

BIBLIOGRAPHIC INFORMATION SYSTEM

Journal Full Title: [Journal of Biomedical Research & Environmental Sciences](#)

Journal NLM Abbreviation: J Biomed Res Environ Sci

Journal Website Link: <https://www.jelsciences.com>

Journal ISSN: 2766-2276

Category: Multidisciplinary

Subject Areas: [Medicine Group](#), [Biology Group](#), [General](#), [Environmental Sciences](#)

Topics Summation: 133

Issue Regularity: [Monthly](#)

Review Process: [Double Blind](#)

Time to Publication: 21 Days

Indexing catalog: [IndexCopernicus ICV 2022: 88.03](#) | [GoogleScholar](#) | [View more](#)

Publication fee catalog: [Visit here](#)

DOI: 10.37871 ([CrossRef](#))

Plagiarism detection software: [iThenticate](#)

Managing entity: USA

Language: English

Research work collecting capability: Worldwide

Organized by: [SciRes Literature LLC](#)

License: Open Access by Journal of Biomedical Research & Environmental Sciences is licensed under a Creative Commons Attribution 4.0 International License. Based on a work at SciRes Literature LLC.

Manuscript should be submitted in Word Document (.doc or .docx) through

Online Submission

form or can be mailed to support@jelsciences.com

**IndexCopernicus
ICV 2022:
83.03**

 **Vision:** Journal of Biomedical Research & Environmental Sciences main aim is to enhance the importance of science and technology to the scientific community and also to provide an equal opportunity to seek and share ideas to all our researchers and scientists without any barriers to develop their career and helping in their development of discovering the world.

REVIEW ARTICLE

Metadata and Sentiment Data Analytics on Social Media Tweets

Alice S Etim*

Winston-Salem State University, USA

Abstract

The outbreak of COVID-19 pandemic in early 2020 led to decisions by both the federal and state governments in the United States of America (hereafter, governments) to implement policies about staying at home to reduce the spread of the COVID-19 virus. With the closure of many businesses, people worked from home and schools moved face-to-face classes to online or remote learning. People were concerned about the policies and the impact of such policies on their livelihood. Social media websites such as Twitter (X) were used to voice opinions about the challenges posed by the stay-at-home orders. People expressed positive and negative sentiments about the closures and reopening of offices, schools, restaurants, and other public places as well as the impact of the government stay-at-home policies. This article examines both the positive and negative sentiments expressed using the tweets from the Twitter (X) platform. The metadata and sentimental data collected via the social media site, Twitter (X) for three states – North Carolina, Pennsylvania, and California on the pandemic, stay-at-home and reopening policies were analyzed and discussed. The study adds value to existing literature about COVID-19 in understanding people's opinions, better information sharing by governments, scientists and others that influence policy decisions in cases of future public health crisis.

*Corresponding author(s)

Alice S Etim, Winston-Salem State University, USA

Email: etima@wssu.edu

DOI: 10.37871/jbres2143

Submitted: 21 June 2025

Accepted: 15 July 2025

Published: 16 July 2025

Copyright: © 2025 Etim AS, Distributed under Creative Commons CC-BY 4.0 © ⓘ

OPEN ACCESS

Introduction

In early March 2020, the United States of America (USA), like many countries experienced the first set of cases of COVID-19 virus. It spread very rapidly in the USA and many parts of the world that by mid-2020, the virus had become a global pandemic with the scientific name, SARS-CoV-2. There was large resource allocation to clinics, hospitals, and pathology laboratories in the USA and around the world to engage quickly both to understand the disease as well as control it with vaccines [1,2]. Many businesses were shut down; schools were closed and some that had technological capabilities turned into virtual learning.

Recently, researchers have begun to examine large datasets that were created and collected because of the COVID-19 pandemic, mainly to analyze, report findings and discuss relevant lessons for the future. Studies that analyzed people's attitude towards policy formulation at the start and throughout the COVID-19 pandemic of 2020 – 2022 are limited in the literature. Important societal experiences and lessons learned from the COVID-19 pandemic need to come from data. Businesses, governments,

VOLUME: 6 ISSUE: 7 - JULY, 2025



Scan Me

How to cite this article: Etim AS. Metadata and Sentiment Data Analytics on Social Media Tweets. J Biomed Res Environ Sci. 2025 Jul 16; 6(7): 884-890. doi: 10.37871/jbres2143, Article ID: JBRES2143, Available at: <https://www.jelsciences.com/articles/jbres2143.pdf>

and society should learn about COVID-19 impact using the data that were collected when the virus was ravaging through our communities. This paper provides a worthy addition to the literature with the analysis of tweets. The author investigated the COVID-19 reopening movement using social media data (tweets) that were collected during the pandemic in three states – North Carolina, Pennsylvania, and California. The analysis of the large data and the results presented explain people's sentiments on protests due to closures and reopening during the COVID-19 pandemic for the three states. The paper provides knowledge that can guide future policy and decision making in cases of future pandemics and serious public health outbreaks. The following three research questions guided the study:

Research Questions

1. Using key metadata created for the large dataset for the selected opinions and sentiments for the three states, what were the counts for North Carolina, Pennsylvania, and California?
2. How do the three states compare on "Average Followers Count" and "Average Status Count" for Supporting Reopening?
3. How do the three states compare on "Average Followers Count" and "Average Status Count" for Opposing Reopening?

These research questions will be answered in the Data Analysis and Results section, after a brief review of literature.

Literature

The spread of the SARS-CoV-2 (COVID-19, for short) virus was very rapid and by October 2021, more than 5 million deaths were confirmed globally. In the United States of America (USA), there were 500,000 deaths reported during the same period [3]. The COVID-19 global pandemic started spreading in the USA in early 2020 and there were no vaccines or medication at the time to combat the disease [1,2,4]. In early 2021, a few COVID-19 vaccines were made available in the USA by large pharmaceutical companies like Johnson & Johnson, and Pfizer. However, one of the lessons that was learned very quickly from the COVID-19 pandemic was that people resisted taking the COVID-19 vaccines, amplifying the long-standing history of vaccine hesitations in the USA. The resistance to taking COVID-19 vaccines

became more pronounced even with the availability of new COVID-19 vaccines and their boosters in the later part of 2021 [5-8].

The COVID-19 spread, vaccine hesitancy and high death rates led to government's' mandatory stay-at-home policies. However, there were tweets particularly in Tweeter (X) against closure and some against reopening. The common threads in tweets that opposed reopening were mostly about public health and safety concerns, but some had political undertones. Based on selected studies reviewed, the following summary provides the common themes about the concerns, complaints, sentiments expressed mostly as tweets about stay-at-home policies, remote work and reopening [9-13].

- Public health concerns
 - o Concerns about the increase in COVID-19 cases and deaths
 - o Concerns about accuracy & adequacy of contact tracing, testing and quarantine
 - o Limited testing infrastructure particularly in remote and poor regions
 - o Concerns about limited hospital beds, facilities & staff due to many COVID-19 cases
 - o Worries about public spaces like schools being the potential arena for the fast spread of the virus among groups like students and staff and the challenges of social distancing and implementation of safety measures.
- Workplace safety, mobility and public transportation concerns
 - o Workers exposure to COVID-19 and the risk of death at the workplace
 - o Calls for mask mandates, social distancing, and other preventive health measures
 - o Concerns about public transport systems becoming high-risk spaces
 - o Anxiety about returning to offices for fear of getting infected.
- Social and economic concerns
 - o Concerns about job and income losses and rising unemployment rates



- o Impact of the pandemic on emotional and mental health.
- Political undertones
- o Complaints about the stay-at-home mandates as being politically motivated
- o Accusations of government for prioritizing economic concerns over people's health.

Methodology

The tweets on the government's stay-at-home orders were extracted for three states - North Carolina, Pennsylvania and California - using specific hashtags from Twitter (X) Application Programming Interface (API). The collected data had 2,032 tweets for North Carolina, 2,300 tweets for Pennsylvania and 2,360 tweets for California. Using stratified sampling, classification techniques and basic functions such as COUNT and COUNTA, the tweets were further analyzed and grouped based on relevant metadata. The two categories used were positive (in favor of reopening group) and negative (against reopening group). For North Carolina, there were 963 positive tweets and 1,069 negative tweets. Pennsylvania's data included 801 positive tweets and 1,499 negative cases. California had 1,716 positive tweets and 644 negative tweets. Some key taggings were ReOpenNC, ReOpenPA, and ReOpenCA.

The state-by-state comparison was summarized with the following relationship terms or metadata.

- ▶ **Tweets with mentions count:** number of tweets that mention another Twitter platform user.
- ▶ **Tweets from verified profiles:** number of tweets that are produced from accounts with verified "checks". This usually notes that the account has a large following and may be a profile of a celebrity or public figure.
- ▶ **Average tweet length:** average number of characters in each tweet.
- ▶ **Favorite count:** number of times the tweets were liked by Twitter users.
- ▶ **Quoted favorite count:** number of times the tweets were liked by Twitter users who have quoted the original tweets.
- ▶ **Average quoted favorite count:** average number for quoted tweets that were liked.

- ▶ **Retweet count:** number of times that original tweets were retweeted by Twitter users.
- ▶ **Average retweet count:** average number of times that tweets were retweeted by Twitter users.
- ▶ **Quoted retweet count:** number of tweets that were retweeted by Twitter users who have quoted the original tweet.
- ▶ **Average quoted retweet count:** average number of tweets that were retweeted by Twitter users who have quoted the original tweet.
- ▶ **Average followers count:** mean number of followers of the tweets' owners.
- ▶ **Average status count:** mean number of tweets posted by the users.
- ▶ **Average list count:** average number of public lists in which users claim membership.
- ▶ **Average friends count:** average number of friends of the tweet owner.

Data Analysis & Results

Research question #1

Using key metadata created for selected opinions and sentiments, what were the counts for North Carolina, Pennsylvania, and California in the collected dataset?

The sentiment analytics led to the creation of several metadata such as average tweet length, tweets with mentions, tweets from verified profiles, etc. The first research question was answered by using the COUNT function and data grouping techniques to analyze "Favoring Reopening" and "Against Reopening."

Table 1 provides a summary of the results on the protests due to closures and reopening for North Carolina. The tweet count for North Carolina was 963 in favor of reopening and 1069 against reopening the state.

Table 2 provides a summary of the results on the protests due to closures and reopening of Pennsylvania. The tweet count for Pennsylvania was 801 in favor of reopening and 1499 against reopening the state.

Table 3 provides a summary of the results on

**Table1:** Metadata and analytics for North Carolina data.

Metadata	Favoring Reopening (Yes)	Against Reopening (No)
Tweet Count	963	1,069
Percentage	47.39%	52.61%
Tweets with Mentions	463	233
Tweets from Verified Profiles	28	26
Average Tweet Length	147	153
Favorite Count	8.982	29.709
Quoted Favorite Count	282.173	177,861
Average Quoted Favorite Count	2,015.52	1,347.43
Retweet Count	354	330
Average Retweet Count	3.28	5.76
Quoted Retweet Count	106.675	55.103
Average Quoted Retweet Count	761.96	417.45
Average Followers Count	5.994	3.418
Average Status Count	31.783	20,612.41
Average List Count	62.8	47.98
Average Friends Count	3.023	2.237

Table 2: Metadata and analytics for Pennsylvania data.

Metadata	In Favoring of Reopening (Yes)	Against Reopening (No)
Tweet Count	801	1,499
Percentage	34.83%	65.17%
Tweets with Mentions	479	466
Tweets from Verified Profiles	59	90
Average Tweet Length	144	160
Favorite Count	3.685	38.079
Quoted Favorite Count	1,758.772	527.443
Average Quoted Favorite Count	2,198.465	352.098
Retweet Count	1.704	11.963
Average Retweet Count	2.13	7.986
Quoted Retweet Count	417.836	176.060
Average Quoted Retweet Count	522.295	117.53
Average Followers Count	2,917.716	4,615.986
Average Status Count	19,506.5	25,756.097
Average List Count	30.915	63.617
Average Friends Count	2,292.911	2,251.748

the protests due to closures and the reopening of California. The tweet count for California was 1716 in favor of reopening and 644 against reopening the state; it was the least number against reopening of the three states.

Research question #2

How did the three states compare on “Average Followers Count” and “Average Status Count” for Supporting Reopening?

When comparing the number of online interactions about supporting and opposing reopen protests across the three states, there were some similarities

and differences. One of the metadata used in the comparison and reported in table 4 was the Average Followers Count. While North Carolina showed the highest Average Followers Count at 5994, California and Pennsylvania followed closely at 3245 (CA) and 2918 (PA).

The Average Status Count was also collected for the three states. On Twitter (X), Status is a feature that helps users to add context and updates about their tweets to indicate the current activity or mood/

Table 3: Metadata and analytics for California data.

Metadata	In Favoring of Reopening (Yes)	Against Reopening (No)
Tweet Count	1716	644
Percentage	72.71%	27.29%
Tweets with Mentions	1095	172
Tweets from Verified Profiles	16	8
Average Tweet Length	165	165
Favorite Count	39.673	3.338
Quoted Favorite Count	4,009.255	791.366
Average Quoted Favorite Count	2,336.39	1,228.82
Retweet Count	574	153
Average Retweet Count	6	1.14
Quoted Retweet Count	1,153.988	157.296
Average Quoted Retweet Count	672.48	244.24
Average Followers Count	3.245	3.934
Average Status Count	22.474	23.629
Average List Count	43.07	44.29
Average Friends Count	2.596	2.511

Table 4: Comparison of tweets supporting reopen protests for the three states.

Metadata	North Carolina	Pennsylvania	California
Tweet Count	963	801	1716
Percentage	47.39%	34.83%	72.71%
Tweets with Mentions	463	479	1095
Tweets from Verified Profiles	28	59	16
Average Tweet Length	147	144	165
Favorite Count	8.982	3.685	39.673
Quoted Favorite Count	282.173	1,758.772	4,009.255
Average Quoted Favorite Count	2,015.52	2,198.465	2,336.39
Retweet Count	354	1.704	574
Average Retweet Count	3.28	2.13	6
Quoted Retweet Count	106.675	417.836	1,153.988
Average Quoted Retweet Count	761.96	522.295	672.48
Average Followers Count	5.994	2,917.716	3,245
Average Status Count	31.783	19,506.5	22,474
Average List Count	62.8	30.915	43.07
Average Friends Count	3.023	2,292.911	2,596

sentiment about specific tweets. For example, some people added “Hot Take” status to tweets to indicate potentially controversial ideas or opinions that were being expressed. The status options that were counted in tweets were linked to public health, economic impact such as job losses, mobility & transportation concerns.

The Average Status Count for Supporting Reopening as shown in Table 4 stood at 31,783 (NC), 19,507 (PA) and 22,474 (CA), making NC the state with the highest Average Status Count for Supporting Reopening.

The tweets supporting Reopen California (ReopenCA) were almost the total of Reopen North Carolina (ReopenNC) and Reopen Pennsylvania (ReopenPA). A major factor in this disparity may be the population of each state and California is the largest state of the three. There were more tweets from verified profiles about ReopenPA compared to the other two states, and the assumption is that more government personnel in Pennsylvania, either living in the state, or representing the state were engaged in the online conversations about the protests. Across each of the three states, tweets that were quoted received more liking and retweeting compared to tweets that were not quoted.

Research question #3

How did the three states compare on “Average Followers Count” and “Average Status Count” for Opposing Reopening?

While Pennsylvania showed the highest Average Followers Count at 4,616 (PA) for Opposing Reopening, California and North Carolina followed closely at 3,934 (CA) and 3,418 (NC).

The Average Status Count for Opposing Reopening as shown in table 5 stood at 20,612 (NC), 25,756 (PA) and 23,629 (CA), making PA the state with the highest Average Status Count for Opposing Reopening.

Discussion and Conclusion

Three datasets were analyzed based on people’s tweets about the stay-at-home orders in the states of North Carolina, California, and Pennsylvania. The first task in the data analysis was to determine the data size, how many tweets, retweets and how often people retweet, whether tweets mentioned peers, public officials, status or included links to other websites to support the tweeter’s opinion. Most of the

Table 5: Comparison of opposing reopen protests for the three states.

Metadata	North Carolina	Pennsylvania	California
Tweet Count	1.069	1.499	644
Percentage	52.61%	65.17%	27.29%
Tweets with Mentions	233	466	172
Tweets from Verified Profiles	26	90	8
Average Tweet Length	153	160	165
Favorite Count	29.709	38.079	3.338
Quoted Favorite Count	177.861	527.443	791.366
Average Quoted Favorite Count	1,347.43	352.098	1,228.82
Retweet Count	330	11.963	153
Average Retweet Count	5.76	7.986	1.14
Quoted Retweet Count	55.103	176.060	157.296
Average Quoted Retweet Count	417.45	117.53	244.24
Average Followers Count	3.418	4,615.986	3.934
Average Status Count	20,612.41	25,756.097	23.629
Average List Count	47.98	63.617	44.29
Average Friends Count	2.237	2,251.748	2.511

effort was spent on mining the data, particularly on lessons that could be relevant for future pandemics. The findings showed that people’s attitudes about the reopening protests were different in California compared to North Carolina and Pennsylvania. In North Carolina and Pennsylvania, many tweets were against reopening while in California, many tweets were in favor of reopening.

When comparing the data across each of the three states, the tweets opposing the Reopen protests had a different trend (Table 5). North Carolina and Pennsylvania had the larger number of opposers with Pennsylvania being highest at 1,499 tweets and North Carolina closely following at 1,069. California had a very low number of opposing tweets (644). The data on favorite counts and their averages tells us about the dissatisfaction that many people had towards why people were protesting in the first place. The retweets and favorite counts were high in Pennsylvania, and it might explain the momentum that was growing on both sides with those that were passionate about wanting to reopen facilities in the state against the protesters that opposed it.

Social media platforms such as the one used for this study (Twitter or X) are mechanisms for spreading information (and sometimes, misinformation), and for connecting people, for both positive and negative motives. Each of the Reopen protests datasets contain various features that were analyzed and compared.

The social features of tweets investigated the sentiment of people and how much momentum was built on Twitter (X) for those supporting the protests compared to those opposing the protests. The mean values and counts were reported in multiple tables in the Data Analysis section. In North Carolina, it was observed that there were slightly more tweets that were against the Reopen protests compared to those supporting. Those that were supporting the protests could be described as adamant in how they were mentioning other users and encouraging more people to join their support groups. Although there were more opposing tweets, there were more interactions with retweeting and liking tweets of Reopen NC protest supporters and even a large following. Those in favor of reopening the state of North Carolina were also very active online in their Reopen NC protests.

Reopen Pennsylvania protests showed that there were more against reopening than supporting reopening. While the average tweet length and average friends count were close in number among both groups, it was observed that there was a lot of activity with those against reopening through the retweets and favorite numbers.

In California, there were significantly more people that were in favor of reopening compared to those against the protests. In California, particularly about mid-2020, several counties showed that positive cases of the virus were trending lower due to people obeying the stay-at-home order. Those that were in favor of reopening were vastly more active on Twitter (X) compared to those that were against. This is interesting to observe because out of the three states chosen for the study, California was the only one that online social interactions and protests were mostly in favor of reopening.

An important lesson to share from this study is that the public stay-at-home and return-to-office policies implemented in CA worked better than those in NC and PA because people accepted and tweeted to reopen CA using scientific evidence/results that supported reduced cases and the need to reopen. At the same time in NC and PA, there was a surge in tweets against reopening the two states. As discussed in [12], scientific evidence and the participation of scientists in tweets could aid in both accurate information to the public as well as policy formulation.

Another key lesson-learned is the large organization of protesters via social media, in this

case, Twitter (X) about the pandemic and the impact of the stay-at-home policies. Based on the findings in the study, people tweeted their opinions and expressed sentiments about the COVID-19 pandemic policies and it led to a large following, and friendship among those tweeting about similar views. Policy makers in government need to be aware of the importance of social media platforms as a resource for spreading information and proactively use it early in a public health crisis like that of COVID-19 to inform people, share policies and provide actionable items for the public to follow to save lives and protect people's livelihoods.

References

1. Delorey TM, Ziegler CGK, Heimberg G, Normand R, Yang Y, Segerstolpe Å, Abbondanza D, Fleming SJ, Subramanian A, Montoro DT, Jagadeesh KA, Dey KK, Sen P, Slyper M, Pita-Juárez YH, Phillips D, Biermann J, Bloom-Ackermann Z, Barkas N, Ganna A, Gomez J, Melms JC, Katsyv I, Normandin E, Naderi P, Popov YV, Raju SS, Niezen S, Tsai LT, Siddle KJ, Sud M, Tran VM, Vellarikkal SK, Wang Y, Amir-Zilberstein L, Atri DS, Beechem J, Brook OR, Chen J, Divakar P, Dorceus P, Engreitz JM, Essene A, Fitzgerald DM, Fropf R, Gazal S, Gould J, Grzyb J, Harvey T, Hecht J, Hether T, Jané-Valbuena J, Leney-Greene M, Ma H, McCabe C, McLoughlin DE, Miller EM, Muus C, Niemi M, Padera R, Pan L, Pant D, Pe'er C, Pfiffner-Borges J, Pinto CJ, Plaisted J, Reeves J, Ross M, Rudy M, Rueckert EH, Siciliano M, Sturm A, Todres E, Waghray A, Warren S, Zhang S, Zollinger DR, Cosimi L, Gupta RM, Hacohen N, Hibshoosh H, Hide W, Price AL, Rajagopal J, Tata PR, Riedel S, Szabo G, Tickle TL, Ellinor PT, Hung D, Sabeti PC, Novak R, Rogers R, Ingber DE, Jiang ZG, Juric D, Babadi M, Farhi SL, Izar B, Stone JR, Vlachos IS, Solomon IH, Ashenberg O, Porter CBM, Li B, Shalek AK, Villani AC, Rozenblatt-Rosen O, Regev A. COVID-19 tissue atlases reveal SARS-CoV-2 pathology and cellular targets. *Nature*. 2021 Jul;595(7865):107-113. doi: 10.1038/s41586-021-03570-8. Epub 2021 Apr 29. PMID: 33915569; PMCID: PMC8919505.
2. Etim A, Yarber L. COVID-19 vaccine hesitancy: Analyzing risk factors impacting minority populations' acceptance/adoption and ICT-based solutions. In: Etim A, editor. *Adoption and use of technology tools and services by economically disadvantaged communities: Implications for Growth and Sustainability*. IGI Global; 2024.
3. US coronavirus vaccine tracker. USAFacts.
4. Charumilind S, Craven M, Lamb J, Singhai S, Wilson M. Pandemic to endemic: How the world can learn to live with COVID-19. McKinsey & Company. 2021.
5. Chou WS, Budenz A. Considering Emotion in COVID-19 Vaccine Communication: Addressing Vaccine Hesitancy and Fostering Vaccine Confidence. *Health Commun*. 2020 Dec;35(14):1718-1722. doi: 10.1080/10410236.2020.1838096. Epub 2020 Oct 30. PMID: 33124475.



6. Cowan SK, Mark N, Reich JA. COVID-19 vaccine hesitancy is the new terrain for political division among americans. Socinus: Sociological Research for a Dynamic World. 2021;7:237802312110236. doi: 10.1177/23780231211023657.
7. Bansal A, Gupta C, Muralidhar A. A sentimental analysis for youtube data using supervised learning approach. International Journal of Engineering and Advanced Technology (IJEAT). 2019.
8. Kates J, Orgera K. The red/blue divide in COVID-19 vaccination rates. KFF. 2021.
9. Habib MA, Anik MAH. Impacts of COVID-19 on Transport Modes and Mobility Behavior: Analysis of Public Discourse in Twitter. Transp Res Rec. 2023 Apr;2677(4):65-78. doi: 10.1177/03611981211029926. Epub 2021 Aug 10. PMID: 37153163; PMCID: PMC10149523.
10. Airak S, Sukor NSA, Rahman NA. Travel behaviour changes and risk perception during COVID-19: A case study of Malaysia. Transportation Research Interdisciplinary Perspectives. 2023 Mar;18:100784. doi: 10.1016/j.trip.2023.100784.
11. Chintalapudi N, Battineni G, Amenta F. Sentimental Analysis of COVID-19 Tweets Using Deep Learning Models. Infect Dis Rep. 2021 Apr 1;13(2):329-339. doi: 10.3390/idr13020032. PMID: 33916139; PMCID: PMC8167749.
12. Biermann K, Taddicken M. Visible scientists in digital communication environments: An analysis of their role performance as public experts on Twitter/X during the COVID-19 pandemic. Public Underst Sci. 2024 May 21;34(1):9636625241249389. doi: 10.1177/09636625241249389. Epub ahead of print. PMID: 38771041; PMCID: PMC11673311.
13. Xing Y, He Y, Zhang Z. Examining themes of social media users' opinion on remote work during COVID-19 pandemic: A justice theory perspective. Library Hi Tech. 2023;43(1):249-273.