\* [ ] Create a Jupyter Notebook describing the \*\*data exploration and cleanup\*\* process

\* [ ] Create a Jupyter Notebook illustrating the \*\*final data analysis\*\*

\* [ ] Use Matplotlib to create a total of 6-8 visualizations of your data (ideally, at least 2 per "question" you ask of your data)

\* [ ] Save PNG images of your visualizations to distribute to the class and instructional team, and for inclusion in your presentation

\* [ ] Optionally, use at least one API, if you can find an API with data pertinent to your primary research questions

\* [ ] Create a write-up summarizing your major findings. This should include a heading for each "question" you asked of your data, and under each heading, a short description of what you found and any relevant plots.

# Presentation Guidelines

You are free to structure your presentations to your liking, but students tend to have success with the following format.

\* Title Slide

\* Include the name of the Project and Group Members

\* Motivation & Summary Slide

\* Define the core message or hypothesis of your project.

\* Describe the questions you asked, and \_why\_ you asked them

\* Describe whether you were able to answer these questions to your satisfaction, and briefly summarize your findings

\* Questions & Data

\* Elaborate on the questions you asked, describing what kinds of data you needed to answer them, and where you found it

\* Data Cleanup & Exploration

\* Describe the exploration and cleanup process

\* Discuss insights you had while exploring the data that you didn't anticipate

\* Discuss any problems that arose after exploring the data, and how you resolved them

\* Present and discuss interesting figures developed during exploration, ideally with the help of Jupyter Notebook

\* Data Analysis

\* Discuss the steps you took to analyze the data and answer each question you asked in your proposal

\* Present and discuss interesting figures developed during analysis, ideally with the help of Jupyter Notebook

\* Discussion

\* Discuss your findings. Did you find what you expected to find? If not, why not? What inferences or general conclusions can you draw from your analysis?

\* Post Mortem

\* Discuss any difficulties that arose, and how you dealt with them

\* Discuss any additional questions that came up, but which you didn't have time to answer: What would you research next, if you had two more weeks?

\* Questions

\* Open-floor Q&A with the audience

# Presentation Requirements

The presentation requirements for Project 1 are as follows.

Your presentation must:

\* [ ] Be at least 8-10 min. long

\* [ ] Describe the core message or hypothesis for your project.

\* [ ] Describe the questions you and your group found interesting, and what motivated you to answer them

\* [ ] Summarize where and how you found the data you used to answer these questions

\* [ ] Describe the data exploration and cleanup process (accompanied by your Jupyter Notebook)

\* [ ] Describe the analysis process (accompanied by your Jupyter Notebook)

\* [ ] Summarize your conclusions. This should include a numerical summary (i.e., what data did your analysis yield), as well as visualizations of that summary (plots of the final analysis data)

\* [ ] Discuss the implications of your findings. This is where you get to have an open-ended discussion about what your findings "mean".

\* [ ] Tell a good story! Storytelling through data analysis is no different than in literature. Find your narrative and use your analysis and visualization skills to highlight conflict and resolution in your data.