Trump Tweets Scientific Report

Maika Francis (UID: 605973753), Chenhao Fan (UID: 006043201),
Ivan Guan (UID: 505512696), Ruyi Lu (UID: 405592759), Jesse Rocha (UID:405590190)

I. Abstract/Context

Donald John Trump served as the 45th President of the United States, leaving a long-lasting mark on American politics. Notably, his presidency was characterized by a significant and impactful social media presence. Trump, known for his background in acting, brought a showman's flair to his communication style, and this was vividly expressed through his active engagement on various platforms.

Remarkably, Trump's social media activities were not confined to his time in the White House; he had been an active participant in political discourse on Twitter long before his presidential campaign. Following his term as president, Trump faced a ban from Twitter, a platform where he had garnered immense attention and influence. Despite this, Trump's legacy as a public figure remains significant, with his tweets continuing to impact the political landscape and resonate with those closely following political affairs.

Our interest lies in investigating the influence of Trump's tweeting habits on user engagement. By analyzing factors such as the timing of his tweets, the correlation between device usage and tweet content, and the levels of engagement in response to positive and negative tweets, we aim to reveal insights into digital behaviors and trends. Through this analysis, we seek to understand the dynamics of social media interactions in the context of political figures, shedding light on the lasting impact of Trump's digital communication strategies.

II. Data Introduction

A. Original Data

The dataset encompasses the period from 2009 to 2021, providing a comprehensive snapshot of Trump’s Twitter activity before, during, and after his presidency. It contains various features

such as tweets' content, timestamps, the number of retweets, favorites (likes), devices used to tweet, etc. With over 56,000 tweets, this dataset offers a valuable resource for us to analyze Trump's tweeting habits, sentiment, and engagement levels of one of the most prominent political figures in recent history.

B. Sentiment analysis added

To further analyze the textual data, we introduce new columns in our dataset using three sentiment analysis packages: NLTK's VADER, NRClex, and TextBlob. Each package offers a unique approach to understanding the underlying sentiments in tweet data.

- **VADER Sentiment Analysis (from NLTK)**
    - This package focuses more on the general attitude of the text, generating the negativity, neutrality, and positivity scores as well as an overall compounded sentiment score of the text analyzed.
    - Negativity, Neutrality, Positivity Scores: We added a column with vectors of 3 entries that encapsulated how negative, neutral, and positive the texts were. Each of these scores ranges between 0 and 1, providing a nuanced view of the sentiment landscape within each text entry.
    - Compound Sentiment Score: Another column generated with the VADER package contains the compound sentiment score, ranging from -1 to 1. This score integrates the three sentiment dimensions into a single metric, offering a concise sentiment summary. To further understand the compounded score, a score above 0.05 indicates relatively positive sentiment, below -0.05 indicates relatively negative sentiment, and a score within [-0.05, 0.05] suggests neutrality.

- **NRClex Sentiment Analysis**
    - This package measures the unique sentiments in the text.
    - Eight-Dimensional Sentiment Analysis: This method adds an array with eight elements to each text entry, representing different emotional dimensions: fear, anger, anticipation, trust, surprise, sadness, disgust, and joy. Each element has a discrete value ranging from 0 to 8. The higher the value is, the stronger the corresponding emotion is for the tweet. This approach provides a comprehensive emotional profile of the text.

- **TextBlob Sentiment Analysis**
    - This package measures the polarity and subjectivity of the text, providing an overview of the compounded sentiment as well as how subjective the text could be.
    - Polarity and Subjectivity Scores: We include a new column using TextBlob analysis. The first measure in the column shows the polarity of the text, which ranges from -1 (negative sentiment) to 1 (positive sentiment) and indicates the overall sentiment tone. The second entry measures subjectivity, ranging from 0 (objective or factual) to 1 (subjective or personal), enabling us to differentiate between factual and opinion-based content.

III. Main Findings

Our main purpose in this project is to examine whether the sentiments expressed in Donald Trump's tweets have an impact on user engagement. We aim to determine whether the level of engagement, measured by retweets and likes, varies based on the overall tone of the tweet – whether positive or negative.

Despite Trump's reputation for delivering unfiltered content in his tweets and frequently touching upon controversial topics, our analysis using the NLTK Sentiment Analysis Analyzer has revealed a noteworthy trend. Surprisingly, Trump's tweets tend to lean towards the positive side, adding an interesting layer to our investigation.

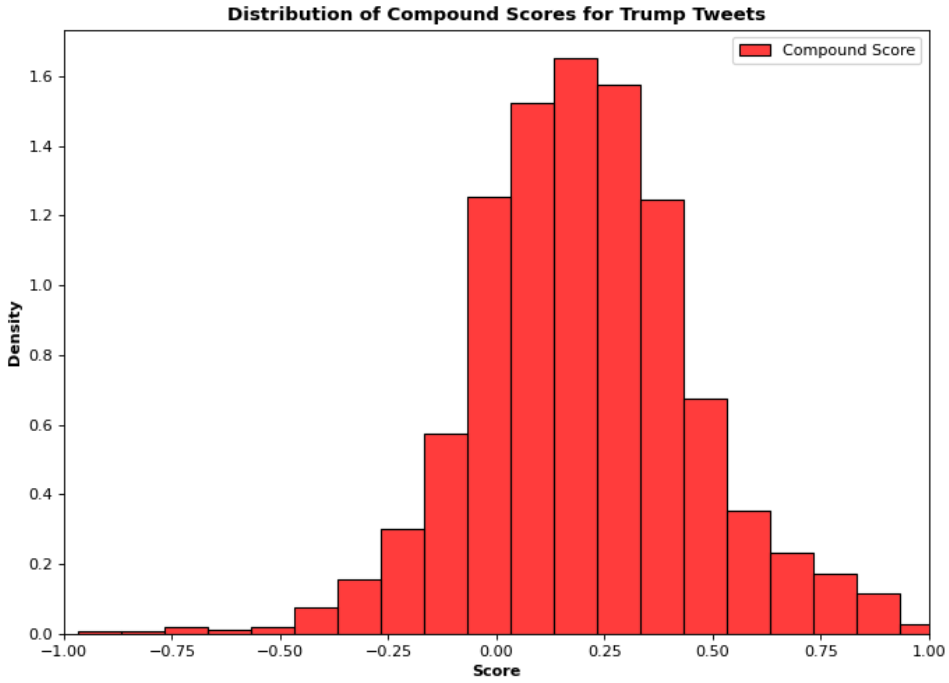

Figure 1: Distribution of Compound Scores for Trump's Tweets

Analyzing the histogram depicted in this graph, it is evident that the distribution is slightly left-skewed. The median, positioned around 0.2, indicates a tendency for Trump's tweets to lean slightly toward the positive side.

Additionally, leveraging the NLTK Sentiment Analysis Analyzer, we categorized Trump's tweets into distinct sentiment labels – Positive and Negative – based on the compound score. A positive sentiment label is assigned when the compound score exceeds 0, while a negative label is applied for compound scores less than 0. If the compound score is precisely 0, the sentiment is labeled as neutral.

Subsequently, we visualized the median number of retweets and likes per tweet, revealing patterns based on the assigned sentiment labels. The outcomes of this analysis are detailed below:

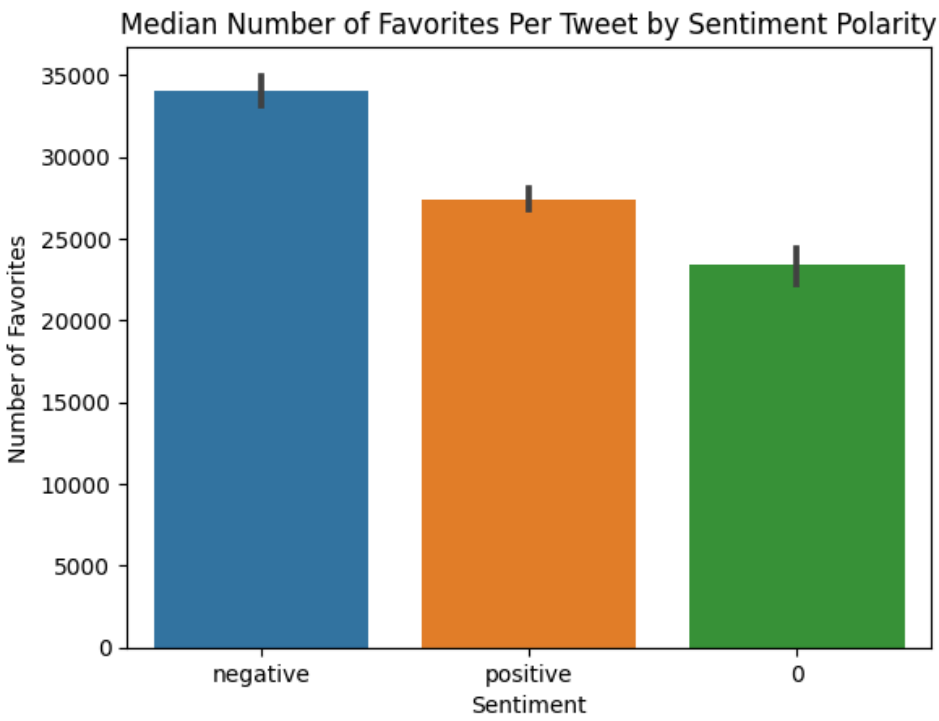

Figure 2: Median Number of Likes per Tweet by Sentiment Polarity

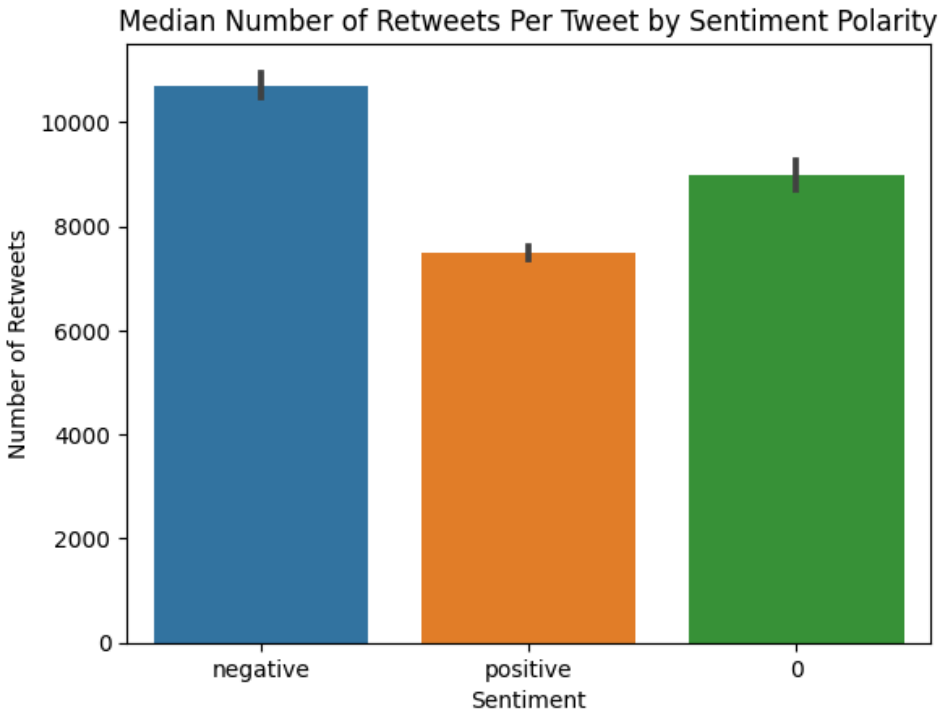

Figure 3: Median Number of Retweets per Tweet by Sentiment Polarity

From these two graphs, we concluded that Trump's negative tweets tend to receive slightly higher amounts of retweets and favorites (as known as, likes), even though the difference is not by a lot.

This prompted us to explore the contents of Trump's negative and positive tweets to identify common words, sorted from the most to the least likes and retweets. The objective was to understand whether certain terms, especially in negative tweets, contribute to higher engagement levels from Twitter users. In this case, we used WordCloud to help us with this process.

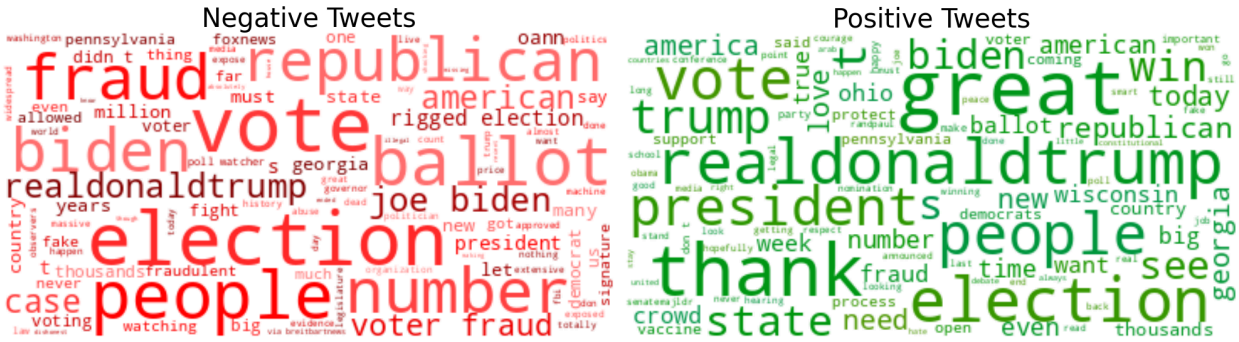

Figure 4: Common Words in Trump's Negative and Positive Sentiment Tweets

The analysis reveals that prevalent terms in Trump's negative tweets include "fraud", "rigged election", "Biden", and "vote", whereas common words in his positive tweets are "realdonaltrump", "election", "thank", and "great".

In conclusion, it appears that Trump's tweets attract more likes and retweets when he discusses controversial political topics that capture people's interest, irrespective of the sentiment being positive or negative. Notably, there is a slightly higher engagement level for Trump's tweets on contentious issues compared to his positive ones.

IV. Other Findings

Besides our main discovery, we were curious about how Trump used Twitter over time and if there were any noticeable trends. We started by checking the total number of tweets he sent from 2009 to 2021. Interestingly, it looks like he wasn't very active on Twitter before becoming

president, especially in 2017 when he got elected. But something changed in 2019, and he started tweeting a lot more. His peak tweeting time seems to be in 2020.

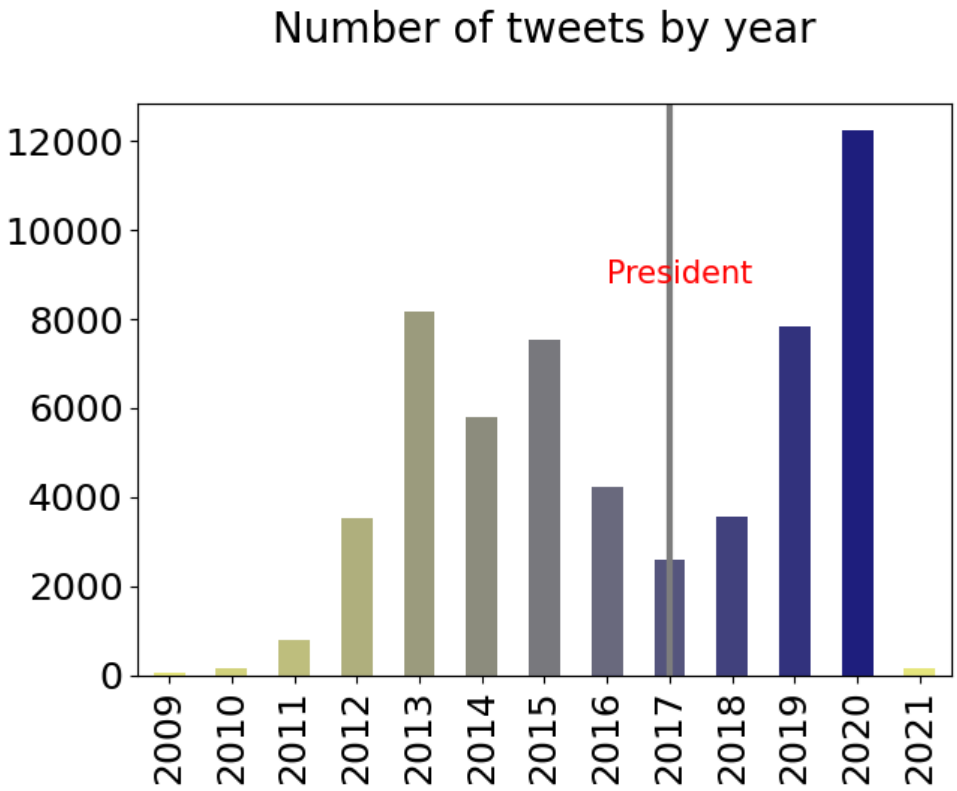

Figure 5: Number of Tweets by Year

Moreover, we discovered that Trump remained highly active on Twitter throughout his presidency. As depicted in the graph below, Trump tended to do most of his tweeting around 12:00 PM, 7:00 PM, and 8:00 PM. Impressively, he maintained a consistent tweeting frequency every hour, showcasing his dedication to engaging on the platform, even with the demands of his busy schedule as a public political figure.

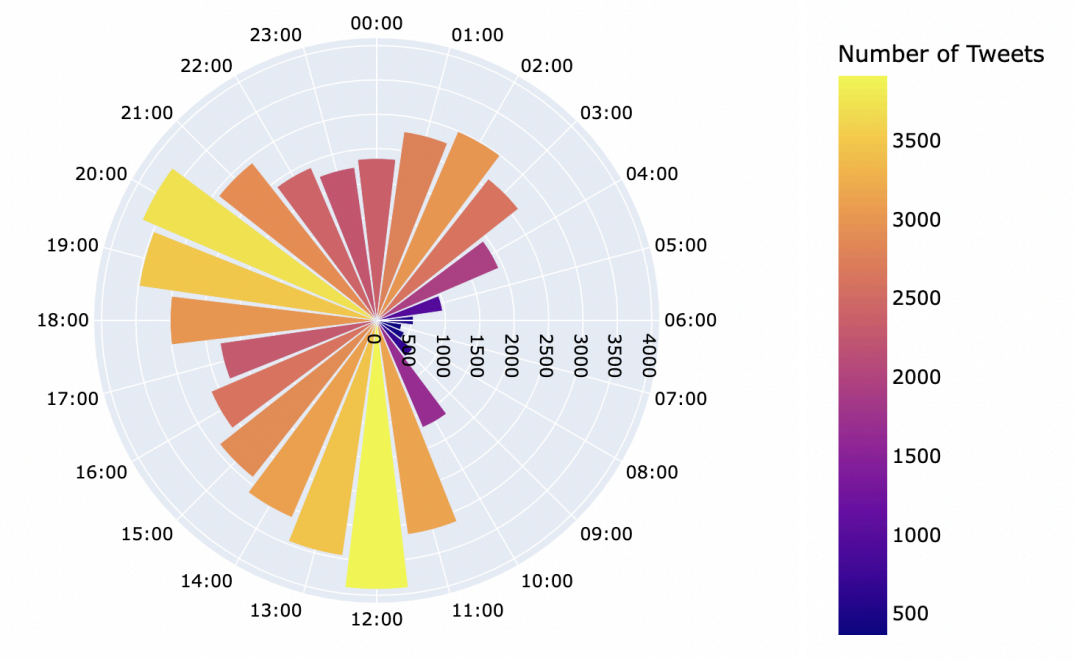

Figure 6: Total Number of Trump Tweets by Hour

Additionally, based on our previous awareness from [an article from The Washington Post](), which suggested that there are two individuals responsible for crafting Trump's tweets – with Trump himself “writes the angrier ones” primarily from his Android phone, we set out to investigate the accuracy of this claim.

Initially, we examined the frequency of tweets originating from various devices on Trump's account. As anticipated, the majority of tweets were from both iPhone and Android devices.

Twitter for iPhone	27967
Twitter for Android	14545
Twitter Web Client	12182
TweetDeck	482
TwitLonger Beta	405
Twitter Media Studio	375
Instagram	133
Facebook	105
Twitter for BlackBerry	97
Twitter Ads	97
Twitter Web App	64
Twitter for iPad	60
Twitlonger	23
Twitter QandA	10
Vine - Make a Scene	10
Periscope	7
Neatly For BlackBerry 10	5
Media Studio	2
Twitter for Websites	1
Twitter Mirror for iPad	1

Table 1: Number of Tweets by Device

The claim was that Trump had assistance in posting positive tweets from someone else, particularly tweeted from an iPhone. Conversely, tweets posted directly by Trump himself, primarily from his Android phone, tended to lean towards the negative side. While we can't definitively prove the involvement of another person in crafting positive tweets, we can leverage our technical skills to identify common words used in tweets from both devices – Android and iPhone. Subsequently, we can classify these common words to determine whether they generally convey a positive or negative sentiment.

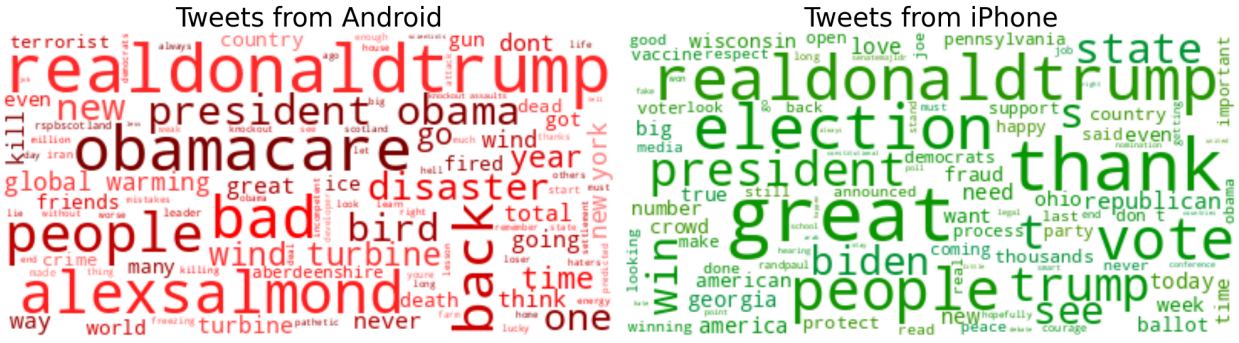

Figure 5: Common Words from Tweets Posted by Android and iPhone

Fascinatingly, our hypothetical question resonates with The Washington Post's assertion that "Trump writes the angrier ones." Upon examination, common words in tweets from Android include terms like "bad," "disaster," "fired," and "gun." In contrast, tweets from iPhone commonly feature words such as "thank," "great," "president," and "win." Consequently, our observation leads us to the conclusion that the common words in tweets from Android tend to convey a negative sentiment, while those from iPhone predominantly express a positive sentiment.