First Visualization & Analysis

2025-11-24

Nkemdi and Vasia

```
library(haven)

climatedata <- read_sav("CCAM SPSS Data 2008-2024.sav")
```

RESEARCH QUESTION AND HYPOTHESIS - RESEARCH QUESTION: How does personal concern about climate change differ by age, region, income, education, and political party? How can campaigns leverage these differences in order to appeal to different groups of voters? This research question is important to political analytics because political opinions are largely influenced by the framing of issues in the media, and discovering how this opinion changes can help explain what types of framing are working, and potentially help contribute to more effective policy actions and media considerations on the issue of climate change. - HYPOTHESIS: We intend to create a climate concern index that yields conclusions about how individuals in both political parties relate to, think about, and show concern regarding climate change in the United States. We hypothesize that while republican voters will score lower on the climate concern index compared to democrats, these scores may vary by education, region, age and income.

DATA DESCRIPTION – CCAM SPSS Data 2008-2024 - This data set is from a research project called "Climate Change in the American Mind," which puts together data about how opinions on climate change have changed over time. The data table is mostly time-series observational survey data, and includes questions about how people feel about climate change. The survey was administered over multiple years, granting us the ability to see how participant answers evolved over time, specifically from 2008 to 2024. The unit of analysis for this table is individuals, as the data set tabulates survey information from a sample size of approximately 35,000 U.S adults. The data set measures "global warming beliefs and attitudes, risk perceptions, policy preferences, and information acquisition behaviors" (Yale Change Communication, Explore Climate Change in the American Mind). The data set also includes information about other countries, but we would focus on the United States. Demographic information such as gender, age, generation, education, and income are also included.

There are many relevant variables for our analysis. Specifically on the harm aspect, some relevant variables include "harm_plants_animals, harm_future_gen, Harm_dev_countries, harm_US, harm_personally" which seek to get at different ways in which participants react to the harm done to the environment, which of course impacts their level of concern. Other relevant variables for our study specifically, include "worry", "year", and then some of the demographic variables. The data has sampling weights, and were typically weighted by respondents' age, gender, race/ethnicity, Census region, metropolitan status, education, and income. These are represented by the variables "weight_wave" and "weight_aggregate."

PRELIMINARY VISUALIZATION AND ANALYSIS - The actual visualizations we created is in a separate document titled "firstdatavisfinalf25.rmd" in our Github repository, and are titled figures 1 through figure 5 - Link: https://github.com/vat2129/finalprojectf25/blob/main/firstdatavisfinalf25.Rmd - Visualization analysis - FIGURE 1: CLIMATE CONCERN BY REGION AND PARTY Here we are visualizing climate concern which is an index that we created utilizing six survey questions from the data set. The highest concern score someone could achieve is 4, meaning that were very worried about climate change, while the lowest score possible would be a 1 which would mean you were not concerned about climate change in any way. In Figure 1 we are visualize concern by region of the country separated out by party. For this visual

we took the previous cleaned data (climate data) that included our questions of interest and demographic information, and that created the new climate concern index. We then took this cleaned data and grouped it by region and party, as well as created a weighted mean with the aggregate weights including in the data set. We also filtered out any NA values. When we wanted to take a closer look at region, the only visual that felt appropriate was a map. As visible in the table above, even dividing by region and party we do not have a subsection with n less than 366 (Independents in New England). What we visualize here aligns very closely with what our hypothesis was. Party seemed to be a much stronger indicator of climate concern than region with Democrats being most concerned, then Independents, and then Republicans. There was only very slight concern variations by region.

Next Steps - As seen on our github repository, we have already gone through and greated quite a few visuals that help to address our research question. - Besides these figures, we intend to look specifically at republicans and independents to evluate whether climate concern evolves over time when impacted by variables such as age, education, income, and region/location. While we assume that the results for the democratic party can largely be predicted, we will also include democrats in the analysis for comparision purposes. - Our main point of analysis will come from comparing results on the climate concern scale and the evolution of climate concern by region, which will allow us to begin to understand how public opinion on climate change has changed over the course of 3 presidencies and multiple nationwide stances on climate change. - When it comes to the presentation and the final paper, we also intend to include analysis about the harm or consequences of certain climate change proposals and opinions that were discussed in this survey and dataset, as well as the ethics and implications of this study on efforts to address climate change in the United States.