Project Title: Election Preferences in Ohio, 2016

Team Ohio is Ohio: Arash, Mike C., Tessa Pinon, Kristina

Why are we interested in Ohio?

\*Math: third largest swing state (18 electoral votes)

\*Competitive state to win due to its diversity - religion, race, age, and economic

\*No Republican has won the candidacy without Ohio ; JFK the only president to win the presidency without Ohio

\*Considered a Republican state

\*Trade is an issue, immigration is not

Project Description: Breakdown of voter preferences by county data in Ohio compared between 2012 and 2016 in Ohio; further characterized by voter income, race, medicare reimbursement, health insurance status, and education level.

Research questions: Is there a correlation between industry type and candidate preference? Did registered parties vote the same in 2012 and 2016? Did demographics affect voter preference between elections? Is there a correlation between voter preference in 2016 by percentage of people receiving Medicare reimbursement and health insurance status? How does voter demographics (such as race, education, and party affiliation) affect voting results? Using the Twitter API and sentiment analysis, what was the mood of the campaign speeches and tweets?

Data sets: Ohio L2 voter files ; datausa.ohio, census bureau, The American Presidency Project, countyhealthrankings.org

Discussion:

Overall, our data analysis characterizes the voter population in Ohio in 2012 and 2016; noting any significant changes in the population, which may have shifted the outcome.

Electoral Map Correlation

There were more Republican voters in 2016 vs 2012. The electoral map demonstrates that voters stayed within their respective parties between the two election years, but there were more Republican voters.

Medicare

We investigated the relationship between the percentage of Ohio voters receiving Medicare enrollments who voted for Trump and Clinton by county. The low calculated p values for these data sets indicate that among Trump voters, there is a significant positive relationship between Medicare enrollment and election outcome. Conversely, there is a negative relationship among Hillary voters and election outcome. However, there is not a strong linear regression trend (low R2 value) overall between Medicare enrollment and voter preference.

Uninsured

The low calculated p values for these data sets indicate that there is not a significant correlation among those who are uninsured among Trump and Clinton voters. Additionally, there was no general trend when these data sets were plotted.

Median Income by County 2016

There is a slight relationship between voter income by county for Trump voters. There is no significant relationship between voter income among Clinton voters.

Vader Sentiment Scores

We wanted to examine some information about the presidential candidates, since we had a lot of data on Ohio voter information (income, education, ethnicity, etc.) We were interested the campaigning of each candidate in the state of Ohio. We used pandas to locate the city in in common in which each candidate had a campaign rally. That city was Cincinnati, we pulled the transcript from each speech and used vader sentiment analyzer. Since vader is typically used for short text, we looked in the vader github for a solution for using vader sentiment for longer text and used vader sentiment in conjunction with nltk to run sentiment analysis on each sentence in the speech and average the all sentences.

According to vader compound scores Republicans speeches were more positive than the Democrats. Trump had the most positive speech and Obama had the lowest compound score.

We wanted to see if the speech sentiment had a correlation to how the candidate is in real life, so we used a twitter API for pulling all candidates last 100 tweets a ran a sentiment analysis on those tweets. Obama was the most positive tweeter and Trump was the most negative according to compound vader scores.

L2 Voter Demographics

Examining race, there was a higher probability of voting among white voters. However, there was a decrease in African American voters by 1% between 2012 and 2016. This is also affected the probability of African American voters in 2016. By party affiliation, there was an increase in independent voters, which decreased the number of voters in the both the Democratic and Republican parties. Interestingly, Republicans were more likely to vote in 2016 vs 2012, in comparison to Democrats. Education levels among voters did have any significant changes between the two election years.