The Emotions of 280 Characters

A quantitative sentiment analysis of how politicians use emotions on Twitter

David Wolfe Bender*

Indiana University

November 13, 2023

Abstract

Social media is now ubiquitous in our national landscape, and it's surely penetrated our politics. Most research on social media focuses on its effect on democracy or political polarization; this study focuses on how politicians use emotion on social media. Separating accounts into Democratic entities and Republican entities, this analysis reviews five million tweets and uses two sentiment analysis models to determine whether or not politicians change their sentiment on Twitter (X) when the White House changes parties. Model 1 focuses on a positive-negative AFINN analysis, while Model II uses NRC sentiments to examine just the emotion of fear. The difference in sentiment between the two parties from 2018 to 2022 is statistically significant in both Model 1 and Model 2.

Introduction

Moore's Law suggests the number of transistors on processing chips will double every two years. It is the well-known theory developed by Intel co-founder Gordon Moore arguing that computational progression will become faster, smaller, and more efficient over time.

The academy offers far less commentary on the lesser-known Second Law named for Moore (or Intel investor Athur Rock, depending on who you ask). That law suggests the cost of a semiconductor plant *doubles* every four years. At its beginnings, the suggestion was that someday these two laws would collide.

This report has little — if anything — to do with semiconductors. Still, those two laws merit discussion because the same two laws could be said for social media in politics. In the ten years between 2006 and 2016, more and more people joined the various social media platforms. Some platforms rose while others fell, but user numbers rose drastically over that time frame.

*Class of 2024, Indiana University (Bloomington, Indiana)

At the same time – similar to the experience of the semiconductor plants – costs to society quickly rose in tandem. In the United States, social media networks quickly went from *almost* irrelevant to political campaigns in 2006 to potentially flipping its result in 2016.

After 2016, journalists and scholars alike spent much effort to discuss how social media may or may not have shifted votes in the 2016 presidential election. While it's certainly an interesting and important discussion, there is far less produced work on how politicians leverage social media.

It should come as no surprise to readers that then-candidate Donald J. Trump drastically changed the landscape of social media in politics. From his campaign announcement in 2015 to his Twitter account's "permanent" suspension in 2021, he used the platform to spread lies, espouse conspiracies, and even convey news about the next cabinet member he would fire.

During the course of his administration, more Democrats started leveraging Twitter as well. Anecdotally, it felt like the Democrats on Twitter responded to President Trump's use of Twitter with anger or negativity. This leads to the question this paper aims to answer: **do political parties change their tone — or sentiment — significantly on social media based on whether or not their party occupies the White House?**

Data Universe and General Assumptions

I turn to the most politically active social media network for answers to this question: Twitter.[1]

This paper employs two separate models to answer that core question. There are general assumptions in place for both models. In this section, I will explain from where I get the necessary data and also the general assumptions in place for the study.

The data for this study is attributed to Alex Litel's Tweets of Congress database.

All of the data from this project is sourced from Mr. Litel's publicly available database. It includes the text and metadata for each tweet, as well as the user data used to examine the necessary data metrics in this report.

I would like to offer my deepest appreciation to Mr. Litel for his work in creating this database and making it publicly available.

[1]It would be more accurate to refer to this network as the platform **formerly** known as Twitter. After Elon Musk's 2023 purchase of the company, he renamed the platform to X. This paper will use the term "Twitter" because it reviews data from before it was renamed. Also, for what it's worth, Twitter was a better name.

This paper examines 824 different entities.

For the purposes of this study, we differentiate between "entities" and users. An entity is an overarching term that refers to different groups or people. Some entities include multiple users, while others could include just a single user. Here are the four types of entities included in this report:

- **Politician**: accounts belonging to an elected representative's office and accounts belonging to political campaigns
- **Caucus**: accounts belonging to a campaign arm of a caucus and accounts belonging to a caucus itself
- **Committee**: accounts belonging to an official committee (controlled by the majority party) and accounts belonging to a party's caucus within a committee
- **Party**: accounts belonging to national committees and accounts belonging to their campaign arms

You notice from this list that a single entity can have more than one user associated with it. For example, take Congresswoman Jennifer Wexton, a Democrat who represents the Virginia 10th Congressional District. The entity Jennifer Wexton (the Politician) contains two users: her campaign account and her office account.

Take the Congressional Black Caucus as another example. That entity contains two users: the account representing their caucus and the account representing its political action committee (PAC).

In total, there are 1,612 total Twitter accounts reviewed across the 824 different entities.

Five data points are attributed to each entity.

For each entity, we track five data points.

- **Name**: the name of the entity[2]
- **Chamber**: whether the entity is based in the House, Senate, or Joint[3]
- **Type**: the type of entity, based on the aforementioned list of entity types
- **Party**: the political party to which the entity's owner belongs[4]
- **User(s)**: the user or list of usernames associated with the entity.

[2] Example: Jennifer Wexton

[3] Joint committees are made up of members from both chambers of the legislature. An example is the Joint Economic Committee

[4] In most cases, this is either the Democratic or Republican parties. However, there are instances in which an entity is listed as "N/A", as ownership might change as the majority party changes in the chamber.

This study tracked more than five million tweets, but "retweets" are removed.

The original batch of five million tweets is cut down to just over 3.5 million after retweets are removed. After a cursory examination of the tweets, it became clear that retweets may not be a good representation of the entity's own sentiment.[5] Additionally, as political Twitter accounts frequently retweet other politicians or their parties, permitting retweets might lead to "double counting" of specific tweets, skewing our data one way or another.

"Quote tweets," however, are counted.[6] For clarity, only the new content – the portion of the quote tweet that was "added" by the quote-tweeting account – is analyzed. This is, once again, done to avoid double counting. See Figure 1.

Figure 1: Quote Tweet Example

In Figure 1, the part of the tweet written by Congressman Gerry Connolly would be analyzed and tagged to the Gerry Connolly entity; the section written by Congresswoman Jennifer Wexton's account would not be attributed to Connolly's entity, though it would be for Wexton's on its own.

[5] As a Twitter professional might say, "Retweets do not necessarily equal endorsements."

[6] Quote tweets allow an account to both retweet and add an additional comment – operating as a new tweet on its own.

Only tweets sent between 2018 and 2022, inclusively, are reviewed.

This is roughly two years worth of tweets from President Trump's term and roughly two years of President Biden's term. The first tweet analyzed was sent January 1, 2018, and the last one analyzed was sent December 31, 2022.

Tweets are divided into two groups: Trump-era tweets and Biden-era tweets

Each tweet is put into one of two buckets based on when the tweet was sent. If it was sent before January 20, 2021, it placed into the Trump-era bucket. If it was sent after January 20, 2021, it is placed in the Biden-era bucket. Tweets sent **on** January 20 were also placed in the Biden-era bucket, **regardless of time – even if they were sent while Biden was not yet President.**

This was due to a limitation in storage space. I made efforts to save as much storage as possible, and while removing time as a variable only saved marginal space, it did speed up runtime a tad. As the tweets sent that morning were a small fraction of the total tweets reviewed, it is unlikely to skew data too much.[7]

Five data points are tracked for each tweet

- **Entity**: the entity from which the tweet was sent
- **Username**: the user/userID from which the tweet was sent
- **Tweet Text**: the text of the tweet
- **Link**: the link of the tweet
- **Date**: the date the tweet was sent

Model 1: Numeric Analysis

Lexicon

Sentiment analyses nearly always employ a lexicon, a predefined list of words all associated with some identifier. This identifier could be an emotion, or it can be numeric. For example, a positive number could be associated with more "positive" emotions (and vice versa for negatives).

There are numerous lexica publicly available for use. Model 1 uses the AFINN Sentiment Lexicon, developed by Finn Årup Nielsen.

[7]This is a situation where, as they say, the marginal benefit of considering time did not exceed the marginal cost of handling the additional storage for each tweet.

The AFINN Lexicon includes a list of nearly 2,500 words and gives each a score between -5 and 5, the latter being the most "positive" sentiment. Words such as "breathtaking" and "thrilled" are given scores of 5. Few words with a score of -5 are devoid of vulgar language.[8]

Those who study sentiment analysis utilize multiple tactics in their study. One method is to use Natural Language Processing (NLP) to derive an overall "score" for a document (in this paper's case, a tweet). Another method is to use the lexicon and complete a word search: find the number of times a word in the lexicon is used and average the scored words to derive an overall sentiment score for the tweet.

Initially, before cleaning the data for this project, I selected the latter method. Upon reflection – and I will expand upon this in this paper's conclusion – the former "NLP path" would've likely been more effective. Still, I used the word search method for this report. I wanted to keep the methodology constant once the data collection began to avoid even the mere appearance that I was changing the methodology to get a different result.

Methodology

A function reviews each tweet individually, matching any words in the tweets to any words listed in the AFINN Lexicon. Any words identified in both the tweet and the lexicon are assigned the appropriate sentiment score. The algorithm calculates a sum from all of the sentiment scores identified in a given tweet.

For example, take this sentence: "The funny friend ran to the store. His mother never approved, telling him to stop." See the example in Table 1.

To achieve the tweet's "sentiment score," we divide the sum by the number of words that matched in the AFINN Lexicon. This is done to ensure that no single tweet is given a higher or lower sentiment score just because it had more words that matched in the lexicon.

Words not matched with the lexicon are considered neutral. Additionally, there are no words in the AFINN lexicon that are given a score of 0.[9]

Entities are given an overall average for sentiment

After all the tweets were examined individually, all of an entity's tweets are used to derive an entity sentiment score. This is done by taking the mean of all that entity's tweet's sentiment scores.

For example, if an entity were to have sent five tweets with the sentiment scores of -3, -1, 2, 4, and 3, the entity's overall sentiment score would be a 1.

[8] In fact, "prick" is the only word with a score of -5 that I was willing to use in this paper as an example. It appears that even in the realm of linguistic scrutiny, some words maintain a certain – let's say – prickly charm.

[9] There is one phrase that is given a score of zero: "some kind". To my knowledge, it is the only phrase in the lexicon, suggesting to me that there was some sort of mistake when I downloaded the lexicon initially. We don't review phrases in this sentiment analysis, only words. This is a limitation I will discuss in this paper's conclusion.

Word	**Score**
The	N/A
funny	4
friend	N/A
ran	N/A
to	N/A
the	N/A
store	N/A
His	N/A
mother	N/A
never	N/A
approved	2
telling	N/A
him	N/A
to	N/A
stop	-1
SUM	**5**
SCORE	**1.66**

Table 1: Word Score Example

An entity's sentiment score is calculated twice, giving us two variables for each entity: one for each Presidential era

Because I compare entities based on when their party occupies the White House, I calculate the sentiment of each entity both during the Trump administration and during the Biden administration.

Entities are grouped together by party

In order to compare whether the political left and right change their sentiment on Twitter, we consider them together as a party. For example, the following would be included in the Democratic Party:

- Democratic members of the House of Representatives
- Democratic members of the United States Senate
- The Democratic National Committee (DNC)
- The Democratic Senatorial Campaign Committee (DSCC)
- The Democratic Congressional Campaign Committee (DCCC)

Of course, there are other accounts included on the Democratic side as well. This same principle works for the GOP, sometimes just replacing a "D" with an "R" (the RNC, NRCC, NRSC).

In the end, we find four values:

- **Republican Entities During the Trump Administration** (RdT)
- **Democratic Entities During the Trump Administration** (DdT)
- **Republican Entities During the Biden Administration** (RdB)
- **Democratic Entities During the Biden Administration** (DdB)

The last step is to run a two-group t-test evaluating our results.

Results

The means for the four groups were calculated in R after the data was cleaned. This is displayed in Table 2.

Group	**President**	**Mean**
Democrats	Biden	0.7720921
Democrats	Trump	0.5158242
Republicans	Biden	0.3822512
Republicans	Trump	0.8874733

Table 2: Party Sentiment Values for Each Presidential Era

Hypothesis Testing

I use two different two-group t-tests to determine whether or not the differences were statistically significant. I assume equal variance between the two groups.

$$H_0 : \overline{democratsBiden} - \overline{democratsTrump} = 0$$
$$H_0 : \overline{republicansBiden} - \overline{republicansTrump} = 0$$

$$H_1 : \overline{democratsBiden} - \overline{democratsTrump} \neq 0$$
$$H_1 : \overline{republicansBiden} - \overline{republicansTrump} \neq 0$$

I evaluate these hypothesis tests at the 95% confidence interval. You can see the results for both hypothesis tests in Table 3 and Table 4, respectively.

Statistic	Value
Data	`democratsBiden` and `democratsTrump`
t-value	9.735
Degrees of Freedom	628
95% Confidence Interval	(0.2046, 0.3080)
Sample Estimates	
Mean of `democratsBiden`	0.7721
Mean of `democratsTrump`	0.5158
p-value	$< 2.2 \times 10^{-16}$

Table 3: Democratic Entities During the Trump and Biden Administrations (Model 1)

Statistic	Value
Data	`republicansBiden` and `republicansTrump`
t-value	-15.681
Degrees of Freedom	662
95% Confidence Interval	(-0.5684836, -0.4419606)
Sample Estimates	
Mean of `republicansBiden`	0.3822512
Mean of `republicansTrump`	0.8874733
p-value	$< 2.2 \times 10^{-16}$

Table 4: Republican Entities During the Trump and Biden Administrations (Model 1)

Model 1 Conclusions

There are a few interesting conclusions to make here. First, **both Democrats and Republicans change their sentiment to a statistically significant degree based on whether President Trump or Biden occupied the White House, using Model 1's methodology.** For both of them, the p-value was not only less than .05 but less than .01.

Interestingly, both parties had a positive mean score, meaning they were — on the whole — more positive than negative.

Additionally, at a raw level, it seems the Republicans are more polar. Their range is higher than the Democrats, meaning they shifted more from administration to administration.

Model 2: Fear Analysis

Lexicon

Model 2 utilizes a different lexicon than Model 1. This time, we use the National Research Council Canada Word-Emotion Association Lexicon. The credit for such lexicon goes to Dr.

Saif M. Mohammad.

Unlike the AFINN Sentiment lexicon — which operates in numbers — the NRC Sentiment lexicon uses emotions. There are seven of them: anger, anticipation, disgust, fear, joy, sadness, surprise, and trust. It also offers two sentiments: positive and negative. As such, this lexicon is often colloquially referred to as the *EmoLex*.

As I researched the NRC lexicon, I became interested in one emotion above the others: fear. We're constantly told that politicians use fear to entice voters toward their side. Does this translate into their Twitter accounts? **Do politicians use more fear when they aren't in power versus when they are?**

Methodology

Model 2 works very similarly in process to Model 1. The general assumptions for the study, of course, maintain for both models. This time, we look to match words in the NRC fear lexicon with words in the tweets. For each word the algorithm finds, it adds one to that tweet's sum.

Then, an algorithm takes the mean score of all an entity's tweets. That number is called a *fear score*.

As was true before, parties are compared to parties. All entities from a single party are considered in the analysis.

Results

The means for the four groups (Republicans and Democrats during both administrations) were again calculated in R after the data was cleaned. That data is displayed in Table 5.

Group	President	Mean
Democrats	Biden	0.7809604
Democrats	Trump	0.7276105
Republicans	Biden	0.7857485
Republicans	Trump	0.5844433

Table 5: Party Fear Score Average for Each Presidency

Hypothesis Testing

Again, I run a two-group t-test, assuming equal variances between the two groups, to determine if the differences are statistically significant. See below for the two hypothesis tests.

$$H_0 : \overline{democratsBiden} - \overline{democratsTrump} = 0$$
$$H_0 : \overline{republicansBiden} - \overline{republicansTrump} = 0$$

$$H_1 : \overline{democratsBiden} - \overline{democratsTrump} \neq 0$$
$$H_1 : \overline{republicansBiden} - \overline{republicansTrump} \neq 0$$

Once again, I evaluate these tests at the 95% confidence interval. The results for Democratic entities and Republican entities are displayed in Table 6 and Table 7, respectively.

Statistic	Value
Data	`democratsBiden2` and `democratsTrump2`
t-value	3.3804
Degrees of Freedom	628
95% Confidence Interval	(0.02235789, 0.08434173)
Sample Estimates	
Mean of `democratsBiden2`	0.7809604
Mean of `democratsTrump2`	0.7276105
p-value	0.0007686

Table 6: Democratic Entities During the Trump and Biden Administrations (Model 2)

Statistic	Value
Data	`republicansBiden2` and `republicansTrump2`
t-value	14.967
Degrees of Freedom	662
95% Confidence Interval	(0.1748961, 0.2277144)
Sample Estimates	
Mean of `republicansBiden2`	0.7857485
Mean of `republicansTrump2`	0.5844433
p-value	$< 2.2 \times 10^{-16}$

Table 7: Republican Entities During the Trump and Biden Administrations (Model 2)

Model 2 Conclusions

In evaluating the emotion, we can conclude that both Republicans and Democrats change their use of fear on social media depending on who occupies the White House to a statistically significant degree. Interestingly, this did not go exactly the way I expected. Democrats, for example, used more fear on average during Biden's term than during Trump's term.

One suggestion for this is that the use of fear is simply becoming more common. Perhaps it doesn't have much to do with who is in power; it's just that, as culture wars become more prevalent, parties are using more fear.

Another suggestion for why the Democratic entities still used a great deal of fear words after Biden took office has to do with the political discussions of the day. If we make an assumption that Democrats were likely to express anger toward President Trump, it's not like Biden's victory took President Trump off the map. Many of the January 6th Commission Hearings also took place during this time, when new evidence was being released about the attack.

One final possibility was the *Dobbs v. Jackson* Supreme Court case in 2022. After the opinion leaked, it quickly became clear that abortion was going to become a notable political issue ahead of the midterm elections that year. Perhaps Democrats chose to utilize fear as a campaign strategy. Anyway, those are mere guesses based on a cursory look at the tweets from the time. Another report could follow with answers to this question.

The Republicans followed the expected path: they used fewer fear words during Trump's term but increased their usage when their party was not in power.

Conclusion

Limitations and Future Ideas

I froze the methodology of this report once the data collection, wrangling, and cleaning began. Still, as I worked on this report, I noticed a number of limitations and ways in which its next iteration could lead to more thoughtful and reasoned conclusions:

- **Abandoning the entity analysis may be preferable.**. I chose this route because I thought it would be important to consider all of a politician's accounts. It didn't end up being worthwhile. A future study may find entities useful, however, if they were to conduct a difference-in-difference analysis. For example, one could look at the members of Congress who were in office before and after the 2020 election and compare their statistics — almost as if the transition of power was a "treatment." Additionally, an entity analysis may not have been helpful because some accounts tweet more than others. If one entity has a single tweet with an overly positive sentiment, it can have a notable effect on the data, while a more-active entity would be under-captured by the data. In my methodology, they would be weighted the same; that may not be reasonable.

- **The word-search method cannot capture the full story.** Were I to do this study again, there is no doubt that I would employ some natural language processing platform to assist in this study. Artificial intelligence would have made this study far more effective. The word-search method, for example, cannot distinguish between the words *delight* and *delighted.* If only one such word is listed in the lexicon, only one is considered in the study, even though they both hold the same sentiment, generally speaking. The word-search method can't consider those differences, but an AI platform certainly could.

- **Data size matters.** I was initially shocked at how long it took to iterate through more than five million tweets. It took an impressive technological combination of a super computer, a remote desktop, and my personal laptop to complete this data analysis over a three-month period. Because there are size limitations, the AI may not be as effective as one might hope. Doing API calls for an AI to analyze each individual tweet could still take a very long time to obtain the data.[10]

The Take-Home Points

If you can only take one thing away from this report, I hope you take this away: **sentiment expressed from political Twitter accounts do change. It shifts. It is not static.** Too frequent do I hear people say that "Congress is always yelling about something." That may be true in the chamber, but sentiment really does change on social media. Congresspeople respond to world events. Fear is not constant.

In addition, this does give some limited insight into the effect of social media algorithms. When people scroll through Twitter, I think they tend to find that the most negative tweets (or the tweets "attacking" someone) are the ones most prevalent on the timeline.[11] Seldom do I see the tweet of a Congressperson celebrating some random piece of good news from their district, but I do think do this with some frequency; we just don't see it. Sentiment is, on the whole, positive.

Thanks and Appreciation

There were numerous people who helped make this report possible. I must thank Professor Christopher DeSante, PhD, of Indiana University, who first introduced me to the idea of sentiment analyses. I never hated mathematics. I also never loved it. All of that changed on day one of Professor DeSante's political mathematics course. He made me fall in love with mathematics and quantitative methods. Thanks for that, Professor DeSante.

Also, I mentioned earlier that I used a supercomputer to complete some of the work on this project. Indiana University's technology team helped me get started on its high-capacity computing systems after I initially made an attempt to cycle through five million tweets on my personal laptop. It turns out that even 32 gigabytes of RAM won't help you get there; it took a computer with 72 general-computing nodes, each of which held 256 GB of RAM.[12] Therefore, I must give attribution to those who made that possible: this research was

[10] Yes, this is a limitation on a limitation. I understand the irony.

[11] This, by the way, is another potential piece of future inquiry: is there a correlation between negative sentiment and impressions on social media? I don't know, but it sure seems like it sometimes.

[12] Of course, I only used a small portion of this available RAM. Nevertheless, I am so appreciative to Indiana University's technology division for helping make this possible.

supported in part by Lilly Endowment, Inc., through its support for the Indiana University Pervasive Technology Institute.