Using Twitter Sentiment Analysis to Analyze Self-Sentiment of the POTUS

Gwendolyn White
whiteg@xavier.edu
Information Systems
Xavier University
Cincinnati, Ohio

Hayden Wimmer
hwimmer@georgiasouthern.edu
Information Technology
Georgia Southern University
Statesboro, Georgia

Carl Rebman
carlr@sandiego.edu
Supply Chain Management
University of San Diego
San Diego, California

Christian Nwankwo
cn02099@georgiasouthern.edu
Information Technology
Georgia Southern University
Statesboro, Georgia

Abstract

This article provides a discussion of different tools that can be used to perform research on Twitter. The study then utilizes sentiment analysis to demonstrate self-sentiment of the POTUS and compare it with popular news sources. A comparison in 2 time periods and 4 months apart was made to determine if there is a change in the self-sentiment of the POTUS versus common news sources. To perform sentiment analysis, we utilize Python and the Vader and Pandas libraries, and statistical analysis was performed for each of the datasets. The first round of tests were based on a November dataset and revealed the means between the public, FOX, CNN and the President were not equal and that the POTUS had a higher self-sentiment than the sentiment of the news sources. The second round of tests were based on the following April data collection and revealed that the self-sentiment of the POTUS was not significantly different from the public or FOX. However, it was significantly different from CNN indicating a possible clash between the POTUS and CNN news.

Keywords: Sentiment Analysis, Trump, Self-Sentiment, POTUS

1. INTRODUCTION

The creation or rather the explosion of the Internet has brought much innovation and disruption to many industries and affecting virtually everyone in the world. Perhaps one of the strongest disruptive results of the Internet is social media, which many people consider as a normal part of their lives. Social media is defined as web-based and mobile-based Internet applications that allow the creation, access and exchange of user-generated content that is ubiquitously accessible (Kaplan & Haenlein, 2010). Two of the largest and most recognized social media applications are Facebook and Twitter. Social media has been an instrument for research in many different areas. For example, Lazer et al. (2009) used social media to investigate questions in computational social science using quantitative techniques (e.g., computational statistics, machine learning and complexity). Cioffi-Revilla (2010) focused on social media big data for data mining and simulation modeling. Batrinca and Treleaven (2015) noted that analyzing social media, in particular Twitter feeds for sentiment analysis, has become a major research and business activity due to the availability of web-based application programming interfaces.

Twitter has changed the social landscape with 140 characters. According to Parra et al. (2016) Twitter has gained a lot in popularity as one of the most used and valuable microblogging services in recent years. Twitter had exponential growth of 1400% between 2006 and 2009 at its peak. In 2016, there were over 500 million tweets per day and approximately 200 billion tweets per year (Internet Live Statistics, 2018). It is easily accessible, and anyone can express his or her ideas. Part of the appeal is the ability of users to follow any other user with a public profile, enabling users to interact with celebrities who regularly post on the social media site. The enormous amount of data generated by the Twitter users has been shown to be very useful in many ways, e.g., to predict the stock market (Bollen, Mao, & Zeng, 2011), to support government efforts in cases of natural disasters (Abel, Hauff, Houben, Stronkman, & Tao, 2012), and to assess political polarization in the public (Conover et al., 2011).

According to Statista, Twitter has also become an important communications channel for governments and heads of state; former US President Barack Obama claimed a runaway first place in terms of Twitter followers, with Indian Prime Minister Narendra Modi and Turkey's President Recep Tayyip Erdoğan ranking second and third, respectively (Statista, 2018b). One item that has been of interest to academics and practitioners is the use of Twitter by the current President Donald Trump and his level of self-sentiment. A few illustrative examples of Trump's tweets are shown below in Figures 1-3.

Figure 1. Trump Tweet Showing His High Self-Sentiment

Trump uses the Twitter platform to express his views regardless of political correctness. One of his favorite targets is Hillary Clinton as seen in Figure 2.
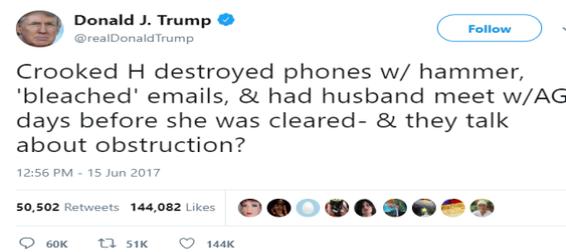

Figure 2. Trump Tweet Regarding Hillary Clinton and His High Self Sentiment

Televisions shows are not safe from his expression of self-sentiment. Saturday Night Live performed a skit that included Trump and he expressed his dissatisfaction (Figure 3).

Figure 3. Trump Tweet of Dissatisfaction on Saturday Night Live Skit

An often and perhaps interesting notation is the difference in Twitter use between former President Barack Obama and current President Donald Trump. For example, Quigley (2017) from Newsweek noted that the current president has significantly less twitter followers (34.1 million versus 92 million for Obama)

http://www.newsweek.com/obama-more-twitter-followers-trump-638890) yet Donald Trump tweets more frequently, is more controversial and thus makes more news headlines. Monaco (2017) from Fortune also reinforced this statement by noting that Obama had more impact at least according to Twiplomacy which is a yearly study/report on how world leaders tweet.

The Twitter platform in combination with Python has the ability analyze the sentiment of a particular set of tweets produces revealing characteristics that are either positive or negative. This paper investigates the sentiment of President Donald Trump explaining his level of self-sentiment in comparison to the opinions of the public, Fox News and CNN.

The format of the study is as follows. First is a discussion of a relevant literature followed by methodology and research hypothesis. Included with the methodology is detailed discussion of system overview and development. The manuscript concludes with results, limitations, and future research.

2. LITERATURE REVIEW

Social media data is clearly the largest, richest and most dynamic evidence base of human behavior, bringing new opportunities to understand individuals, groups and society (Batrinca & Treleaven, 2015). Three areas that have utilized social media for research include business, bioscience, and social science. Within business, the early adopters were the areas of finance and marketing. These areas were interested in measuring brand awareness, fraud detection, and news propagation. The biosciences were interested in using social media to collect data for behavioral change initiatives or impact monitoring of spread of diseases. Computational social science applications include: monitoring public responses to announcements, speeches and events especially political comments and initiatives; insights into community behavior; social media polling of (hard to contact) groups; early detection of emerging events (Batrinca & Treleaven, 2015).

Social sciences analyze social media using content analysis thematic analysis, and sentiment analysis. Twitter is ideal for sentiment analysis based on the availability of text and language (Kouloumpis, Wilson, & Moore, 2011).Twitter is used to express ideas through "tweets" of no more than 140 characters (test markets of 280 characters started in September 2017). On average, there are approximately 313 million Twitter users (Internet Live Statistics, 2018). A special feature of Twitter is the use of hastags (#) that are used to emphasize a topic or keyword of interest which categorize the tweet to make it searchable. In comparison to other social media platforms, Twitter is number 11 in popularity (Statista, 2018a). Each tweet is a way for businesses to analyze customers or individuals to communicate with friends. Tweets are in real time and the information is ideal to determine the feelings and opinions of individuals or businesses. Figure 4 shows Social Media outlets ranked by popularity.
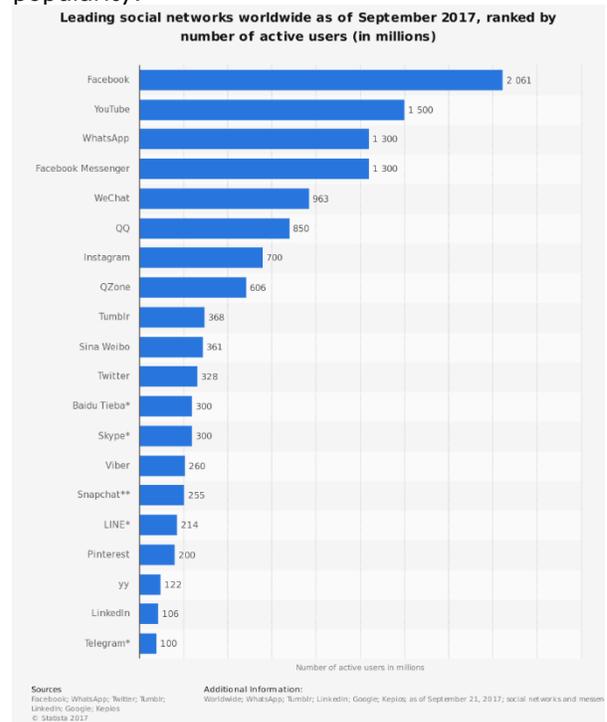

Figure 4. Social networks ranked by popularity (Statista, 2018a).

Sentiment Analysis
Sentiment analysis is "the process of computationally categorizing opinions expressed in a piece of text to determine whether the writer's attitude towards a particular topic, product, etc. is positive, negative or neutral (dictionary.com, 2018). Semantic analysis is divided into two areas: word sense disambiguation and disambiguation. Word sense disambiguation is analyzing a word for its intended meaning. Disambiguation is analyzing how sentence and text are cohesive.

The most common uses of sentiment analysis include the analysis of public opinion, Twitter, Facebook, market research and social sciences. Sentiment analysis has become popular when

used in combination with Twitter to determine the direction of opinions. The combination of Twitter and sentiment analysis makes it easy to complete real-time analysis.

For example, a tweet that proclaims Donald Trump is horrible and Barak Obama was fantastic is interpreted only using the sentiment of each topic and not the entire tweet. There for each topic is treated separately: Trump – negative and Obama – positive.

Originally, sentiment analysis when combined with Twitter was relegated to a crude proprietary focus and was primarily used with the Spanish language. It only used hashtags and emoticons, which were limited in scope. Advances in semantic analysis improved when SemEval started to evaluate Twitter in 2013. SemEval is the International Workshop on Semantic Evaluation that is part of SIGLEX, a special interest group with the Lexicon of the Association for Computational Linguistics. The SemEval performed tasks at each convention that were designed to evaluate text and classify it as negative, positive or neutral sentiment using computers. There are now several tools available for retrieving and analyzing social media data on Twitter (Ahmed, 2017). They are listed below in Figure 5.

Tool	OS	Download and/or access from
Audiense (offers 14 day trial)	Web-based	https://buy.audiense.com/trial/new
Boston University Twitter Collection and Analysis Toolkit (BU-TCAT)	Web-based	http://www.bu.edu/com/research/bu-tcat
Chorus (free)	Windows	http://chorusanalytics.co.uk/chorus/request_download.php
COSMOS Project (free)	Windows	http://socialdatalab.net/software
DiscoverText (offers 3 day trial)	Web-based	http://discovertext.com
Echosec	Web-based	https://www.echosec.net
Followthehashtag	Web-based	http://www.followthehashtag.com
IBM Bluemix	Web-based	https://www.ibm.com/cloud-computing/bluemix
Mozdeh	Windows	http://mozdeh.wlv.ac.uk/installation.html
Netlytic	Web-based	https://netlytic.org
NodeXL	Windows	http://nodexl.codeplex.com
NVivo	Windows and MAC	http://www.qsrinternational.com/product
Pulsar Social	Web-based	http://www.pulsarplatform.com
SocioViz	Web-based	http://socioviz.net
Trendsmap	Web-based	https://www.trendsmap.com
Twitonomy	Web-based	http://www.twitonomy.com
Twitter Arching Google Spreadsheet (TAGS)	Web-based	https://tags.hawksey.info
Visibrain	Web-based	http://www.visibrain.com
Webometric Analyst	Windows	http://lexiurl.wlv.ac.uk

Figure 5. Tools for Twitter Retrieval (Different Twitter Data Retrieval/Analysis Tools, (Ahmed, 2017))

There have been some interesting previous studies using Twitter and sentiment analysis. Bollen et al. (2011) measured sentiment of random sample of Twitter data, finding that Dow Jones Industrial Average (DJIA) prices are correlated with the Twitter sentiment 2–3 days earlier with 87.6 percent accuracy. Wolfram (2010) used Twitter data to train a Support Vector Regression (SVR) model to predict prices of individual NASDAQ stocks. Bae and Lee (2012) investigated sentiment analysis of audiences and Zainuddin, Selamat, and Ibrahim (2018) investigated hybrid sentiment classification. In the realm of health sciences, Oscar et al. (2017) examined Twitter sentiment analysis for stigma of Alzheimer's disease, while Clark et al. (2018) explored perceptions of breast cancer.

These analyses have laid the groundwork for our study using Python as an analysis tool. The programming language is used to identify and categorize opinions through text. Python was chosen due to its flexibility and diverse number of tools that can be used for sentiment analysis. A few programs are used to extract and analyze tweets. Tweepy is an API that is used with Twitter to extract tweets that are later dissected for analysis. TextBlob is a Python library used to process the text data. The natural language tool kit (ntlk) is used along with Python to classify language. LTK has been called "a wonderful tool for teaching, and working in, computational linguistics using Python," and "an amazing library to play with natural language." (NTLK, 2018). The text is introduced to the Python and the tools analyze and create an output for data explanation.

3. METHODS

For this project, we are collecting tweets about Donald Trump from the official twitter handles of CNN and Fox News and also the official handle of the President of United States, Donald Trump. Tweets from all twitter users will also be collected and analyzed together with those from the official twitter accounts. The tweets were collected twice: from November 2017 and in April 2018 using Python programming tool. Figure 6 below gives an overview of the entire process and the steps of data collection, cleaning and sentiment analysis are listed below.
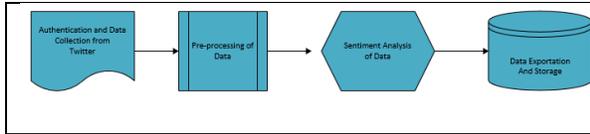
Figure 6: Proposed Methodology

Authentication and Data Collection

To get the tweets, we have to create a twitter developer account, so we can have access to pulling tweets directly from twitter. After creating the account, we get for four authentication keys namely; consumer key, consumer secret, access token and access token secret. Some of the Python libraries and functions used are explained in the following paragraph.

Tweepy.OAuthlander which is a function in the Python tweepy library was used to authenticate the keys generated with twitter while the stream function was used to establish a continuous connection for collecting tweets from twitter users about Trump. The Cursor function was used to collect tweets from the official accounts of the news sources and Donald Trump.

Data Preprocessing

For the data cleaning, tweet-processor which is also a Python package was used for data cleaning. The package was imported into the Python interpreter and applied on every tweet collected. This includes the removal of stop words, urls and other irrelevant features found in the tweets. This process is very important for the results of the sentiment analysis to be correct. The output from the data cleaning process was used for the text analysis phase.

Sentiment Analysis

After cleaning the data, we used Python Vader library for sentiment analysis of the tweets extracted. This library was build off of the Python Natural Language Processing Toolkit and returns the positive, negative, neutral and the compound score of the tweets. The pandas Python library was used to compile each tweet and the corresponding sentiment score category for the tweet. The entire data set is then outputted in csv formatted for further statistical analysis. Listed below this paragraph is the Python script used for implementing the methodology. The Python packages used in this project include; nltk, numpy, tweepy, textblob, preprocessor and string.

Figure 5 details creating the Python class and authentication keys to connect to twitter. Figure 6 shows how the tweepy. Cursor object is used to pass in the parameters through the api.search

function. The items function returns a view of object that displays a list of the tweets returned. The datatime.strptime was used to specify the days of the tweets to be collected. The collected tweets was put in list format and returned as a csv document.

```python
class TwitterClient(object):
 def __init__(self):
 #set access token keys
consumer_key                        =
"bf9W6KVDcDiq8pi1XXXXXXX"
consumer_secret                     =
"29p8EDMVRJL2YrLJGQLK4cNYz4Wy1cPsbGXX
XXXXXXX"
access_token       =       "2492500904-
r1BfiSBMHorQNY6cJPxDXXXXXXXXXXXXXXXX"
access_token_secret                 =
"QTxRBpljyYX9BkyUWIKcByP5Sn5amHXXXXX
XXXXXXXXXXX"
```

Figure 5: Create Python class and set the authentication keys.

```python
for    tweet    in    tweepy.Cursor(api.search,
q='Nigeria',   since='2017-11-13   10:10:50',
until='2017-11-14 14:00:00').items():
 processed_tweet = p.clean(tweet.text)
 d                                  =
datetime.strptime((str(tweet.created_at)),
'%Y-%m-%d %H:%M:%S')
 day = d.strftime('%Y-%m-%d')
 range     =     pd.date_range('2017-11-13',
periods=14, freq='D')
 if day in range:
  t = [processed_tweet, d]
  result.append(t)
```

Figure 6: Tweet Retrieval

4. RESULTS

Round 1 With Dataset from November 2017

Since the number of observations in the groups are not uniform, we have to use a robust test, namely the Brown-Forsythe test which is a one-way ANOVA on the deviations from the mean. The null hypothesis states there is no difference between a normal distribution and the distribution of our data. The alternative hypothesis states there is a difference; therefore, rejecting the null indicates the data are statistically different from a normal distribution. Results from our testing indicate a significant difference between our data and a normal distribution (Table 1 in Appendix); therefore, we are required to circumvent a standard ANOVA for a more robust non-parametric alternative. Table 2 illustrates the

results of the Brown-Forsythe test for equality of means.

	N	Mean	Std. Deviation	Minimum	Maximum
Trump	38	.301732	.5310215	-.9328	.8777
Public	653	-.051336	.4154201	-.9552	.8720
FOX	84	.035154	.4056491	-.8020	.8074
CNN	108	-.017331	.3363191	-.6486	.7579
Total	883				

Table 1: Descriptive Statistics for Twitter Dataset November 2017

	Statistic	df1	df2	Sig.
Brown-Forsythe	8.559	3	136.003	.000

Table 2: Robust Tests of Equality of Means

Our selection for the non-parametric alternative to a one-way ANOVA is the Kruskal-Wallis H Test which will inform us if there is a significant difference in our groups. The Kruskal-Wallis (Figure 7 and Table 3) does not reveal the difference in the groups; therefore, we are required to use post-hoc analysis. Normally, Tukey is the preferred method of post-hoc analysis; however, Tukey requires the data do not violate the assumption of normality. The non-parametric alternatives to the Tukey test are the Tamhane T2 and Games-Howell test.
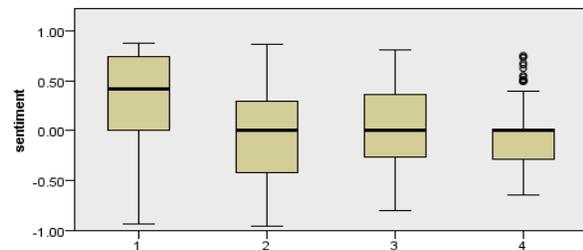
Figure 7: Kruskal-Wallis Test

	sentiment
Chi-Square	23.57
df	3
Asymp. Sig.	0

Table 3: Kruskal-Wallis Test Statistics

Figure 7 illustrates the distribution of sentiment among groups ordered as Trump (1), Public (2), Fox (3), CNN (4). Table 2 shows a significant difference exists between groups; therefore, we

advance to the Tamhane and Games-Howell tests as illustrated in Table 4. As demonstrated by both of the aforementioned tests, a significant difference exists between Trump and all other sources. Moreover, no significant difference exists between the public, Fox New, and CNN's sentiment of Trump. Based on the mean values from each group, we can see that Trump self-sentiment is statistically higher than the sentiment generated by the public, Fox News, or CNN.

			Mean Difference (I-J)	Std. Error	Sig.	Interval Lower Bound	Interval Upper Bound
Tamhane	Trump	Public	.3530671*	.0876636	.001	.110349	.595785
		Fox	.2665780*	.0968482	.047	.002672	.530484
		CNN	.3190621*	.0920214	.007	.066528	.571596
	Public	Trump	-.3530671*	.0876636	.001	-.595785	-.110349
		Fox	-.0864891	.0471510	.350	-.212890	.039911
		CNN	-.0340050	.0362160	.924	-.130440	.062430
	Fox	Trump	-.2665780*	.0968482	.047	-.530484	-.002672
		Public	.0864891	.0471510	.350	-.039911	.212890
		CNN	.0524841	.0548294	.917	-.093581	.198549
	CNN	Trump	-.3190621*	.0920214	.007	-.571596	-.066528
		Public	.0340050	.0362160	.924	-.062430	.130440
		Fox	-.0524841	.0548294	.917	-.198549	.093581
Games-Howell	Trump	Public	.3530671*	.0876636	.001	.118011	.588124
		Fox	.2665780*	.0968482	.038	.010316	.522840
		CNN	.3190621*	.0920214	.006	.074132	.563993
	Public	Trump	-.3530671*	.0876636	.001	-.588124	-.118011
		Fox	-.0864891	.0471510	.263	-.209553	.036575
		CNN	-.0340050	.0362160	.784	-.127994	.059984
	Fox	Trump	-.2665780*	.0968482	.038	-.522840	-.010316
		Public	.0864891	.0471510	.263	-.036575	.209553
		CNN	.0524841	.0548294	.774	-.089866	.194834
	CNN	Trump	-.3190621*	.0920214	.006	-.563993	-.074132
		Public	.0340050	.0362160	.784	-.059984	.127994
		Fox	-.0524841	.0548294	.774	-.194834	.089866

Multiple Comparisons. Dependent Variable:
*. The mean difference is significant at the 0.05 level.

Table 4: Results of Tamhane T2 and Games-Howell Tests – Round 1 November 2017

Round 2 with Dataset from April 2018
The second round of data revealed the following descriptive statistics in which a similar difference in means before additional testing. Based on results, Trump's self-sentiment in November 2017 was significantly different from the public, Fox, and CNN's sentiments. In April of 2018, we note that the same statistical tests needed to be employed due to the violation of the assumption of homogeneity of variance as shown in Table 6. Proceeding to the Kruskal-Wallis test we see that a difference exists in our data and therefore, we advance to the post-hoc analysis of Tamhane and Games-Howell (Table 8).

	Statistic[a]	df1	df2	Sig.
Brown-Forsythe	3.383	3	105.236	.021

Table 6: Robust Test of Equality of Means

	score
Chi-Square	9.373
df	3
Asymp. Sig.	.025

Table 7: Kruskal-Wallis Test Statistics

.

Multiple Comparisons

Dependent Variable:

(I) whonum			Mean Difference (I-J)	Std. Error	Sig.	Interval Lower Bound	Upper Bound
Tamhane	Trump	Public	.20135	.08776	.159	-.0445	.4472
		Fox	.11377	.10946	.885	-.1843	.4119
		Cnn	.26047	.09563	**.054**	-.0028	.5237
	Public	Trump	-.20135	.08776	.159	-.4472	.0445
		Fox	-.08758	.06981	.770	-.2815	.1063
		CNN	.05913	.04515	.725	-.0623	.1805
	Fox	Trump	-.11377	.10946	.885	-.4119	.1843
		Public	.08758	.06981	.770	-.1063	.2815
		CNN	.14671	.07948	.353	-.0697	.3631
	CNN	Trump	-.26047	.09563	**.054**	-.5237	.0028
		Public	-.05913	.04515	.725	-.1805	.0623
		Fox	-.14671	.07948	.353	-.3631	.0697
Games-Howell	Trump	Public	.20135	.08776	.120	-.0363	.4389
		Fox	.11377	.10946	.727	-.1757	.4032
		Cnn	.26047*	.09563	**.044**	.0053	.5157
	Public	Trump	-.20135	.08776	.120	-.4389	.0363
		Fox	-.08758	.06981	.597	-.2752	.1001
		CNN	.05913	.04515	.559	-.0590	.1773
	Fox	Trump	-.11377	.10946	.727	-.4032	.1757
		Public	.08758	.06981	.597	-.1001	.2752
		CNN	.14671	.07948	.263	-.0635	.3569
	CNN	Trump	-.26047*	.09563	**.044**	-.5157	-.0053
		Public	-.05913	.04515	.559	-.1773	.0590
		Fox	-.14671	.07948	.263	-.3569	.0635

Table 8. Results of Tamhane T2 and Games-Howell Tests – Round 2 April 2018

In the new post-hoc analysis, we see that Trump's self-sentiment did not significantly differ from the public or Fox; however, there was a significant difference between Trump's self-sentiment and CNN's sentiment of Trump illustrated in Table 8. One can only speculate on whether Trump's self-sentiment changed or that of the public and Fox. Trump has battled with many news sources, one of which is CNN, on his fake news campaign.

5. CONCLUSION

Social media is becoming increasingly pervasive in society. As social media has become more prevalent, it is only natural for it to extend into politics. Investigating the changing sentiments in social media provides a mechanism to understand how society, politicians, and news outlets are reacting to the political climate. Politicians employ social media as a way to communicate directly with the constituents. While President Obama used social media, for the first time in history President Trump, the President of the United States, is an avid user of social media, namely Twitter. As more politicians follow the direction and influence of the POTUS, Twitter will become more common among politicians. The method presented in this paper details how sentiment analysis can be combined with statistical analysis to compare the sentiment of politicians with news sources, the public, and even each other. To illustrate our method we compared the sentiment of the POTUS with common news sources and the public. The first round of statistical analysis shows a significant difference between the self-sentiment of the POTUS and common news outlets combined with the public. A second round of analysis illustrates changing sentiment between the POTUS and news outlets with the self-sentiment of the POTUS, the public, and news sources aligning with one exception. Our results indicate a change in political climate since the election of the POTUS.

6. REFERENCES

Abel, F., Hauff, C., Houben, G.-J., Stronkman, R., & Tao, K. (2012). *Twitcident: fighting fire with information from social web streams.* Paper presented at the Proceedings of the 21st International Conference on World Wide Web.

Ahmed, W. (2017). Using Twitter as a data source: an overview of social media research tools (updated for 2017). *Impact of Social Sciences Blog*.

Bae, Y., & Lee, H. (2012). Sentiment analysis of twitter audiences: Measuring the positive or negative influence of popular twitterers. *Journal of the American Society for Information Science and Technology, 63*(12), 2521-2535.

Batrinca, B., & Treleaven, P. C. (2015). Social media analytics: a survey of techniques, tools and platforms. *Ai & Society, 30*(1), 89-116.

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science, 2*(1), 1-8.

Cioffi-Revilla, C. (2010). A methodology for complex social simulations.

Clark, E. M., James, T., Jones, C. A., Alapati, A., Ukandu, P., Danforth, C. M., & Dodds, P. S. (2018). A Sentiment Analysis of Breast Cancer Treatment Experiences and Healthcare Perceptions Across Twitter. *arXiv preprint arXiv:1805.09959*.

Conover, M., Ratkiewicz, J., Francisco, M. R., Gonçalves, B., Menczer, F., & Flammini, A. (2011). Political polarization on twitter. *ICWSM, 133*, 89-96.

dictionary.com. (2018). Dictionary.com. Retrieved from http://www.dictionary.com

Internet Live Statistics. (2018). Internet Live Statistics. Retrieved from http://www.internetlivestats.com/

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons, 53*(1), 59-68.

Kouloumpis, E., Wilson, T., & Moore, J. D. (2011). Twitter sentiment analysis: The good the bad and the omg! *ICWSM, 11*, 538-541.

Lazer, D., Pentland, A. S., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., . . . Gutmann, M. (2009). Life in the network: the coming age of computational social science. *Science (New York, NY), 323*(5915), 721.

Monaco, N. (2017). How Obama is Owning Trump on Twitter. Retrieved from http://fortune.com/2017/08/17/barack-obama-donald-trump-twitter/

NTLK. (2018). NLTK 3.2.5 Documentation. Retrieved from http://www.nltk.org

Oscar, N., Fox, P. A., Croucher, R., Wernick, R., Keune, J., & Hooker, K. (2017). Machine learning, sentiment analysis, and tweets: an examination of Alzheimer's disease stigma on Twitter. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 72*(5), 742-751.

Parra, D., Trattner, C., Gómez, D., Hurtado, M., Wen, X., & Lin, Y.-R. (2016). Twitter in academic events: a study of temporal usage, communication, sentimental and topical patterns in 16 computer science conferences. *Computer Communications, 73*, 301-314.

Quigley, A. (2017). THE FAVORITE PRESIDENT? OBAMA BEATS TRUMP WITH TWITTER FOLLOWERS. Retrieved from http://www.newsweek.com/obama-more-twitter-followers-trump-638890

Statista. (2018a). Most popular social networks worldwide as of July 2018, ranked by number of active users (in millions). Retrieved from https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/

Statista. (2018b). Number of monthly active Twitter users worldwide from 1st quarter 2010 to 2nd quarter 2018 (in millions). Retrieved from https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/

Wolfram, S. (2010). *Modelling the stock market using Twitter.* (Masters of Science), University of Edinberg.

Zainuddin, N., Selamat, A., & Ibrahim, R. (2018). Hybrid sentiment classification on twitter aspect-based sentiment analysis. *Applied Intelligence*, 1-15.
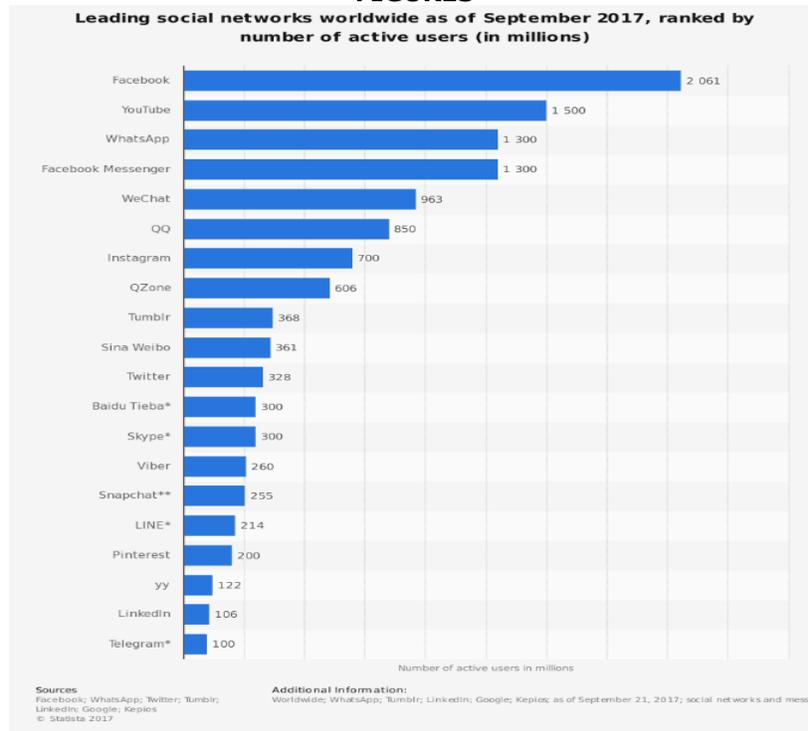
7. APPENDICES AND ANNEXURES

FIGURES

Figure 4. Social networks ranked by popularity (Statista, 2018a)

Figure 5. Tools for Twitter Retrieval (Different Twitter Data Retrieval/Analysis Tools, (Ahmed, 2017)

TABLES

	N	Mean	Std. Deviation	Minimum	Maximum
Trump	38	.301732	.5310215	-.9328	.8777
Public	653	-.051336	.4154201	-.9552	.8720
FOX	84	.035154	.4056491	-.8020	.8074
CNN	108	-.017331	.3363191	-.6486	.7579
Total	883				

Table 1: Descriptive Statistics for Twitter Dataset November 2017

Multiple Comparisons

Dependent Variable:

(I) whonum			Mean Difference (I-J)	Std. Error	Sig.	Interval Lower Bound	Interval Upper Bound
Tamhane	Trump	Public	.3530671*	.0876636	.001	.110349	.595785
		Fox	.2665780*	.0968482	.047	.002672	.530484
		CNN	.3190621*	.0920214	.007	.066528	.571596
	Public	Trump	-.3530671*	.0876636	.001	-.595785	-.110349
		Fox	-.0864891	.0471510	.350	-.212890	.039911
		CNN	-.0340050	.0362160	.924	-.130440	.062430
	Fox	Trump	-.2665780*	.0968482	.047	-.530484	-.002672
		Public	.0864891	.0471510	.350	-.039911	.212890
		CNN	.0524841	.0548294	.917	-.093581	.198549
	CNN	Trump	-.3190621*	.0920214	.007	-.571596	-.066528
		Public	.0340050	.0362160	.924	-.062430	.130440
		Fox	-.0524841	.0548294	.917	-.198549	.093581
Games-Howell	Trump	Public	.3530671*	.0876636	.001	.118011	.588124
		Fox	.2665780*	.0968482	.038	.010316	.522840
		CNN	.3190621*	.0920214	.006	.074132	.563993
	Public	Trump	-.3530671*	.0876636	.001	-.588124	-.118011
		Fox	-.0864891	.0471510	.263	-.209553	.036575
		CNN	-.0340050	.0362160	.784	-.127994	.059984
	Fox	Trump	-.2665780*	.0968482	.038	-.522840	-.010316
		Public	.0864891	.0471510	.263	-.036575	.209553
		CNN	.0524841	.0548294	.774	-.089866	.194834
	CNN	Trump	-.3190621*	.0920214	.006	-.563993	-.074132
		Public	.0340050	.0362160	.784	-.059984	.127994
		Fox	-.0524841	.0548294	.774	-.194834	.089866

*. The mean difference is significant at the 0.05 level.

Table 4: Results of Tamhane T2 and Games-Howell Tests – Round 1 November 2017

	N	Mean	Std. Deviation	Minimum	Maximum
Trump	31	.2143	.47911	-.77	.93
Public	599	.0130	.42201	-.95	.96
Fox	34	.1006	.39443	-.81	.78
CNN	68	-.0461	.34409	-.85	.73
Total	732	.0201	.41873	-.95	.96

Table 5: Descriptive Statistics for Twitter Dataset from April 2018

Multiple Comparisons

Dependent Variable:

(I) whonum			Mean Difference (I-J)	Std. Error	Sig.	Interval Lower Bound	Upper Bound
Tamhane	Trump	Public	.20135	.08776	.159	-.0445	.4472
		Fox	.11377	.10946	.885	-.1843	.4119
		Cnn	.26047	.09563	**.054**	-.0028	.5237
	Public	Trump	-.20135	.08776	.159	-.4472	.0445
		Fox	-.08758	.06981	.770	-.2815	.1063
		CNN	.05913	.04515	.725	-.0623	.1805
	Fox	Trump	-.11377	.10946	.885	-.4119	.1843
		Public	.08758	.06981	.770	-.1063	.2815
		CNN	.14671	.07948	.353	-.0697	.3631
	CNN	Trump	-.26047	.09563	**.054**	-.5237	.0028
		Public	-.05913	.04515	.725	-.1805	.0623
		Fox	-.14671	.07948	.353	-.3631	.0697
Games-Howell	Trump	Public	.20135	.08776	.120	-.0363	.4389
		Fox	.11377	.10946	.727	-.1757	.4032
		Cnn	.26047*	.09563	**.044**	.0053	.5157
	Public	Trump	-.20135	.08776	.120	-.4389	.0363
		Fox	-.08758	.06981	.597	-.2752	.1001
		CNN	.05913	.04515	.559	-.0590	.1773
	Fox	Trump	-.11377	.10946	.727	-.4032	.1757
		Public	.08758	.06981	.597	-.1001	.2752
		CNN	.14671	.07948	.263	-.0635	.3569
	CNN	Trump	-.26047*	.09563	**.044**	-.5157	-.0053
		Public	-.05913	.04515	.559	-.1773	.0590
		Fox	-.14671	.07948	.263	-.3569	.0635

Table 8. Results of Tamhane T2 and Games-Howell Tests – Round 2 April 2018