

# Political Ads in the 2016 Presidential Election



Bronte Baer, Jean-Luc Jackson, Christian Montecillo, and Richard Robbins

# Agenda

Project Description

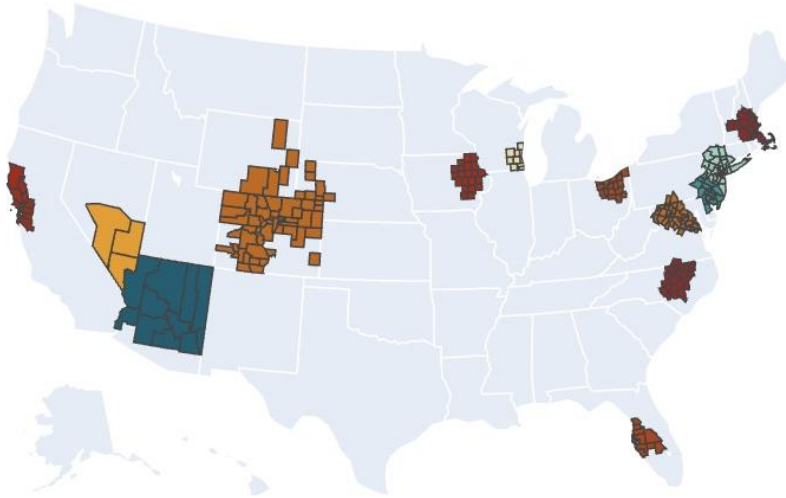
Cleaning and Refining the Data

Data Insights and Questions

Conclusions

# Project Description

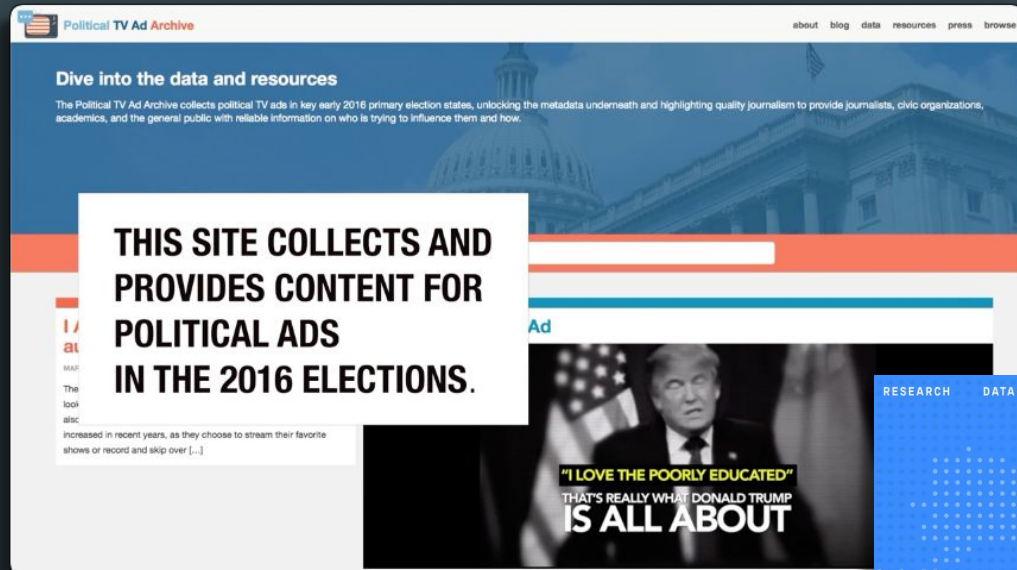
## Regional Broadcast Markets



### Regions

- Boston Region
- Cedar Rapids Region
- Cleveland Region
- Denver Region
- Las Vegas Region
- Milwaukee Region
- New York City Region
- Philadelphia Region
- Phoenix Region
- Raleigh Region
- San Francisco Region
- Tampa Region
- Washington Dc Region

# Cleaning and Refining the Data



The screenshot shows the homepage of the Political TV Ad Archive. The header includes the site logo and navigation links: about, blog, data, resources, press, browse. A main heading reads "Dive into the data and resources". Below it, a paragraph states: "The Political TV Ad Archive collects political TV ads in key early 2016 primary election states, unlocking the metadata underneath and highlighting quality journalism to provide journalists, civic organizations, academics, and the general public with reliable information on who is trying to influence them and how." A large white text box is overlaid on the page with the text: "THIS SITE COLLECTS AND PROVIDES CONTENT FOR POLITICAL ADS IN THE 2016 ELECTIONS." Below this, a video player shows a clip of Donald Trump with the subtitle: "I LOVE THE POORLY EDUCATED" THAT'S REALLY WHAT DONALD TRUMP IS ALL ABOUT.

Political TV Ad Archive

about blog data resources press browse

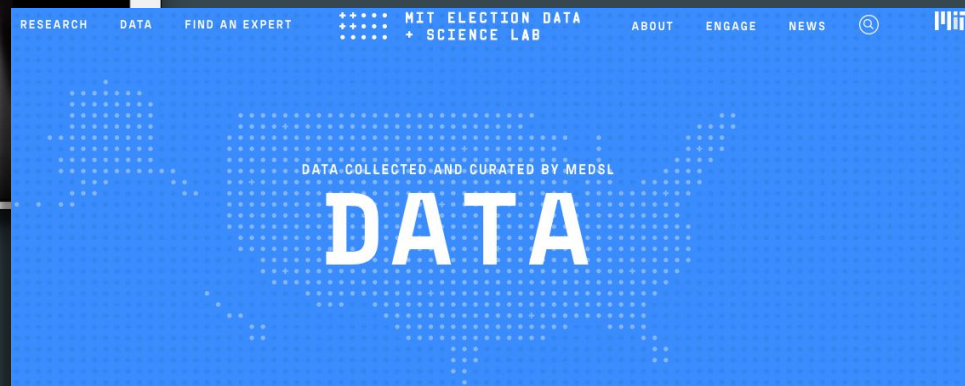
**Dive into the data and resources**

The Political TV Ad Archive collects political TV ads in key early 2016 primary election states, unlocking the metadata underneath and highlighting quality journalism to provide journalists, civic organizations, academics, and the general public with reliable information on who is trying to influence them and how.

**THIS SITE COLLECTS AND PROVIDES CONTENT FOR POLITICAL ADS IN THE 2016 ELECTIONS.**

Ad

"I LOVE THE POORLY EDUCATED"  
THAT'S REALLY WHAT DONALD TRUMP  
IS ALL ABOUT



The banner features a blue background with a white dotted map of the United States. The word "DATA" is prominently displayed in large white letters. Above it, the text "DATA COLLECTED AND CURATED BY MEDSL" is visible. The top navigation bar includes links: RESEARCH, DATA, FIND AN EXPERT, MIT ELECTION DATA SCIENCE LAB, ABOUT, ENGAGE, NEWS, and the MIT logo.

RESEARCH DATA FIND AN EXPERT

MIT ELECTION DATA  
SCIENCE LAB

ABOUT ENGAGE NEWS

DATA COLLECTED AND CURATED BY MEDSL

**DATA**

# Cleaning and Refining the Data

```
1 # We are only looking at the presidential race.
2 df = df[df.race == "PRES"]
3
4 # We are only considering ads that mention Trump or Clinton.
5 df = df[df.candidates.str.contains("Trump|Clinton", case=False, regex=True)]
6
7 # Eliminate that are redundant or otherwise not relevant to our project.
8 df = df.drop(columns=["id",
9                       "wp_identfier",
10                      "race",
11                      "cycle",
12                      "type",
13                      "date_created"])
```

```
1 # Limit votes data to 2016 results
2 # remove year, office and version columns as being uniform across the dataset
3 # remove state_po column as it provides the same information as state
4 # remove totalvotes column from source dataset we substitute our own below
5 # remove mode column as we do not use it
6
7 votes_df = votes_df[votes_df.year == 2016]
8 votes_df = votes_df.drop(
9     columns=['year', 'state_po', 'office', 'totalvotes', 'version', 'mode'])
10
11 # We shorten and simplify the region_id values.
12 # We limit the dataset to rows with the region_ids of interest.
13
14 region_ids = ["boston_region",
15               "cedar_rapids_region",
16               "cleveland_region",
17               "denver_region",
18               "houston_region",
19               "los_angeles_region",
20               "miami_region",
21               "minneapolis_region",
22               "new_york_region",
23               "oakland_region",
24               "philadelphia_region",
25               "portland_region",
26               "san_francisco_region",
27               "seattle_region",
28               "washington_dc_region",
29               ]
```

# Data Insights

## Four Main Topics

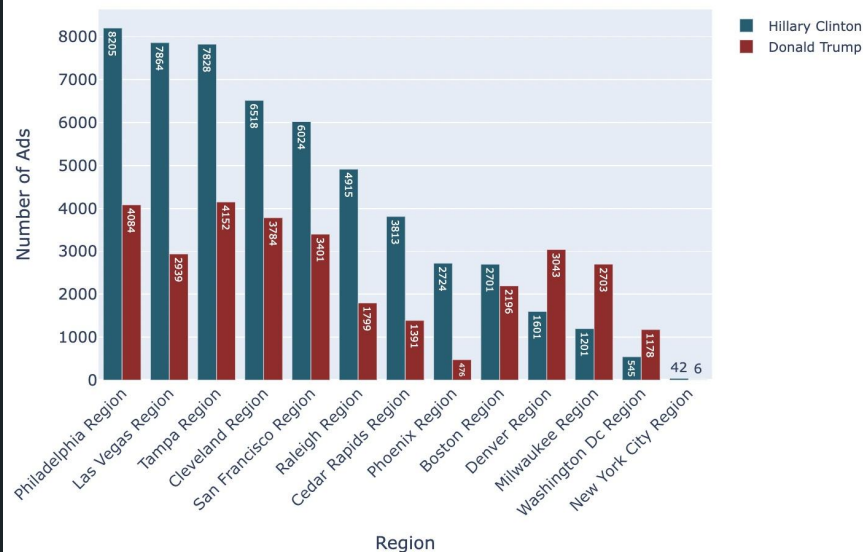
1. Volume
2. Subject Matter
3. Tone
4. Timing

---

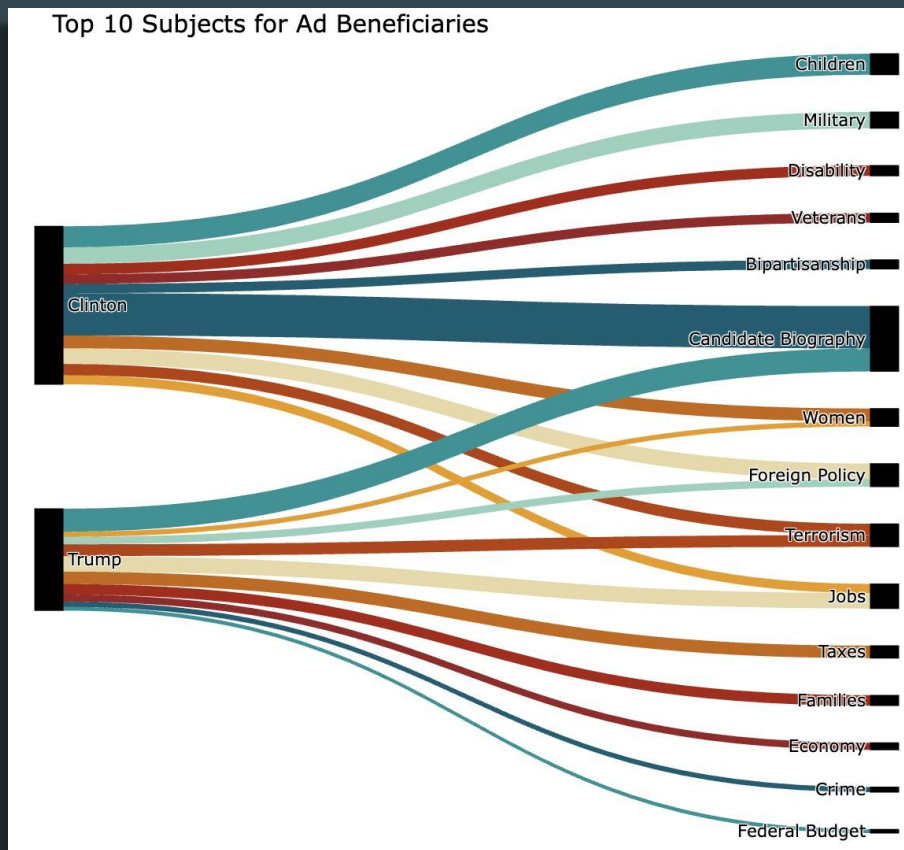
# Data Insights: Volume

Region	Most Votes	Most Ads
Boston Region	democrat	democrat
Cedar Rapids Region	democrat	democrat
Cleveland Region	democrat	democrat
Denver Region	democrat	republican
Las Vegas Region	democrat	democrat
Milwaukee Region	democrat	republican
New York City Region	democrat	democrat
Philadelphia Region	democrat	democrat
Phoenix Region	republican	democrat
Raleigh Region	democrat	democrat
San Francisco Region	democrat	democrat
Tampa Region	republican	democrat
Washington Dc Region	democrat	republican

Number of Ads for Ad Beneficiary by Region



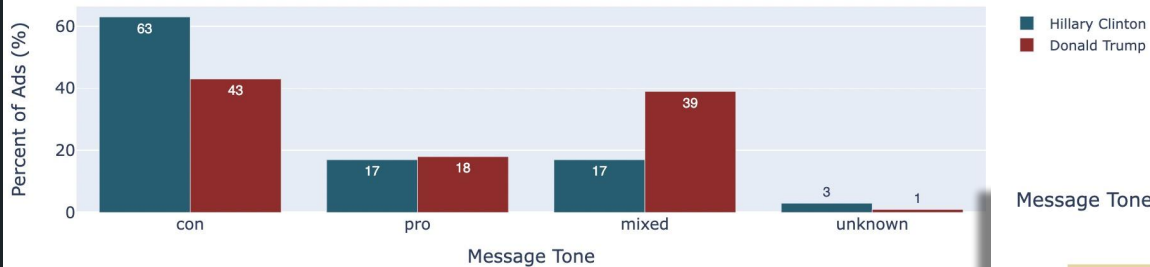
# Data Insights: Subject Matter



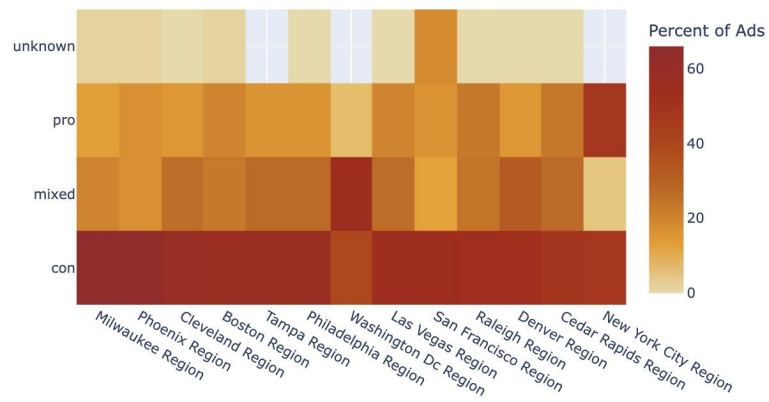


## Data Insights: Tone

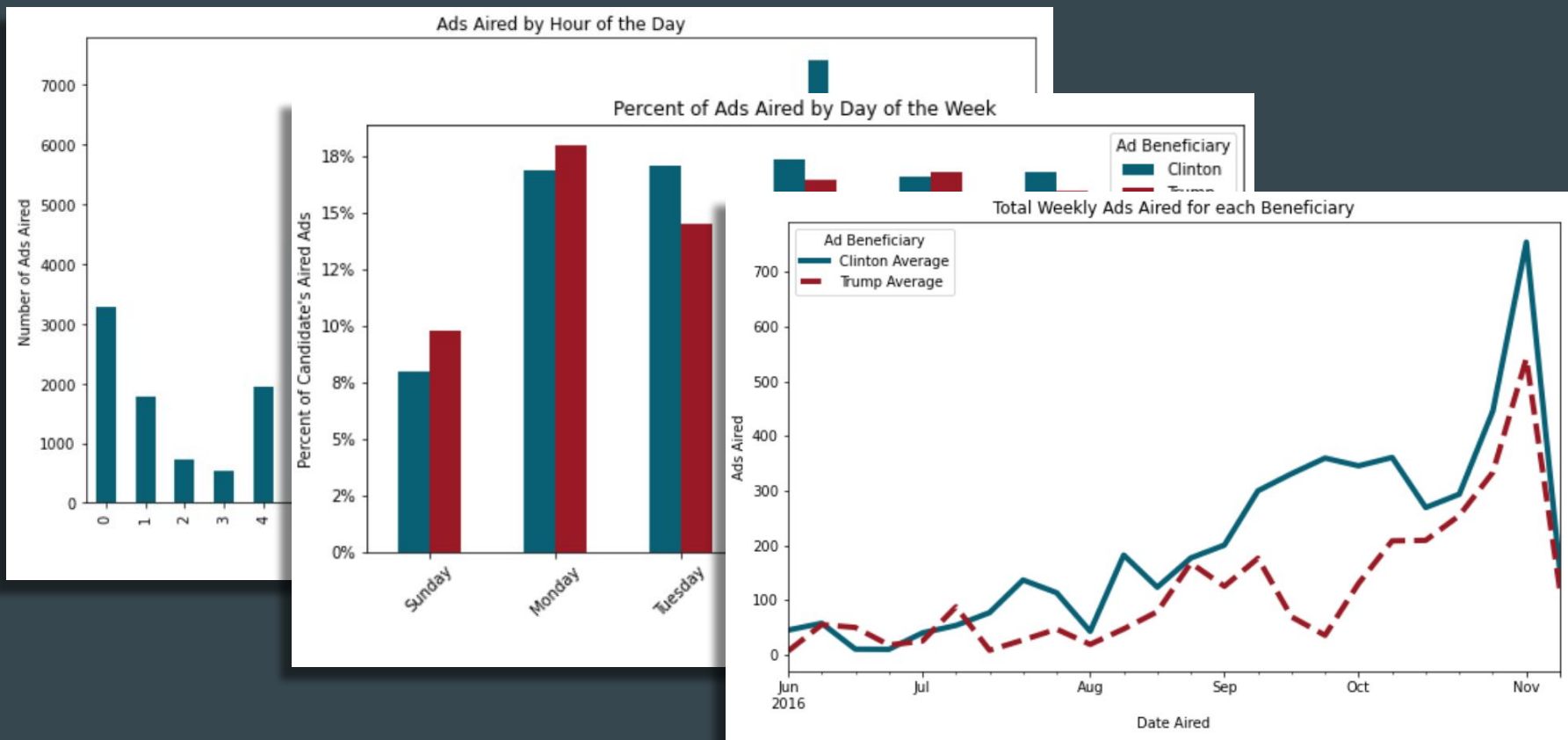
### Percent of Ads by Message Tone for Ad Beneficiary



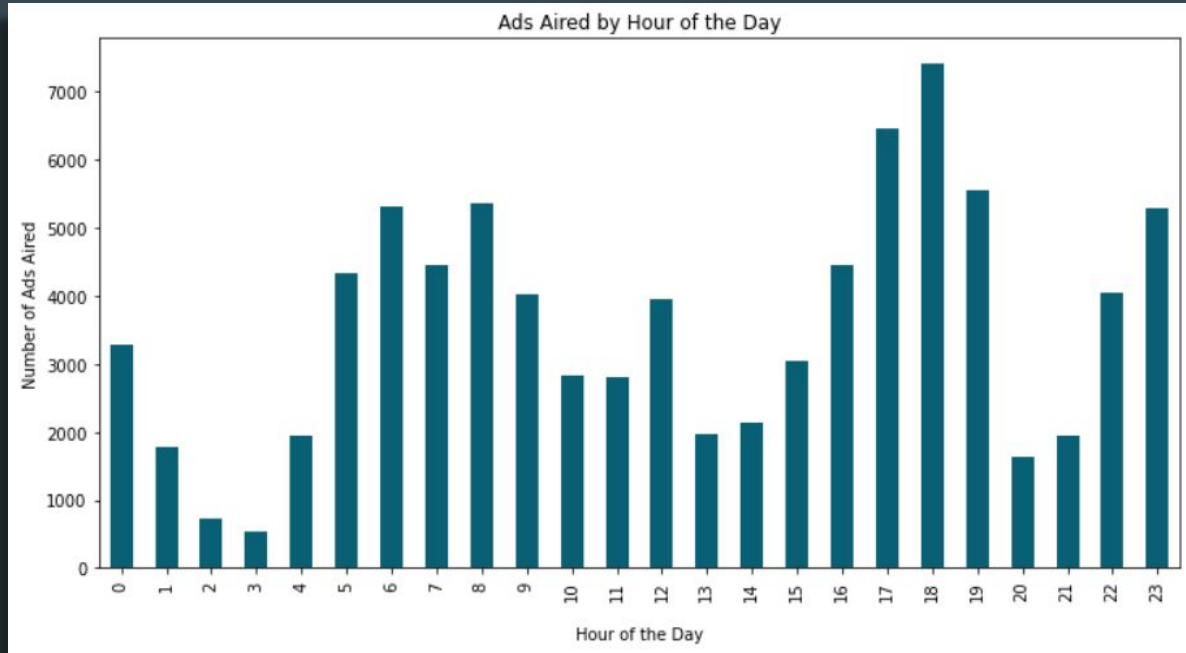
### Message Tone by Region



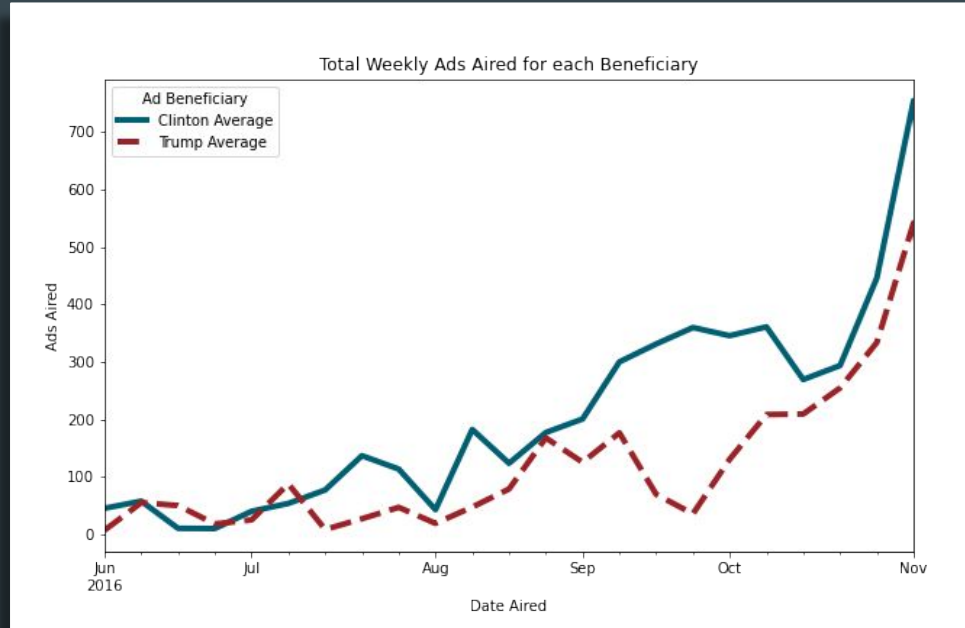
# Data Insights: Timing



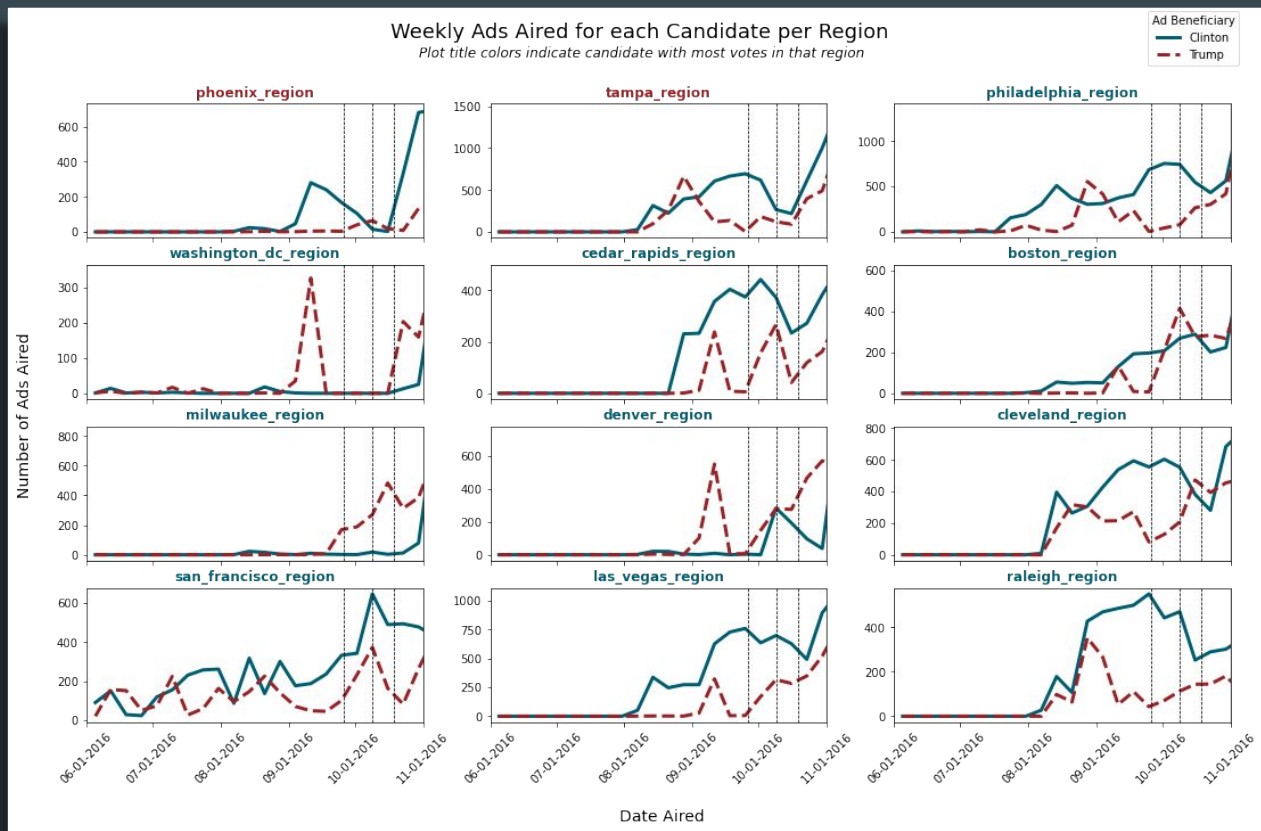
# Data Insights: Timing



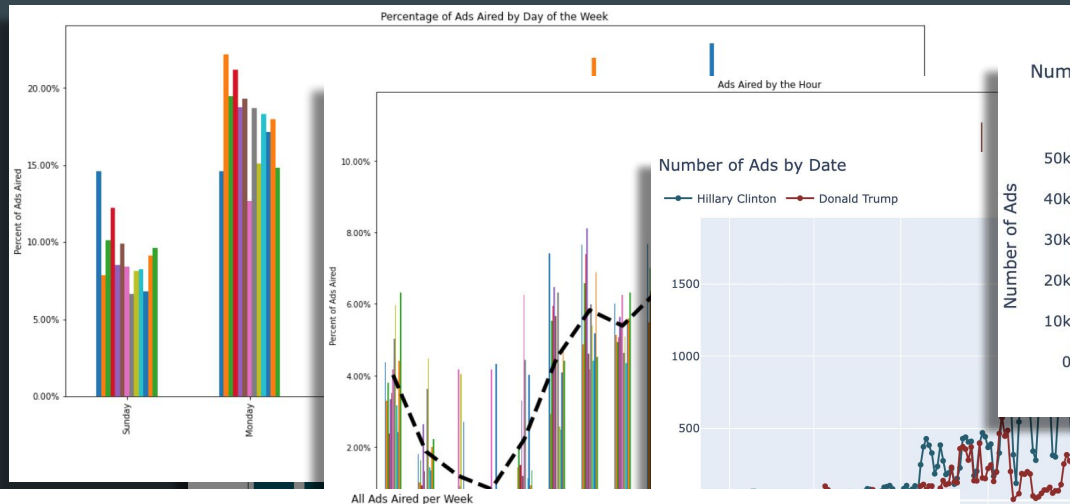
# Data Insights: Timing



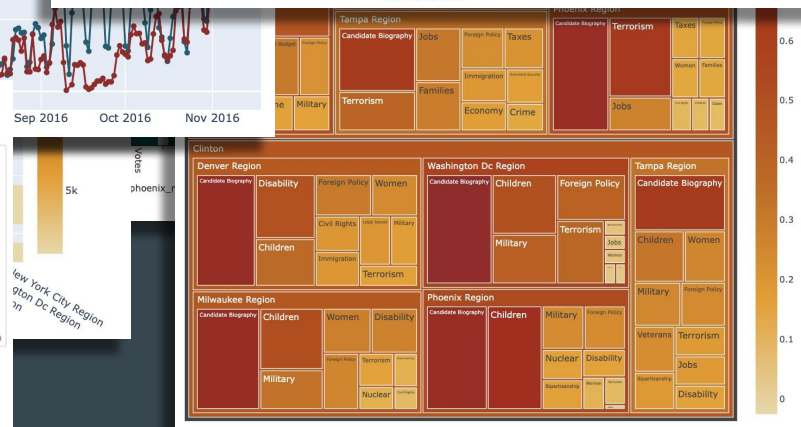
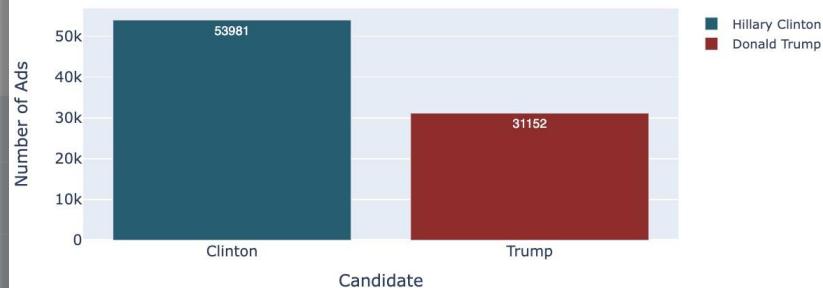
# Data Insights: Timing



# Conclusion



Number of Ads by Candidate



**Thank you**