Tableau Visualization: Newspaper Coverage of Clinton/Trump Tweets During the 2016 Presidential Race

Joshua Timm

PhD Candidate, University of Southern California

October 15, 2019

1 Josh Knows Tableau and Data Visualization

This document should provide a basic demonstration of data visualization abilities. First, I'll describe a brief research question and then display data visualizations that answer that research question[1].

Research Question

Of course, everyone is quite familiar with the 2016 presidential election and all the surprises and strong feelings that went along with it. After the election, I was particularly curious about the candidates' use of Twitter during the 2016 presidential campaign and mainstream media coverage of the two candidate's twitter accounts. I personally perceived much more media coverage of Donald Trump's tweets than Clinton's tweets during the 2016 campaign, which caused me to be surprised when I learned Hillary Clinton actually tweeted more than Donald Trump. I questioned if my perception that Trump's tweets received more coverage than Clinton's was actually correct, so I collected Twitter data and newspaper article data to make an empirical finding. The basic result of this research is that while Hillary Clinton tweeted much more than Donald Trump during the 2016 election, the media didn't seem to care as much about her tweets. In other words, Donald Trump's tweets were simply deemed more newsworthy by publications.

I isolated the period of study to the general presidential election of 2016. While the specific date boundary for when the general election started is somewhat flexible given factors like nominating conventions and delegate votes, I consider the general election period as June 8, 2016 to November 9, 2016. I chose this date because June 8 is the date Hillary Clinton was considered the presumptive nominee for the Democratic Party and November 9 is the day the election was officially decided. I collected the tweets made during this period from both Hillary Clinton and Donald trump using the Twitter API. Hillary Clinton tweeted 3,963 times during this period while Donald Trump tweeted 1,804 times. To measure newspaper coverage of candidate's twitter usage, I searched for newspaper articles across the Proquest U.S. Newsstream, a database containing articles from 1109 unique publications. Specifically, I searched for the words "Trump" and "Clinton" located within 5 words of the terms "Twitter" or "Tweet", all terms being case insensitive. This is of course not a perfect measure, but it works as a broad tool to collect newspaper articles that mention the candidate's twitter usage or dicuss the candidate's tweets.

For the period of the general election, the searches yielded 4,504 articles for Trump and 1,550 articles for Clinton. I isolated the articles published by the state they were published in and created state-by-state coverage maps displaying the newspaper coverage of tweets by Clinton and Trump, as well as a figure displaying the difference in coverage to visualize which candidate received more coverage. Those figures were generated using Tableau and are reported below.

---

[1] I do not include data or code source files here simply for sake of parsimony, but I would be quite happy to share any and all data sources/code files upon request! Email me @ jtimm@usc.edu

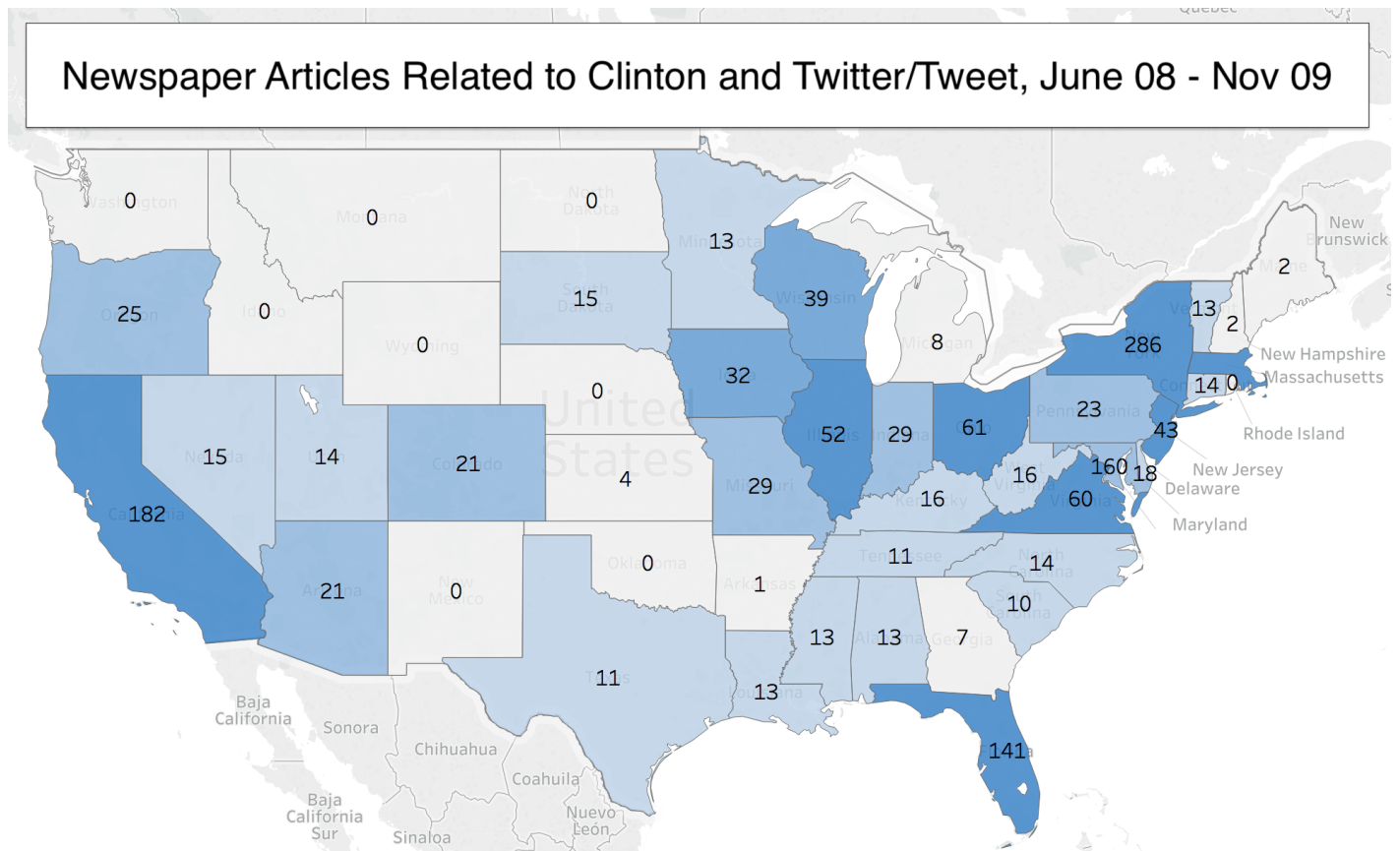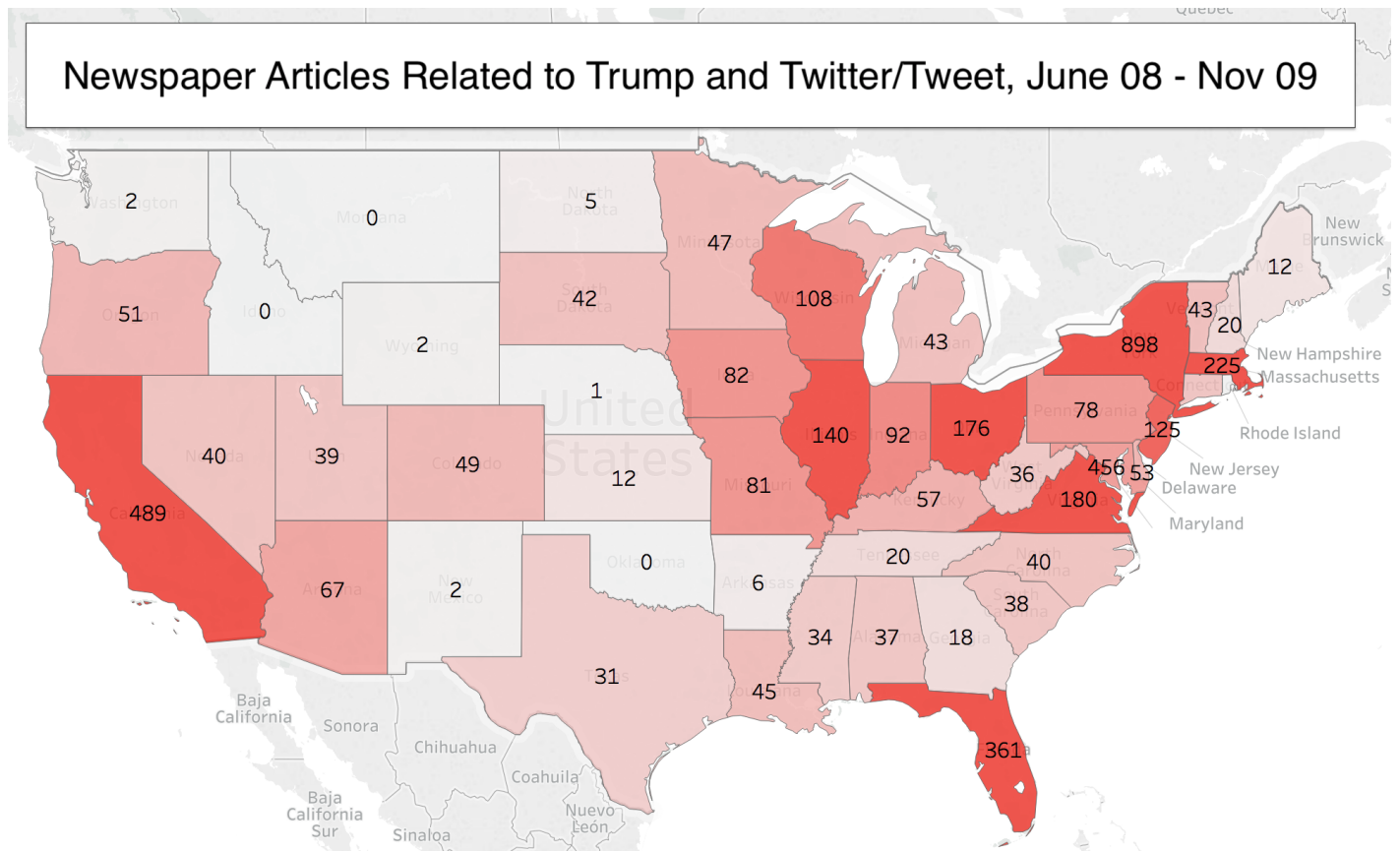Figure 1: Clinton Tweet Coverage Map, June 08, 2016 - Nov 09, 2016

Newspaper Articles Related to Clinton and Twitter/Tweet, June 08 - Nov 09

Figure 2: Trump Tweet Coverage Map, June 08, 2016 - Nov 09, 2016

Newspaper Articles Related to Trump and Twitter/Tweet, June 08 - Nov 09

These maps may be interpreted as follows: Take the state of California as an example. The map for Clinton displays the number 182 to show that there were 182 articles published in California publications that contained the word "Clinton" within 5 words of the term "Twitter" or "Tweet" (case-insensitive). Similarly for Trump, the number displays 489 which means there were 489 articles published in California publications that contained the word "Trump" within 5 words of the term "Twitter" or "Tweet" (case-insensitive). In the map that displays the difference between the two candidates, the number for California is 307, representing that there were 307 more newspaper articles ostensibly about Trump's tweets than about Clinton's tweets.

While it is abundantly clear from these two maps that the media covered Donald Trump's tweets more than Hillary Clinton's, I also include a map to clearly display how much more coverage Donald Trump received.

Figure 3: Trump Tweet Coverage Map, June 08, 2016 - Nov 09, 2016

Difference Between Trump/Clinton Newspaper Articles Related to Twitter/Tweet

Additionally, I downloaded the full text of each of these articles and used regular expressions to track if any candidate tweets were quoted within the articles. This data makes it even more clear that Newspapers were not including Clinton's tweets as often as they were including Trump's tweets.

		Clinton	Trump
1	Total # of Tweets	3963	1804
2	Total # of unique tweets covered in news stories	11	173
3	Total # of news stories on unique tweets	98	604
4	Total # of news stories on candidate and Twitter/tweet	1550	4504

Hillary Clinton tweeted 3,963 times and newspapers quoted 11 of her tweets across 98 stories. Donald Trump tweeted 1,804 times and newspapers quoted 173 of his tweets across 604 stories. Of course this means that newspapers were covering Donald Trump's tweets at a much higher rate than Hillary Clinton's tweets, offering some insight as to why Donald Trump became so popular: his twitter activity was deemed more newsworthy than Clinton's.

I include similar maps that track the number of tweets posted by each candidate that were covered in newspaper articles across the 48 contiguous states of the U.S. These maps may be interpreted as follows: Again, take the state of California as an example. On the map for Clinton's tweets quoted in newspapers, the number 17 displays across the state of California. That indicates that 17 articles in California-based publications quoted at least one of Clinton's tweets. On the map for Trump's tweets quoted in newspapers, the number 82 indicates that 82 articles in California-based publications quoted at least one of Trump's tweets. And again the difference map shows the number 65, indicating that 65 more articles quoted Trump's tweets than articles quoting Clinton's tweets.

Figure 4: Number of Clinton Tweets Quoted in Newspaper Articles
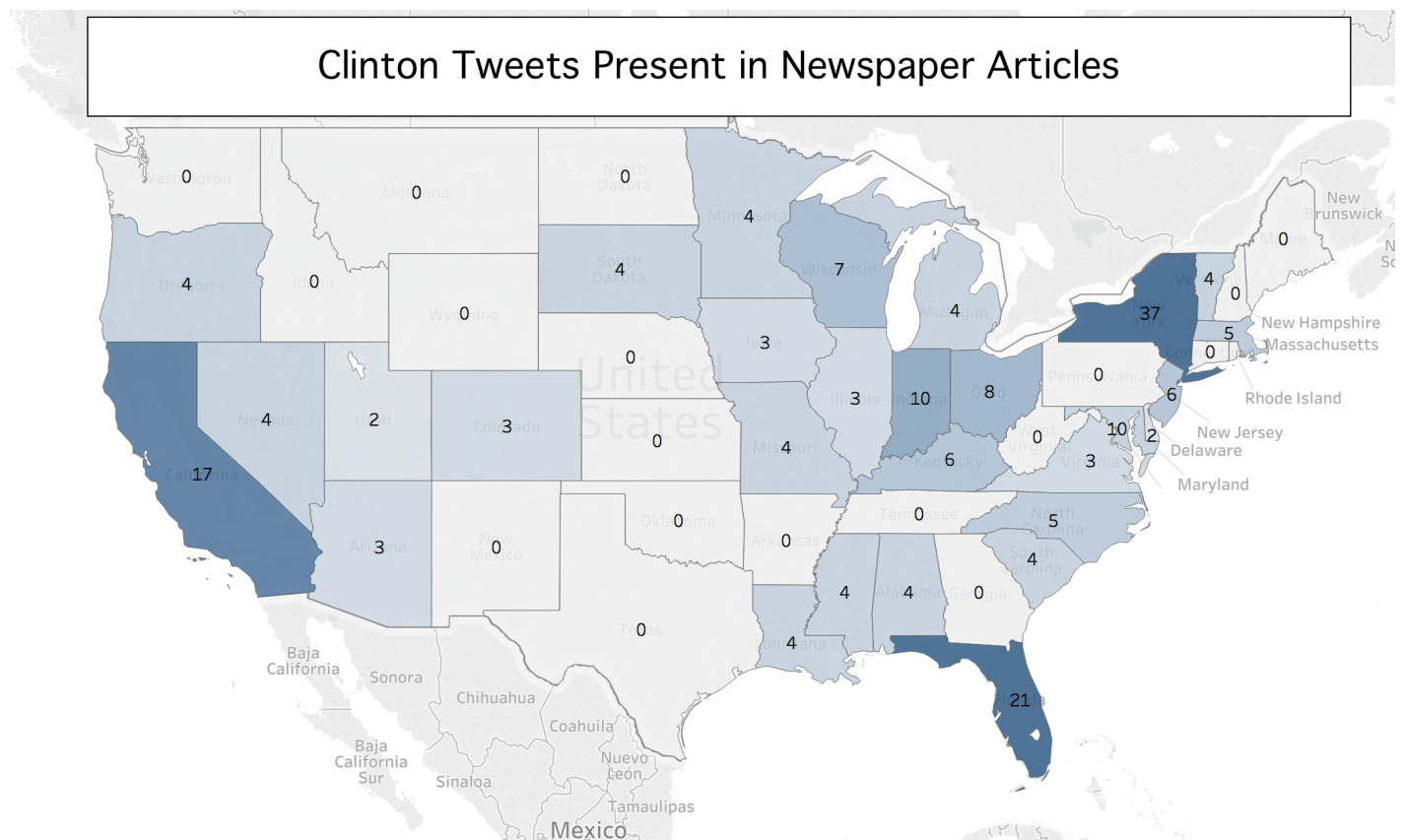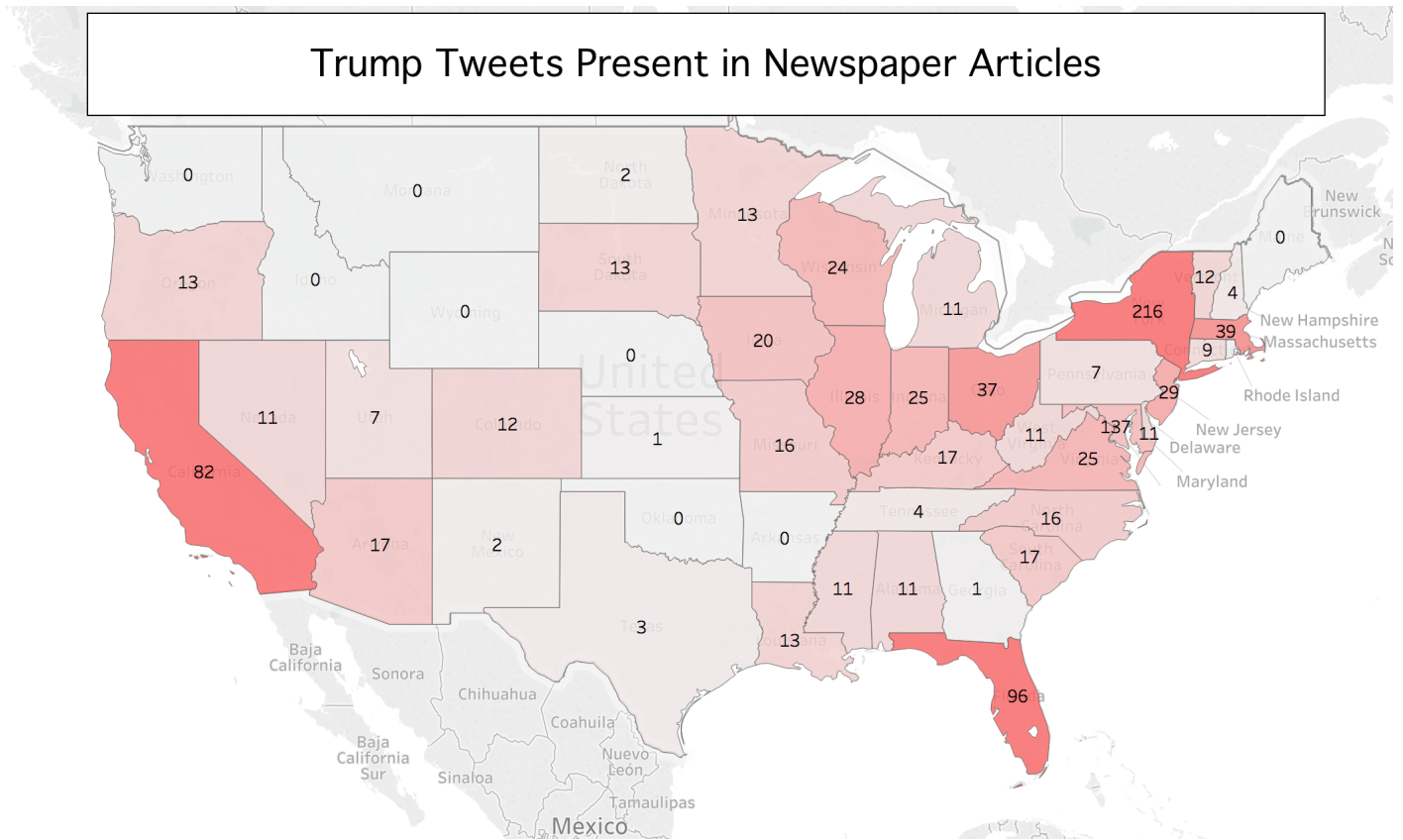
Figure 5: Number of Trump Tweets Quoted in Newspaper Articles

The picture here is quite clear and I won't spend much time clearly describing the results. In every single state, newspaper articles quoted tweets posted by Donald Trump at considerably higher rates than they quoted tweets posted by Hillary Clinton. This offers further support to the notion that Donald Trump was better at handling the media during his presidential campaign, which contributed to his likelihood of being elected.

Figure 6: Difference in Number of Tweets Quoted in Newspapers, Trump-Clinton

Difference Between Trump/Clinton Tweets Present in Newspaper Articles