

Data Visualization Projects

We will have three projects in our course, all of them the same except for the following set of differences:

1. **Maps:** The first project will focus on mapping and must include GIS data. **The “main” visualization that you create must include mapping elements** (including the possibility of it “just being” a map... when I say “mapping elements” I am thinking of [this graphic](#) as an example). The source of the data can be anything, except that at least some of the data is GIS data, obviously. Some ideas:
 - a. Map the migration of a people, currently or historically,
 - b. Visualize the distortion in old maps (alongside Dr Blunt and the Knafel Map Collection),
 - c. Pick a political issue of interest and plot the votes on a county-level, compared to other, more standard votes
 - d. Construct a global description of another planet (e.g. “where should humans set up camp on Mars?”)
 - e. Find a topic of interest on our campus and plot it on a new version of [this map](#). For example, working with the BIO100 classes to plot all the trees of campus and their health status.
2. **Text: The source of the data for this project must be natural language.** That means, you must process text that was written in plain, “normal” language, by a human (or by many humans), and then visualize it however you’d like. Note that we’ll take some time to talk about tools to analyze text. Some ideas:
 - a. By analyzing all the words in a novel, find quantities which allow you to plot the [dramatic structure](#), or some other theme of interest to you.
 - b. Download a massive collection of tweets relating to a topic of interest and visualize the similarities and differences between people based on some criterion.
 - c. Analyze news articles from a source to try and discover bias.
3. **Open:** You may create the project of your choice. Feel free to check with me regarding the scope of this project. You may, if you’d like, plan out previous projects to be stepping stones toward a larger topic of interest, so that this project serves as a culmination of the previous. That is not necessary for you to be successful, however.

Rules

1. Be mindful of self-plagiarism, but feel free to try to formulate connections between your classes!
2. If you plan to have your projects chain together or connect thematically, please know that each project must stand on its own.
3. Group size is 1, 2, or perhaps 3 for all projects. I'm expecting many solo projects, and if you do group up, expect a *much* more strict collection of expectations from me in terms of each element of the rubric.

Specifications

For each project, you will submit both of the following.

- Data Acquisition and Preparation Notebook: You'll need to collect data for this project, and it's almost certain that whatever data you collect will be not in a proper format upon receipt. Construct a Jupyter Notebook which documents the process. It should:
 - Describe the goal of the project and *why* you wanted to tackle it,
 - Describe what data you'll need to collect,
 - Be well organized with headers/sections, links to important resources, and other organizational tools that you deem necessary,
 - Contain the actual code that collects the data, if necessary,
 - Contain the actual code that prepares your data for use,
 - Contain enough markdown cells to weave together a story that goes from concept to clean and prepared datasets.
 - Be reproducible: if I read your notebook, follow your instructions, and run your cells (in order!), I should be able to end up in the same place you do.
- Visualization and Presentation: You may produce one single visualization, if that is what makes sense, or you may decide to produce many. You should submit an html page and all surrounding files (css, javascript, json, csv, etc.) which:
 - Contains (all) your visualization(s),
 - Describes in sufficient detail what your visualization is covering so that a layperson can understand what you've created (without reading the Notebook!),
 - Be designed thoughtfully and creatively, in a way that complements your visualizations,
 - Adheres to all the graphical design standards we've learned during this course (especially in the reading!)

See the rubric for more information on how I will assess the projects.