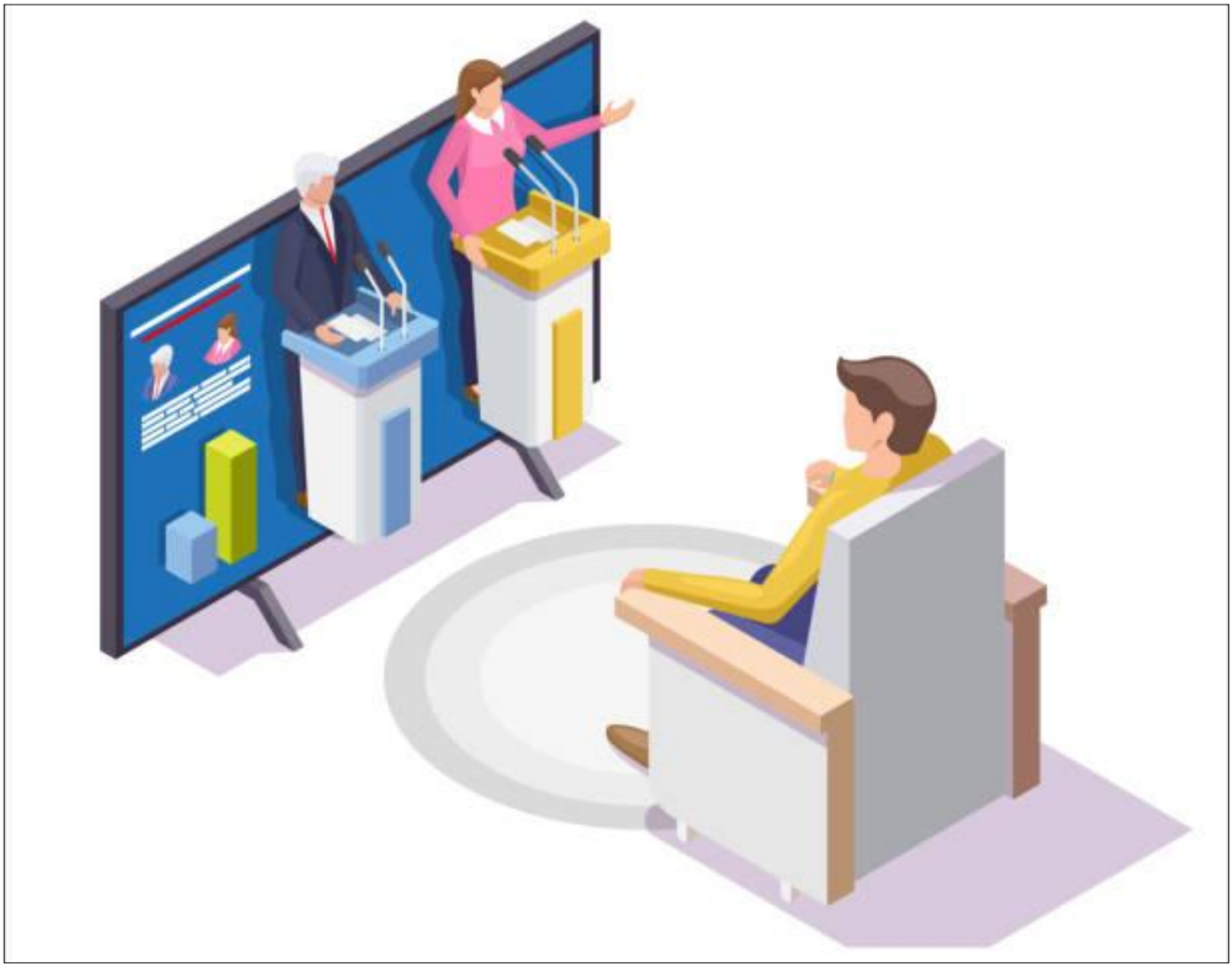
PROJECT-2 REPORT: *ANALYSIS OF POLITICAL TV ADS FROM 2016 ELECTION*



Prepared for: Gerry Benoit

Prepared by: Gerardo Gonzalez, Steven Lu, Snehal Desai

Date: December 10, 2022

Primary Dataset: http://politicaladarchive.org/data/

GitHub Repo Name: Project2_Gonzalez_Lu_Desai

GitHub Repo URL: https://github.com/UC-Berkeley-I-School/Project2_Gonzalez_Lu_Desai

Overview

In the 2016 election, Donald Trump beat Hillary Clinton 304 to 227 in the electoral college vote, with the remaining seven choosing a different candidate other than their party's candidate. The results of the election were surprising to many, as most models and analysts had predicted a clean Clinton victory. This remains one of the closest elections in United States history; a couple of swing states decided the election, and even those swing states were decided by mere fractions of a percentage point. For example, Michigan was decided by a mere 10,704 votes, or a difference of 0.3% of the population (Trump winning with 47.3%, Clinton with 47.0%.) Many analysts criticized Clinton for several aspects of her campaign strategy, including her lack of attention towards several important states. For example, Clinton did not visit Wisconsin whatsoever in the last year of her campaign, perhaps due to overconfidence in her ability to win the traditional "blue-wall" state. Michigan and Wisconsin were key states that ultimately led to Trump's victory in the election. This neglect is also apparent in the graphic to the right, which displays the relative ad volume of several included states for her campaign: the deeper the shade of blue, the more Democratic ads were aired.

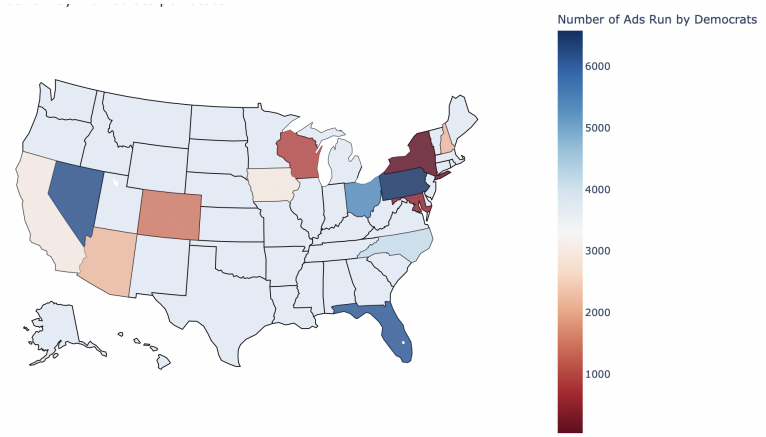


Figure 1

We can see that Wisconsin was relatively ignored, while states such as Nevada, Florida, and Pennsylvania were given more attention. On the other hand, Trump's campaign strategy successfully targeted the same states that Cilnton had neglected, as shown by the graphic on the right (the deeper the shade of red, the more ads were shown in that state.) Trump focused on states

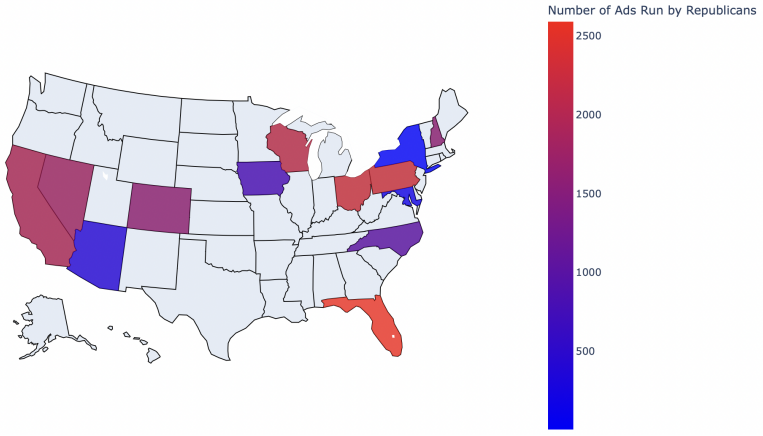


Figure 2

such as Wisconsin, Ohio, Pennsylvania, and Florida, all of which proved instrumental in his victory.

These graphs were created using a dataset of advertisements aired during the 2016 electoral campaign. Given previous analyses of Clinton's campaign, we wanted to see how the advertisements shown in each state could further break down the discrepancies in strategy and how those led to the final, shocking outcome. Thus, our overall goal was to analyze this political TV ads dataset and look at the proportion and total number of advertisements run by Clinton and Trump in preparation for the election. We chose to analyze the five states included in the data that had less than a 3% margin of victory: Florida, New Hampshire, Nevada, Pennsylvania, and Wisconsin.

Research Question

Main Question:

- Was each party's (Republican and Democratic) TV ad campaign strategy successful in the 5 states where the margin of victory was less than 3%?

Sub-questions:

1. Did the message tone (pro or con) of a party's television ads impact the party's results in these five states?
2. Did each party's advertisement frequency vary by show-type (news vs. non-news)?
3. Did each party's advertisement frequency vary by swing state?

Before we dive deeper into the data and analysis, here is a quick look at the margin of victory (MoV) in the results of the five swing states discussed in this report[1]:

- New Hampshire (MoV = 0.4%) - Clinton 47.6 percent, Trump 47.2 percent - Difference: 2,701 votes
- Wisconsin (MoV = 1%) - Trump 47.9 percent, Clinton 46.9 percent - Difference: 27,257 votes
- Pennsylvania (MoV = 1.2%) - Trump 48.8 percent, Clinton 47.6 percent - Difference: 68,236 votes (99 percent reporting)
- Florida (MoV = 1.2%) - Trump 49 percent, Clinton 47.8 percent - Difference: 114,455 votes
- Nevada (MoV = 2.4%) - Clinton 47.9 percent, Trump 45.5 percent - Difference: 26,434 votes

Data Cleaning

The Political TV Ad archive collected and provided content for the 2016 election political ads. It was a free, open-source service for journalists, civic organizations, academics and the general public to track the ads.

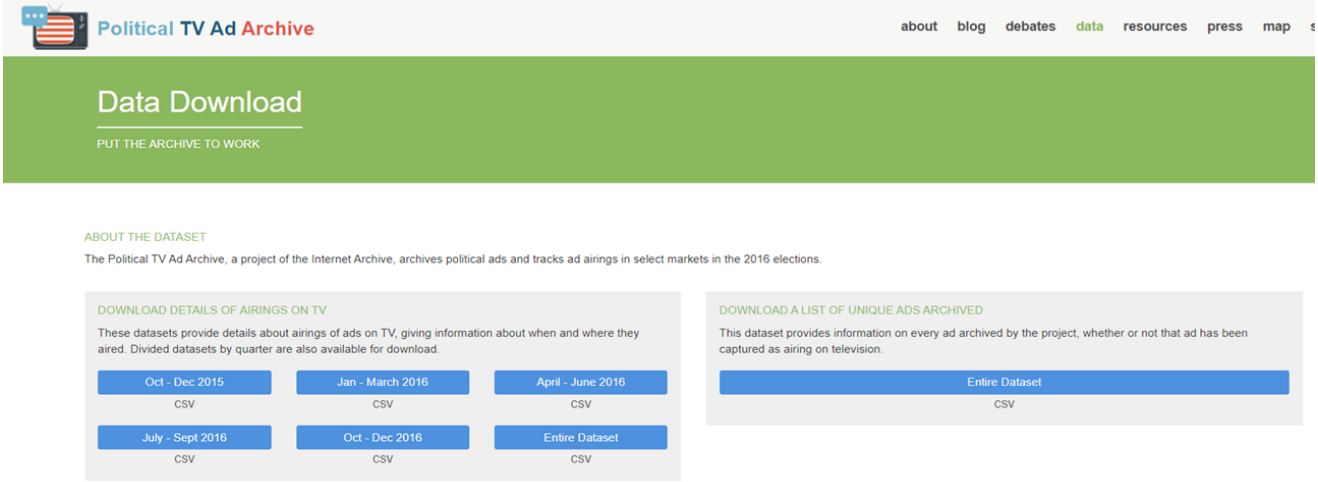


Figure 3: Dataset collected from Political TV Ad Archive

The original dataset had 375,350 rows by 19 columns. Some of the fields that had missing values
were location, program, race, cycle, subjects, candidates, sponsors, and sponsor types.

Race and Date fields:

The dataset has ads that start from September 2014 onwards and includes ads for the state senate, presidential primaries and the presidential race. We decided to concentrate on the presidential general election. Because the Democratic National Convention finished at the end of July 2016 and election day was November 8, 2016, we decided to filter the dates between these two dates.

```
1  #Looking at earliest and latest date of Presidential general election
2  print(f"Earliest: {presidential_ge_df.start_time.min()}")
3  print(f"Latest: {presidential_ge_df.start_time.max()}")
✓ 0.1s
Earliest: 2016-08-01 01:19:59 UTC
Latest: 2016-11-07 23:59:41 UTC
```

Figure 4: Filtering dataset to include ads between 8/1/2016 and 11/8/2016

Additionally, because the dataset included ads for the state-level races as well as on the primary election, we decided to filter on the race field and only use the ads with the value = 'PRES'.

```
1  #Using race field to select ads marked as "PRES" for the presidential election
2  presidential_df = pol[(pol['race'] == "PRES")]
```

Figure 5: Filtering dataset to include ads where race is equal to 'PRES'

```
3  # Filtered dataframe to only include the top two candidates: Hillary Clinton and Donald Trump
4  presidential_ge_df=presidential_ge_df[presidential_ge_df["candidates"].isin(["Hillary Clinton","Donald Trump","Donald Trump, Hillary Clinton"])]
```

Figure 6: Filtering dataset to only include ads for the top 2 candidates

Also, for our analysis, we decided to only focus on the top two candidates: Hillary Clinton and Donald Trump and filtered the dataset to reflect this.

In order to gather state-wide data, we split the 'location' column and parsed its data into two additional columns called 'city' and 'state'. Furthermore, we derived the party behind a given ad by filtering the 'candidate' field with the 'message' field. For instance, if the 'candidate' field had "Hillary Clinton" and the corresponding 'message' field was con, then the 'benefactor' field was populated as "Republican".

For the bar charts we further filtered the above dataframe to eliminate entries from all states except the 5 states under analysis, namely, NH, WI, PA, FL, and NV. These data cleaning steps resulted in a final count of 42774 rows, down from 375,350 rows.

Data Analysis

Sub-Question 1: Did the message tone (pro or con) of a party's television ads impact the party's results in these five states?

The plots show that in 4 of the 5 states analyzed, there were more negative ads targeting candidate Trump. Only in Wisconsin, were there more negative ads for Hillary Clinton. Furthermore, there were far more negative than positive ads in these five states in general. The obvious conclusion that we can draw from these plots is that the 2016 Presidential election was significantly negative in the five states, from a television-ad perspective:

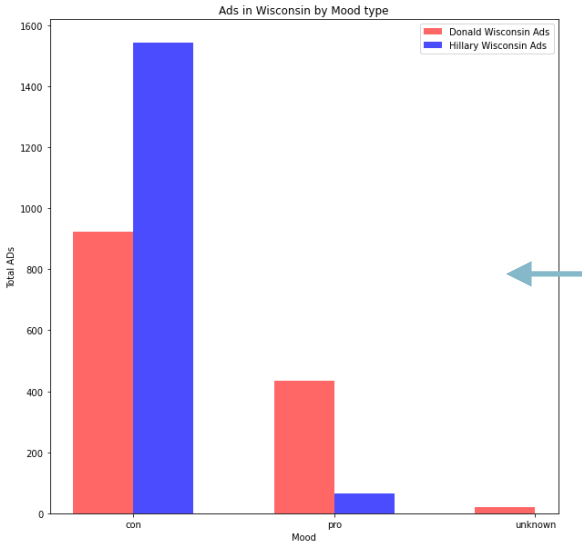


Figure 7: Mood Tone for ADs shown in Wisconsin

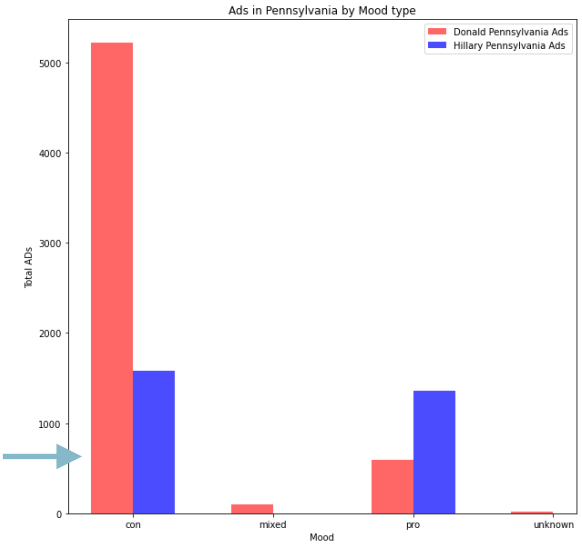


Figure 8: Mood Tone for ADs shown in Pennsylvania

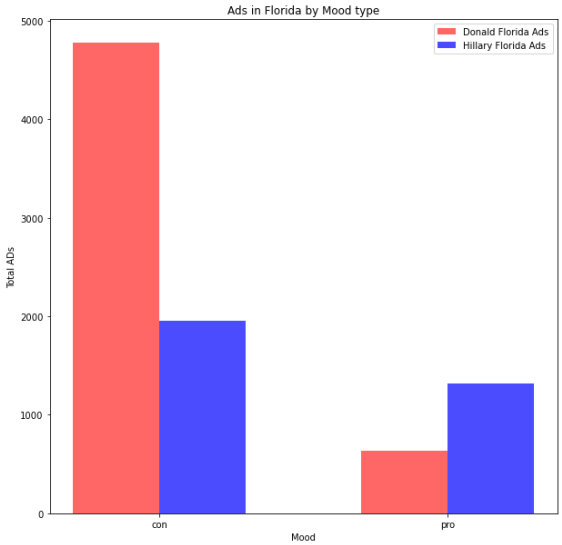


Figure 9: Mood Tone for ADs shown in Florida

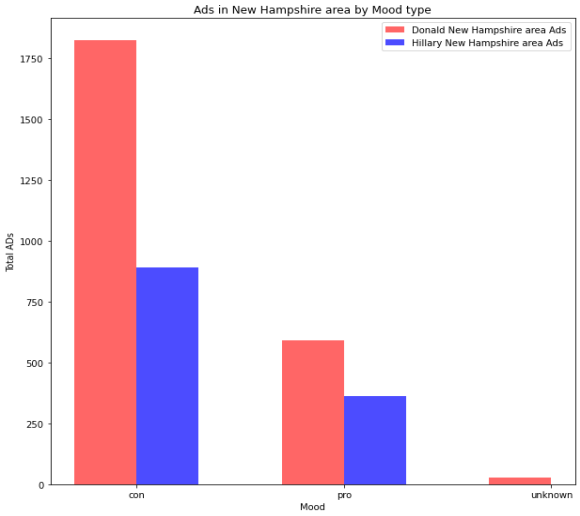


Figure 11: Mood Tone for ADs shown in New Hampshire

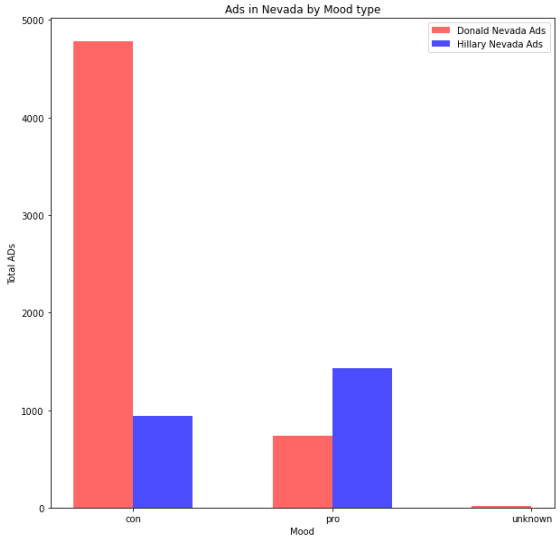


Figure 10: Mood Tone for ADs shown in Nevada

Filtering this data against the final results of 2016, we observe that the negative tone adopted by the Democratic party aligned with their victories in New Hampshire and Nevada, but not in Pennsylvania or Florida. This would imply that adopting a negative tone did aid the Democrats in their chance of winning.

Wisconsin was the only state amongst these five where more negative ads were shown targeting Hillary Clinton, and the Republican party did win that state.

Going a step further and looking at the total Electoral College (EC) votes from these 5 states' dataset we observe that though Democrats outspent Republicans in terms of negative ad campaign dollars, they only won 10 of the 69 EC votes represented by these 5 states. Therefore, the response to this sub-question is that increased negative tone in TV ads did not impact results in their favor for the Democrats in these 5 states. For the Republican party the observed effect is too small for any well-formed conclusions.

Sub-Question 2: Did each party's advertisement frequency vary by show-type (news vs. non-news)?

The next sub-question we wanted to explore was whether the two parties targeted different show types for advertisements. For example, if Republicans had happened to target Soap Operas or TV stations such as TNT during midday, it could be inferred that they are targeting older populations. However, a big limitation of the dataset was that it only split the program

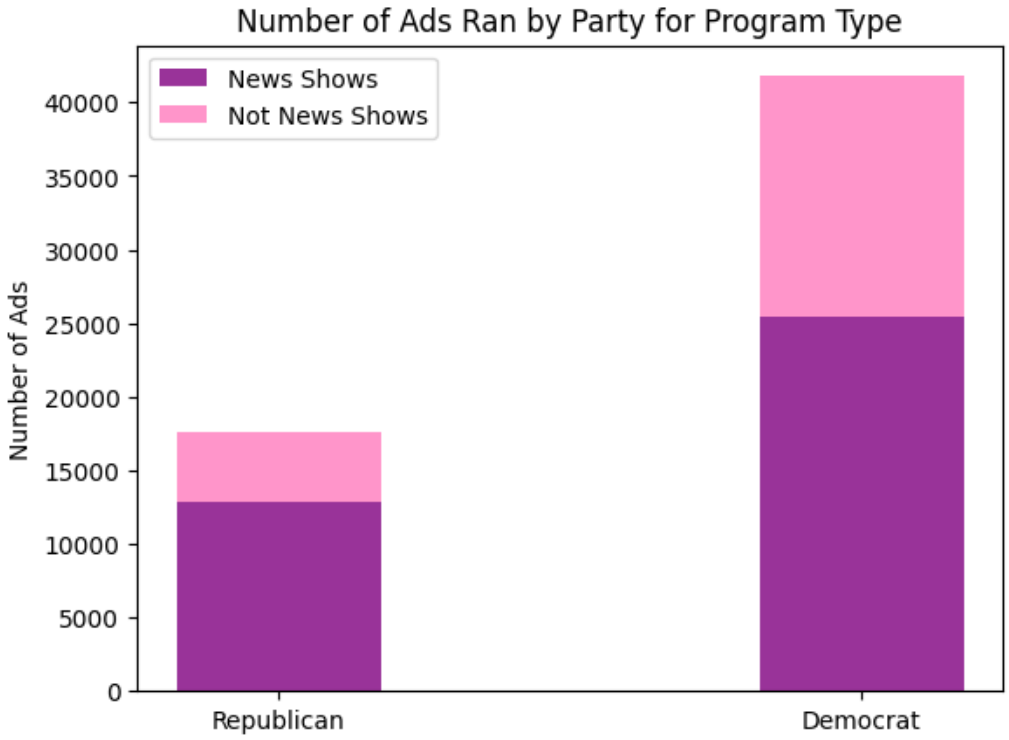
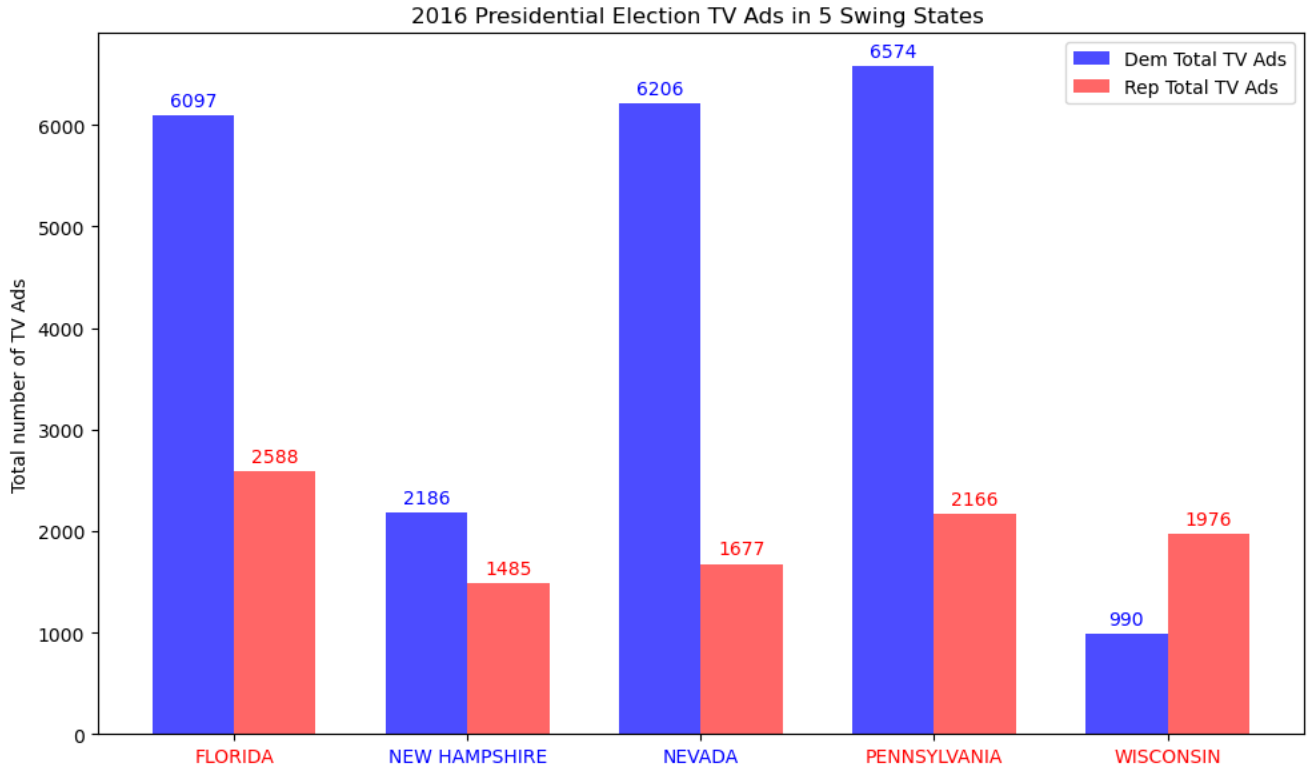


Figure 12: Number of ads run in various types of TV shows by each party.

category into “News” and “Non-news”, and thus we were not able to go quite in-depth as we would have liked. However, we were still able to glean some insights into each party’s strategy with regards to the type of show, as demonstrated below

It seems that Republicans tended to feature ads primarily on news shows, as they had almost three-quarters of ads focused on news shows. On the other hand, Democrats had a more even split between news shows and non-news shows, with around 60% of ads targeted towards news shows as opposed to non-news shows.

Sub-Question 3: Did each party's advertisement frequency vary by swing state?

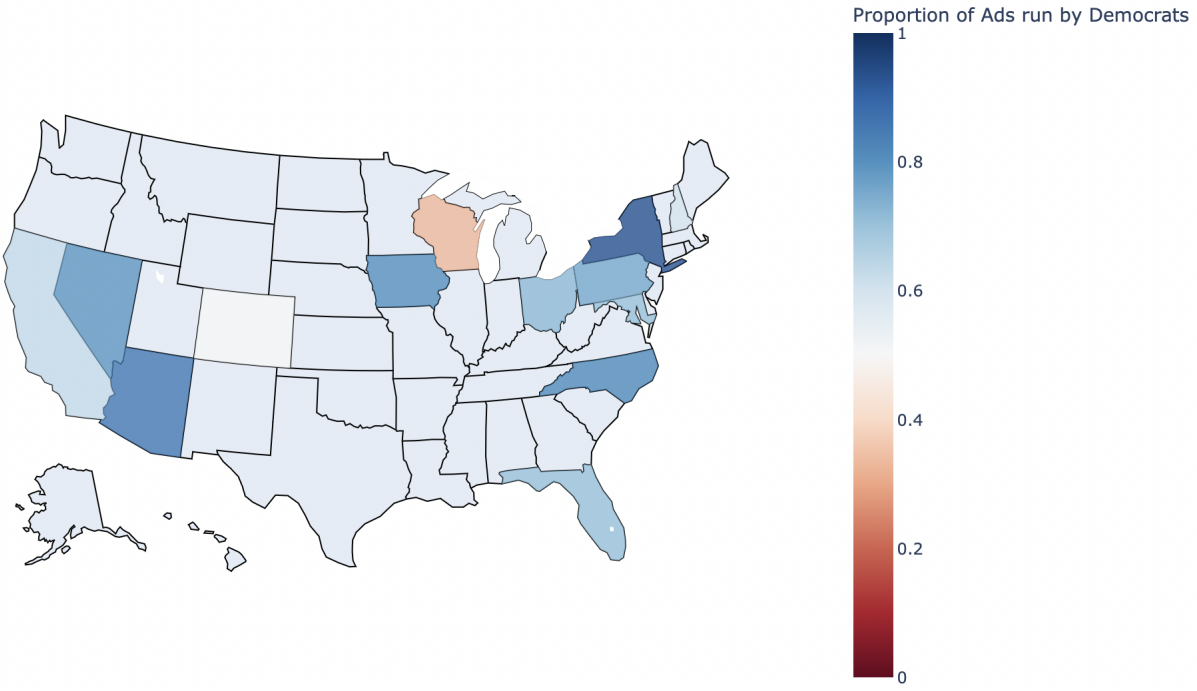


[Please note that the state names are color-coded to the final winners of that state in the 2016 Presidential Election. This is a neat hidden feature in Matplotlib]

It is clear from this bar graph that Republicans were fairly consistent with spending on TV ads regardless of the size of the state, as compared with Democrats (note our assumption that on equal cost per ad, on average). The max-to-mean ratio for Republicans was 1.74 as compared to 6.64 for Democrats.

The data here shows a targeted approach by Democrats in the states of Florida, Nevada, and Pennsylvania. However, the outcome was not in their favor in the large EC-vote states of Florida and Pennsylvania.

Conclusions



From this dataset, we observed that Democrats outspent Republicans by a margin of 2.2 to 1 in the 5 swing states of NH, WI, PA, FL, and NV.

Summarizing our analysis of our 3 sub-questions we can conclude from this dataset that in these 5 states Democrats ran a largely negative TV ad campaign against their opposing candidate, Donald Trump, and that those ads were more widely distributed over news and non-news shows, and that their TV ad campaign was targeted one, focusing more on 3 states and less on the other 2.

However, the final result of the 2016 Presidential Election was not in the favor of Democrats in terms of EC votes in these 5 states. They won only 10 of the 69 possible EC votes possible from these states. The results of this data analysis answers the main research question as follows: Democratic party's TV ad campaign strategy was not successful in these 5 states and that the Republican party received better returns, in terms of political capital, from their own strategy.

References:

1. The 10 Closest States in the 2016 Election: https://www.usnews.com/news/the-run-2016/articles/2016-11-14/the-10-closest-states-in-the-2016-election

2. 2016 Presidential Election Results: https://www.270towin.com/maps/2016-actual-electoral-map

3. Political TV Ad Archive: http://politicaladarchive.org/data/