

# Public Opinions toward COVID-19 Vaccine Mandates: A Machine Learning-based Analysis of U.S. Tweets

Yawen Guo, MISM<sup>1</sup>, Jun Zhu<sup>1</sup>, Yicong Huang<sup>1</sup>, Lu He<sup>1</sup>, Changyang He<sup>2</sup>, Chen Li, PhD<sup>1</sup>, Kai Zheng, PhD<sup>1</sup>

<sup>1</sup>University of California, Irvine, Irvine, CA, USA; <sup>2</sup>Hong Kong University of Science and Technology, Hong Kong SAR, China

## Abstract

*While it has been scientifically proven that COVID-19 vaccine is a safe and effective measure to reduce the severity of infection and curbing the spread of the SARS-CoV-2 virus, skepticism remains widespread, and in many countries vaccine mandates have been met with strong opposition. In this study, we applied machine learning-based analyses of the U.S.-based tweets covering the periods leading toward and after the Biden Administration's announcement of federal vaccine mandates, supplemented by a qualitative content analysis of a random sample of relevant tweets. The objective was to examine the beliefs held among twitter users toward vaccine mandates, as well as the evidence that they used to support their positions. The results show that while approximately 30% of the twitter users included in the dataset supported the measure, more users expressed differing opinions. Concerns raised included questioning on the political motive, infringement of personal liberties, and ineffectiveness in preventing infection.*

**Keywords:** Twitter; public opinion; COVID-19 vaccine mandate policy; sentiment analysis

## Introduction

Since its outbreak, the COVID-19 pandemic has resulted in significant morbidity, mortality, and economic fallout around the global<sup>1</sup>. Due to the highly contagious nature of the SARS-CoV-2 virus, in addition to its diverse variants and the general public's lack of ability or willingness to follow public health guidelines, the disease remains under- or uncontrolled in many parts of the world<sup>2</sup>. COVID-19 vaccine, even though developed and approved for emergency use through an expedited process, has been scientifically proven to be safe and highly effective<sup>3</sup>, and it remains the best strategy presently available to mitigate the pandemic and restore normalcy<sup>4</sup>. A total of 16 COVID-19 vaccines to date have been approved by the World Health Organization (WHO), with abundant supplies for initial shot(s) and booster(s) particularly in western counties. However, there continues to be widespread skepticism regarding vaccine safety and its potential unknown long-term health effects. As of March 2022, a third of the U.S. population have not been fully vaccinated, and this rate varies to a great extent across different states<sup>5</sup>, making it difficult to achieve nationwide community immunity<sup>6,7</sup>.

To accelerate vaccination uptake, many countries have implemented measures such as vaccine passports and vaccine mandates. In the U.S., the Biden Administration announced federal policies in September 2021 to enforce mandatory vaccination, including a requirement for employers with 100 or more employees to ensure each of their workers is fully vaccinated or tests for COVID-19 on at least a weekly basis, and a requirement for healthcare workers at facilities participating in Medicare and Medicaid are fully vaccinated, which is expected to affect more than 17 million workers at approximately 76,000 healthcare facilities<sup>8</sup>. Several states, counties, and municipal authorities introduced similar mandates, and the Secretary of Defense issued an order to require full COVID-19 vaccination of all servicemembers of the U.S. Armed Forces<sup>9</sup>. While these vaccine mandates are an effective means to increase the vaccination rate to control the pandemic, there has been strong pushback from certain states, employers, civil liberties organizations, and religious groups. In some countries, including the U.S., vaccine mandates have been met with furious protests, such as the "Freedom Convoy" that caused weeks-long blockage of cities and key economic passages that led to the invocation of the Emergencies Act in Canada<sup>10</sup>.

Understanding the public's opinions toward vaccine mandates is thus critical for policy makers, public health agencies, and health organizations to develop proper measures such as education campaigns, alternative incentives, and special exemption considerations<sup>12</sup>. Polling public opinions using conventional surveying methods are however costly and difficult to reach a representative sample involving all relevant stakeholders. In recent years, social media platforms such as Twitter have become an invaluable resource for researchers to collect large volumes of data to reveal prevailing public opinions at a very low cost<sup>12</sup>. While there have been social media-based studies on other COVID-19 measures such as lockdowns and mask mandates<sup>13-17</sup>, to date, very few studies have specifically looked

into the social media discussions around COVID-19 vaccines and vaccine mandates. To fill this gap, we analyzed U.S.-based tweets collected from middle July 2021 to December 2021 using machine-learning and qualitative content analysis methods. The objectives of the study were threefold: (1) to explore how the general public reacted to the vaccine mandates in the U.S., (2) to examine the evolution of the public opinions over time, and (3) to investigate the justifications or concerns that Twitter users used to support their differing opinions.

## Methods

### Data collection and preprocessing

The dataset analyzed in this study was collected in real time using the Twitter Covid-19 Stream API<sup>18, 19</sup>, covering all COVID-19 relevant tweets across a 5-month timeframe from July 14 to December 31, 2021. The beginning of this period was marked by the first public call for vaccine mandates issued by the Society for Healthcare Epidemiology of America<sup>20</sup>. The data collection ended three months after the Biden Administration's announcement of federal COVID-19 vaccine mandates.

Because this study focused on the public's reaction to the vaccine mandate policies in the U.S., only the tweets with a U.S.-based geocode (continent U.S., Alaska, and Hawaii) were included. We did not use the location information included in the user profiles as the location where the tweet was posted may be different<sup>21</sup>. Further, as retweets do not have a geocode, we did not include retweets in our analyses<sup>22</sup>.

Next, we manually reviewed a sample set of the data to develop a comprehensive keyword list to identify tweets relevant to COVID-19 vaccine mandates. This analysis arrived at 20 pairs of keywords: {vaccine, vax, vaxxine, immunization, mRNA} \* {mandate, requirement, policy, force}. We then used a collaborative big-data analytics platform called Texera developed by the co-authors (YH, CL) to facilitate keyword search among the tweets<sup>23</sup>. The Texera platform provides a full-text search capability enabling us to identify tweets containing a keyword pair of different parts of speech. It also provides the capability of performing additional preprocessing steps such as stemming and removal of stop words, punctuations, and hashtags to prepare the data for the subsequent analyses. A word2vec algorithm was then applied to represent words as 100-dimensional vectors<sup>24</sup>.

### Two-stage machine-learning classification

A substantial proportion of the tweets retrieved from the Twitter Covid-19 Stream API, while relevant to COVID-19, contained no personal opinions toward the COVID-19 vaccine mandates; for example, there were many tweets that merely shared news articles reporting the federal mandatory vaccination requirements. Thus, we developed a supervised machine-learning model to remove irrelevant tweets as the first stage of the analysis. Then, we developed another supervised sentiment classification model to classify the positivity of the public opinions expressed in the tweets identified from the first stage. We comparatively examined the performance of four machine-learning models that have been commonly used in text classification tasks. These include logistic regression, AdaBoost<sup>25</sup>, Xgboost<sup>26</sup>, and Long Short-Term Memory (LSTM) Network<sup>27</sup>. These machine-learning models were programmed in Jupyter Notebook 4.8 using scikit-learn libraries.

Two authors (YG and JZ) manually annotated a random sample of 1,000 tweets to prepare training data. For each tweet, they determined whether it contained an expression of personal opinion(s), followed by whether the opinion expressed was positive, neutral, or negative. They first coded 200 tweets independently to calibrate the results. Differences were resolved through consensus development meetings. The interrater reliability was 0.91 for determining whether a tweet expressed a personal opinion, and 0.93 for determining the positivity of the opinion. The two coders then went on to annotate the remaining 800 tweets with an even split between them.

### Longitudinal analyses and topic modeling

To investigate the longitudinal evolution of public opinions toward COVID-19 vaccine mandates, we calculated weekly tweet counts and average positivity ratings over the study period. We also divide the data into three distinct phases: the Incubating phase (7/14–9/8/2021) when discussions took place on whether the U.S. federal government or local authorities should consider instituting vaccine mandate policies; the Promulgation phase (9/9–11/4/2021) that immediately followed the Biden Administration's announcement of federal vaccine mandates; and the Aftermath phase (11/5–12/31/2021) marked as lawsuits being filed by conservative groups against the mandates. For the data from each phase, we studied the top features using Tf-idf<sup>29</sup> and top words using Latent Dirichlet

Allocation<sup>30</sup>, a popular topic modeling approach, to identify commonly discussed topics and how they evolved over time.

**Table 1.** The three distinct phases of study

Phase	Dates	Milestone Event
Incubating	7/14–9/8/2021	A statement issued by the Society for Healthcare Epidemiology of America on July 14, 2021 that called for healthcare facilities to require their workers to receive the COVID-19 vaccine, noting that a sufficient rate of vaccination will not be achieved without a mandate.
Promulgation	9/9–11/4/2021	The Biden Administration announced on September 9, 2021 federal policies mandating COVID-19 vaccination for employers with 100 or more employees and for healthcare workers at facilities participating in Medicare and Medicaid.
Aftermath	11/5–12/31/2021	Lawsuits filed by at least two conservative groups on November 5, 2021 to challenge the legality of the federal vaccine mandates.

### Qualitative content analysis

With the results obtained through the above steps, we carried out a qualitative content analysis based on the grounded theory<sup>31</sup> to examine the evidence that the Twitter users used to support their positions regarding the COVID-19 vaccine mandates. Further, we chose not to use a computational method for this analytical task due to the complexity and ambiguity of tweet data. It has been shown that the state-of-the-art natural language processing techniques are still inadequate to reliably derive a proper understanding of complex human expressions (e.g., through sarcasm), especially through short social media texts such as tweets<sup>32</sup>.

We sampled a random set of 200 tweets for open coding in order to generate the initial coding schema. Two coders (YG and JZ) independently coded the data; differences were resolved through consensus development meetings. The interrater reliability of the opening coding was 0.87. Theoretical saturation<sup>33</sup> was achieved after coding approximately 100 tweets. The two coders then separately coded an additional set of 300 tweets (150 each). Thus, we analyzed a total of 500 tweets in this qualitative content analysis.

### Results

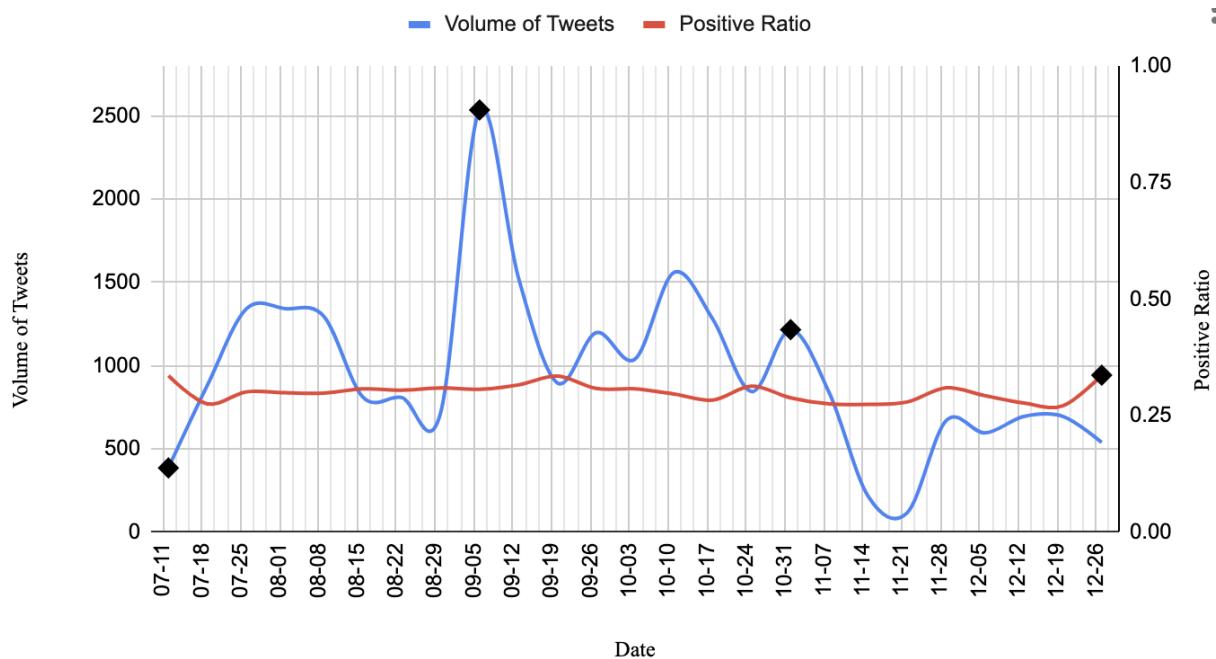
During the study period (7/14/2021–12/31/2021), we retrieved a total of 1,466,879 tweets from the Twitter Covid-19 Stream API. Of them, 41,421 were the U.S.-based tweets according to the recorded geo-location information.

In the first step to remove irrelevant tweets, the best performing machine-learning model is LSTM, with a dropout layer added (rate: 0.1) to avoid overfitting<sup>28</sup>. The accuracy achieved for this binary classification (i.e., relevant vs. irrelevant) was 76.0%. After applying this classifier, a total of 24,063 tweets remained in the dataset which contain expressions of personal opinions. For the second-stage machine learning-based analysis to classify the positivity of the expressed opinions, Xgboost demonstrated the best performance, with an overall accuracy of 81.8%.

Table 2 reports the average weekly volume of relevant tweets and average positive ratio aggregated within each of the three study phases. Figure 1 exhibits the week-by-week trends. Overall, a larger number of tweets were circulated in the Promulgation phase and fewer in the Incubating and Aftermath phases. The positive ratio remained steady at the 30% level during the first two phases, and dropped below 30% in Aftermath.

**Table 2.** Descriptive analysis of tweets containing personal opinions towards vaccine mandate

Phase	Average Weekly Volume	Average Positive Ratio
Incubating	949	30.4%
Promulgation	1,345	30.6%
Aftermath	485	29.0%



**Figure 1.** Week-by-week tweet volume and positive ratio

We experimented with topic models using different sizes of clusters and found that a 3-cluster model yielded the most meaningful topic separation. The results are reported in Table 3. Initially, before the Biden Administration's announcement of the federal vaccine mandates, the relevant discussions on Twitter largely focused on general vaccination requirements for travel, public events attracting large crowds (e.g., NFL games), and business and school reopening. After the announcement, more negative topics started to emerge, voicing opinions on how the federal mandates might restrict freedom and liberty and affect businesses and other populations, in addition to the potential side effects of vaccination. In the Aftermath phase, the discussions intensified on the legality of the vaccine mandates, the anti-vaccine and anti-mandates protests, and the effects of the mandates on domestic traveling.

**Table 3.** Evolution of frequently discussed topics over time

Phase	Topics	Top Words
Incubating	<ul style="list-style-type: none"> <li>• Current and potential mandate policies</li> <li>• Vaccination requirements for business reopen and public events</li> <li>• Vaccination requirements for school reopening</li> </ul>	<ul style="list-style-type: none"> <li>• <i>delta, government, right, forcing, choice, body</i></li> <li>• <i>business, company, workers, players, NFL</i></li> <li>• <i>students, teachers, schools, safe, keep, wear, masks</i></li> </ul>
Promulgation	<ul style="list-style-type: none"> <li>• Liberty restriction and legality</li> <li>• Populations affected</li> <li>• Effectiveness and safety</li> </ul>	<ul style="list-style-type: none"> <li>• <i>free, freedom, illegal, immigrants, border, Biden, force, state</i></li> <li>• <i>workers, staff, businesses, companies, citizens, military, federal</i></li> <li>• <i>diseases, experimental, herd, side, effects, immunity</i></li> </ul>
Aftermath	<ul style="list-style-type: none"> <li>• Legality</li> <li>• Anti-vaccine / anti-mandates protests</li> </ul>	<ul style="list-style-type: none"> <li>• <i>military, federal, choice, employee, Biden, unconstitutional, right, government, illegal</i></li> <li>• <i>anti, stop, right, transmission, force, right, health,</i></li> </ul>

	<ul style="list-style-type: none"> <li>Effects on travel</li> </ul>	<p><i>natural, immunity</i></p> <ul style="list-style-type: none"> <li><i>travel, flights, infection, spread, domestic, Trump</i></li> </ul>
--	---	--

Table 4 reports the themes that emerged from the qualitative content analysis, focusing on evidence that the Twitter users used to support their pro- or anti-mandates beliefs. The percentage column was calculated by dividing the number of tweets containing a particular opinion and the supporting evidence over the total number of tweets analyzed ( $N=500$ ). Some tweets were counted more than once if they offered multiple supporting arguments. The sample tweets included in the table and the rest of this paper were paraphrased to protect the identity of the user.

**Table 4.** Themes from the qualitative content analysis

Positivity	Theme	%	Description	Sample Tweet
Positive	General positive attitude	15.5	The tweet expresses a general supporting attitude without specific reasons provided.	<i>Get Vaccinated! Vaccine mandates work!</i>
	Health risks of not being vaccinated to self and others	4.5	The tweet supports the mandates because remaining unvaccinated poses elevated risks both for self and others.	<i>Nothing consistent in supporting vaccine mandates and abortion rights. The decision of whether or not to have an abortion affects one person and that's the woman. Refusing to get vaccinated poses a risk to public health- learn the difference.</i>
	Historical precedents	4.5	The tweet supports the mandates because there have been historical precedents of vaccine mandates.	<i>Vaccine mandates have a long history when in the interest of public health. In the 70s it was required to get vaccinated when traveling from some countries.</i>
	Safe and effective	3.5	The tweet supports the mandates because COVID-19 vaccines are effective in reducing severity, hospitalization, and mortality, and they are safe both near- and long-term.	<i>Vaccines are safe and effective. This policy protects the health and safety of BART riders and workers.</i>
	Necessary for reopening and economic recovery	2.0	The tweet supports the mandates in order to accelerate reopening and economic recovery.	<i>The tension between rebuilding positive school culture and keeping kids safe, so we can stay at school is real. Can we get a vaccination mandate and start opening things up?</i>
	No other better alternatives	0.5	The tweet supports the mandates because there are no other better alternatives currently available.	<i>It is a standard principle of decision theory that the expected utility of a proposed option must be compared with the expected utility of relevant alternatives.</i>
Negative	General negative attitude	34.5	The tweet holds a general negative attitude toward vaccine mandates with no specific reasons provided.	<i>They gotta stop tryna force people to take that vaccine.</i>
	Political motive	12.0	The tweet opposes the mandates because of political beliefs that the policies are manipulation tactics by certain politicians and special interest groups.	<i>They all lied then pushed a Federal Vaccine Mandate in sync and obviously on purpose. So the democrat party and deep state corruption actors use lying as a tactic. Awake yet?</i>
	Human rights and personal liberties	7.0	The tweet opposes the mandates because they infringe upon	<i>Vaccine mandates are un-American and anti-freedom. Give people the information,</i>

			individuals' liberty of making their own decisions. <i>and let them make an informed decision.</i>
Ineffectiveness in preventing infection	6.5	The tweet opposes the mandates because even vaccinated, one can still get infected.	<i>Got covid on a family trip. This omicron is literally crazy. It feels like literally a cold and most of us are vaccinated too. The vax is not working at all!</i>
Involuntary unemployment	4.0	The tweet opposes the mandates due to the concerns over losing their jobs as a sequence of declining vaccination.	<i>Soon I probably won't have a job due to the mandated covid vaccine. My quality of care that I give does not depend on a vaccine California is not going to choose what goes in my body!</i>
Medical side effects	3.0	The tweet opposes the mandates because of potential side effects that may be associated with vaccination.	<i>So if my husband ends up having side effects from the J&amp;J he was forced to take in order to work ... And he ends up not being able to work due to the side effects from Covid Vax.</i>
Legality of vaccine mandates	2.5	The tweet opposes the mandates because it is deemed illegal or unconstitutional.	<i>There is no legal precedent for requiring vaccines approved under emergency use, which may be why some colleges are reluctant to announce mandates now.</i>
Unnecessary due to natural immunity	1.5	The tweet opposes the mandates because of the belief that individuals who have contracted COVID-19 already have natural immunity and thus do not need to be vaccinated.	<i>Covid won't go away and some people already have covid so they don't need the vaccine.</i>
Unknown health effects on children or the elderly	1.5	The tweet opposes the mandates because of unknown health effects on small age children or the elderly.	<i>I hope it's not intentional. But I think until they've been through all testing and not rushed the mandate, kids should wait.</i>
Mask requirements despite vaccination	1.0	The tweet opposes the mandates because even if vaccinated they are still required to wear a facial mask.	<i>Now the Absurd mask mandates vaccination. Earn credibility!</i>
Religious reasons	1.0	The tweet opposes the mandates because of religious beliefs.	<i>So your solution is to vaccinate everyone. How do you implement that without using force? Also, what about the folks who have religious or other objections to vaccination.</i>
Medical reasons	0.5	The tweet opposes the mandates because some people have medical reasons that prevent them from being vaccinated.	<i>FYI, if you force this FDA full approval for vaccines, you again can't force the entire population to take it! Because you have some people for medical reasons who can't and we have a constitution as well!</i>

As shown in Table 4, most tweets only expressed an opinion without providing specific reasoning. These constituted 15.5% of the positive tweets and 34.5% of the negative tweets. The most commonly offered reasons arguing for the necessity of vaccine mandates included "Risks of not being vaccinated to self and others" (4.5%), for example:

*"That's why it's important to mandate the vaccine at all schools. It's to protect everyone. There could always be one person who isn't vaccinated that can spread it to everyone."*

*"I'm so sorry to read this. We do need mandates. Noticing too many nurses that are also anti-vaxx. I work side by side w/ them daily, in very small spaces w/o physical distancing. I'm very concerned. Working in healthcare for so long, I took for granted everyone was vaccinated."*

4.5% of the tweets used historical precedents to back up the support for vaccine mandates:

*“Multiple vaccines have been required over the years. Remember polio? We would still have polio today if vaccines were not required. Your personal freedom is guaranteed except when it infringes on the safety of others. And not getting vaccinated against Covid does exactly that.”*

*“LAUSD already requires vaccinations for measles, mumps, rubella, polio, et al. Why not require Covid-19 vaccinations for all students 12 and up?”*

In addition, some pro-mandates tweets argued that the COVID-19 vaccines were safe and effective (3.5%), were necessary for reopening and economic recovery (2%), and they needed to be mandated because there were no other better alternative means for controlling the pandemic (0.5%).

Among the anti-mandates tweets, a substantial proportion (12%) expressed a belief that these mandates were politically motived rather than for controlling the pandemic, and 7% cited reasons for violation of human rights and infringement upon personal liberties of making free decisions, for example:

*“This vaccine mandate thing is bullshit. It’s Biden’s tactic to distract republicans from the audits and 2020 election fixing.”*

*“How many more vaccines are yall going to try to mandate until you can control everyone.”*

*“Vaccine mandates are un-American and anti-freedom. Give people the information, and let them make an informed decision.”*

6.5% of the tweets were against the vaccine mandates because their lack of effectiveness and the fact that vaccinated individuals could still be infected and spread the virus:

*“Since July I told YOU all about the Ineffectiveness of the Vaccines. Yet YOUR GOVERNMENT is still pushing vaccine MANDATES. So how you all feel after you fully Vaccinate & boost yet in less than 6 months (some people 2 months) ketching Covid and in d hospital.”*

*“CDC stops counting COVID cases among the vaccinated. Gov’t says vaccines will save us but vaxxed can contract and pass on the virus. Biden wants to mandate employment requirements to private businesses.”*

4% of the anti-mandates tweets expressed concerns on unvaccinated individuals being forced out of work and thus could not provide for their families:

*“No your brain dead president is the one that’s getting thousands of them fired because I’m f\*\*\*\*\* vaccine mandate.”*

*“But my theory is we gone be forced into the vaccine because they gone spin it like oh if you not vaccinated no unemployment or assistance of any sort while jobs are closed which leads to people not being able to provide for their families so now they in panic and controlled.”*

The remaining negative tweets cited reasons related to potential medical side effects (3%) and legality of instituting vaccine mandates by the federal government (2.5%). Other evidence used to support the anti-mandates opinions included the lack of consideration for those who already had natural immunity through infections (1.5%), unknown health effects on smalls-age children and the elderly (1.5%), no lifting of mandatory mask requirements despite of vaccination (1%), and religious (1%) or medical reasons (0.5%).

## **Discussion**

The purpose of this research was to analyze public opinions expressed in Twitter about the COVID-19 vaccine mandates in the U.S. We studied the tweets circulated between July 14 and December 31, 2021 to examine the trends of relevant Twitter discussions. In particular, we investigated the reasons why Twitter users supported or opposed the mandatory measure using both machine-learning and qualitative content analysis methods.

The results show that there were strong oppositions voiced on the Twitter platform against the COVID-19 vaccine mandate policies. In fact, a majority of the relevant tweets—approximately 70% throughout the study

period—expressed negative opinions. This rate is alarming, as there has been robust scientific evidence showing that the COVID-19 vaccines are effective in reducing symptom severity, hospitalization, and mortality, and that adverse incidents associated with the vaccination are extremely rare. This rate is also considerably higher than that reported in other studies. For example, an online survey conducted by researchers from the University of Pennsylvania in September 2020 found that 44.9% of the respondents supported state vaccine mandates among adults, and 47.7% deemed employer-enforced mandates acceptable<sup>34</sup>. Similarly, another survey study conducted in late October and early November 2020 in educational settings found that a majority of students and teachers supported vaccine mandates<sup>35</sup>. It is possible that Twitter users, an inherently self-selected sample, are more likely to hold negative opinions toward the COVID-19 vaccine mandates, or other government mandates at large. It is also possible that the results of our study are more representative than those obtained through conventional surveys, which may be limited by smaller sample sizes and lower response rates.

Among the reasons for opposition, questioning the political motives that underlay the federal vaccine mandates was represented in a substantial proportion (17%) of the negative tweets. This shows that the lack of trust in the government, or the governing party, is a major driver of the anti-mandates attitude in the U.S. This opinion is also supported by other related arguments challenging the legality of the federal government imposing mandatory vaccine requirements and on the constitutional implications of the mandates in infringing upon personal liberties. Further, 6.5% of the negative tweets spoke about the lack of effectiveness of the COVID-19 vaccines in preventing infection. This may be reflection of some inaccurate depictions of the purpose of the vaccination in early promotion campaigns. Notably, health experts who have reservations on mandating COVID-19 vaccination often used reasons such as natural immunity acquired by contracting with the virus, inadequate data on effectiveness and risks among the pediatric population, and religious and medically related considerations. These however only constituted a small percentage of the negative opinions expressed on Twitter.

The results of this study offer several implications for researchers, public health practitioners, and policy makers. First, the week-to-week changes in the volume of relevant tweets corresponded very well to the key events defining the three distinct study phases such as the announcement of the federal mandate policies. This suggests that analysis of social media data can provide timely insights into public opinions, and longitudinal evolution thereof, which offers a unique advantage over conventional survey methods. Second, some of the negative opinions were misguided or based on outdated information. For example, while COVID-19 vaccines reduce severity and the likelihood of infection, it does not eliminate the possibility of getting infected and for vaccinated individuals to spread the virus. Some Twitter users however appeared to base their negative opinions on unrealistic expectations or overpromised benefits. Thus, communication from public health authorities needs to be precise and reflective of the best available scientific evidence in order to avoid any misinterpretation of the anticipated effect of public health measures. Further, we also found that some oppositions highlighted by many (e.g., pundits in news media) were not fully reflected in the Twitter data. For example, religious beliefs are often quoted as a top reason for individuals to refuse vaccination. This concern however appeared infrequently in relevant Twitter discussions. This finding suggests that public health practitioners and policy makers should avoid making presumptions or overly rely on experts' opinions. Instead, they should solicit information and feedback more widely, e.g., through channels such as social media and alike.

This study has several limitations. First, in order to limit the scope to U.S.-based tweets, we filtered out tweets of which the geolocation could not be determined. This substantially reduced the usable sample size. Second, because of the large volume, we could not manually annotate all tweets in the dataset. Instead, we drew random samples both to train the machine-learning models and to use in the qualitative content analysis. This may affect the machine-learning performance and limit the generalizability of our study findings. Third, our analysis was conducted prior to the surge of COVID-19 cases due to the omicron variant. Some Twitter users' attitude toward the vaccine mandates might have changed because of it, and because of more recent developments such as the widespread lifting of mask mandates, which were however not captured in our study.

## Conclusion

COVID-19 vaccine is an essential measure to mitigating the global pandemic. However, mandatory vaccination requirements have provoked significant controversies in many countries. In this study, we analyzed a large Twitter dataset to examine the public opinions toward the COVID-19 vaccine mandates in the U.S. and the evidence used by Twitter users to support their opinions. We found that 70% of the tweets circulated during the study period were negative, voicing strong oppositions due to political, economic, and health-related reasons. We believe that much

work is needed to improve public health communication and restore the trust between certain segments of the population and the government, which are critically needed to bring the COVID-19 pandemic to an end and to better prepare us for future public health crises.

### Acknowledgement

The data collection part of this work was supported by the NSF Award III 2107150.

### References

1. Centers for Disease Control and Prevention. COVID-19 Cases, Deaths, and Trends in the US. CDC COVID Data Tracker. <https://covid.cdc.gov/covid-data-tracker>. Accessed January 30, 2022.
2. Brown TS, Walensky RP. Serosurveillance and the COVID-19 Epidemic in the US: Undetected, Uncertain, and Out of Control. *JAMA*. 2020;324(8):749-51. PMID: 32692350.
3. Daniel W, Nivet M, Warner J, Podolsky DK. Early Evidence of the Effect of SARS-CoV-2 Vaccine at One Medical Center. *N Engl J Med*. 2021;384(20):1962-1963. PMID: 33755374; PMCID: PMC8008752.
4. Hagan K, Forman R, Mossialos E, Ndebele P, Hyder AA, Nasir K. COVID-19 vaccine mandate for healthcare workers in the United States: a social justice policy. *Expert Rev Vaccines*. 2022;21(1):37-45. PMID: 34709969.
5. Centers for Disease Control and Prevention. (n.d.). CDC Covid Data tracker. Centers for Disease Control and Prevention. Retrieved March 8, 2022, from <https://covid.cdc.gov/covid-data-tracker/#datatracker-home>
6. Meeker JK. The political nightmare of the plague: The ironic resistance of anti quarantine protesters. In COVID-19 (pp. 109-121). Routledge; 2020.
7. Cornwall W. Officials gird for a war on vaccine misinformation. *Science*. 2020;369(6499):14-5. PMID: 32631873.
8. The United States Government. (2021, November 4). Fact sheet: Biden Administration announces details of two major vaccination policies. The White House. Retrieved March 8, 2022, from <https://www.whitehouse.gov/briefing-room/statements-releases/2021/11/04/fact-sheet-biden-administration-announces-details-of-two-major-vaccination-policies/>
9. Mandatory Coronavirus Disease 2019 Vaccination of Department of Defense Service Members. (n.d.). <https://media.defense.gov/2021/Aug/25/2002838826/-1/-1/0/MEMORANDUM-FOR-MANDATORY-CORONAVIRUS-DISEASE-2019-VACCINATION-OF-DEPARTMENT-OF-DEFENSE-SERVICE-MEMBERS.PDF>.
10. Austen, D. B. and I. (2022, March 3). Canada News: Canada initiates emergency act, expanding measures to end protests. The New York Times. Retrieved March 9, 2022, from <https://www.nytimes.com/live/2022/02/14/world/canada-protests-news>
11. Cornwall W. Officials gird for a war on vaccine misinformation. *Science*. 2020;369(6499):14-5. PMID: 32631873.
12. He L, Yin T, Hu Z, Chen Y, Hanauer DA, Zheng K. Developing a standardized protocol for computational sentiment analysis research using health-related social media data. *J Am Med Inform Assoc*. 2021;28(6):1125-1134. PMID: 33355353; PMCID: PMC8200276.
13. Hussain A, Tahir A, Hussain Z, Sheikh Z, Gogate M, Dashtipour K, et al. Artificial Intelligence-Enabled Analysis of Public Attitudes on Facebook and Twitter Toward COVID-19 Vaccines in the United Kingdom and the US: Observational Study. *J Med Internet Res*. 2021;23(4):e26627. PMID: 33724919.
14. Jang H, Rempel E, Roth D, Carenini G, Janjua NZ. Tracking COVID-19 Discourse on Twitter in North America: Infodemiology Study Using Topic Modeling and Aspect-Based Sentiment Analysis. *J Med Internet Res*. 2021;23(2):e25431. PMID: 33497352.
15. Engel-Rebitzer E, Stokes DC, Buttenheim A, Purtle J, Meisel ZF. Changes in legislator vaccine-engagement on Twitter before and after the arrival of the COVID-19 pandemic. *Hum Vaccin Immunother*. 2021;1-5. PMID: 33970786.
16. Germani F, Biller-Andorno N. The anti-vaccination infodemic on social media: A behavioral analysis. *PLoS One*. 2021;16(3):e0247642. PMID: 33657152.
17. Guntuku SC, Purtle J, Meisel ZF, Merchant RM, Agarwal A. Partisan Differences in Twitter Language among US Legislators during COVID 19: A Cross-Sectional Study. *J Med Internet Res*. 2021 Apr 16. PMID: 33939620.
18. Twitter. (n.d.). twitter developer platform. Twitter. Retrieved March 3, 2022, from <https://developer.twitter.com/en/docs/twitter-api/tweets/covid-19-stream/overview>

19. Twitter. (n.d.). Filtered stream introduction | docs | twitter developer platform. Twitter. Retrieved March 3, 2022, from <https://developer.twitter.com/en/docs/twitter-api/tweets/filtered-stream/introduction>
20. Weber, D., Al-Tawfiq, J., Babcock, H., Bryant, K., Drees, M., Elshaboury, R., . . . Young, H. (2022). Multisociety statement on coronavirus disease 2019 (COVID-19) vaccination as a condition of employment for healthcare personnel. *Infection Control & Hospital Epidemiology*, 43(1), 3-11. doi:10.1017/ice.2021.322
21. He L, He C, Reynolds TL, Bai Q, Huang Y, Li C, Zheng K, Chen Y. Why do people oppose mask wearing? A comprehensive analysis of U.S. tweets during the COVID-19 pandemic. *J Am Med Inform Assoc*. 2021;28(7):1564-1573. PMID: 33690794; PMCID: PMC7989302.
22. Twitter. (n.d.). Filtering tweets by location | docs | twitter developer platform. Twitter. Retrieved March 3, 2022, from <https://developer.twitter.com/en/docs/tutorials/filtering-tweets-by-location>
23. Texera. (n.d.). Texera/texera: Big Data Analytics using interactive workflows. GitHub. Retrieved March 3, 2022, from <https://github.com/Texera/texera>
24. Mikolov T, Sutskever I, Chen K, Corrado GS, Dean J. Distributed representations of words and phrases and their compositionality. In: Burges CJC, Bottou L, Welling M, Ghahramani Z, Weinberger KQ, eds. *Advances in Neural Information Processing Systems*. Red Hook, NY: Curran Associates; 2013:3111-9.
25. T. -K. An and M. -H. Kim, "A New Diverse AdaBoost Classifier," 2010 International Conference on Artificial Intelligence and Computational Intelligence, 2010, pp. 359-363, doi: 10.1109/AICI.2010.82.
26. Chen, T., & Guestrin, C. (2016). XGBoost. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. <https://doi.org/10.1145/2939672.2939785>
27. Cheng, J., Dong, L., & Lapata, M. (2016). Long short-term memory-networks for Machine Reading. Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. <https://doi.org/10.18653/v1/d16-1053>
28. Gal Y, Ghahramani Z. A theoretically grounded application of dropout in recurrent neural networks. In: *Proceedings of the 30th International Conference on Neural Information Processing Systems*; 2016;29:1027-35.
29. Q. Liu, J. Wang, D. Zhang, Y. Yang and N. Wang, "Text Features Extraction based on TF-IDF Associating Semantic," 2018 IEEE 4th International Conference on Computer and Communications (ICCC), 2018, pp. 2338-2343, doi: 10.1109/CompComm.2018.8780663.
30. David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent dirichlet allocation. *J. Mach. Learn. Res.* 3, null (3/1/2003), 993-1022.
31. Corbin J, Strauss A. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Thousand Oaks, CA: Sage; 2014.
32. Tang J, Meng Z, Nguyen X, Mei Q, Zhang M. Understanding the Limiting Factors of Topic Modeling via Posterior Contraction Analysis. In: *Proceedings of the 31st International Conference on Machine Learning*; 2014; 32: I-190-8.
33. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, Burroughs H, Jinks C. Saturation in qualitative research: exploring its conceptualization and operationalization. *Qual Quant*. 2018;52(4):1893-1907. PMID: 29937585; PMCID: PMC5993836.
34. Largent EA, Persad G, Sangenito S, Glickman A, Boyle C, Emanuel EJ. US Public Attitudes Toward COVID-19 Vaccine Mandates. *JAMA Netw Open*. 2020;3(12):e2033324. doi:10.1001/jamanetworkopen.2020.33324
35. Simon F. Haeder, Joining the herd? U.S. public opinion and vaccination requirements across educational settings during the COVID-19 pandemic, *Vaccine*, Volume 39, Issue 17, 2021, Pages 2375-2385, ISSN 0264-410X, <https://doi.org/10.1016/j.vaccine.2021.03.055>.