CSE 564 - Final Project Proposal

An Analysis of U.S. General Elections
By Pranav Sista (Group 90)

Background

As I've grown older, I've come to appreciate the results of elections as a way to look into the general sentiments of a country. The United States of America has been unique in this regard considering the extensive amount of data that is available. With a history of more than 200 years, there have been many election cycles to observe and contrast how states have voted. The recent election cycles have been interesting in this regard considering how much more volatile they are compared to the past. I've always been interested in knowing more about how partisan states are, how states have shifted over the years, the turnout in various states and how that affects their results, and how often people vote for certain parties. I've additionally developed a fascination with how demographics have generally voted over time.

Various websites have helped in answering these questions by representing the data in unique visual forms. The now defunct website FiveThirthyEight was a valuable resource, and I attribute it to making me interested in election results. So much so that it has motivated me to visualize the data in a way which I find interesting. With this project, and the skills I have learned throughout this course, I believe that I now have the opportunity to do this.

Problem Statement

The goal of this project is to develop an interface which will allow a user to understand how states and demographics have voted in the general elections which have recently taken place in the United States. They will be able to do so by interacting with a map of the U.S. that displays the results of the elections. There will also be visualizations that help quantify the margin by which a candidate has won. At a state level, the user should be able to observe how each county (and possibly congressional district) voted, and to what extent there was any split ticket voting. The tool should also be able to highlight the turnout in these elections.

Technical Implementation

This project will be implemented in a server-client manner. The server will be implemented using Python and Flask. The client will be implemented using HTML, CSS, Bootstrap, and D3.js. When the user interacts with the client interface, the server will process the data accordingly and provide the view the user wants. If the need arises, other tools will be used.

Datasets

The following datasets will be used for this project -

- **Results of past U.S. Presidential elections in the 21st century.** This has been collected from Harvard Dataverse. The data for the 2024 election has not been made available on their website yet, so I will be relying on a dataset in GitHub which has scraped the election results from official cable news networks. The results will be collected at the county level.

- **Turnout data in U.S. general elections.** This data has been collected from the University of Florida's election lab.
- **The Demographic breakdown of counties in the U.S.** This data will include a breakdown by age, race, and sex. Information on the median household income in each county will also be used. This data was collected during inter censual surveys and has been downloaded from the U.S. census bureau's website.
- **Results of past U.S. House and Senate elections.** This data will be needed to understand to what extent split ticket voting occurs. This data has been taken from the MIT election website, and the results from the 2024 election have been scraped from Fox News.
- **Results of the U.S. Presidential elections by congressional district.** This data will be used to check how each congressional district voted for a candidate at the presidential level. It has been collected from a Kaggle dataset.

Apart from the listed datasets, I will try to use any other datasets I find that can assist with the visualizations.

Approach

For representing the results at the national level, the following plots will be present -

1. A geographical map of the U.S. with each state colored by the party that won it. The shade of the color will convery by what margin it was won by a particular party.
2. A bar chart that displays the popular vote (a bar chart)
3. Either a pie or donut chart for displaying the electoral college vote (a pie or donut chart). The electoral college plot will also have the margin of victory in each state reflected in it.
4. A parallel coordinate plot will be used to show the characteristics of a county and how they voted in the election.
5. A scatter plot that maps the turnout of each state with the margin of victory.

The geographical map will be interactive. The user can select a state of interest and learn more about how that state voted. If a state is chosen, the following options will be available -

1. There will be a bar plot that shows how the state as a whole voted for president.
2. There will be another bar plot that shows the results of each House race and the senate race if applicable.
3. The parallel coordinate plot from before will be present but only with the counties from the state highlighted.
4. There will also be some way to compare the house, senate, and presidential elections with each other (I'm still unsure about the best way to implement this).

Another aspect I would like to visualize is how much voters have shifted either away or towards a political party compared to another election cycle. I will probably convey this information by representing the shift in the map or in a separate plot. It is also possible that I don't display the plots above should I find a better plot to replace them.