Health and Voting

Group 1

Research Question and Hypothesis

Empirical Puzzle: The Healthcare Paradox

People who would benefit from state provided health insurance oppose policies which would expand it

Key questions for the project:

1) Does the paradox exist?
2) How can we explain this paradox?
    a) Does health insurance actually improve health outcomes?
    b) Do highly uninsured counties even have access to healthcare?
    c) Are people unaware of the benefits of having insurance?
    d) Is healthcare considered a key voting issue?

Hypothesis: the health paradox does exist however it is viewed as a less important voting issue than other topics which explains why people who are uninsured vote against policies like the ACA

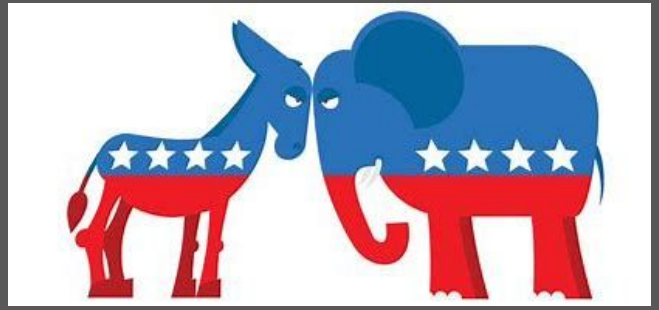
Health and Socioeconomic Data

Data sets were pulled from-

https://www.countyhealthrankings.org/explore-health-rankings/rankings-data-documentation

We decided to analyze our data set on the county level. This website was helpful because we were able to split up this data into multiple variables. This was a good way to compare demographics within each county. For example, we were able to get in depth analysis of specific data sets such as median income, which we could split across races. In total, this data set had over 700 health and socioeconomic variables.

Voting Data



Our voting data set comes from-

https://github.com/tonmcg/US_County_Level_Election_Results_08-20

This repository combines voting data from from: The Guardian, townhall.com, Fox News, Politico, and the New York Times.

Dimensions: The data is 10 columns and 3152 rows. The columns represent the names and identifiers of the states as well as the voting data, and the rows represent each of the voting districts.

Analyses Individual

Health: Currently we have analyzed measures of overall health, mental health, physical health, life expectancy. Each of these measures have a strong correlation with each other. As overall health increases, as does mental health, physical health, and life expectancy. Between these variables mental and physical health have the highest correlation to each other.

Voting: The data showed a staunch association between how rural a county was and how Republican it voted. There were far more counties that voted Republican than Democrat, and those that did were more polarized. However, counties that voted Democrat were considerably larger on average. Alaska is not pictured because it records its data differently than other states.

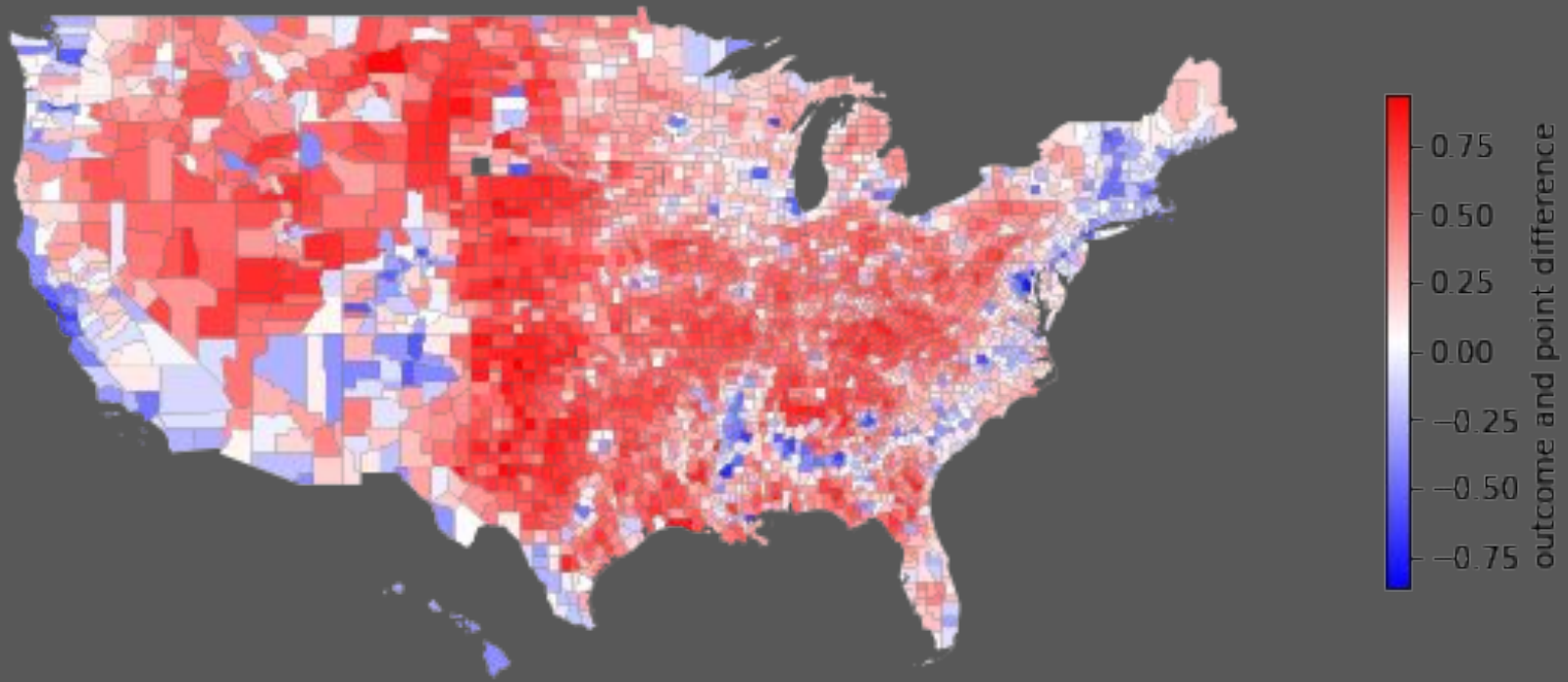
Socioeconomic: In this notebook, we explored a variety of socioeconomic indicators such as income, rural population, housing problems, and, crucially, levels of insurance. We hope that this analysis will help us to explain the healthcare paradox.

Analyses Combined

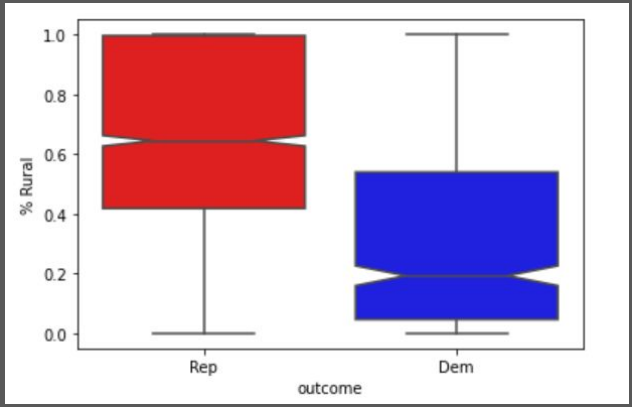
To combine and analyze our data we are essentially looking at the variables we have selected separated by political party, and analyzing how multiple variables compare when separated by political party.

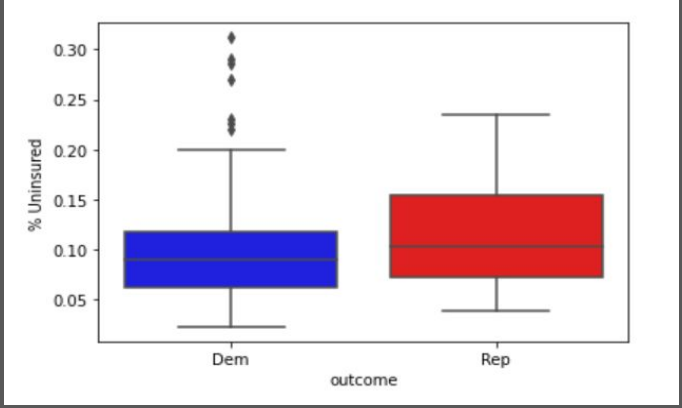
First we have started to look at the spread of our variables split up by political party using violin and box plots. Then we started to look at how insurance rates affect health to establish a relationship between the two. From there we have started to look at other variables compared to insurance rate, and the correlation between these variables when separated by political party.

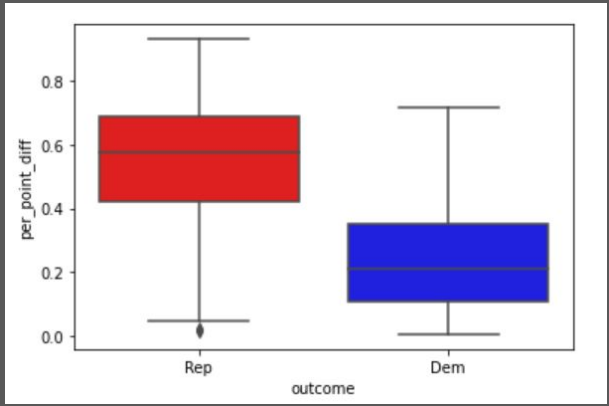
Voting Results

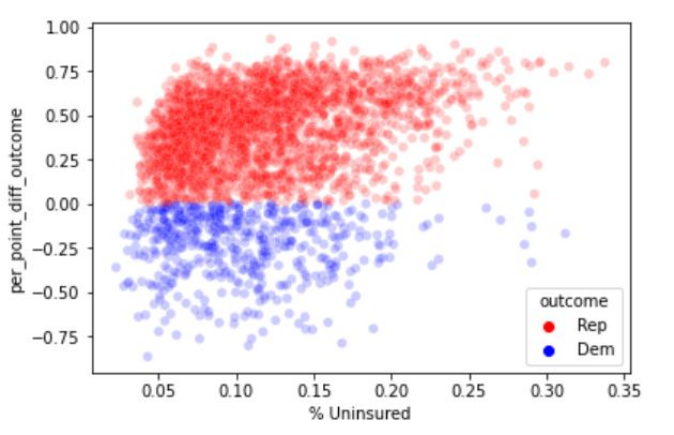


Key Findings Socioeconomic

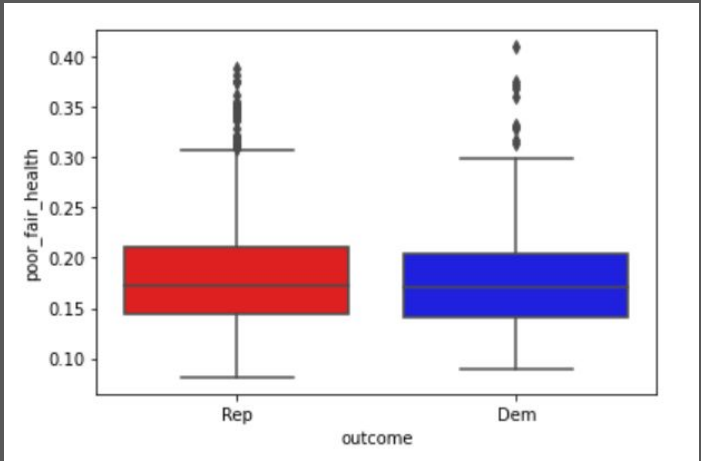


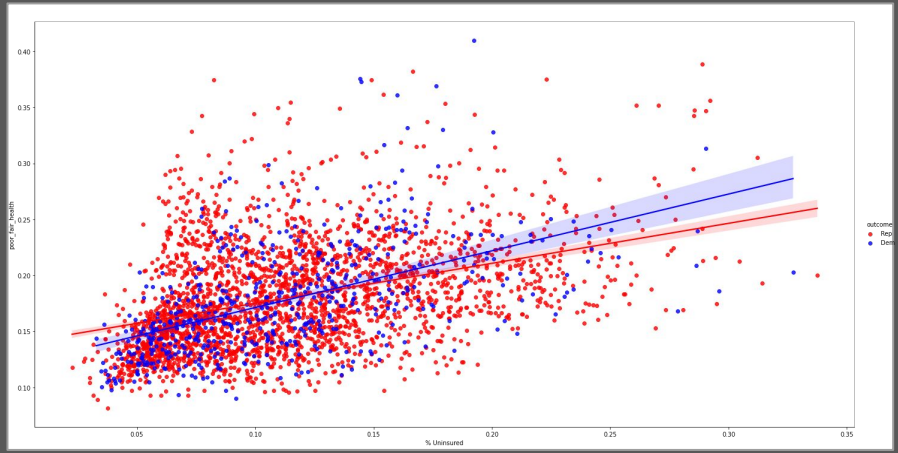


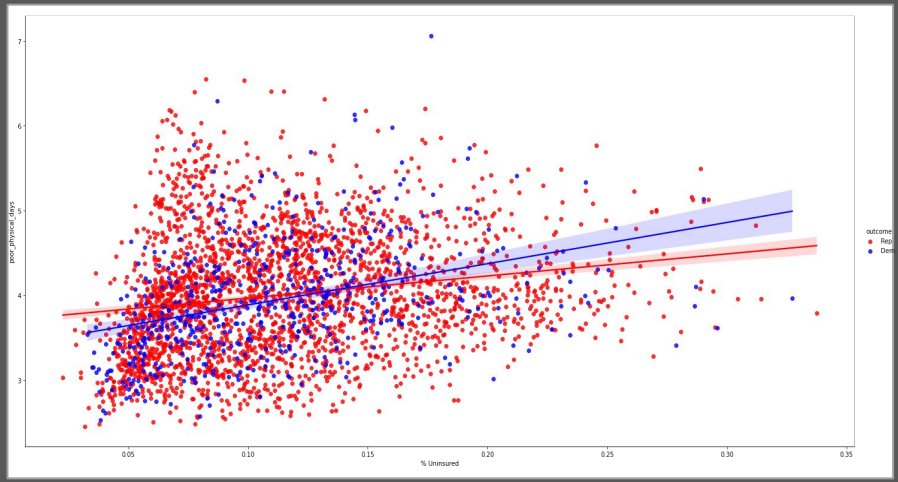




Key Findings Health







Next Steps

Continue combined analysis

Look at health care providers by county to determine the access to healthcare

Machine Learning

Potentially comparing results from 2016