Election Inspection

Name of the group members and CNetIDs
Alejandro Navarrete ID: 12324829
Victor Perez, ID: 12277410
Christian Jordan, ID: 12332110

A brief overview of the final project (200 words maximum)
The aim of our project is to analyze how the different proposed maps from the 2020 Michigan Congressional Redistricting process would affect voter participation. Using the most recent Census and voter file data, we estimate multiple linear regression models of voter participation on voting district demographic characteristics and previous election metrics, like participation and vote share difference between the winning candidate and the second best candidate. We use a forward selection algorithm to pick a model, and then we estimate the predicted turnout in the average voting district (VTD) for the different proposed maps in an attempt to understand how redistricting changes the configuration of the districts and the voting districts within them. These figures are then presented along on our dashboard at the per-district level. Along with those figures, we have an interactive map of a given drafted map associated where you can see what geographic area each possible district encompasses. Using an API to Google's Civic Information hub, selecting a district also provides the most recent information on the district's representation in Congress.

The overall structure of the software (1-page maximum). It would be nice to include a helpful diagram of how the modules are connected with each other but this is not required.
The program is structured between two tiers of packages. The top-most includes the initializing and execution files. Below that are three packages split amongst the three components of this project's analysis, data cleaning, statistical analysis, and visualization. In that order is the flow of data that we gathered through our program. The data cleaning subpackage cleans and orders the data. Then, it outputs the final clean data to the statistical analysis subpackage. The regression and reconfiguration of the data into geo files is done in the statistical subpackage. The output of this process is then placed into the visualization subpackage

A description on the code responsibilities for each group member (i.e., who was responsible for what module, files, tasks, etc.).
Christian:
- Collect Census Bureau data from API
- Clean data and output a csv for analysis
- Create code book for analysis

- Create run script for running the analysis

Alejandro:
- Dashboard and visualization creation.
- Initial data search and collection from static internet sources and for API.

Victor:
- Estimate model
- Generate geojson files for dashboard and visualization

Short description on how to interact with the application and what it produces.

The first dropdown menu, for 'map version', allows you to choose what map draft version will be displayed. Along with that, its predicted turnout figures and demographic information will be shown in the table to the right.

The second dropdown menu allows you to choose the statistic for that district that will be visible in the tooltip resulting from hovering or selecting a district within the map. For further visualization, this attribute of the district will also be used as the metric determining the shading of the 'choropleth map' of each district.

If the option to 'reveal VTDs' is selected, the 'choropleth' map will be done at the per-VTD level to show the composition of the districts as it pertains to the selected attribute. Given that there are thousands of VTDs in the state, selecting this option will be slower than loading other versions of the map.

At the bottom of the page is space for displaying the up-to-date information of the selected district's representative information from Google's Civic Information API. Selecting the district from the table will update that information to that specific district.

What the project tried to accomplish and what it actually accomplished (200 words)

The project started out with an aim to create an educational tool that would allow citizens to compare the impact of the different congressional district configurations on ultimate voter turnout. We identified that similar tools focus mainly on how redistricting affects the winner of the election, but not the potential effect on voter participation which could be more relevant to citizens and activists. Understanding that voter participation is an outcome dependent on environmental variables, we sought to model and visualize the different expected outcomes possible by shifting district boundaries around. For this purpose, we collected data from the US Census Bureau and Redistricting Data Hub to estimate the determinants of voter turnout in Michigan. We used a forward selection algorithm to find the best linear model. Then we designed a dashboard to visualize the relationship between voter turnout and the different congressional districts, along other socio-demographic variables. Additionally, our visualization shows the current Congress Representative and their contact information. Due to the lack of recent data at the VTD level, we had to focus our analysis at the average VTD within a district level, which isn't what we initially intended. Although that limits the prediction of our model, we still