Project 2 Proposal

Anamika Sinha, Chet Gutwein, Nathaniel Velarde

Summary:

One of the most intriguing and unusual presidential elections in the history of the United States occurred in 2016. While the world's dynamic technological innovations have dictated significant changes in marketing and advertising, an amazing amount of campaign money continues to be spent on print and television advertisements. Thanks to The Political TV Ad Archive, we have data for campaign advertising in the most critical markets for the 2016 election. With this data in hand we intend to pursue answers to the following questions:

- Did television advertisements have a quantifiable impact on the 2016 presidential election?
- How did each presidential candidate use positive and negative messaging, how did this vary across markets, and how did the advertising strategy evolve throughout the campaign?
- How does each candidate's advertising spending correlate geographically?
- How does advertising volume fluctuate throughout the timeline of the election?
- What proportion of ads were fact-checked?

Dataset:

2016 Election Data (http://politicaladarchive.org/data/)

Critical Parameters

'Location' - market in which TV ad was aired

'Start time' - date and time advertisment began

'Stop time' - date and time advertisement began

'Sponsors' - funding source for advertisement

'Candidates' - candidates featured in ad

'message' - positive or negative

Preliminary data cleaning will include removing unused data series, filtering out any ads that do not relate to presidential election, removing ads that aired prior to Republican National Convention. In addition, dates and times will be reformatted and total airtime for each ad in units of second will be calculated.

Supplemental Datasets:

- 2016 Presidential Election Electoral College Results (http://uselectionatlas.org/RESULTS/data.php?year=2016&datatype=national&def =1&f=0&off=0&elect=0)
- Unique ads Archive - Ad specific data which provides fact checking info among other items (http://politicaladarchive.org/api/v1/ads?output=csv)

Approach:

Each ad will be quantified in terms of how many seconds it was aired. Although certain networks, programs, and timeslots are more expensive than others - all add time will be treated equally.

The 2016 Election Results for each state included in the dataset will be compared to total ad time in order to determine if there is a correlation between TV advertising and final vote margin. This will be done cumulatively and also on a state by state basis. A preliminary bar graph plot of number of ads aired in each state is shown below:
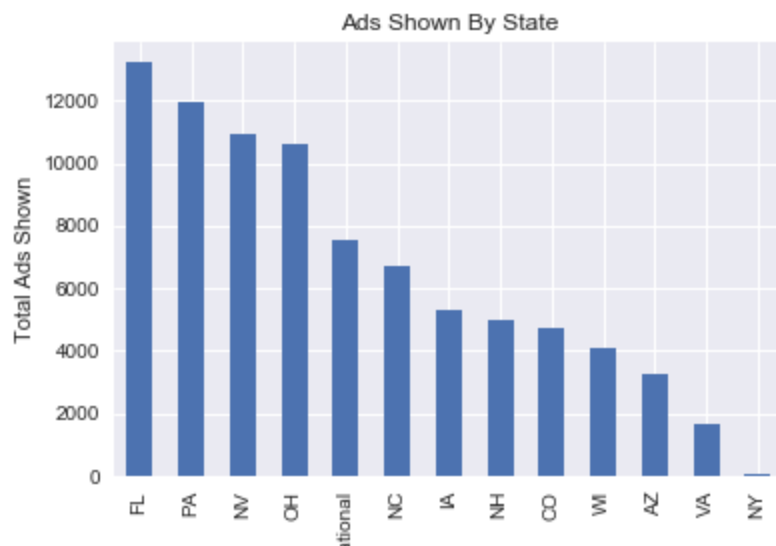
We will be performing a similar analysis based on message type - a preliminary bar graph plot of messaging type for all presidential election ads is shown below:
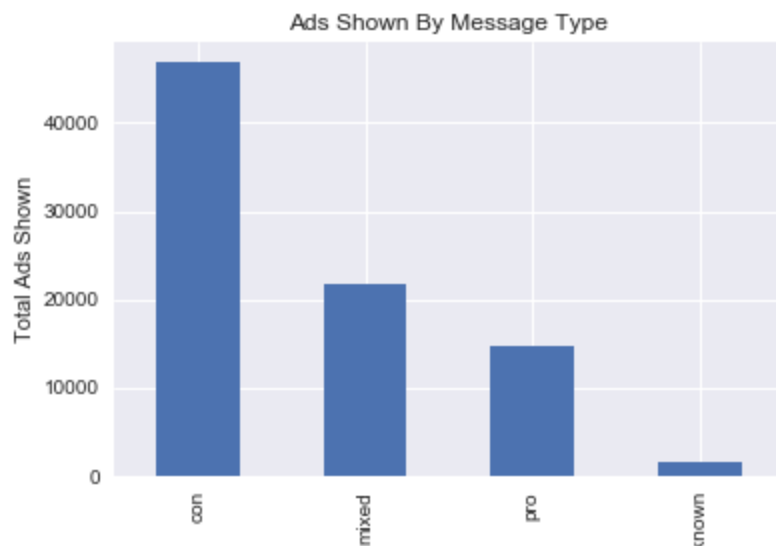
Some data will be plotted as a function of time in order to determine how advertising increased throughout the length of the campaign. Also, the type of message will be plotted versus time and also versus location to determine potential correlations.

Finally, specific ad data will be integrated into our analysis taking into account attributes of individual ads such as whether an ad was fact-checked in order to determine if there is any significant correlation. The specific ad factors will be factored into our geographical and time series analyses to determine what discernable effects they may have had.

We expect to gain some insight and answers to our questions and look forward to reporting our results!