

Leigh Wright

CSC 2730

Project

Dr. Zhang

## METHOD

For my project, I wanted to really use something that would allow me to get some experience with data visualization and using pandas. The problem I considered had to deal with Hate Groups as designated by the Southern Poverty Law Center (SPLC); given the increasing political polarization within the US that we are recently experiencing, I wanted to see if there were any correlations between various hate groups and any demographic.

The data I used was taken from the census bureau; this gave me the percentages of different demographic measurements for each state: poverty, breakdown by race, unemployment rates, etc. The other data was taken off the SPLC's website; I had to figure out how many hate groups each state had, *and* how many hate groups of each type a state had (for example, how many black nationalist groups a state had, or how many Neo Nazi groups a state had). Getting that information from the SPLC was a time consuming; I had to make my own CSV by hand to get the data.

The majority of my time spent on programming was just cleaning the data; getting rid of NaN's, and combining the two CSV's into one data frame. Once that was finished, I was able to begin making data visualizations. In particular, I wanted to examine how poverty had an effect on the total number of hate groups a state had.

## RESULTS

There was definitely a correlation between poverty and the overall number of hate groups a state had. I used a lmplo from the seaborn library for the data visualization. The data visualization worked perfectly, but I would have like to have gone back and examine further how to better predict causes of hate groups, so I'm not really satisfied with how the model predicts.