American Electoral Politics

CS171: Final Project Proposal

Team: Ryan Kerr, Evan Sandhoefner, Mengting Zhang

Background and Motivation

Discuss your motivations and reasons for choosing this project, especially any background or research interests that may have influenced your decision.

We initially brainstormed several areas of interest in trying to choose this project, then attempted to find data for those target areas. Politics, especially those involved in the United States presidential election, was one of those areas where our interest intersected with public data that was publicly accessible. The history of the presidential election is not a familiar topic, and with this project, we hope to educate the general public about how the process and the parties involved have changed over time.

Project Objectives

Provide the primary questions you are trying to answer with your visualization. What would you like to learn and accomplish? List the benefits.

The primary goal of our visualization is to display voting trends in the United States, as well as the changes in political party support through the history of the country. In addition to currently existing parties, the data we visualize will include parties that have faded out of existence. A benefit of this visualization is that we will also be able to visualize the history of the United States and view the growth of not only the states as they exist currently, but also how parties and the election process have evolved over time. Through the creation of this visualization, we will also be able to learn about how certain demographic factors relate to both voting and voter turnout. Not only will we be able to view how each state changes through elections, but also how states differ with each other.

Data

From where and how are you collecting your data? If appropriate, provide a link to your data sources.

We will be combining data from a number of sources to obtain a comprehensive understanding of changes in American electoral politics over the centuries, and of differences between states for a given election year. A list of sources and the specific data located at site follows:

http://www.presidency.ucsb.edu/showelection.php?year=2012

The above link has data for each presidential election from 1789 to 2012. The data include, for each state: total votes and votes for each major candidate.

https://docs.google.com/spreadsheets/d/1bH38j6_e8yA9xq8OMlyLOL6h_iTS7ABQMKNxzFgKBDo/edit#gid=435419492

The above link has national turnout rate for 1789-2012.

https://docs.google.com/spreadsheets/d/1or-N33CpOZYQ1UfZo0h8yGPSyz0Db-xjmZOXg3VJi-Q/edit#gid=1670431880

The above link has the following variables for 1980-2014, by state: voting-eligible population, votes counted, voting-age population, % non-citizen, and population of prison/probation/parole/ineligible felon. Unfortunately, there are some missing values in the spreadsheet. We will attempt to fill in those values from another data set, or find a better data set and not use this one at all.

http://www.presidency.ucsb.edu/data.php

The above link contains a ton of other data that we'll think of as best-case features for now. This includes not only more detailed election data, but also data on the presidencies themselves, which could be an entirely new component of the overall visualization. This includes things like number of vetoes, number of executive orders, and approval ratings. Approval ratings in particular have a direct link to election data - we may be curious to see, for example, if there are presidents who were elected by a landslide but who ended up with low approval ratings. This whole section is a "maybe" feature, because it would be an entirely separate and non-essential wing of the overall visualization.

https://cse.google.com/cse/publicurl?cx=002720237717066476899:v2wv26idk7m

The above link leads to Google's Datasets Search Engine, which we have been using to find a lot of this data. We will continue to use it, in addition to general Google and sites that specifically compile political/electoral/demographic data from various sources, to continue bolstering the above data and introducing the below data.

Other possibilities for detailed state data: income, ethnicity, and age. We would like to find these data for the voting population, but we recognize that it may only be conveniently available for the overall population. Similarly, we would like to find these data for as much of the time-series as possible, but it may only be conveniently available for recent years. Due to time constraints, we have not finished searching for these data yet. Because of apparent limitations in the likelihood of their availability, we consider them to be best-case, non-essential features.

Data Processing

Do you expect to do substantial data cleanup? What quantities do you plan to derive from your data? How will data processing be implemented?

Two of the data sources above are Google Doc spreadsheets. We will save those as .csv and convert them to JSON using a simple online converter such as this one.

The one non-Google-Doc source, which happens to be the most important, is in the form of an HTML table. Our basic plan is to simply copy and paste the tables manually into Excel, and then follow a similar csv-to-JSON procedure as above.

Overall, we will have to do some data cleanup and conversion - because our data isn't already in JSON format - but we don't expect it to be a substantial time commitment. We cannot say much more about the data processing at this time because, simply put, we haven't tried it yet. In terms of specific quantities derived, we want to track all the

variables specifically mentioned in the above Data section. There's not a whole lot that we're going to throw away.

Visualization

How will you display your data? Provide some general ideas that you have for the visualization design. Include sketches of your design.

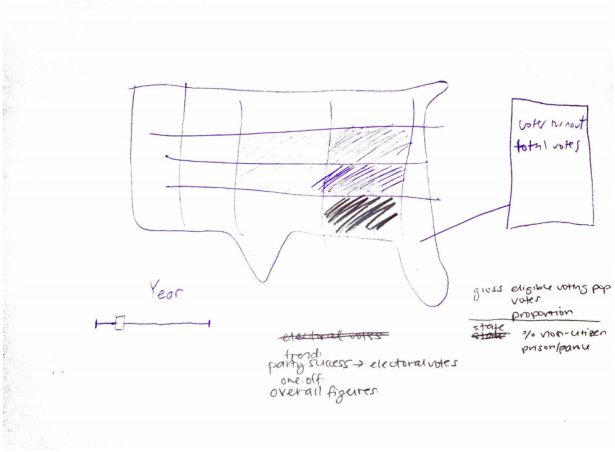
We plan to encode our data geographically. The main visualization will be a map of the United States with each state's information for every presidential election. There will be a slider to represent time, as the data spans all presidential elections. As the user changes the slider, the map will update to show how each state voted for that election.

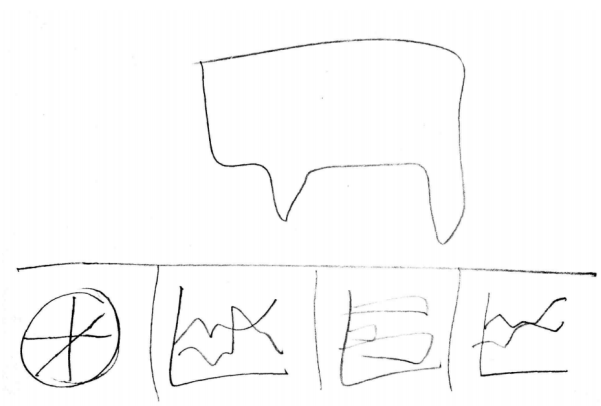
Each state will be encoded by color based on the party of the candidate that they voted for. The intensity/hue of the color will be dependant on what percentage of the state voted for the candidate, where greater intensity will indicate a larger margin of victory.

Additionally the user will be able to click on a state to pull up more detailed statistics for that state during that particular election. For example, this will show statistics like voter turnout, population demographics and income levels. This will be visualized below the graph of the United States in smaller views of pie charts, line graphs, and bar charts (as appropriate for each statistic we want to show).

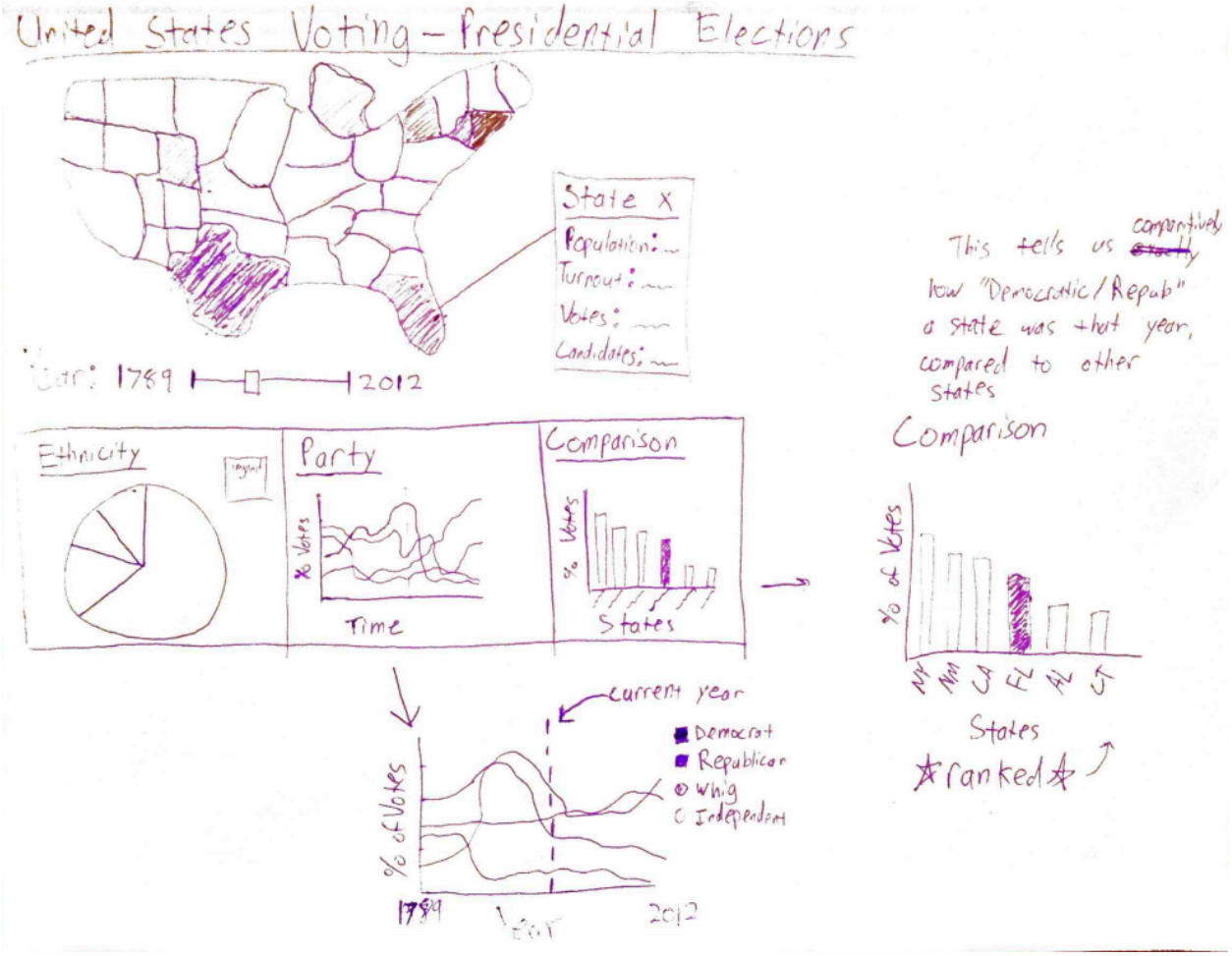
Below are sketches to show the spatial relationships for the visualization. The main view is the map of the united states with a slider below and state-specific data in smaller graphs on the bottom of the page.

Prototype Sketches





Final Sketch



Must-Have Features

These are features without which you would consider your project to be a failure.

1. A primary feature we would want in our visualization is the slider to show change over time for each state's result in the presidential election. The intensity of the color of the state will represent the winning party and by how much they won. This feature will also include political parties that had previously been represented in elections.
2. The user can click on a state to show more detailed statistics for that state during that particular election, which will be represented visually below the United States map.

Optional Features

Those features which you consider would be nice to have, but not critical.

1. Largely dependent on availability of data, we would like to implement the ability to filter voters by statistics (ex. gender, race, age) to show how elections would turn out solely based on certain voters.
2. Another optional feature will be a timeline that will allow the user to enter a story mode, which will display a certain number of significant events in the history of the United States and select the corresponding elections. This will ideally communicate how voting was potentially influenced in that period of time.
3. An option to show only one party (color) on the map would help compare states that voted for the same party.
4. Post-election stats would be another neat optional feature to visualize. Statistics like approval ratings per state could be visualized on the map in a similar way to how each state voted, with the magnitude of the approval rating reflected in the intensity of the color.

Project Schedule

Make sure that you plan your work so that you can avoid a big rush right before the final project deadline, and delegate different modules and responsibilities among your team members. Write this in terms of weekly deadlines.

4/3: Submit project proposal

4/10: Finalize data, logistical aspects, and establish prototype.

- Evan: Collect, process and format all data.
- Ryan: Set up Github pages/other hosting.
- Mengting: Establish prototype for USA map with each state.

4/17: **Milestone 1 due**. Basic functionality of visualization.

- Evan: Year slider functionally updates map.
- Ryan: States can be encoded by color.
- Mengting: Map and all states are functionally present.

4/24: Complete local version of visualization and its back-end.

- Teamwide: Complete functionality for each state on-click (state-specific displays).
- Address corner cases (states lacking data).
- Work on implementing optional features.

5/1: Complete front-end and establish hosting.

5/5: **Final project due**. Code cleanup. Publicize website.