

# Winning the Airways: An Analysis of Political TV Advertising in the 2016 Election

**Team Members:** Haoyue Zhang, Julia Lau, Christie Liu

**GitHub Repository:** [https://github.com/UC-Berkeley-I-School/Project2\\_Zhang\\_Lau\\_Liu](https://github.com/UC-Berkeley-I-School/Project2_Zhang_Lau_Liu)

## Background

The 2016 election featured unprecedented elements including Trump's unconventional campaign, Clinton's historic nomination, and social media's rising influence, as this study examines trends in political TV advertising during that period.

## Primary Question

How did political advertising trends differ between leading presidential candidates, and how did these trends impact electoral outcomes? To answer this question, we explore what our leading candidates did with their airtime in terms of content, tone, and reach.

## Primary Dataset

The Political TV Ad Archive maintained by the Internet Archive includes political TV ads and ad airings across 20 markets in nine states related to the 2016 general elections. The [Unique Ads Archive dataset](#) provides information on every ad archived by Internet Archive whether or not that ad has been captured as airing on television. A separate dataset provides details about [airings of ads on TV](#) and provides additional information about when and where they aired.

## Secondary Dataset

To glean insights into correlations with election results, we supplement the primary dataset with [2016 presidential election results](#) sourced from MIT Election Data and Science Lab. This provides votes per candidate and total votes at the state level for presidential elections from 1976 to 2020.

## Data Cleaning and Assumptions

The political ads dataset contains ads information for both the 2016 presidential race and ballot-down races. Since we are focusing on the presidential race and because only two candidates remained in the race for the entire campaign cycle, we filtered for the presidential candidates with the most air time. This left us with a dataset for political ads for Trump, Clinton, and Sanders. Though Sanders' primary loss meant a shorter campaign and more limited reach than the other two, information pertaining to Sanders would allow us to glean some insights by

comparing candidates within the same party. For our analysis related to election results, Sanders was omitted from the dataset.

#### Fields dropped or augmented:

- **transcript:** Since over 50% of records contained null values in the transcript field, we dropped this column.
- **subjects:** This column contains comma separated lists of subjects covered in an ad. Since we intend to take count of subjects covered in ads, we used the `explode()` function to separate this list into multiple rows.
- **location** and **network:** In order to create a region/state dimension to analyze against election results, a **state** column was created from the location field which contained city and state information for where an ad aired. We were able to infer values for several null fields for the location field by cross referencing the network field.
- **date\_ingested:** We converted this column from UTC to a timezone-aware pandas datetime object. This would help us standardize data across multiple time zones (ex. Ads airing in various regions).

#### Additional Sanity Checks:

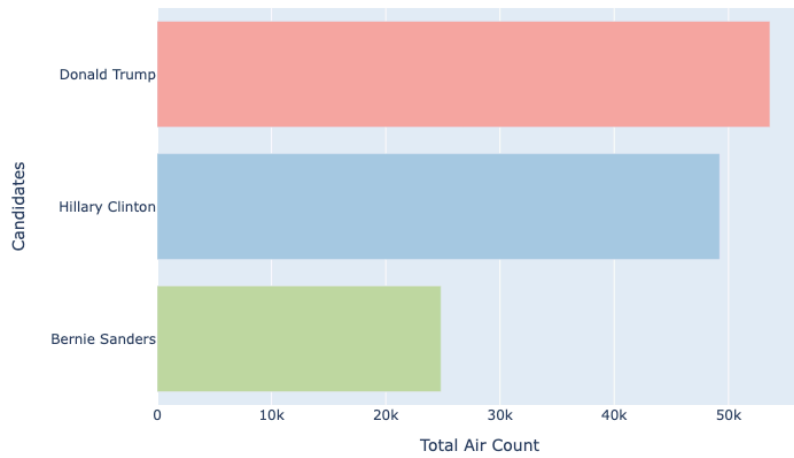
- Duplicate Records: Checked for duplicate entries but found none
- Missing Data: Analyzed null values across all fields and changed to 'Unknown' when fit
- Consistency Checks: Ensured all ads had valid airtime date values.

We first wanted to perform EDA to get a better understanding of the data. To do that, we came up with a set of questions that spanned across different columns and segmentation. The questions are as follows:

#### **Which candidates had the most air time?**

The top 3 candidates were Donald Trump, Hillary Clinton, and Bernie Sanders. For this analysis, we will focus primarily on these three candidates. Note that since Bernie Sanders lost the Democratic Primary to Hillary Clinton, his advertising campaign would have been shorter in duration, resulting in fewer ads and a more limited reach compared to the other two candidates, who continued their campaigns through the general election.

Top 3 Candidates with Highest Airtime

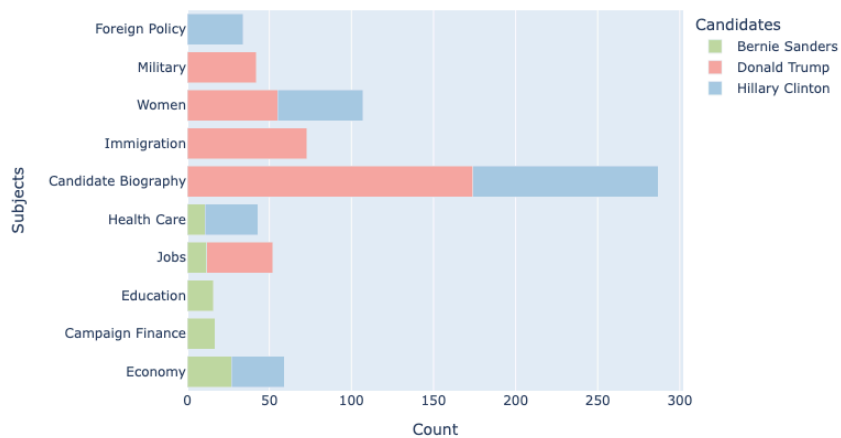


## What are they doing with their airtime?

### Content

The content of their ads are spread across various topics. Hillary had the most overlap with the other two candidates, while Bernie had the least. The most popular subject across their ads were `Candidate Biography`

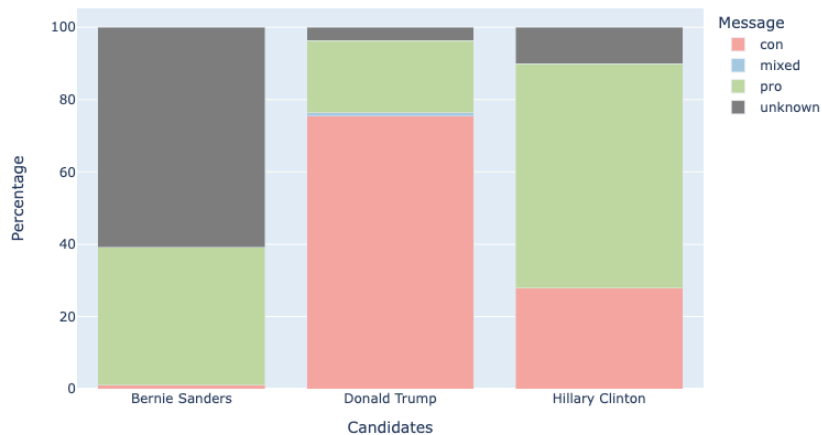
Top Subjects in Ads for Top 3 Candidates



### Tone

The tone of their ads were very different. Negative ads made up 75% of Trump's ads. Positive ads made up 62% of Hillary's ads. Unknown ads made up 60% of Bernie's ads.

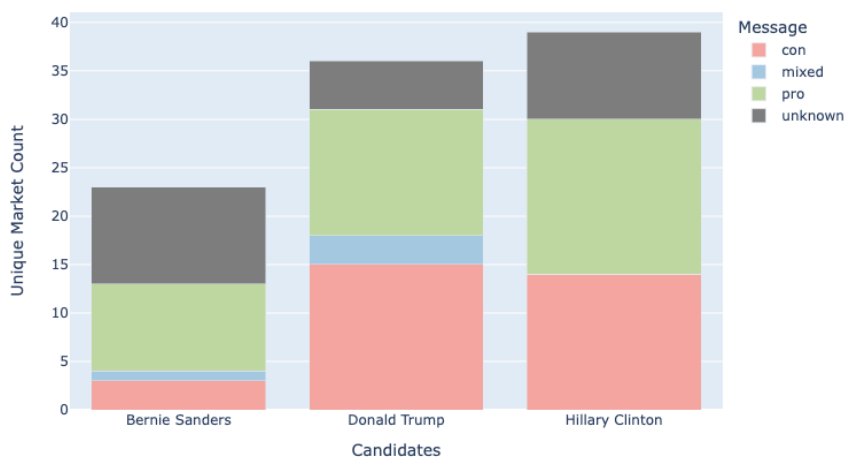
Tone Analysis of Ads for Top 3 Candidates



### Geographic Reach

Hillary Clinton had the largest reach. She reached 39 unique markets while Bernie only reached 23 unique markets.

Unique Geographic Reach & Tone of Ads for Top 3 Candidates



### Sponsor Types

Donald Trump had the most diverse set of sponsors. 98.7% of Bernie's ads were funded by his candidate committee Bernie 2016. Whereas Trump and Hillary's were funded by multiple sponsors and Super PACs.

**Sponsor Type**

- Candidate Committee
- Corporation
- Hybrid Super PAC
- Joint Fundraising Committee
- Multiple
- Non Profit
- PAC
- Super PAC
- Unknown

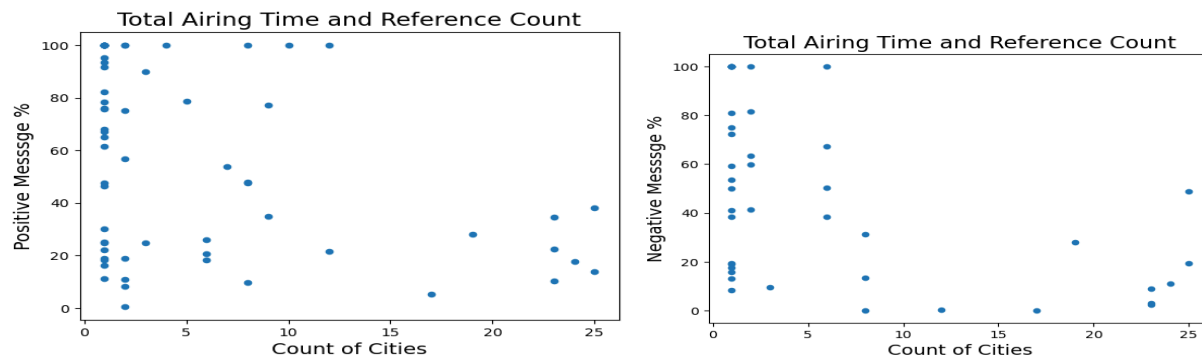
Candidate	Candidate Committee	Corporation	Hybrid Super PAC	Joint Fundraising Committee	Multiple	Non Profit	PAC	Super PAC	Unknown
Bernie Sanders	24,500	0	0	0	0	0	0	0	0
Donald Trump	5,000	0	1,000	0	24,000	0	0	19,000	5,000
Hillary Clinton	7,000	0	1,000	0	36,000	0	5,000	0	0

## Potential Impacts and Patterns

The scatter plot displays the relationship between the total airing time in minutes and the reference count. The x-axis, labeled 'Airtime(total minutes)', ranges from 0 to 2500. The y-axis, labeled 'Reference Count', ranges from 0.0 to 4.0. The data points are as follows:

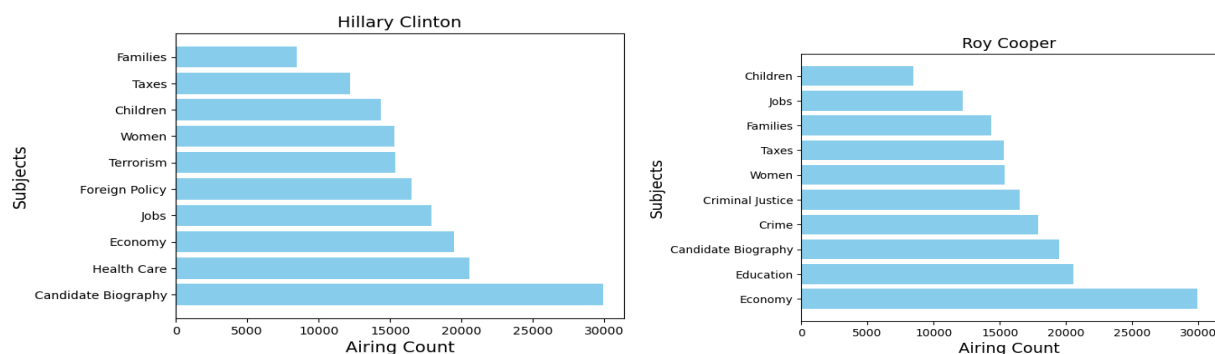
Airtime (total minutes)	Reference Count
0	1.0
0	2.0
0	4.0
100	1.0
100	2.0
100	4.0
200	1.0
200	2.0
300	1.0
300	2.0
400	1.0
400	2.0
500	1.0
500	3.0
600	1.0
700	1.0
800	1.0
800	2.0
900	1.0
1000	1.0
1100	1.0
1200	1.0
1200	2.0
1300	1.0
1400	1.0
1500	1.0
1600	1.0
1700	1.0
1800	1.0
1800	2.0
1900	1.0
2000	1.0
2100	1.0
2200	1.0
2300	1.0
2400	1.0
2500	1.0

### Candidates with greater geographic reach balance tone strategies



### Candidates with broader geographic exposure diversify their messaging:

Candidates with border geographic reach tend to focus on 'Candidate Biography', while candidates with lower geographic reach may focus on specific topics.



Lastly, we wanted to address our primary question: How do airtime patterns correlate with election results? By focusing on this relationship, we aimed to connect the observed trends in advertising strategies to their tangible impact on electoral outcomes.

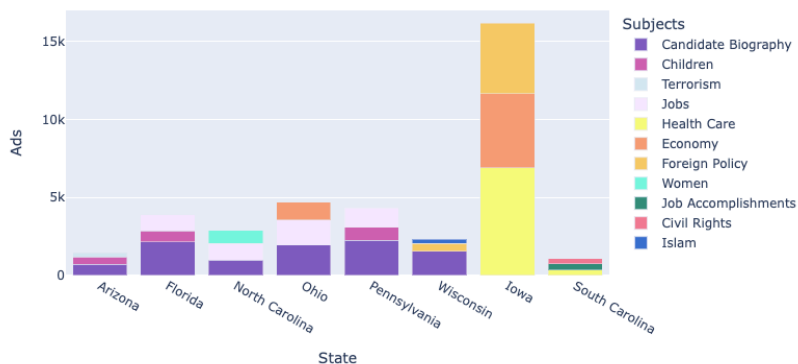
### **What correlation does airtime patterns have with election results?**

The ads maintained by the Political Ads Archive related to the 2016 presidential race were limited to 15 states, but this includes swing states Wisconsin, Pennsylvania, New Hampshire, Arizona, Georgia, Virginia, Florida, Michigan, Nevada, Colorado, North Carolina, and Maine, which are arguably the most important markets to extract campaign strategy insights from. Political campaigns have long focused on states where candidates have the greatest opportunity to sway voters off the fence and win Electoral Votes. This analysis of political campaign messaging reveals that while candidates with broader geographic reach like Donald Trump used moderately negative messaging (49%), those with limited reach exhibited extremes - some like Kelly Ayotte and Roy Blunt opted for heavily negative campaigns (67-100% negative), while others like Gerald Daugherty and Tim Huelskamp maintained exclusively positive messaging.

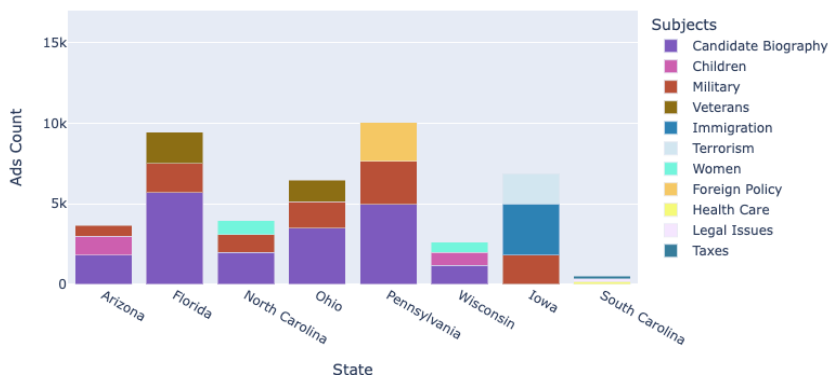
## Impact of Topics on Election Results

States Trump Won: Looking at the top three topics covered in political ads across states where Trump won, candidate biographies were consistently a top topic in ads. Both Trump and Clinton took a more targeted approach in Iowa. Trump focused on the military in most of these states, but for Iowa, double down on immigration. Clinton invested a lot more airtime in Iowa and focused on healthcare, economy, and foreign policy. Despite the heavy investments, she still lost Iowa.

Clinton Ads

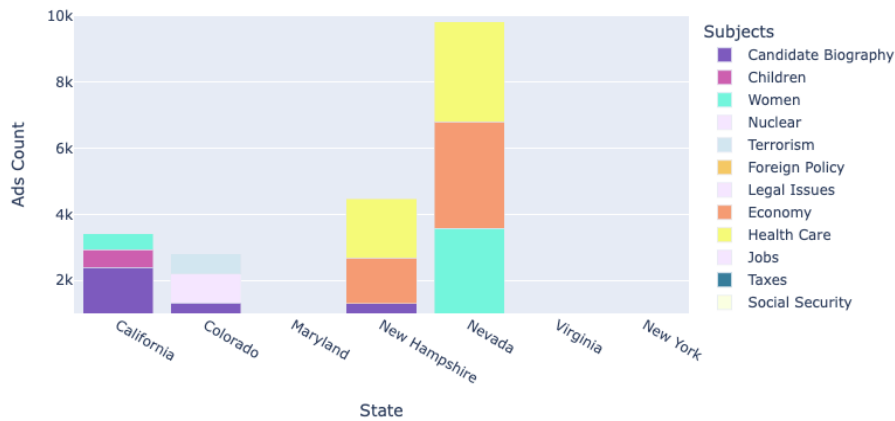


Trump Ads

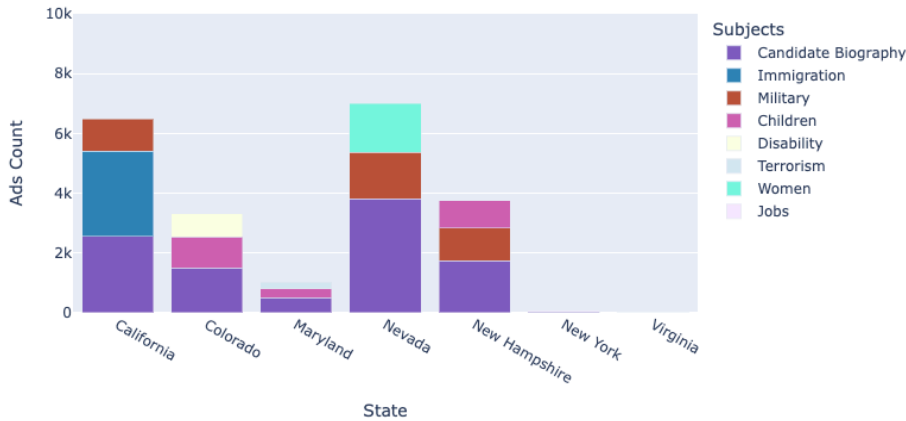


States Clinton Won: In states where Clinton won the popular vote, we see the same blend of topics in Trump ads with heavy focus on candidate biography and the military. Clinton again doubles down on one key state - Nevada. For Nevada, her focus on women, the economy, and health care and the amount of air time invested made a positive return on investment.

Clinton Ads



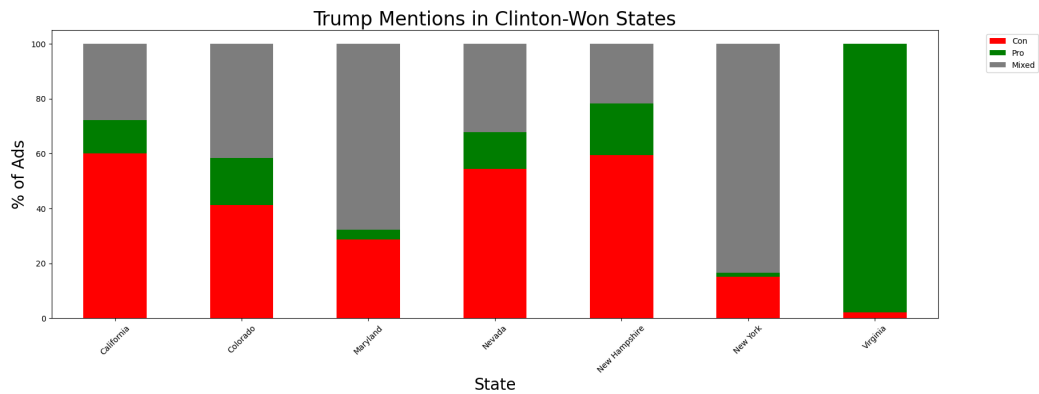
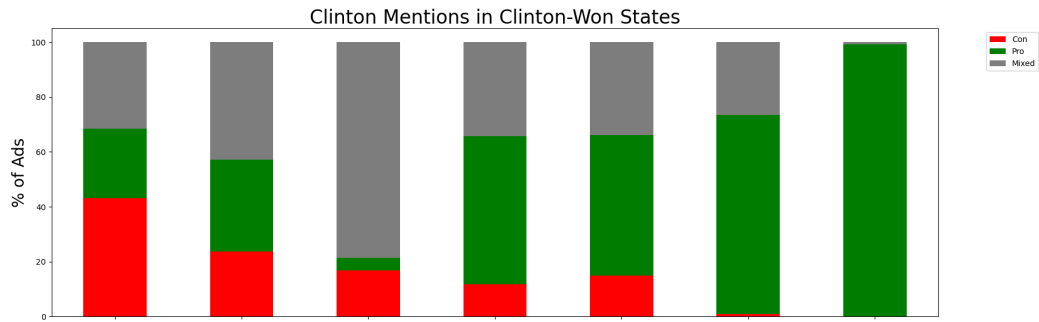
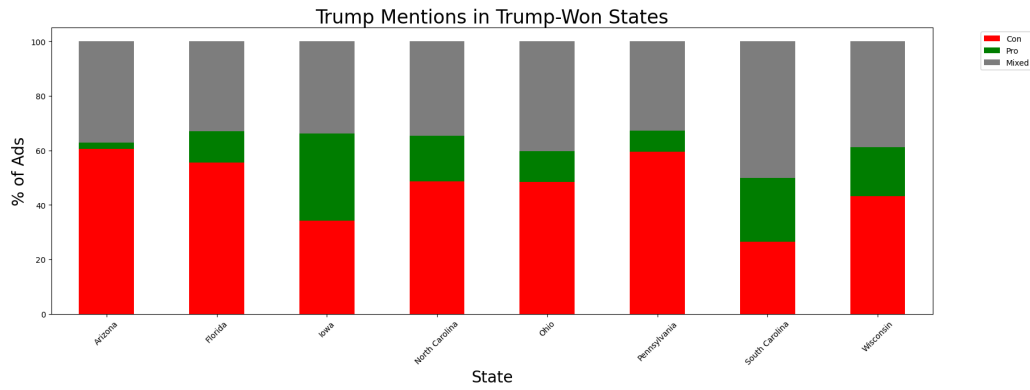
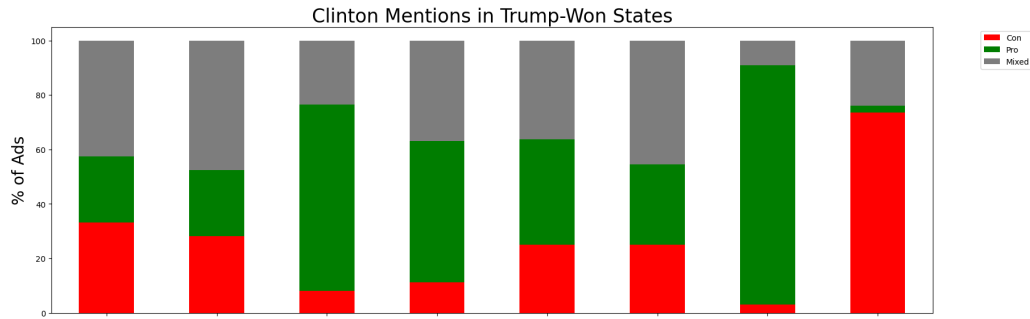
Trump Ads



## Impact of Message Type on Election Results

The unique values of the message include pro, con, mixed, and unknown. While we can infer that the “mixed” type includes both “pro” and “con” sentiments, we cannot form any conclusions about the “unknown” type. An “unknown” can very well be completely positive or completely negative. Therefore, we will omit this category for this section. Not surprisingly, we see more negative messaging towards Trump in states where Clinton won majority votes. However, in states where Trump won, we still see the majority of advertising targeting Trump negatively. This seems to imply that Trump’s opponents leveraged negative attacks on Trump through TV advertising than Clinton’s opponents and negative attacks have little impact on election results.





## Key Takeaways

Looking at topics ads covered, we see similarities within party lines. Looking at topics leading candidates focus on across states, we see more consistency with Trump. This was the case in both states he won and in states where Clinton won. While both Trump and Clinton were more targeted in topics covered in Nevada and Iowa, Clinton made significantly more investments in these states. This strategy did not sway election results in Iowa, but it did in Nevada. Perhaps the high percentage of negatively toned ads against Trump across states he won and lost were the most interesting insights gleaned. Clinton had a much greater percentage of airtime devoted to positive messaging yet failed to win the election.