Ethan ElHamahmy, Ella Grady, Eve Harmon, Ali McKiernan

Research Methods

Professor Scoggins

4 March, 2022

How Do You Feel About Women?: An Analysis of Tweets About the 2017 Women's March

Throughout history, marches have been used as a disruptive form of protest to apply pressure and call attention to the need for political change. The Women's March of 2017, following the historical precedent seen in the Woman Suffrage Parade of 1913 and the Women Strike for Equality of 1970, protested misogynistic ideals. The sexism in American society that was highlighted in the ideals espoused by President Donald Trump served as the catalyst for the nationwide protests that saw millions of protestors take to the streets across the United States. Our group, motivated to conduct this research by our beloved Professor Scoggins, decided to ask the question: how did the sentiment of tweets around the 2017 Women's March change when President Trump was mentioned versus when he was not mentioned?

The Women's March and Donald Trump's inauguration were both talked about a lot on Twitter during the time they happened in 2017. We decided to focus the research on both of these topics because of the role the election played in the organization of the Women's March. The purpose of using content analysis and Twitter specifically was to gauge people's opinions and overall feelings when these events were taking place. Twitter is a place where people typically express strong opinions, therefore it was not difficult to find tweets from people who had strong opinions about both the Women's March and Trump's inauguration.

We analyzed Tweets that referenced or mentioned Trump and his inauguration, and compared these to those that did not. Prior to gathering data, we hypothesized that the tweets referencing Trump would generally be negative, and those that did not would generally be more positive. We assumed that most people tweeting would be in support of the March itself, which is why it made sense to us that tweets referencing both the Women's March and Trump would have negative connotations towards Trump.

The unit of analysis that we used was the sentiments of the tweets that we chose to focus on. We analyzed tweets from December 2016 through January 2017 that referenced the Women's March and Trump's inauguration, and made decisions about what words or phrases we were going to classify as having a positive or negative connotation. These mainly included words and phrases that were used commonly by people who either attended or supported the Women's March in some capacity to reference it and explain why they supported it, such as inauguration, administration, and presidency.

When hand coding, it was easier to determine the sentiment of several tweets based on our knowledge of context and connotation that we have from living in the world, using Twitter, and being aware of attitudes towards the Women's March in 2017. The computer, however, is unable to apply those same intricacies of attitudes in the tweets. This resulted in many tweets, especially in the dataset that mentioned Trump, being coded as neutral despite a positive or negative coding when done by hand. When run through the computer, the computer coded these tweets as neutral because they may not have included words that were in either of the provided Liu dictionaries, or used an equal amount of positive and neutral words that summed to zero. When hand coding these same tweets, we are able to determine, based on our context awareness, that the sentiment of these tweets are either negative or positive. While computer coding makes

sentiment analysis easier when examining large datasets, it is important to understand the

limitations of this method as computers do not have the same awareness and understanding of

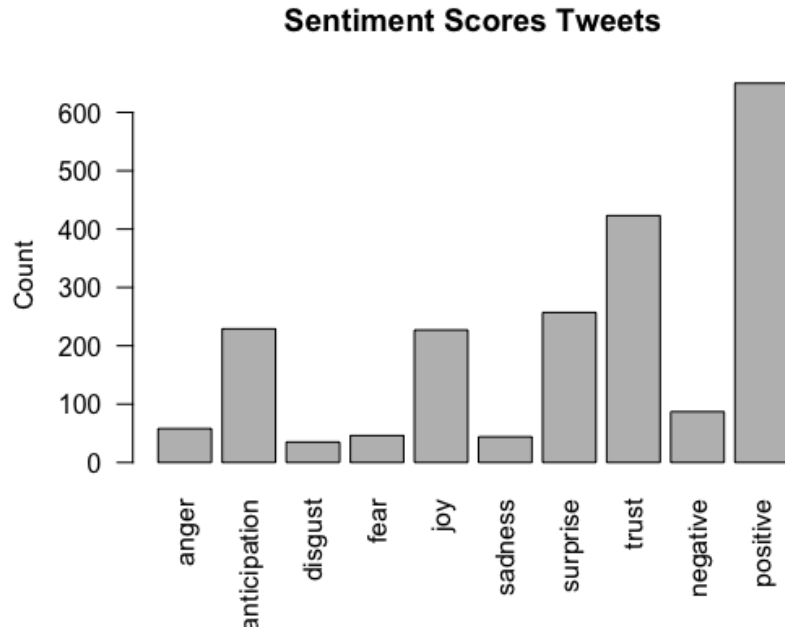the world and society.

Results Obtained from Our Research

Sentiment Scores Tweets

Figure 1. Emotions of Tweets that mention Trump, his inauguration, or administration
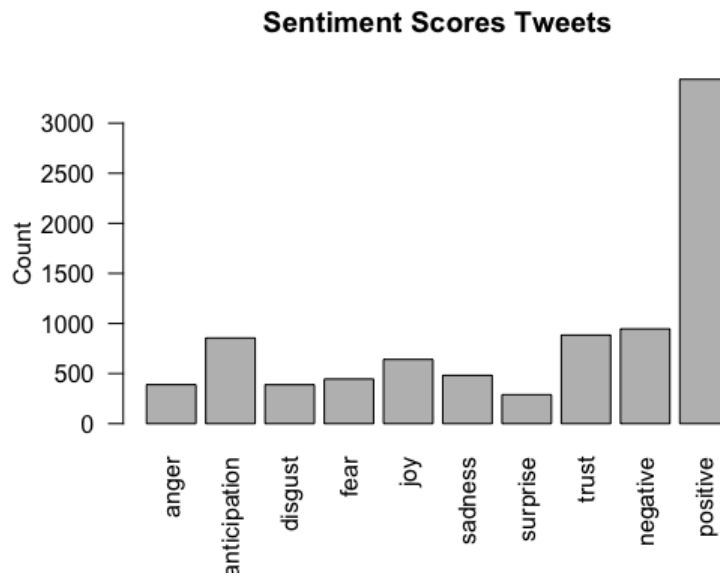

Sentiment Scores Tweets

Figure 2. Emotions of tweets that do not mention Trump, his inauguration, or administration

We originally thought that tweets that directly mentioned President Trump, his inauguration, administration, or platform would have more negative sentiments, but after coding the data with the provided dictionaries, we found that many of the tweets from the Trump dataset were coded as neutral. The Trump dataset contained about 350 tweets, restricted to those that mention words such as Trump, presidency, administration, and inauguration. Of these tweets, the majority were coded by the program as positive, as shown in Figure 1, likely because many of the tweets that we had coded by hand as negative, included words that were not in the negative words dictionary. With differences like these that make it more difficult to determine with the computer. For the dataset of tweets that specifically excluded tweets referencing Trump, his presidency, administration, and so on, there were about 2,500 tweets included. This dataset more closely followed our hypothesis, with the majority of tweets coded as positive, as shown in Figure 2.

Comparing the data from both sets, the Trump dataset was coded with a greater variety of emotional sentiments, as shown in Figure 1, while the Trump excluded tweets majority were coded mostly as positive, as shown in Figure 2, with little variety amongst the other emotions. In the Trump dataset, with the numbers of about 650 positive counts to 100 negative counts, versus the Trump excluded tweets counts of 3,500 positive counts to 800 negative counts, this may provide further proof that our hypothesis was incorrect, however, it is important when examining this data to keep in mind that the computer is unable to apply the same context and understanding when coding tweets that we, as humans, are able to see in the meaning of the tweets.

Our computer generated data tells us that the attitude on social media regarding the Women's March, regardless of whether President Trump was mentioned, was positive and focused on fighting for equality in all aspects of society, although there are certainly many outlying examples of hateful and mean spirited tweets. Looking past the immediate timing of the Women's March, these attitudes towards equality and women's rights movements have remained prevalent and similar throughout the entirety of Trump's presidency and beyond, with many focusing on the fight for equality, while opponents still remain openly vocal on social media. Throughout this research project we encountered several challenges, however, the primary challenge we faced was adopting, understanding and using the different programs, including R and the tweet hydrator. With past experience in programming as a computer science major, one of our group members, Ella, was able to guide us through these tasks so we could better understand and analyze the data. In addition to our findings on the research topic and working through technical challenges, this research project also taught us to be aware of the differences that occur between hand coding and computer generated coding, and the way context can be a major determining factor on a tweet's sentiment.

Works Cited

Benoit K, Watanabe K, Wang H, Nulty P, Obeng A, Müller S, Matsuo A (2018). "Quanteda: An
R package for the quantitative analysis of textual data." *Journal of Open Source Software*,
3(30), 774. doi: 10.21105/joss.00774, https://quanteda.io.

"Hu, M. & Liu, B. (2004). Mining and Summarizing Customer Reviews. In Proceedings of the
ACM SIGKDD International Conference on Knowledge Discovery and Data Mining
(KDD-2004), Aug 22--25, 2004, Seattle, Washington, USA.
https://www.cs.uic.edu/~liub/publications/kdd04-revSummary.pdf

Littman, Justin; Park, Soomin, 2017, "Women's March Tweet Ids",
https://doi.org/10.7910/DVN/5ZVMOR, Harvard Dataverse, V1