Political Ad Effectiveness in the U.S. Swing States

W200 Final Report
August 13th, 2021
Joyce Li, Jeremy Yeung, Amy Jung, and Noor Gill

I. Abstract:

The 2016 presidential election was one of the closest and most polarizing elections in United States history. Donald Trump defeated Hilary Clinton due to a majority in the electoral vote, although he lost the popular vote. The election was also monumental because of the harsh and unconventional rhetoric that candidates used to get ahead, especially through marketing campaigns. To provide some background, the United States presidential election works through an Electoral College system, in which each state gets a certain number of electors based on its total number of representatives in Congress. There are a total of 538 electoral votes, and the candidate with more than 270 votes wins the election regardless of the result of the popular vote[1]. Due to the "winner-takes-all" method most states use to determine their presidential electors, candidates often campaign extensively in competitive states, resulting in a majority of presidential election advertisements flooding broadcast television in select states. The states that are most crucial to a president's victory are swing states, which typically have very competitive polling results and do not historically side towards any particular political party. In 2016, the swing states were Nevada, Arizona, Colorado, Iowa, Wisconsin, Michigan, Ohio, Pennsylvania, Virginia, North Carolina, Florida, New Hampshire[2]. Our research aims to explore relationships between presidential election ads, voting statistics, and election results. We hope that this analysis can be used to better understand the unique outcome of the 2016 election, and possibly guide or inform future election advertisement campaign strategies.

II. Data Overview:

We used two main data sources to conduct our analysis: the Political TV Ad Archive Airings[3] Dataset and the 2016 Census Voter Registration Dataset[4]. The Ad Dataset contains information on ads aired during the election cycle, while the census data that we pulled contains information on state populations and voter registration rates. Since our supplementary dataset was collected by the U.S. Census Bureau, it can be deemed credible. To join these two datasets, we manually added a state column in the Ad Dataset, slicing text from the location column. We joined the tables on state, allowing us to access all the data in one table.

In order to continue with our research, we made key assumptions regarding swing states and features of political advertisements:

1. Although Michigan was determined to be a swing state in the 2016 election since we did not have

[1] Presidential election process. USAGov. (n.d.). https://www.usa.gov/election#:~:text=In%20the%20Electoral%20College%20system,(270)%20wins%20the%20election

[2] Presidential battleground States, 2016. Ballotpedia. (n.d.). https://ballotpedia.org/Presidential_battleground_states,_2016.

[3] "Political TV Ad Archive» Data." 2016. Politicaladarchive.org. 2016. http://politicaladarchive.org/data/.

[4] Bureau, US Census. n.d. "Voting and Registration in the Election of November 2016." The United States Census Bureau. https://www.census.gov/data/tables/time-series/demo/voting-and-registration/p20-580.html.

advertisement data for Michigan, it was excluded from our analysis. We assumed having eleven of twelve swing states was enough to make a conclusion about swing states in general.

2. "Pro" and "Con" advertisement messages are most useful in determining advertisement sentiment. Although advertisement messages were labeled as either pro, con, mixed, or unknown, for the purpose of our research question, we focused on only the pro and con message categories as they were more revelatory and easily distinguishable.
3. The time frame chosen was from October 2015 to November 2016 (the day of the election) as we assumed that aired advertisements would be most effective within this period.
4. All ads aired within this time period are captured in the dataset so there were no missing ads.

III. Research Question:

In the 2016 presidential election, was there an association between the number and type of ads displayed in a swing state and the percentage of voter registration or election winner?
Note: Ad types include Pro/Con-Candidate Message, News/Not News Source, and Ad Duration.

IV. Methodology:

The first step that we took to begin our analysis was to clean our data. We dropped unnecessary columns not needed for analysis, such as id, wp_identifier, archive_id, cycle, and race. We then trimmed the ads aired dataset to dates before the election, since the original dataset included ads aired between October 2015 and December 2016. Then, we filtered the candidates to include only presidential election candidates Hillary Clinton and Donald Trump. We also dropped null values and deleted rows with values that didn't exist, since null values cannot be analyzed.

Next we had to clean the voter registration dataset. Again, we dropped unnecessary columns and then converted string columns such as "Total Population" to numerical values. To join the voter registration and ads dataset together, we created a State column and used regular expression and string slicing to retrieve state names and create a common variable in both datasets. After joining, we used numerical encoding to analyze categorical variables. For the "Message" column, we replaced "pro" with 1 and "con" with 0. Similarly, for the "News" column, we encoded "news" with 1 and "not news" with 0.

Lastly, we performed feature engineering to create new columns to analyze such as Duration, State, Swing, and Winner. Duration represents the length of the ad in seconds. To create this column, we subtracted start time from end time to get the duration of an ad. We used information from Ballotpedia to generate a column "Swing", classifying states into Democratic, Republican, or Swing States and also a "Winner" column showing the candidate that won in each state[5].

To gain a better understanding of the dataset, we performed exploratory data analysis on different columns.

[5] Presidential battleground States, 2016. Ballotpedia. (n.d.). https://ballotpedia.org/Presidential_battleground_states,_2016.

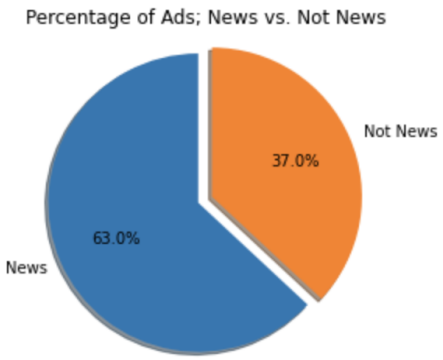

Fig 0.1: Program Type

For program types, we found that around ⅔ of the ads were aired on news programs, with the rest of the ads being aired on other programs. When looking at the distribution of pro/con-candidate messages, we found three times more con-candidate than pro-candidate ads for Trump than Clinton. Clinton also had almost two and a half times more pro-candidate ads than Trump.

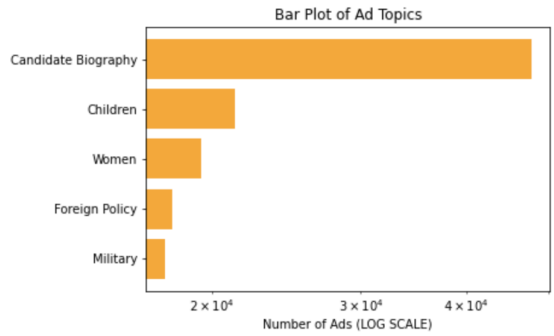

Fig 0.2: Ad topics (log scale)

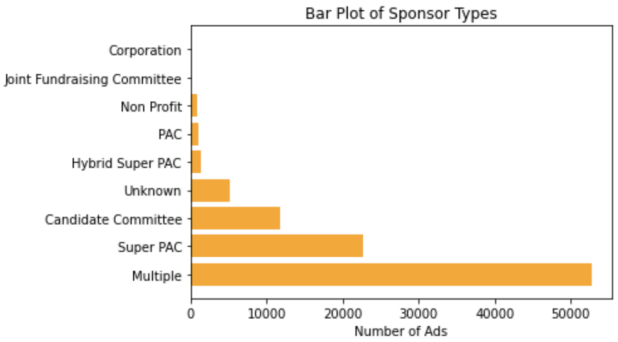

Fig 0.3: Sponsor Types

In regards to ad topics, we plotted the topics (Fig 0.2) and found the top five most discussed were candidate biography, children, women, foreign policy, and the military. In Fig 0.3, we found that the biggest sponsors for ads were Super PACs and candidate committees. We also found that there were no con-candidate ads for Clinton and only two pro-candidate ads for Trump in Virginia.

To answer the research question, we decided to investigate four key variables concerning voter registration, voting percentage, and the election winner. These variables include unique ads per candidate, program type (news vs. not news), pro vs. con ads, and average ad duration per candidate for each swing state. For each variable, we generated visualizations and analyzed statistics such as Pearson's correlation to form a conclusion about the research question.

V. Results:

Section 1: Unique Ads per Candidate per Swing State

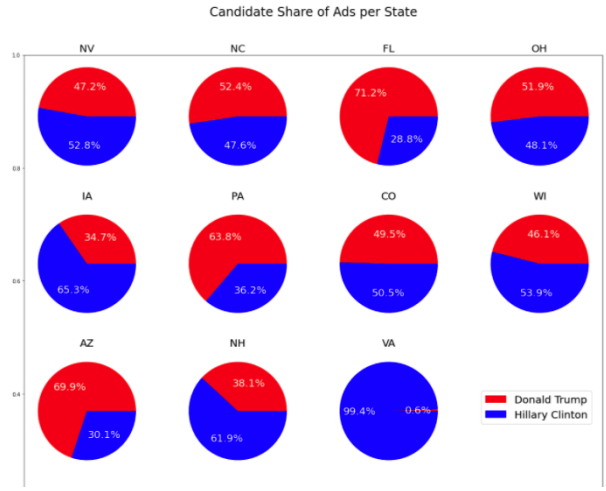

Fig 1.1 Candidate Share of Ads Per Swing State

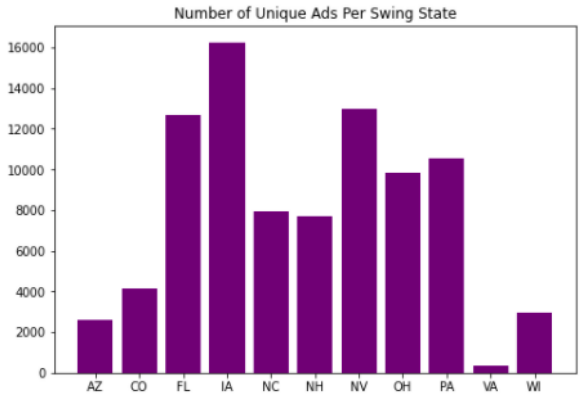

Fig 1.2: Number of Unique Ads per Swing State

Fig 1.1 shows that swing states with the most share of ads about Trump were Florida, Arizona, Pennsylvania, and states with the most share of ads about Clinton were Virginia, Iowa, New Hampshire. Fig 1.2 demonstrates that Iowa and Nevada had the most ads shown pre-election. Virginia had the least by far, which makes sense referring back to Fig 1.1 since there were almost no Trump ads in the state.

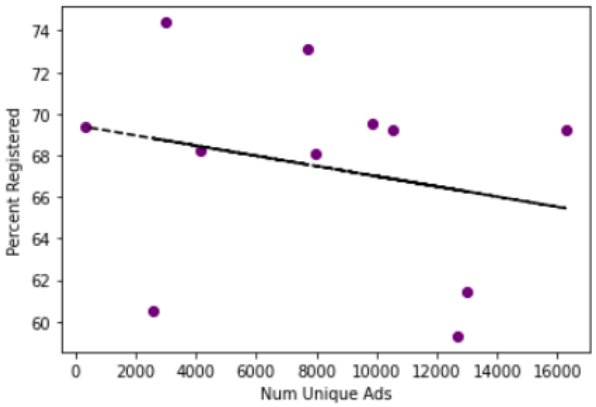

Fig 1.3: Percent Registration vs. Num Unique Ads

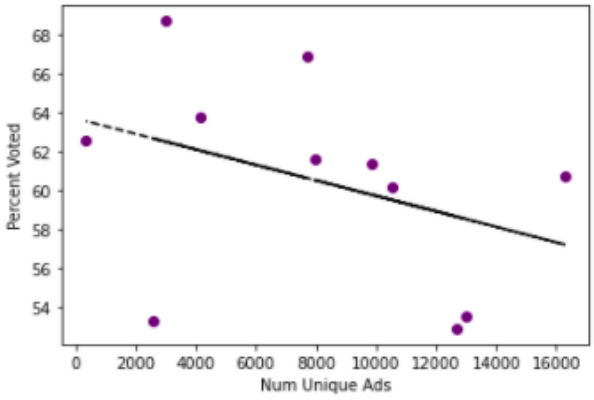

Fig 1.4: Pct Voted vs Num Unique Ads

The graphs above (Fig 1.3 and 1.4) show the negative associations between the number of ads in a swing state and the percentage of people who registered to vote and the percentage of people who voted. This implies that either ads were not effective in convincing candidates to vote, or that candidates were targeting states with low voter registration rates by showing more ads in an effort to increase voter turnout.

Besides the graphs shown, we also observed that in swing states, the candidate with more ads about them won 82% of the time. Breaking this down by candidate, we discovered that in states with more ads about them, Trump won 100% of the time, while Clinton only won 67% of the time.

Section 2: Program Type (News vs. Not News) per Candidate per Swing State

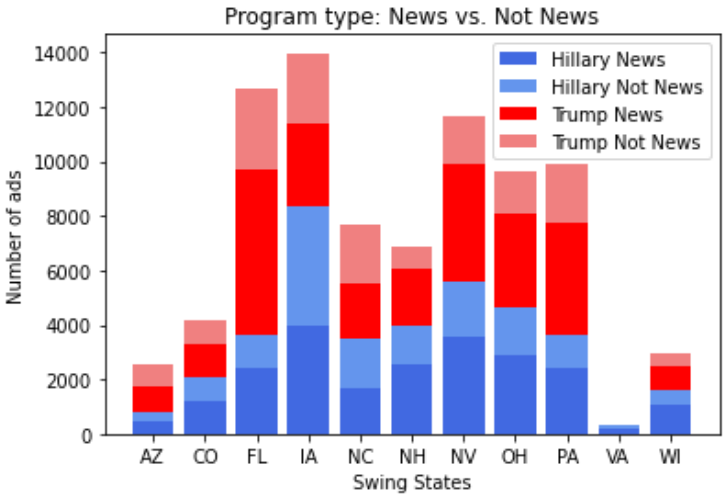

Fig 2.1: News vs. Not News per State

Hillary difference

	state	news - not news counts	percent_voted	total_percent_reg	winner	result
0	AZ	100	53.3	60.5	Donald Trump	-
1	CO	338	63.8	68.2	Hillary Clinton	win
2	FL	1207	52.9	59.3	Donald Trump	-
3	IA	-374	60.7	69.2	Donald Trump	-
4	NC	-95	61.6	68.1	Donald Trump	-
5	NH	1120	66.9	73.1	Hillary Clinton	win
6	NV	1532	53.5	61.4	Hillary Clinton	win
7	OH	1153	61.4	69.5	Donald Trump	-
8	PA	1133	60.2	69.2	Donald Trump	-
9	VA	16	62.6	69.4	Hillary Clinton	win
10	WI	549	68.7	74.4	Donald Trump	-

Fig 2.2: Difference in News vs. Not News for Clinton

Initial analysis indicates that both candidates had advertisement quantities for either program type that did not differ greatly per state. The program type is defined by the program the ad was shown during, regardless of the television network. The graph in Fig 2.1 shows that although the number of ads per state varied from less than 2,000 to almost 14,000 ads, neither candidate ad dominated a certain state through a specific program type. The difference in news vs. not news per state for Clinton, Fig 2.2, interestingly highlights that for the two states where Clinton had more "not news" ads, she did not win the votes, but the difference is very small in comparison to the number of ads that were played for both program types in either state (3.5% of total Clinton ads in Iowa, 2.50% of total Clinton ads in North Carolina). In the rest of Fig 2.2, the numbers show that the program type in which the ads were aired did not show a clear association with the result of who won the electoral votes nor the percent of residents who voted in the presidential election for that state.

Section 3: Pro vs. Con Ads per Candidate per Swing State

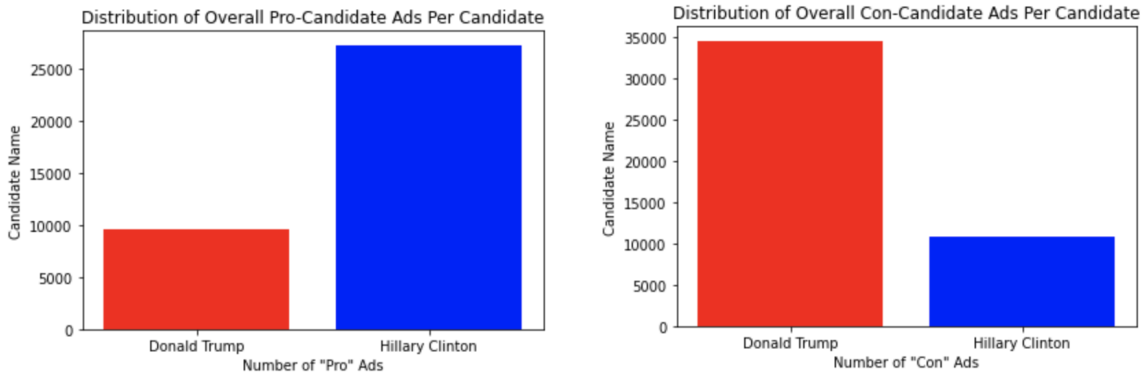

Fig 3.1: Distribution of Overall Pro-Candidate and Con-Candidate Ads for each candidate

Based on the representation of pro and con-candidate ads for both candidates across all states (Fig 3.1), it is evident that Clinton had over 250% more pro-candidate ads about her compared to Trump, whereas Trump had over 300% more con-candidate ads about him compared to Clinton. Collectively,

with regards to the average number of pro and con-candidate ads for the candidates, the average number of con-candidate ads is greater (22,760) than the average number of pro-candidate ads (18,460).

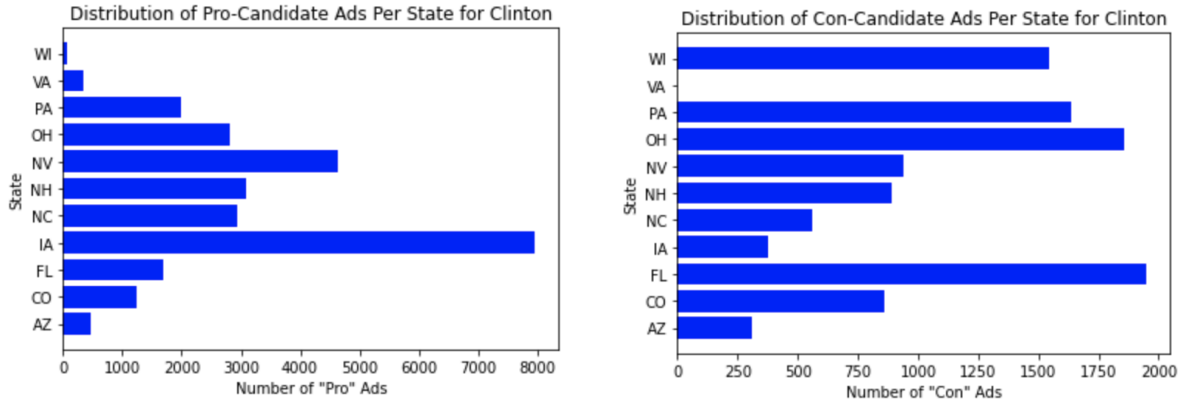

Fig 3.2: Distribution of Pro-Candidate/Con-Candidate Ads Per State for Clinton

Clinton had the least pro-candidate ads about her in the swing states of Wisconsin and Virginia. On the other hand, she had the most positive advertisements in Iowa, followed by Nevada. Clinton had the least con-candidate ads about her in the swing states of Virginia, Arizona, and Iowa. She had the most negative advertisements in Florida, followed by Ohio and Pennsylvania.

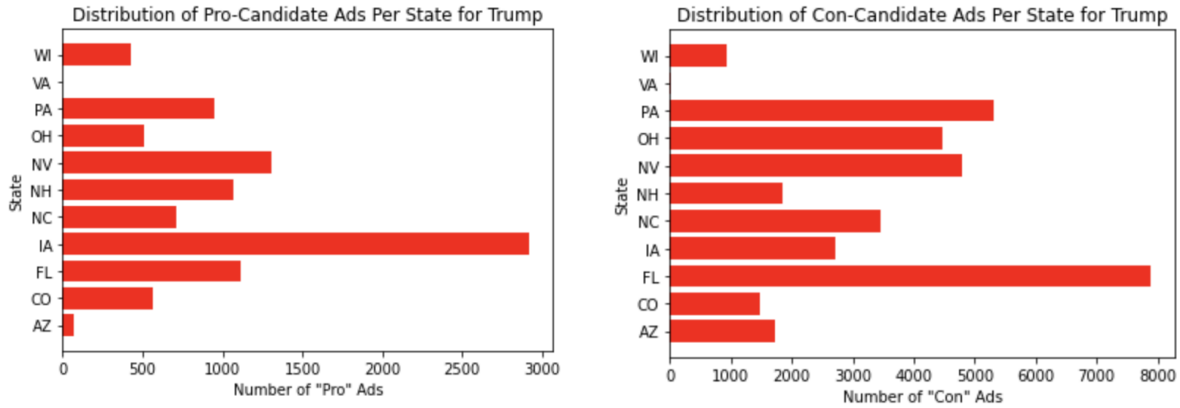

Fig 3.3: Distribution of Pro-Candidate/Con-Candidate Ads Per State for Trump

While Trump had the least pro-candidate ads about him in the swing states of Virginia and Arizona, he had the most pro-candidate ads in Iowa, followed by Nevada and Florida Trump had the least con-candidate ads about him in Virginia and Colorado. Conversely, he had the most con-candidate ads about him in the swing state of Florida, followed by Pennsylvania and Nevada.

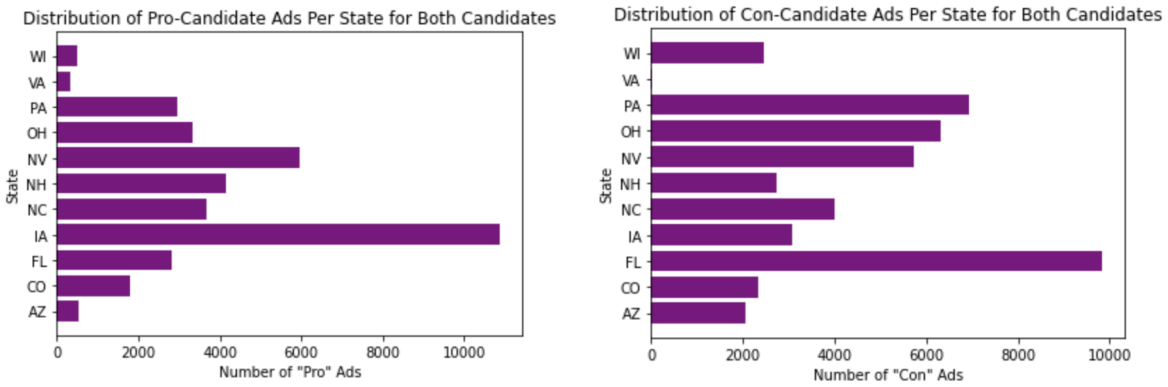

Fig 3.4: Distribution of Pro and Con-Candidate Ads Per State for Both Candidates

Across all states for both candidates, the most pro-candidate ads were aired in Iowa, the most con-candidate ads were aired in Florida, and the least pro-candidate and least con-candidate ads were aired in the swing state of Virginia. As depicted in Fig 3.4, the number of pro and con-candidate ads across states for both candidates was more evenly distributed for con-candidate advertisements and less evenly distributed for pro-candidate advertisements. Like we observed on a candidate-level, on a national level, the average number of combined con-candidate ads (4,138) is also greater than the average number of combined pro-candidate ads (3,356).

Based on our descriptive analysis, the correlation between message polarity and the total percentage of the population registered to vote is negative in both swing and non-swing states. However, the Pearson correlation coefficient for non-swing states is much greater than for swing states with a value of -0.75 compared to -0.29, indicating a weaker relationship between these variables for swing states.

Following a similar pattern, the correlation between message polarity and the total percentage of the population that voted is also negative in swing and non-swing states. The Pearson correlation coefficient for non-swing states is slightly greater than for swing states with a value of -0.40 compared to -0.67, indicating a weaker relationship between these variables for swing states as well.

Section 4: Ad Duration per Candidate per Swing State

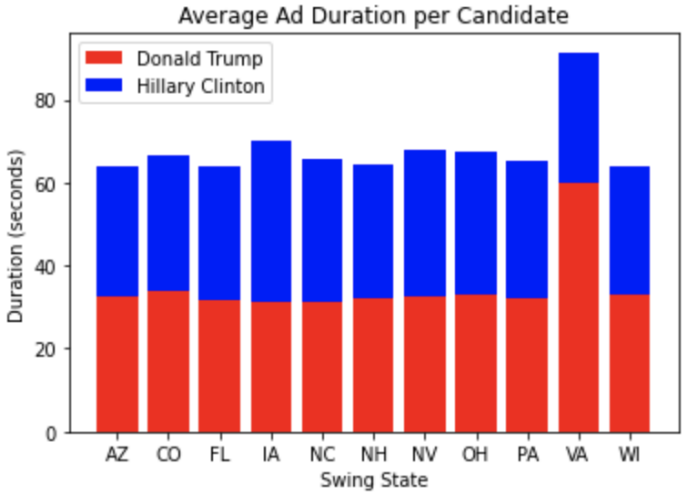

Fig 4.1: Stacked Bar Chart of Average Ad Duration per Candidate

Average ad durations between candidates were very similar in most swing states, except for Virginia. Virginia had only two advertisements related to Clinton, so this seems like an outlier. Clinton had one more state that had a longer average duration than Trump.

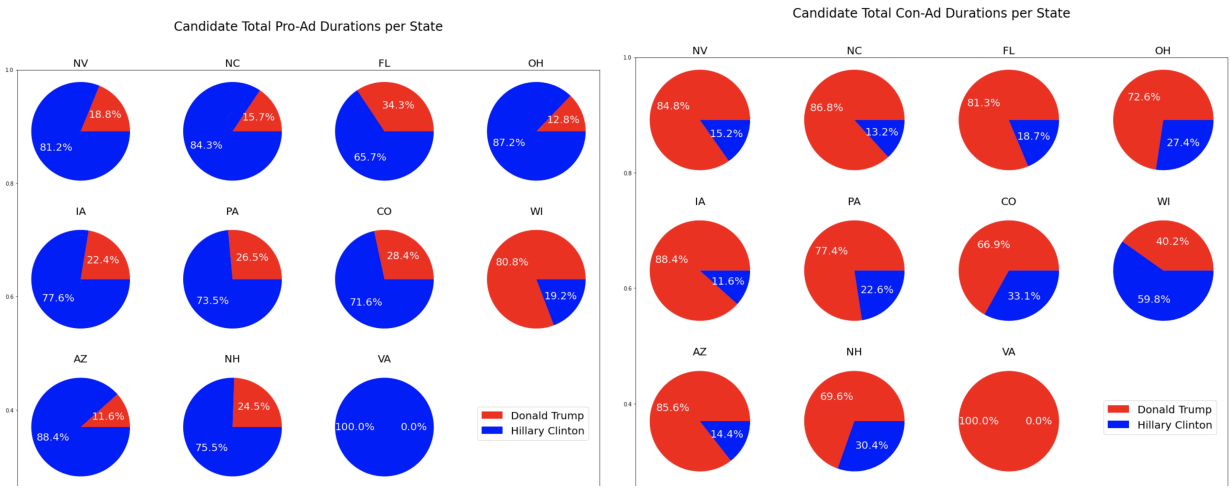

Fig 4.2: Pie Charts of Pro-Candidate/Con-Candidate Ad Durations for each State

Clinton had more total pro-candidate ad duration across all states except Wisconsin than Trump. Trump had more total con-candidate ad duration across all states than Clinton, again except for Wisconsin. When analyzing the pro/con duration for each candidate, we expected the candidate with more pro-candidate ad duration and less con-ad duration would win that state. Four of the states (CO, NH, VA, WI) followed that expected trend. Yet in three states (AZ, FL, IA) that had less pro-ad duration and more con-ad duration for Trump, he still won. In the remaining four states, Trump either had less of both pro-candidate ad duration and con-candidate ad duration (NC), or more of both (NC, NV, OH, PA). Trump won all these states except Nevada.

VI. Discussion of Results:

To summarize our findings, our visualizations and charts allowed us to understand how ads affected the 2016 election results. Candidates had similar average ad durations and aired their news and a similar proportion of news vs. non-news ads in swing states. However, that's the extent of the similarities between the two candidates. Ads played the role of convincing voters to lean a certain way with their political views. There was a negative correlation between the number of ads in a state and voter registration percentage as well as voter turnout percentage. It seems that in states with fewer voters, more ads were played to increase voter turnout. In general, political ads aired more on news platforms than non-news platforms. In the only states where there were more non-news ads than news ads, Iowa and North Carolina, Trump won both.

Trump seemed to benefit from having more ads on TV more than Clinton did. In the swing states that Trump had more ads in, he won 100% of the time. On the other hand, Clinton only won 67% of the time when she had more ads about her in the swing state. This is interesting because Trump had more negative (con) ads about him in general, while Hilary had more positive (pro) ads. However, reflecting the harsh rhetoric of the election, there were on average almost 3x more con-candidate ads than pro-candidate

ads. The con ads for Clinton must have affected her reputation much more than they affected Trump since he won the election despite having more con ads in every state except Wisconsin.

We expected that more pro-candidate ad duration and less con-candidate ad duration would lead to a candidate's win. Only four of the swing states (CO, NH, VA, WI) followed that expected trend. However, in 3 states (AZ, FL, IA), Trump still won even though he had less pro-candidate ad duration and more con-candidate ad duration. In general, it seems that Trump won more states regardless of the type of ads.

Interpreting our findings from background information and research, we hypothesize that ads may have been more beneficial for Trump's 2016 presidential election campaign due to several contributing factors. First, the presidential candidates' backgrounds were drastically different. Trump was involved in the entertainment, real estate, and business industries[6]. In these sectors, he was already known as a controversial figure. On the other hand, Clinton was seen as an experienced female politician who was the previous First Lady of the United States (1993–2001) and former Secretary of State (2009–2013)[7]. To Trump, more negative press was still press[8]. He accepted every interview and was able to get into the headlines and news articles, which may have led to the desensitization of Trump's negative actions and con-ads. He sold the narrative that people shouldn't trust politicians — that he was different, and positioned himself to be the most ideal candidate[9]. On the other hand, Clinton had more trust to lose since she was closely tied to politics even before the presidential election. Thus, negative press affected her more than it did for Trump. The strategy of negative campaigning, which surrounded the entire election, focused on defamation and slandering, ultimately benefited Trump. In addition, whether negative or positive, wherever they were aired, political advertisements may have been more effective for Trump due to the topics he focused on for his campaign. Compared to Clinton, who focused on health care, rights for women, minorities, LGBT, and fair taxes, Trump's presidential campaign was centered around more provocative topics such as "building the Mexican border wall", draining "the swamp" (ending corruption in Washington, D.C.), and opposing free trade deals[10]. These specific points may have been easier to convey in thirty-second ads, givinging a lasting impression and message when contrasted with Clinton's broader arguments. Overall, these differences and factors may have positioned ads to be a key part of Trump's campaign that helped him win the 2016 presidential election.

VII. Conclusion:

Our hope is that our analysis of how advertisements features like number of unique ads, program type, ad polarity, and total number of ads impacted the voting trends and outcome of 2016 election can shed light on future election advertisement campaign strategies and be used to more comprehensively examine the election from a psychological perspective. In terms of research limitations, we did not have access to advertisement data for the swing state of Michigan, and there was no causal analysis conducted

---

[6] Trump, Donald. 2017. "Donald Trump: Life before the Presidency | Miller Center." Miller Center. April 11, 2017. https://millercenter.org/president/trump/life-presidency.

[7] "Hillary Clinton | Biography, Politics, & Facts | Britannica." 2019. In *Encyclopædia Britannica*. https://www.britannica.com/biography/Hillary-Clinton.

[8] Kruse, Michael. n.d. "Trump and the Dark Art of Bad Publicity." POLITICO Magazine. Accessed August 14, 2021. https://www.politico.com/magazine/story/2016/07/donald-trump-2016-convention-melania-trump-speech-dark-art-of-pr-214083/.

[9] Rivero, Cristina. 2016. Review of *How Marketing Helped Donald Trump Win the 2016 Election*. Washington Post. November 17, 2016. https://www.washingtonpost.com/graphics/politics/2016-election/trump-campaign-marketing/.

[10] History.com Editors. 2018. "The Divisive and Stunning U.S. Presidential Election of 2016." HISTORY. November 29, 2018. https://www.history.com/topics/us-presidents/us-presidential-election-2016.

to evaluate the relationship between variables due to the scope of this course so we cannot assume statistical significance due to the lack of inferential statistics.

With regards to our future research goals, we hope to apply this process to all swing states with more accessible data. With this more comprehensive dataset, we also hope to compare the percentage of votes for each candidate with the percentage of ads per candidate. Even more so, diving deeper into sponsor type (the candidate who funded the ad; who the ad was "by") would be interesting and provide meaningful insight to build on our current approach of examining who the ad was "about." We would also like to conduct statistical analysis to analyze causal relationships for key variables, which would strengthen our findings or perhaps add a new perspective to our story. Finally, for further research, by applying methods such as KNN and other clustering algorithms, we can classify voter preferences based on variables such as message polarity if we determine there to be a statistically significant association between two variables. Beyond this, logistic regression can be applied for predictive modeling based on cluster analysis.

VIII. Sources:

1. Presidential battleground States, 2016. Ballotpedia. (n.d.). https://ballotpedia.org/Presidential_battleground_states,_2016.

2. Presidential election process. USAGov. (n.d.). https://www.usa.gov/election#:~:text=In%20the%20Electoral%20College%20system,(270)%20wins%20the%20election.

3. "Political TV Ad Archive» Data." 2016. Politicaladarchive.org. 2016. http://politicaladarchive.org/data/.

4. Bureau, US Census. n.d. "Voting and Registration in the Election of November 2016." The United States Census Bureau. https://www.census.gov/data/tables/time-series/demo/voting-and-registration/p20-580.html.

5. Trump, Donald. 2017. "Donald Trump: Life before the Presidency | Miller Center." Miller Center. April 11, 2017. https://millercenter.org/president/trump/life-presidency.

6. "Hillary Clinton | Biography, Politics, & Facts | Britannica." 2019. In *Encyclopædia Britannica*. https://www.britannica.com/biography/Hillary-Clinton.

7. Kruse, Michael. n.d. "Trump and the Dark Art of Bad Publicity." POLITICO Magazine. Accessed August 14, 2021. https://www.politico.com/magazine/story/2016/07/donald-trump-2016-convention-melania-trump-speech-dark-art-of-pr-214083/.

8. Rivero, Cristina. 2016. Review of *How Marketing Helped Donald Trump Win the 2016 Election*. Washington Post. November 17, 2016. https://www.washingtonpost.com/graphics/politics/2016-election/trump-campaign-marketing/.

9. History.com Editors. 2018. "The Divisive and Stunning U.S. Presidential Election of 2016." HISTORY. November 29, 2018. https://www.history.com/topics/us-presidents/us-presidential-election-2016.