University of Wisconsin Milwaukee
UWM Digital Commons

Theses and Dissertations

August 2021

# Semantic Analysis of Vaccine and Mask Sentiments in COVID-19 Twitter Data

Mohammadreza Sediqin
*University of Wisconsin-Milwaukee*

Follow this and additional works at: https://dc.uwm.edu/etd

Part of the Computer Sciences Commons

Recommended Citation
Sediqin, Mohammadreza, "Semantic Analysis of Vaccine and Mask Sentiments in COVID-19 Twitter Data" (2021). *Theses and Dissertations*. 2835.
https://dc.uwm.edu/etd/2835

This Thesis is brought to you for free and open access by UWM Digital Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of UWM Digital Commons. For more information, please contact scholarlycommunicationteam-group@uwm.edu.

# SEMANTIC ANALYSIS OF VACCINE AND MASK SENTIMENTS IN COVID-19 TWITTER DATA

by

Mohammadreza Sediqin

A Thesis Submitted in
Partial Fulfillment of the
Requirements for the Degree of

Master of Science
in Computer Science

at

**University of Wisconsin-Milwaukee**

August 2021

# *ABSTRACT*

SEMANTIC ANALYSIS OF VACCINE AND MASK SENTIMENTS IN COVID-19 TWITTER DATA

by

Mohammadreza Sediqin

The University of Wisconsin-Milwaukee
Under the Supervision of Dr. Purushottam Papatla and Dr. Hossein Hosseini

SARS CoV-2 (COVID-19) was identified as the cause of severe respiratory disease in China in 2019. It is a virus that will be transferred person-to-person by sneezing, coughing, or talking. This phenomenon not only affects public health and economics but also mental health as well. SARS-CoV-2 vaccines and wearing masks plays significant roles in preventing the spread of the COVID-19 virus, but vaccine hesitancy and anti-mask beliefs threaten the efficacy of the government orders in prevention and immunization against Coronavirus. The impact of the COVID-19 pandemic has been investigated from different aspects, but few large-scale studies focus on the opinion of people toward government orders to wear face mask and get vaccination. The abundant data on online social media however enables researchers to analyze people's attitudes toward vaccination and the use of face mask. In this study, we use twitter API and scrape 340 million COVID-19 tweets posted in the timeline of December 2020 to March 2021. Our goal is to investigate how people respond to tweets about masking and vaccines as a means of understanding sentiments towards both practices. Specifically, we focus on which tweets about the topics tend to become viral relative to those that are neither retweeted nor receive any replies. Toward this end, we split the dataset into three categories: 1) replied tweets 2) retweeted tweets, and 3) no-engagement tweets which are tweets that receive no response. We then deploy topic modeling to identify the most popular tweet topics in each category. Furthermore, we filter tweets for vaccine and mask related hashtags and use the algorithm,VADER to find the sentiment of these tweets. Our analysis indicates a slight difference in the distribution of tweets with positive and negative sentiments with vaccination or mask hashtags, with the dominant polarity of positive sentiments. Despite the overall strength of positive stances, negative opinions about COVID-19 vaccines and masks remain among people who are hesitant towards wearing face masks and vaccination. We also investigate and show that sentiments among Twitter users shift from positive to negative and vice versa over time. The most probable reasons for the domination of positive sentiments in tweets with vaccine and mask hashtags , appears to be the belief that such tweets are providing accurate information and also because of the risks of COVID-19 as discussed by well-regarded organizations. At the same time, however, inaccurate information, mistrust of well-regarded organizations or media and the influence of celebrities on their followers does push a segment of users into hesitancy and negative views about masks and vaccination.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# ACKNOWLEDGEMENTS

Firstly, I would like to express my appreciation to my thesis advisors, Dr.Papatla and my academic advisor Dr.Hosseini for giving me the opportunity to continue my education under their guidance over these years. Their guidance and advice played an essential role in making the accomplishment of this study. They always helped me with their practical suggestions in every aspect of my study that were crucial in making this project. I would also like to thank my committee member, Dr.Mali, for his contribution to this work. I would like to especially thank Dr.Papatla to support me financially and gave me this chance to continue my education by his grant. Finally, I must express my very deep appreciation to Zahra for offering her steady support, love, and her patience when I could not spend much time with her through the process of studying and writing this thesis. I also would thank my parents for providing me with constant support and encouragement throughout my education. This achievement would not have been achievable without them.

# CHAPTER 1

# INTRODUCTION

The Coronavirus (COVID-19) syndrome was reported to be a global pandemic by the World Health Organization (WHO) on March 11, 2020 [1]. It is a virus that spreads by breathing in the air close to an infected person, or through the droplet infection when someone coughs or sneezes.[2]. By March 2021, COVID-19 has caused about 131 million cases of infections and approximately 3 million deaths in the world. The total number of infections and fatalities in the United States was about 33 million and about 600.000 respectively by the same date. While many studies have been reported on the various of influences of COVID-19 on public health prevention measures, there are a few findings of the impact of the pandemic on people's attitude. [1]. For instance, research has improved our knowledge about the spread of virus, self-protection and the effectiveness of different vaccines. Similarly , there are several studies that focus on the role of protective tools like face masks and social distancing in reducing the spread of COVID-19. Pharmaceutical research has also advanced substantially. For instance, three COVID-19 vaccines, Moderna and Janssen (Johnson Johnson) and Pfizer have received authorization in the US, and 155,884,601 people, about 47% of the population in the United States have completed a vaccination series by March 2021 [2]. While many people got pharmaceutical treatment, individual protective measures, such as face masks and practicing social distancing, remain an essential behavior for decreasing the spread of the Coronavirus. The advantages of treatments and facial masks can only be obtained

[1] www.covid19.who.int
[2] www.usafacts.org

however when the majority of the population attend to be vaccinated or wear masks. Hence, it is essential to understand the public's viewpoints for and against vaccination and wearing masks. This is where social media can provide an insight into opinions. Since people share their beliefs and opinions on social media, such posts can be investigated using text mining for such insights. Text mining is a technique to investigate and analyze sentiments [3].

There are two sources to examine public health opinions, traditional surveys [4], and social media where people share their attitudes about disease outbreaks [5, 6]. Due to the ease in the accessibility to huge amount of online social network data, different studies leverage such a data to evaluate their proposed approach, conduct experiments and analyze people's attitudes toward a specific topic [7, 8]. At the current pandemic, people discuss related topics to COVID-19 vaccines and protective manners in social media. Hence, users will be led to vaccine and mask hesitancy and negative trends when they are exposed to wrong information [9]. According to the World Health Organization (WHO) [3], vaccine hesitancy is named as one of the 10 main threats to global health [10]. For stopping COVID-19 transmission, Jeremy Howard et al. [11] recommend wearing face masks, as an efficient form of source control and reduce the transmissibility of COVID-19 and the number of death. Moreover, misinformation will damage trust in public health authorities and lead to decline in vaccine comprehension [12, 13]

In this study, we use Twitter dataset due to its popularity as a social media for discussions and explanations of opinions related to health information [14]. after cleaning the dataset (340 million Tweets), we obtain sentiments and opinions of 359,906 openly available vaccine-related Twitter posts with 351,589 U.S based tweets and 114,967 mask-related tweets with 105,494 U.S. based tweets during four months from December 2020 to March 2021. We essentially concentrate on aspects of sentiments about vaccination and wearing masks. We are interested in analyzing the change of sensation and opinions shared on Twitter in different periods about vaccination as a COVID-19 treatment and the commands of respecting social distance and wearing masks for decreasing virus transmission. It is also interesting for us to investigate people's attitudes toward selected

[3] www.who.int

words "masks", and "vaccine". We restrict our analysis to "mask" and " vaccine" words as the most frequent hashtags and their sentiments distribution and changes among all 50 states of the United States. Our analysis show that tweets with mask and vaccine hashtags have mostly positive sentiments indicating people's inclination toward receiving vaccine and wearing face covering during the pandemic.

The sections of the study are organized as follows. In chapter 2, we explain samples of conducted related studies to our article. In chapter 3, we present the dataset and pre-processing steps such as cleaning the data and using topic modeling to find the highest frequency words . In chapter4, we state the study plan and the sentiment analysis methods, which include the procedure of text mining and detecting sentiments by applying the Vader method on our dataset. In this chapter, we also represent the results based on tweets shared in the united state and the frequency of different sentiments and their changes during periods and find the reasons for emotions shifting. Finally, we conclude our study in chapter 5.

# CHAPTER 2

# BACKGROUND

During last decades, we observe numerous studies focused on different aspects of social network analysis from causal perspective [15, 16] to visualization [17, 18] and polarity analysis [19]. Here, we review related work on attitude analysis in social media and COVID-19 data, as well as researches on the sentiment of Tweets about COVID-19 vaccination and masks.

## 2.1 Sentiment analysis in social media

Some studies focus on evaluating different features of COVID-19 in social media, including sentiment dynamics ShahriareSatu et al. [19] design an intelligent clustering-based classification and topic extracting model named TClustVID to obtain important sentiments with high efficiency by analyzing tweets that are related to COVID-19.They use different natural language processing techniques to propose a COVID-19 tweet analytical model to extract important topics from Twitter datasets. Sentiment analysis of COVID-19 Twitter posts found that most of the people, in spite of being emphasized and under lockdown, appreciated the efforts of their governments and heroes like the health workers, and police staff [20]. To classify sentiments on tweets in the more articulated class of emotional strength (weakly positive/negative, strongly positive/negative). Nemes and Kiss [21] propose a model to analyze the emotional nature of different tweets by using the

Recurrent Neural Network(RNN) for emotional prediction and Natural Language Processing to conclude and analyze the sentiments and signs (comments, hashtags, posts, tweets) of the users of the Twitter social media. There are many types of research in the sentiment analysis domain using Machine Learning and Deep learning models[22] uses deep learning models to propose an achievement for sentiment classification. It applies NLP for topic modeling to find the problems related to the Covid-19 dataset collected from social media. They use LSTM Recurrent Neural Networks (LSTM RNN) model for stance classification.In another study Sanders et al. [23] work on a database for over 1 million tweets for five months from March to July in 2020 (march -July)to evaluate public attitude regarding wearing the mask during the COVID-19 pandemic.They apply NLP, clustering, and sentiment analysis methods to determine that the frequency of mask-related positive tweets has increased.

## 2.2 Covid-19 analysis on Social media

Except for stance analysis in social media, there are many studies focus on evaluating other aspects of COVID-19 through social media, such as mental health. Gao et al. [7] measure high outbreak of mental health problems and examines their connection with social media.In another study, Valdez et al. [24] work on English language US tweets collected from a public database obtain the correlation between using social media and outbreak of mental issues during the COVID-19 pandemic. This study concludes that heavy usage of social media may further increase negative feelings in the long term for many users. using Latent Dirichlet allocation (LDA) model they find the topic of the tweets to analyze the hashtags and deploy Valence Aware Dictionary and Sentiment Reasoner (VADER) tool to analyze all timeline tweets.

Wahbeh et al. [25] analyze COVID-19–related tweets posted by medical experts and test their content to examine the gathered data to recognize relevant tweets concepts about the pandemic. He uses machine learning methods and text analysis to identify topics and ideas.In a similar study, Hussain, Amir et al. [26] use natural language processing (NLP) and deep learning models to predict average sentiments and sentiment trends.

## 2.3 Vaccination in Covid-19

Exploring the effect of Vaccination as one of the most successful interventions on preventing the covid-19 disease [27] has been central to many studies like Vaccine hesitancy[28] that presents significant and dangerous effects on health areas all over the world [29, 30].

Thelwall et al. [31] Recognize the types of vaccine hesitancy information shared on Twitter. They analyzed the tweets and obtain the reasons for vaccination hesitancy which is addressing by misleading attitudes and distributing fear on Twitter as a popular social media platform. Regardless of a general belief that points to a strong relationship between liberals and anti-vaccination attitudes in the United States, Baumgaertner et al. [32] conclude that there is strong evidence that mentions conservative people tend toward this trend more. J. Hornsey et al. [33] express that anti-vaccination tweets by President Trump increased their supporters' anxiety about vaccination. He is known as the first US president that declares anti-vaccination attitudes. Some other studies observe shifts in people's viewpoints about vaccination. Erika Bonnevie et al. [34] estimate that vaccination opposition on Twitter by analyzing Tweets related to vaccine opposition. The study shows that the rate of shifted ideas against vaccination increased. They estimate that vaccine opponents are promoting opposition toward a COVID-19 vaccine and promoting distrust in health authorities. There are some more studies in this area [35] for example, H Dodd et al. [36] work on investigation of concerns and motivations about COVID-19 vaccine to the reasons of being vaccinated. Jeffrey V. Lazarus et al. [37] assess the determination of potential acceptance rates and factors that impact on accepting COVID-19 vaccine by apply logistic regression on their dataset. Malik Sallam [38] presents an up-to-date evaluation of COVID-19 vaccination acceptance rates all over the world.In more studies , Goyen.G To et al. [39] work on the Twitter dataset during the COVID-19 pandemic and recognize anti-vaccination tweets by assessing the performance of different natural language processing (NLP) models on Twitter posts and Garay et al. [40] analyze the stance and perception of the anti-vaccine in the social media platform by using K-means clustering algorithm and Vader to find the sentiment of each tweet.There are also Various conducted studies about analyses of vaccine-related Tweets to evaluate people's positions towards vaccination [41], and to obtain and recognize sentiments [9], dominant sentiments[42], and the relationship between Twitter users and most influential users in approaching specific opinions [43].

## 2.4 Mask in Covid-19

Sentiment and topic modeling analysis has been employed in many aspects of COVID-19 to increase the research domains in this phenomenon. While many scientific types of research have improved our awareness of the disease, vaccines are licensed and available for all people especially in the United States, Despite all gains in this area, It still remains crucial to have individual protecting manners, such as wearing the facial mask and following social distancing rules, for lessening the spread of COVID19 [1]. Steffen E.Eikenberry et al. [44] develop a model for evaluating the impact of masks in general by using COVID-19 dataset based on New York and Washington. They approached that face masks may strikingly reduce transmission of COVID-19. Many researchers like A. Davies et al. [45] propose that masks may both protect the people who wear a mask from getting infections. Abraham C. Sanders et al. [46] apply NLP, and sentiment analysis techniques to mask-wearing tweets. They find that topic clustering and visualization-based from relevant data reporting improvement in people perceptions about COVID-19 and its prevention. They also mention that the polarity of mask-related tweets has increased. In another study, J Wu et al. [47] work on SARS patients as a controlled case study and find that wearing masks is strongly protective. On the other hand, some people have beliefs in opposition to the use of facial masks. He.Lu et al. [48] propose classifiers to classify related tweets, and analyze the dataset to understand the reason for those opposite trends of wearing the mask. They observe physical discomfort, lack of effectiveness, and belief of wearing the mask is inappropriate for certain people as the common reasons for the anti-mask trend [48]. Moreover, Qihuang Zhang et al. [49] employ the Vader and NRC models, to estimate the sentiment polarity scores and visualize the data in the pandemic period. They analyze sentiments of tweets based on mask, vaccine, and lockdown. They propose connections between cases who are infected, the relevant tweets, and the sentiment scores of tweets that are related to the Corona

[1]https://covid19.who.int/

Virus.There is proof that a massive number of people about 88% support wearing masks, They believe that wearing a mask should be a compulsory in public areas [50]. Qihuang Zhang et al.indicate that while people have a positive attitude about COVID-19 and masks, they have negative opinions about vaccines and lockdown. There are also some studies to determine the reasons perception changes in mask-related tweets. For example, Xueting Wang et al. [51] present a study that aims to examine public sentiments toward COVID-19 on social media. They employ Vader to analyze the variances in sentiment changes between California and New York or Jun Lang et al. [52] work on the stance of tweets and the patterns of sentiment changes about mask-wearing tweets that are posted in the United States. They present results that show the reason for the predomination of positive sentiments in comparison with anti-mask views.

# CHAPTER 3

# PREPROCESSING

## 3.1 Data

We track online posts on social media regarding COVID-19 for December 2020 January, February , and March 2021.Our primary focus is Twitter entries (tweets). The collected full dataset includes more than three hundred and forty million tweets.There is a large amount of data created per day. Machine learning methods can learn and predict important data from the dataset[53]. The main problem is that the efficacy of learning Machine learning methods can be influenced by a poor quality dataset[54]. As a result, primary data need to be pre-processed before being used with Machine learning models.

## 3.2 Data collection

We collect tweets from December 2020 to March 2021 by using Twitter streaming API to serve particular keywords and accounts that are shared on Twitter. We also use COVID-19 IDs from the repository of the university of southern California that generated and saved the tweets with Twitter's standard search API [55]. The dataset contains more than 340 million tweets from December first of 2020 to the last day of March 2021, which is constructed of about two terabytes of dehydrated raw data.The dataset includes

89detailed tweet attributes such as "ID", "geo", "text",created-at", and etc. The collected tweets carry keywords related to COVID-19 including "Vaccine","Mask","SARS-2","pandemic" ,and etc in a multi-layered selection process. First, we apply for a Twitter developer account to have access to Tweets APIs. Then we generated Tweets with API, and due to make a perfect dataset, we obtain the IDs of tweets from December 01, 2020, to March 31, 2021, that contain COVID-19-related keywords from Chen et al.'s Github repository of tweet IDs [55]. In next step, we use the Tweet Data Retrieving tool (Twarc) Morstatter et al [56] to gather the information and attributes of tweet closely related to these tweet IDs, covering the tweet content (text, Language,Hashtags, etc) and authors' metainformation from these tweet-ids. We use attribute "lang" and "location" to filter this dataset for English language tweets from the United States and obtain approximately 83 million tweets. The dataset contains three models of posting a tweet: (i) retweets (ii) reply(iii) No engagement. Each of these models is used for different purposes. A reply is an answer to a tweet, usually posted by another user, while it is possible to self reply. A retweet is an action that allows other users to re-share another tweet that will appear without any modifications. It is usually posted by other users while it's possible for users to retweet their own tweets and Finally, No engagement is a tweet that is not retweeted and get no reply.In this section, we describe the process of selection and annotation of the dataset.

Table 3.1: Size of dataset based on three categories

| Category | December | January | February | March |
|---|---|---|---|---|
| The dataset | 88,397,381 | 104,968,412 | 58,589,009 | 89,109,196 |
| Number of tweets that received at least one reply | 9,071,851 | 668,301 | 188,655 | 8,392,465 |
| Number of tweets that have been retweeted | 3,615,244 | 3,867,611 | 2,519,934 | 3,388,767 |
| No-engagement | 16,715,808 | 14,668,071 | 10,370,683 | 12,314,591 |

## 3.3 Methodology

In the first step we filter all tweets by "reply-count" ,"retweet-count" features to categorized our data in three categories (i) Retweet (ii) Reply(iii) No engagement.To have a clean dataset, We eliminate some attributes which are less important and reduce the number of features from 89 to 24 features.Then we obtain the top 20 hashtags on each tweet based on the frequency of all hashtags in each category. we use the 20 hashtags because some of the high frequencies hashtags are the same like: COVID-19, coronavirus, pandemic,Mask, and vaccination.

TABLE 3.2: The Frequency of top 10 Hashtags for January

| RetweetHashtags | Frequency | Reply Hashtags | Frequency | No-engagement Hashtags | frequency |
|---|---|---|---|---|---|
| COVID19 | 224055 | COVID19 | 6045 | COVID19 | 419041 |
| coronavirus | 41016 | coronavirus | 925 | coronavirus | 106743 |
| lockdown | 28813 | Covid19 | 625 | lockdown | 72953 |
| Covid19 ) | 26606 | covid19 | 522 | covid19 | 70451 |
| covid19 | 22882 | COVID19. | 411 | covid | 57200 |
| COVID | 18667 | COVID | 378 | Covid19 | 49789 |
| vaccine | 13697 | lockdown | 376 | WearAMask | 37889 |
| StayHome | 13611 | Covid | 320 | pandemic | 34923 |
| covid | 13142 | WearAMask | 273 | vaccine | 33041 |
| pandemic ) | 13005 | vaccine | 248 | CovidVaccine | 22652 |

TABLE 3.3: The Frequency of top 10 Hashtags for February

| RetweetHashtags | Frequency | Reply Hashtags | Frequency | No-engagement Hashtags | frequency |
|---|---|---|---|---|---|
| COVID19 | 145873 | COVID19 | 716 | COVID19 | 249080 |
| coronavirus | 22872 | coronavirus | 93 | coronavirus | 61750 |
| lockdown | 13358 | Covid19 | 70 | covid19 | 43753 |
| Covid19 | 16648 | Covid19 | 60 | covid | 36631 |
| COVID | 12276 | lockdown | 34 | lockdown | 33789 |
| vaccine | 9818 | WearAMask | 31 | Covid19 | 49789 |
| pandemic | 9031 | vaccine | 21 | COVID | 28239 |
| WearAMask | 6424 | COVID19: | 19 | pandemic | 23714 |
| Covid | 6399 | COVID-19 | 17 | vaccine | 22963 |
| CovidVaccine | 3638 | covid | 17 | WearAMask | 18420 |

In the second step, we work on the text of each tweet by employing topic model analysis to incorporate tweets text to appropriate topic and find keywords for the next analyses. For selecting the keywords, We change all words to lower case, then eliminate punctuation, stop words, and URLs. We apply topic modeling to obtain the different topics of COVID-19 tweets and identify the keywords. We employ Latent Dirichlet Allocation (LDA) [57] using Python's scikit learn library on all datasets to extract 10 topics, which results in the top 30 relevant terms with their frequencies in our dataset, and the number of each word being represented in a topic.

By this model, we investigate that vaccine and mask are two of the most frequent words in the most frequent topics in all three categories in each month. For example: For Vaccine word, (i)tweets that got at least one reply: Overall term frequency is about

10.000 in January and about 3000 in February. The estimated term frequency within the selected topic is about 2000 in both months.

(ii) retweeted tweets: Overall term frequency is about 100.000 in January and about 70.000 in February. The estimated term frequency within the selected topic is about 90.000 and 30.000 each month respectively.

(iii)tweets that have not been retweeted or received a reply: Overall term frequency is about 120.000 in January and about 100.000 in February. The estimated term frequency within the selected topic is about 30.000 and 20.000 in each month respectively.

Considering the high-frequency hashtags and words in LDA model we work on "vaccine" and "mask" hashtags for more analysis. We use regular expressions to recognize all hashtags in the text from all tweets. Hashtags usually come with a prefixed  symbol related to the topic correlated with the tweet. We store all hashtags in our dataset and extract all hashtags related to "vaccine" and "mask" to work for further analysis. Then we filter the English dataset for tweets that include at least one of the relevant hashtags to the vaccine. For filtering process , we use more than 4000 relevant hashtags to vaccine. For example: "#vaccine","#vaccination","#covid19vaccine","#Modernavaccine" and etc. This dataset includes 359,906 tweets in all three categories from December 2020 to March 2021.

For analyzing tweets related to hashtag "mask", we filter the English dataset for tweets that contain at least one hashtag related to the "mask". For filtering process, we use 3756 mask-related hashtags. For example: "#wearamask","#facemask","#covidmask","#maskup" ,and etc. This dataset includes 114,967 tweets in all three categories from December 2020to March 2021.

TABLE 3.4: The Number of Tweets in all 3 categories with "vaccine" hashtag

| Month | No-engaged tweets | Tweets got reply | Retweeted tweets |
|---|---|---|---|
| December 2020 | 66,288 | 927 | 23,845 |
| January 2021 | 70,858 | 2,197 | 31,411 |
| February 2021 | 44,435 | 565 | 72,971 |
| March 2021 | 97,268 | 929 | 30,212 |

TABLE 3.5: The Number of Tweets in all 3 categories with hashtag "mask"

| Month | No-engaged tweets | Tweets got reply | Retweeted tweets |
|---|---|---|---|
| December 2020 | 22,911 | 203 | 6,278 |
| January 2021 | 29,832 | 479 | 5,845 |
| February 2021 | 21,192 | 122 | 4,012 |
| March 2021 | 18424 | 155 | 30,514 |

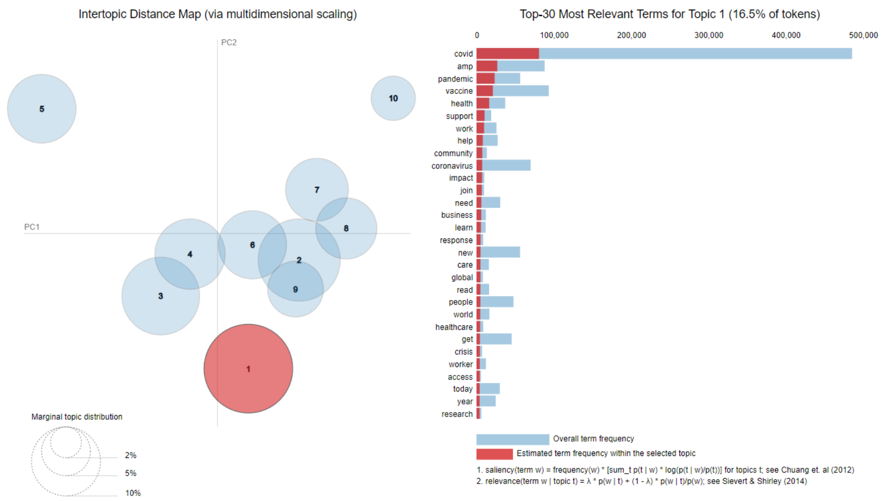


FIGURE 3.1: LDA of retweeted tweets for January

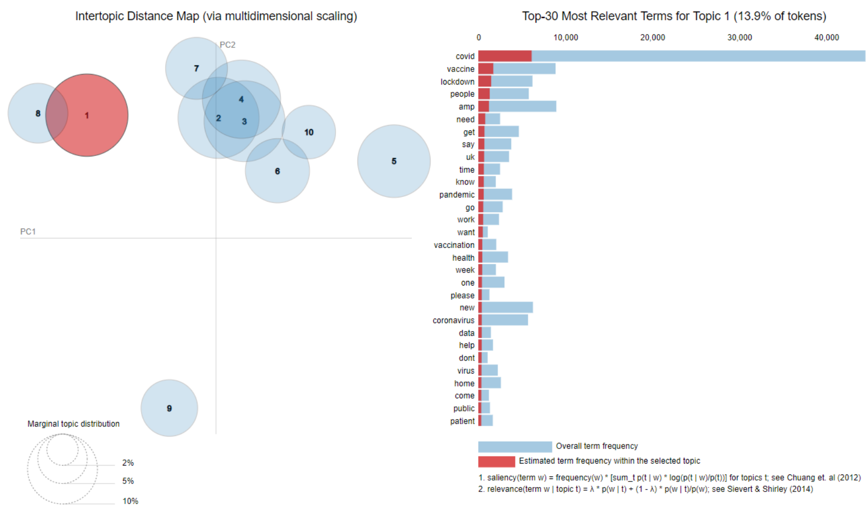


FIGURE 3.2: LDA of tweets that received at least one reply for January

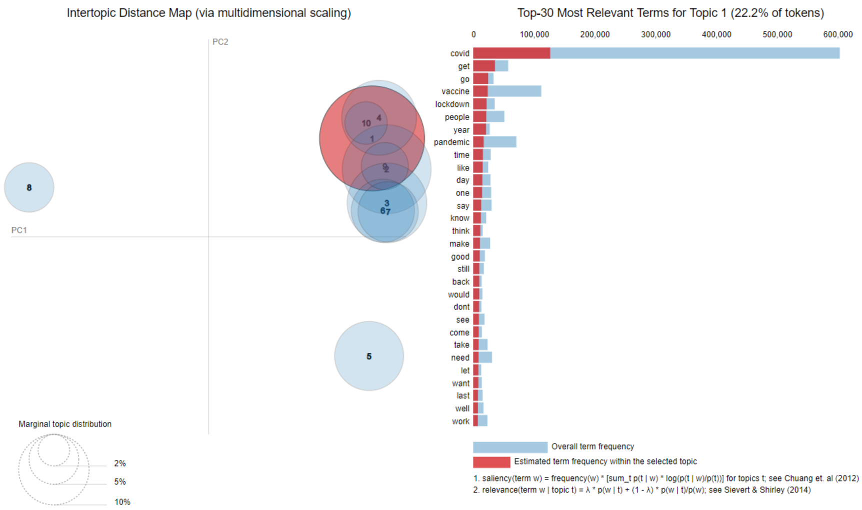


FIGURE 3.3: LDA of tweets that have not been retweeted or received a reply for January

# CHAPTER 4

# SENTIMENT ANALYSIS

Sentiment analysis offers us an opportunity to monitor changing inclinations in users' viewpoints, especially in common topics based on COVID-19 tweets. The Valence Aware Dictionary and Sentiment Reasoner (VADER) model is a lexicon and rule-based sentiment analysis tool for processing text and predicting sentiment [58]. VADER is designed to find the sentiments of shared posts on social media based on their texts and words [59]. In this study, we apply Vader to search and assign the polarity of 'positive', 'negative', or 'neutral' to each tweet. We start Preprocessing on the text of each tweet to drop useless characters and words such as punctuations, links, etc. VADER outputs a normalized value from [-1, 1]. Then we assign sentiments based on compound scores given by VADER. Any tweet with a score of -0.05 or greater describes a positive stance, a score of 0.05 or less describes a negative stance, and any score between those values describes neutral sentiment.

## 4.1 Evolution of sentiments

There are three types of ideas in vaccine-related Twitter(vaccine hesitancy, pro-vaccine, and vaccine opposition). Unlike pro-vaccine, Vaccine opposition ideas refer to beliefs against vaccination, and vaccine hesitancy points to uncertain opinions about receiving vaccines. There are also the same types of feelings based on Tweets related to masks. Pro mask, Anti-mask, and mask hesitancy with positive, negative, and doubtful viewpoints, respectively. Unlike Anti-mask ideas, pro-mask refer to beliefs toward wearing masks while Anti-mask indicates opinions against wearing the mask.

Two bar charts Figure 4.1 and Figure 4.8 represent the monthly distribution pattern of positive, negative, and Neutral sentiments in all three categories to estimate the sentiment toward vaccination and wearing masks respectively among all vaccine and mask related tweets in English-speaking countries such as: United States, Australia, India, the United Kingdom, Canada, and Ireland.The tweets from these countries contain a significant proportion of the total number of tweets that we collect in the study. The figures vividly illustrate the frequency of positive, negative, and neutral sentiments based on three categories in each month.

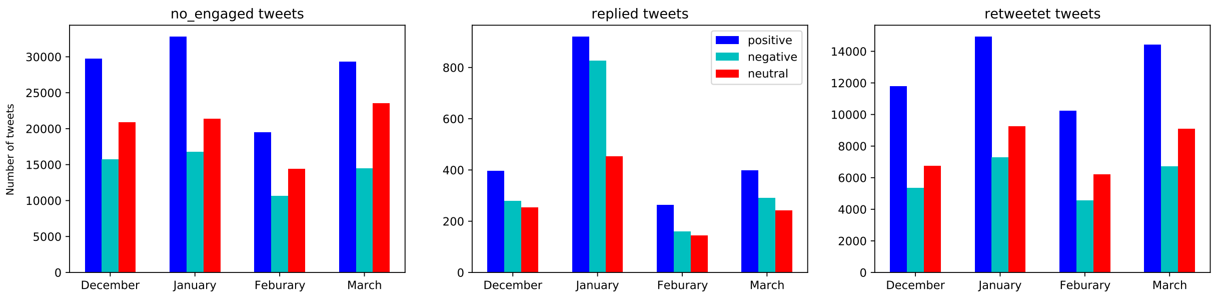


Figure 4.1: Sentiment distribution for hashtag vaccine

## 4.2 Vaccine sentiment analysis

We use VADER to categorize the tweets into three sections of positive, negative, and neutral. As shown in Figure 4.1, the frequency of positive tweets gains higher than tweets with negative and neutral sentiments. The distribution of tweets indicates tweets sentiments for and against vaccination and vaccine hesitancy in all three categories from December to March. The bar chart displays that the vast majority of tweets are in favor of vaccination by the frequency of 164,511 positive stances compare to the frequency of negative sentiments with 82,969 tweets.

Figure 4.2 shows the percentage of tweets in each sentiment category. According to the graph, 46.8% of the vaccine-related tweets are positive sentiments followed by the neutral and negative categories with 29.6% and 23.6% of the tweets, respectively. The negative sentiments are related to a range of concerns about vaccination and caused by vaccine opposition viewpoints such as doubts in vaccine safety and effects, political aspects, and manufacturers. On the other hand, positive tweets are usually about scientific researches, medical advice, and hope matters.

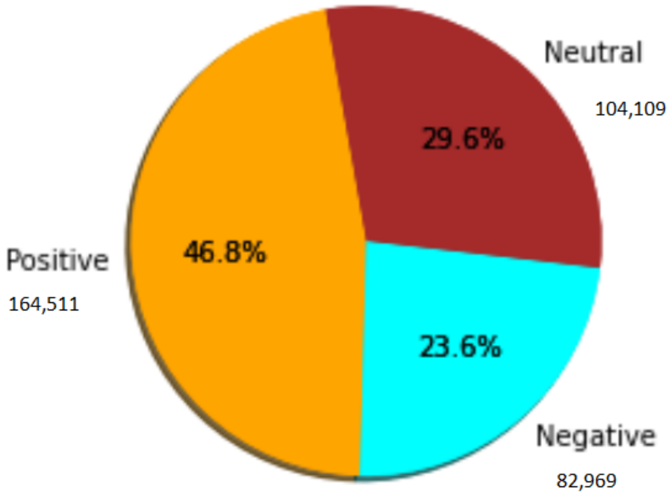


FIGURE 4.2: Sentiment polarities of tweets for hashtag vaccine

## 4.3 Sentiment polarities of tweets for hashtag vaccine in United States

The first case of COVID-19 in the United States was reported on January 19, 2020[60]. All 50 U.S states reported cases by mid of March 2019. [61] For presenting the sentiments based on population, we examine tweets from the United States and exclude tweets with a location outside of the United States. We filter the vaccine dataset based on the United States location and obtain the sentiment frequency of tweets in each state. California, the most populous and the third-largest U.S. state by area, contains the most positive frequency among all 50 states by 2,577 positive stances and followed by New York and Washington DC with 1,801, and 1,379 positive sentiments, respectively. On the other hand, for Anti-vaccine opinions, a similar allocation is consistent with differences in frequency. Wyoming state as the least populated U.S. state with a population of 576,412 people contains the lowest sentiment frequency with only 11 positive sentiments and 3 negative sentiments.

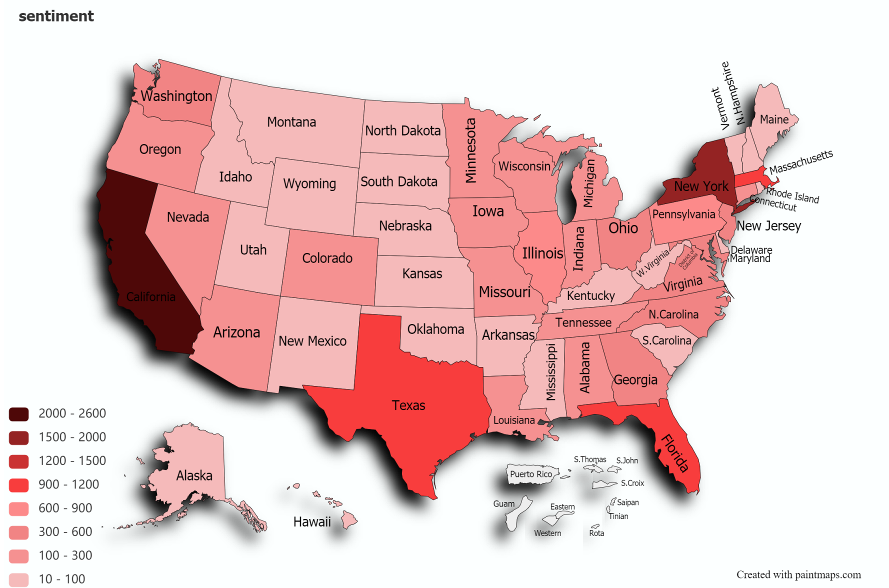


FIGURE 4.3: Map of positive sentiment distribution in the United States

We use Map Maker[1] to have a colorful map of all states of the United States. Figure 4.3 displays a map of positive sentiment distribution scores seen in each state of the United States. The states with darker colors contain the highest distribution of positive sentiments. Conversely, the states with lighter colors show less positive sentiment distribution. California with 14.73%, New York with 10.69%, Dc with 7.48%, and Massachusetts with 6.2% are among states with the most positive sentiment.In addition, these states are in the top 10 democratic states [2]

Figure 4.5 illustrates the map of negative sentiment distribution in each state of the United States. The darker the color of states indicates the highest negative sentiments frequency. In opposite, the lighter the color of state shows the lowest negative sentiments frequency. California with 17.45%, DC with 7.53%, and Florida with 6.31% are among the states with the most anti-vaccine inclinations. Moreover, Florida is a state under Republican control[3].

[1]www.paintmaps.com
[2]www.worldpopulationreview.com
[3]www.atr.org

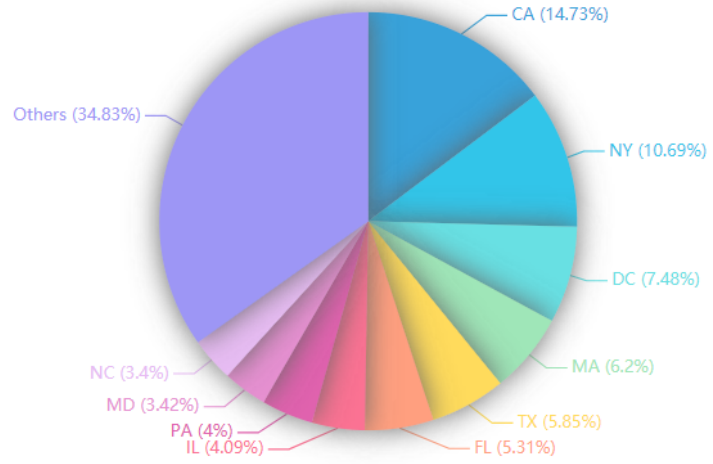


FIGURE 4.4: percentage of positive sentiment distribution in the United States

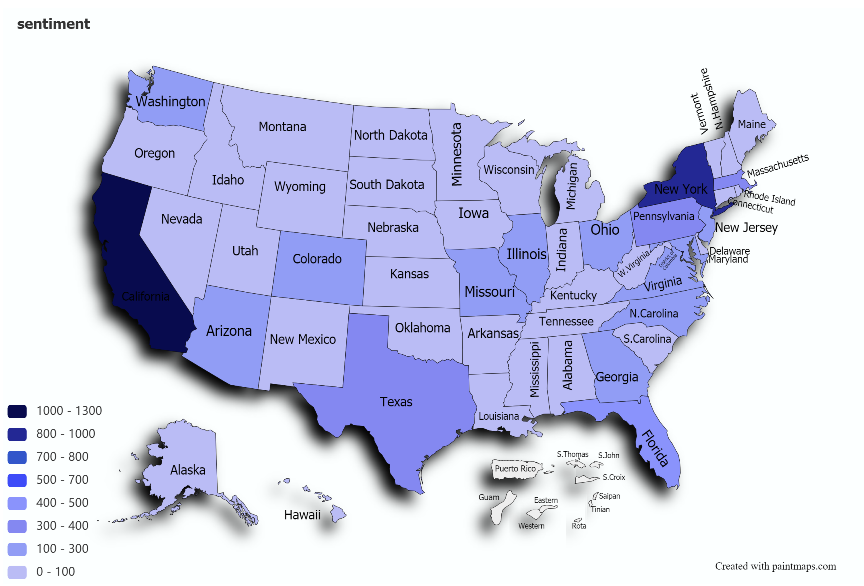


FIGURE 4.5: Map of negative sentiment distribution in the United States

Our results present information for the problems in geographical regions with lower levels of COVID-19 vaccine comprehension that causes vaccine hesitancy and anti-vaccination beliefs. The system of controlling the state (Democrat/Republic), religion, and education have effects on anti-vaccination messages [62, 63].

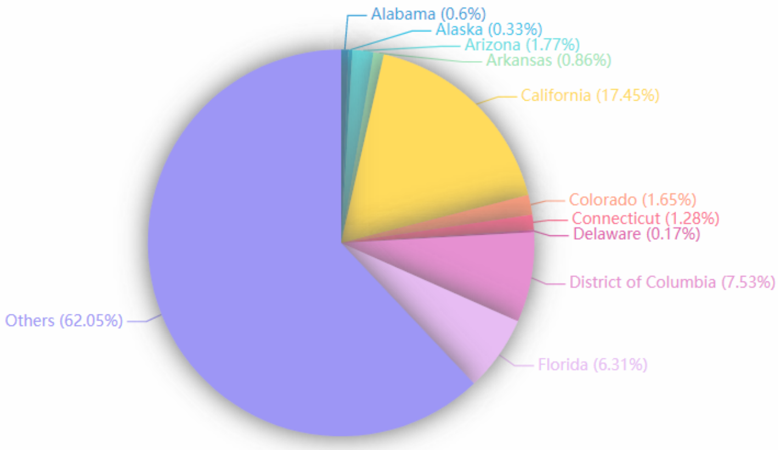


FIGURE 4.6: percentage Negative sentiment distribution in the United States

TABLE 4.1: Top 10 states in positive sentiment

| State | Positive | Negative | Neutral |
|---|---|---|---|
| California | 2577 | 1217 | 1531 |
| New York | 1871 | 848 | 1181 |
| Washington DC | 1309 | 525 | 672 |
| Massachusetts | 1085 | 324 | 485 |
| Texas | 1023 | 391 | 695 |
| Florida | 929 | 440 | 629 |
| Illinois | 715 | 240 | 461 |
| Pennsylvania | 700 | 384 | 444 |
| Maryland | 598 | 127 | 255 |
| North Carolina | 594 | 206 | 399 |

## 4.4 Trends of vaccine-related sentiments in COVID-19 Tweets

The US is now (early July 2021) in an excellent situation globally, where it has more vaccines available for people who are enthusiastic about receiving. New reports have indicated that overall, 155,884,601 people or 47% of the population in the United State got their second vaccine doses, and they are fully vaccinated[4]. figure 4.7 reveals that the patterns of sentiments and beliefs have changed in response to vaccine-related results during four months. In general, the positive stance about the COVID-19 vaccine is the dominant polarity on Twitter. The strong positive sentiment contains tweets that include operations related to vaccination posts that shared on Twitter such as: "What you're looking at here is essentially increasing the vaccination rate for the city from, well, by

[4]https://usafacts.org

TABLE 4.2: Top 10 states with Negative sentiment

| State | Positive | Negative | Neutral |
|---|---|---|---|
| California | 2577 | 1217 | 1531 |
| New York | 1871 | 848 | 1181 |
| Washington DC | 1309 | 525 | 672 |
| Florida | 929 | 440 | 629 |
| Texas | 1023 | 391 | 695 |
| Pennsylvania | 700 | 384 | 444 |
| Massachusetts | 1085 | 324 | 485 |
| Illinois | 715 | 240 | 461 |
| North Carolina | 594 | 206 | 399 |
| Georgia | 474 | 198 | 275 |

about 50% to 70%, which is very, very significant" , " COVID19 vaccine will be given free ".

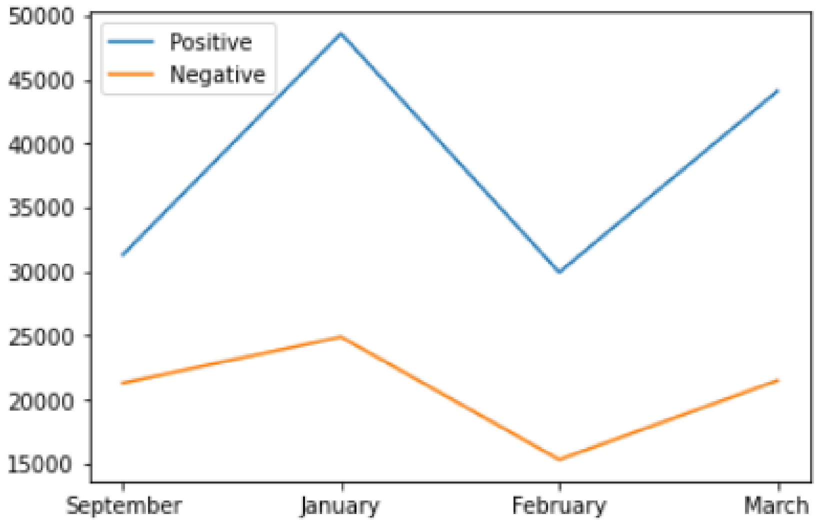


FIGURE 4.7: sentiment pattrens in the United States

Surprisingly, by the variation in the portion of positive tweets, we recognize that positive public sentiment erosion from early Jan to late Feb 2021. We attempt to identify mass perceptions by working on tweets texts to recognize the reasons that may affect users' feelings and change positive sentiments into neutral or negative territory.

For example: "This pretty 32-year-old doctor has suffered sudden inflammation of the brain leading to convulsions and breathing difficulties half an hour after receiving the Pfizer-BioNTech COVID-19 vaccine", "Not all of the COVID19 vaccines being tested have been successful and many have fallen by the wayside", " the COVID19 vaccine doesn't stop you spreading or contracting the "virus", and "vaccines/meds may have unexpected sideeffects not seen in the trials. It's ESSENTIAL to share info of these so we can learn amp; adapt to save lives/reduce risk! ". According to the examples, misinformation effects on losing the enthusiasm or confidence of receiving vaccines which may change the users' opinion from positive to negative and finally reduce the number of positive sentiments. Publishing reports of logical and positive results from governments and validated organizations can be quite helpful to prevent the decay of positive public perceptions about vaccination. In addition, the vaccine must be shown to the public to be tested several times to be authorized and announced [64]. To increase the motivation of receiving vaccines, people need to be convinced about the effectiveness of the vaccine and how strict the conducted clinical trials have been.

## 4.5 Vaccine Sentiment analyzing per IDs

In this step, we work on Negative and positive sentiments per IDs, and the sentiments shift to determine how effective the Twitter discussions related to COVID-19 vaccination are and the pattern of sentiment changes during 4months. We filter the dataset based on id and their sentiments and obtain the number of shifts and how the trends have changed. We gain 256,079 ids with Positive, Negative, and Neutral sentiments.

- *Sentiment Shifts in General*:

  In this step, we filter all tweets with at least one sentiment and calculate the Frequency distribution of sentiment changes, regardless of positive to negative or vice versa, to identify the number of ids and their opinions changes. We gain 256,079 ids with at least one sentiment(Positive, Negative, Neutral). There are 5,318 ids with at least one shift from negative to positive views compare with 12,495 ids with at least one exchange from positive to opposite opinions.

- *Positive to Negative sentiment shifts* :

  Our sentiment scoring method declares that the number of negative sentiments increased from early February to late March in 2021. But it is not alarming because of the growth of the portion of positive sentiment across the same periods.

Table 4.3: Number of sentiment changes pr ID

| | Positive to Negative per ID | Negative to Positive per ID |
|---|---|---|
| No change | 16,582 | 23,758 |
| 1-50 | 12,460 | 5,301 |
| 51-100 | 22 | 15 |
| 101-150 | 6 | 1 |
| 151-200 | 3 | 0 |
| 201-250 | 1 | 0 |
| 251-300 | 0 | 1 |
| 301-350 | 0 | 1 |
| 351-400 | 1 | 0 |
| 401-450 | 0 | 1 |
| 450-500 | 0 | 0 |
| 584 | 1 | 0 |
| 598 | 1 | 0 |
| 3670 | 1 | 0 |

In this step, We obtain that 16,582 IDs that have positive sentiments have not changed their views while 12,460 IDs change their opinions at least for one time until 50 times. In addition, some specific IDs have too many sentiments changes, for instance, 3,670 changes by one Twitter id. These Ids point to the tweets shared with Celebrities and Organizations. There are many factors and features that cause shifts of sentiment from positive to negative, and those shifts are not necessarily reflected in the loss of trust in vaccines. Another reason that has effects on this opinion shifting is fake and inaccurate information. For example: "Beware of inaccurate information says @WHO", "So... Americans who are *so afraid* to get the Chinese Virus are, let me get this straight... going to intentionally infect themselves with the cov? HEALTHY PEOPLE SHOULD NOT GET VACCINATED" , "So what's the Government gonna do when people die from the mandatory COVID19 vaccine? Vaccines are safe for the vast majority, not for all." , and "There seems to be confusion regarding the efficacy of vaccine for COVID19. Frankly it's foolhardy to expect a magic potion so soon. Scientists, research bodies and governments are trying. Meanwhile the onus is on the individual. Remember even god helps those who help themselves". There could also be some irregular emphasis on some examples of people falling ill after taking the first or both COVID-19 vaccine dose/s or deviate results such as deaths after taking the vaccine. Descriptive tweets convinced other people to change their minds about vaccination:"Vaccine notice: I'm feeling the pain in my body especially my arm, a headache, and got fever a day after the first shot of vaccine." , "I would never receive the vaccine if I knew it would put me in sick.","after the second dose of vaccine, I couldn't even walk for a day".While there is a common concern about some anxiety posts, most people seem to agree to be vaccinated. , as demonstrated by tweets such as: "second vaccine dose made me sick. but I am still so thankful and hopeful " . Possibly, a significant challenge has appeared when some news and information were shared by the gossip of people who are affected by vaccinations or even died because of the vaccine. "This is serious. Production of this vaccine must be stopped. The risk of serious reactions, infertility in women or death is far too high in relation to the extremely low risk of death from the virus itself. Proceeding with the rollout of this vaccine is insanity. ","Many asking if Pfizer vaccine on COVID19 is safe. No. It is not. Only fools rush in, 'advisory' agreement only amp; still won't immunise you from getting COVID or side effects. Trail data has been hidden by @BorisJohnson amp; pfizer, trails had to be stopped, amp; had deaths. So then...". Such fearful and hopeless posts spread fast on social media and impact every person's viewpoint about vaccination.

- *Negative to Positive sentiment shifts* :

  Our result illustrates that positive sentiments are dominant from early September 2020 to late March 2021. In this step, We gain that 23,758 IDs have negative sentiments without any shifts in contrast with 5,301IDs that change their sentiments from negative to positive at least once to 50 times. Moreover, It is shown in our result that many sentiment changes are done by some definite users with their unique ids that point to celebrities and Organizations' Twitter accounts.

  The number of positive sentiments and the pattern of its growth demonstrate that people are agreed with vaccination and are willing to take the vaccine. Several factors may impact sentiment switching from negative to positive. We observe positive beliefs towards the COVID-19 vaccine. Tweets with positive opinions spread hope and illustrate positive perceptions derived from the high confidence and trust that people have about vaccines and their effectiveness. For example: "The fastest way to end the COVID19 pandemic is to make safe and effective vaccines available to everyone on the planet. Tell pharma to pool all knowledge, intellectual property amp; data together for the benefit of all. The fastest way to end the COVID19 pandemic is to make safe and effective" , "Much hope has been placed in the COVID19 vaccine being the â€˜silver bulletâ€™ to the pandemic and getting back to normal...vaccines available to everyone on the planet". The process of making vaccines accessible among people and sharing valid positive news from validated sources are other two reasons that may affect increasing the portion of positive sentiments by changing negative views to positive. For instance: "Those who couldnâ€™t get registered on CoWIN amp; are not able to access the Internet can go to any centre amp; opt for walk-in vaccination facility. They just have to carry their identity proof with them","Vaccines in US are free and do not need your Social Security number" , "The Covid-19 vaccination program is making great progress, but scammers are doing the best to take money from people. The vaccine is free. Ignore any requests which say you need to hand over payment information in order to receive it".

## 4.6 Mask sentiment analysis

The coronavirus has caused an unusual public health disaster all over the world. The strong awareness-raising of COVID-19 has made various titles and discussions like considering the usage of facial masks as a necessary individual protective manner for safety and quite spreading the virus. SE Eikenberry et al [44] proved that facial masks have effects on death caused by COVID-19. They explained that wearing masks reduces the mortality rate by more than 20% by the condition of being worn by more than 80% of the general public. So the benefits of facial masks can only be achieved when most people wear them [65]. Wearing facial masks in the US became a major topic while some countries such as Singapore, South Korea, and China have reached this goal[66].

In our study, We use VADER to identify and categorize related tweets, followed by a content sentiment analysis of a subset of the tweets to determine the number of sentiment changes for and against mask-wearing. We provide three categories of positive, negative, and neutral sentiment . Figure 4.8, illustrates that tweets with neutral sentiments earned higher rates than the other two categories of Negative and Positive sentiments. It displays tweets sentiments for, against, and neutral ideas about wearing the mask in all three categories from December 2020 to March 2021. Figure 4.8 demonstrates that a large number of tweets have neutral sentiments about wearing facial masks by the frequency of 45,398 stances and followed by positive sentiments with 41,128 sentiments and negative sentiments by 20,059 tweets.

The higher rate of Neutral stance distribution than Negative sentiments shows that people are more doubtful than being unsatisfied about wearing face masks.

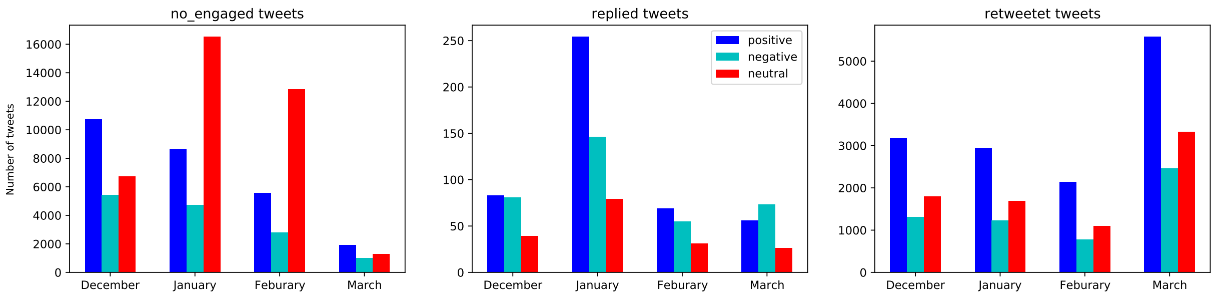


FIGURE 4.8: Sentiment distribution for hashtag mask

Figure 4.9 vividly demonstrates the percentage of tweets in each sentiment section. The neutral sentiments are 42.6% of the tweets followed by the positive and negative feelings with 38.6% and 18.8% of the tweets, respectively. The negative opinions are related to concerns of individual protective manners like wearing face masks caused by mask opposition viewpoints such as doubts in wearing facial masks, their safety, and effects attributes. On the contrary, tweets with positive sentiments are normally based on scientific researches, medical suggestions, and hope matters to stop spreading the virus, especially in public areas. According to Figure4.8, positive sentiments are dominant in all categories except in the no-engagement(tweets that got no reply and no retweet) dataset in January and February.

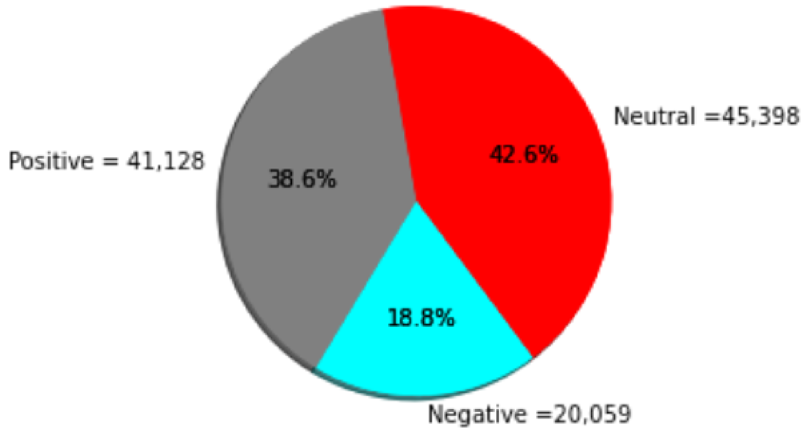


Figure 4.9: Sentiment polarities of tweets for hashtag mask

## 4.7 Sentiment polarities of tweets for hashtag mask in United States

Unlike other measures to fight against Coronavirus, wearing face masks has remained uncertain [67]. At the moment that some countries such as China commanded wearing face masks, the US and UK were in doubt of accepting and using this plan to decrease the rate of virus transmission [66, 68]. Finally, In early April, federal officials modified their guidelines and commanded that people in all states need to wear masks to reduce the COVID-19 transmission because some people are asymptomatic transmitters. Moreover, based on the analysis of COVID-19 cases in Beijing, the chance of virus transmission was 79% lower in families that at least one member had worn a mask, contrasted to the families that were on the opposite side of using face masks and none of them used it [69].

To gain the sentiments based on community, we analyze tweets from the United States and eliminate tweets from other locations. We clean the mask dataset derived from the United States and collect the Sentiment repetition of tweets in each state.

According to our results, California includes the most positive frequency among all 50 states by 1,706 positive stances and followed by New York and Texas with individually 660 and 569 positive sentiments. On the other hand, we obtain similar states at the top for negative sentiments. (California by 587, Texas by 272, and New York with 250 negative sentiments). Wyoming state includes the lowest positive sentiment frequency

Table 4.4: Top 10 states with positive sentiment

| State | Positive | Negative | Neutral |
|---|---|---|---|
| California | 1706 | 587 | 981 |
| New York | 660 | 250 | 380 |
| Texas | 569 | 272 | 362 |
| Florida | 349 | 143 | 182 |
| Massachusetts | 283 | 144 | 131 |
| Georgia | 263 | 86 | 145 |
| Washington | 254 | 102 | 87 |
| Pennsylvania | 252 | 85 | 207 |
| Ohio | 241 | 51 | 79 |
| Illinois | 228 | 93 | 129 |

Table 4.5: Top 10 states with Negative sentiment

| State | Positive | Negative | Neutral |
|---|---|---|---|
| California | 1706 | 587 | 981 |
| Texas | 569 | 272 | 362 |
| New York | 660 | 250 | 380 |
| Massachusetts | 283 | 144 | 131 |
| Florida | 349 | 143 | 182 |
| Washington DC | 225 | 116 | 116 |
| Washington | 254 | 102 | 87 |
| Arizona | 179 | 93 | 144 |
| Illinois | 228 | 93 | 129 |
| North Carolina | 178 | 88 | 111 |

among other U.S states. It contains only two stances in positive and negative opinions. In addition, Alaska, the largest U.S. state by the area( Larger than the total area of the three largest states Texas, California, and Montana), contains four positive sentiments and one sentiment on the opposite side.

We provide a colorful map from the U.S States by Map Maker to illustrate the distribution of positive and negative sentiments in all 50 states. Figure 4.10 shows a map of favorable sentiment allocation in each state of the United States. The states with the darker green color include the highest distribution of positive sentiments. Conversely, lighter colors show the states where have less positive opinions distribution in using face masks.

Among the 50 states, California with 22.49%, followed by New York and Texas with more than 16% of all distribution of positive sentiments, have shown as the states with the most pro-masks trends. Unlike the Democratic-led states, the Republican States were less likely to follow the health order of wearing face masks. By the first order of wearing

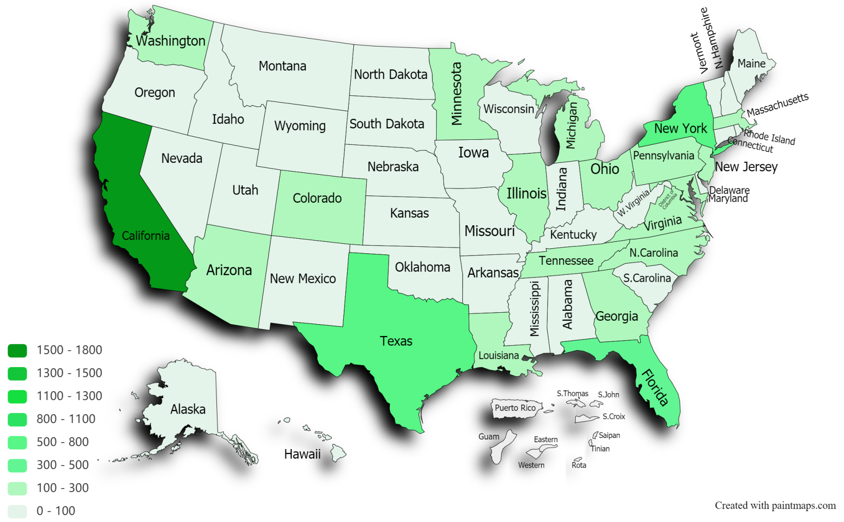


FIGURE 4.10: positive sentiment distribution in the United States in mask

facial masks, some states, such as Arizona, Georgia, and Texas, tried to prevent localized health orders requiring face masks, but later the governors of those states changed their minds and stances to fight with COVID-19[5]. Our outcomes in Figure 4.10 and Figure 4.11 display information about geographical regions and their trends in mask-related comprehension.

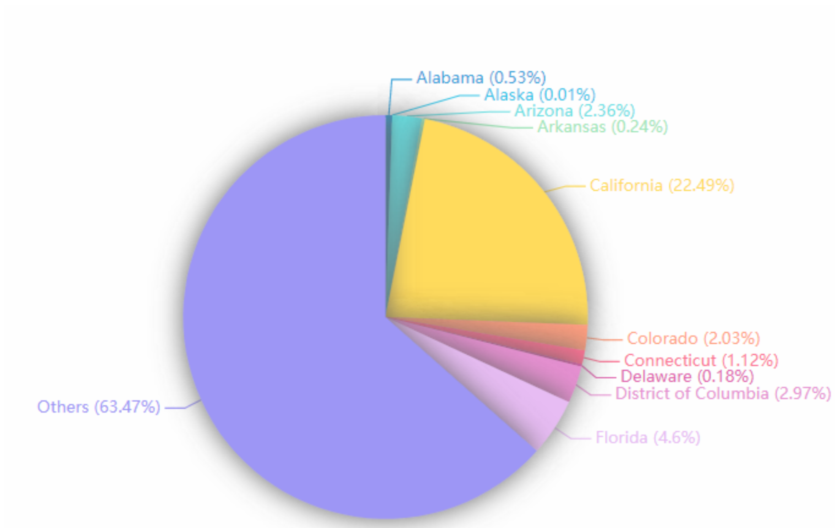


FIGURE 4.11: percentage of positive sentiment distribution in the United States based on wearing mask

[5]https://www.washingtonpost.com/

We also perform the same process to obtain negative sentiments distribution map. Figure 4.12 demonstrates Anti-mask sentiments distribution in each state of the United States. The darker the yellow color of states means the highest negative sentiments frequency. On the other hand, the lighter the color of state points to the lowest Anti-mask sentiments frequency. Depending on a Pew Research survey established in fall 2020, 19% of Republicans consider masks and the act of wearing masks as a pandemic-related hardship, and 27% were doubtful about wearing masks and the hardness of the COVID-19 pandemic while comparing to Democrats 31% declared concerns about politicizing the safety commands and rules. The rests were those who do not care about the pandemic. For example, a republican responded to the survey that “The entire unnecessary shutdown of the country got my husband furloughed for 9 weeks, more government overreach with mask orders, people are just so terrified to live it’s disgusting, so the ones of us like me who aren’t scared get treated like we are awful people” or a democrat declared his opinion “Customers complaining about masks and not wearing them are both the biggest personal issues COVID has caused. The maskless customers are usually ruder than other people”[6].


FIGURE 4.12: Anti-Mask distribution in the United States

The three states of California with 19.37% followed by Texas and New York totally with 18% are the states with the most anti-mask inclinations among all 50 U.S states. Our

[6]https://www.pewresearch.org/

results represent knowledge for the problem of geographical regions with lower levels of COVID19 wearing mask comprehension that causes anti-vaccination beliefs. The system of state government states, physical discomfort and the negative effects of wearing masks have impacts on anti-mask messages[48].

Interestingly, based on our figures, it is vividly demonstrated that California and New York are two states located at the top three states with the most frequency in both negative and positive sentiments, while Alaska and Wyoming are two other U.S states that contain the lowest rate in both negative and positive stances. The reasons for these results would refer to their population, geographic location, and the number of educated, religious people in each state.

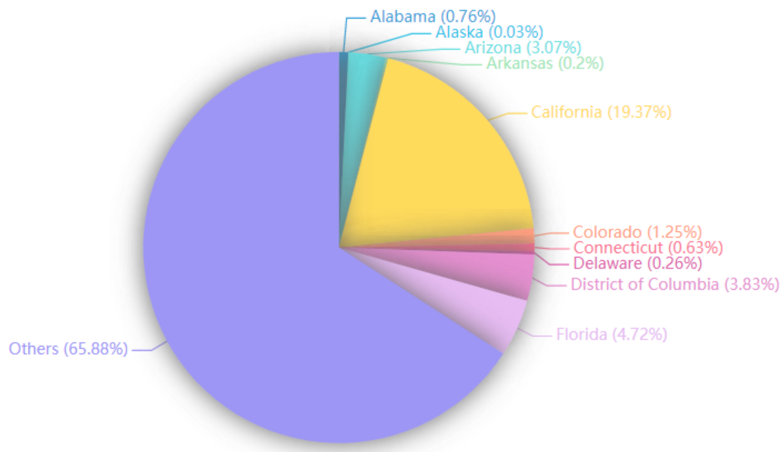


FIGURE 4.13: percentage Anti-Mask distribution in the United States

## 4.8 Trends of mask-related sentiments in COVID-19 Tweets

The United state now (early July 2021) stays in a wonderful position in contrast with other countries, where 47% of the population in the United State are fully vaccinated and wearing mask mandates are coming down over the U.S. Based on the Centers for Disease Control (CDC) announcement, fully vaccinated people no longer need to wear masks or follow the social distance regulation[7].

[7]https://www.nbcnews.com/

Figure 4.14 describes that the patterns of sentiments and opinions have changed in mask-related tweets during four months. In general, the positive sentiments about wearing masks are the dominant polarity on Twitter. The strong favorable trends contain tweets that include viewpoints related to mask posts such as: "CDC recommends the use of cloth face coverings to supplement social distancing in fight against spread of covid19". The Neutral sentiments as the highest rate of sentiments after positive ones include posts like " Has to be fake. I really doubt facial masks work well . can trust CDC?!"

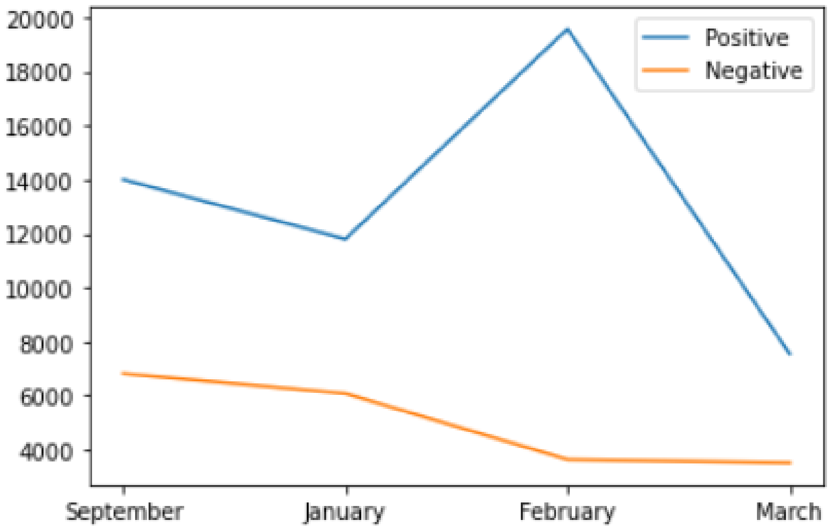


FIGURE 4.14: Mask sentiments pattern in the United States

As Figure 4.14 shows, we realize that positive public sentiment decreases from early September 2020 to late January 2021 and from February to March after a sharp rise from January to late February. We try to recognize the reasons for such perceptions by working on tweets text contents that would have led such a volume of positive opinions into neutral or negative territory. For example: "So why would masking help with Covid? Well... it doesnt! who knows???" , " I cannot breath well while I'm wearing mask" , " we are immuned by defult. No need to wear mask or get vaccine" , " *important notification* masks have more opposite insights that positive point. wearing mask brings out more diseases in a long time of usage."

Some specific factors affect the changes in the number of positive sentiments to negative or neutral stances. For example, a wide range of wrong information and physical problems like uncomfortably breathing would drop confidence and increase uncertainty among people about wearing masks [48]. Moreover, the attitudes and manners of famous people about Wearing face masks can change their followers' ideas. For instance: President Trump displayed opposition about wearing masks in public media [8], and did not order to use face masks at his gatherings and other public campaign events for the 2020 presidential election [9]. Sharing articles of logical and positive consequences from governments and validated organizations working on the COVID-19 area can be beneficial to prevent the failure of positive public perceptions and motivate people to use facial masks in at least public areas.

## 4.9 Mask Sentiment analyzing per IDs

In this step, we work on sentiments per IDs, and sentiments changes on both sides of positive and negative to define how the opinions shared in Twitter to COVID-19 about wearing masks have changed during 4months. We filter the dataset based on ID feature and their sentiments to gain the number of changes and how their trends have changed.

- *Sentiment Shifts in General*:

  In this step, we filter all tweets with at least one sentiment and calculate the total number of sentiment changes without paying attention to specific shifts from positive to negative or vice versa. We try to identify the number of ids and their sentiments changes. We gain a whole number of 106,585 ids with at least one sentiment (Positive, Negative, Neutral). There are 4,762 ids with at least one shift from positive to negative and 2,217 ids with at least one change from negative to positive.

- *Positive to Negative sentiment shifts* :

  Our sentiment estimation method indicates that the number of positive sentiments decreased two times once from early December to late January in 2021 and from the first day of February until the end of March after sharp growth in January. It is not alarming due to two reasons. First, the graph shows the decay in the number of

[8] www.time.com , www.npr.org
[9] www.azcentral.com , www.politico.com

Table 4.6: Number of sentiment changes pr ID

| | Positive to Negative per ID | Negative to Positive per ID |
|---|---|---|
| No change | 5,564 | 8,109 |
| 1-50 | 4,721 | 2,211 |
| 51-100 | 32 | 2 |
| 101-150 | 3 | 1 |
| 151-200 | 3 | 1 |
| 201-250 | 3 | 0 |
| 251-300 | 0 | 0 |
| 319 | 0 | 1 |
| 489 | 0 | 1 |

negative sentiments across the same periods. Second, Figure 4.14 displays that the general opinions are toward wearing masks and dominant polarity on Twitter. In this step, We gain 5,564 IDs that have positive sentiment have not switched their beliefs, while 4,721 IDs exchange their opinions at least one to fifty times. Moreover, three IDs have between 200 to 250 sentiment shifts. These IDs belong to celebrities and COVID-related organization's Twitter accounts. The change of sentiment from positive to negative can be caused by many factors and features such as religion, beliefs, physical issues like specific allergies, and adverse reactions. Other reasons that affect the sentiment changes refer to wrong information shared by users and the manner of famous people against wearing masks like the U.S. Surgeon General mentioned in his tweet that "Seriously people- STOP BUYING MASKS! They are NOT effective in preventing the general public from catching Coronavirus, but if healthcare providers can't get them to care for sick patients, it puts them and our communities at risk! ". More examples in anti-mask opinions: "Masks kill people, breathing in your own carbon monoxide is poison.and it weakens your immune system" No Christian should be living in fear ... only fear of God","Even N95 masks don't make any difference at all, just like all randomized studies show" ,"Yesterday I heard this old lady where I live say she wears a mask everywhere in public. She said that her c0vid19 test is positive. how would it have happened? I never wear mask. I do believe that it doesn't have prevention effects","Cannot get my head round how stupid people are. This poor lady is hospital with covid pregnant. It sucks. But she STILL FUCKING believes that wearing masks work even though she clearly wears hers and yet still caught covid".

- *Negative to Positive sentiment shifts* :

  Our results demonstrate that positive sentiments are dominant in all four months, while the negative stances are decreased gradually from December 2020 to March 2021. In this part, We obtain that 8,109 IDs have negative sentiments without any changes in contrast with 2,211 IDs that change their sentiments from negative to positive less than 50 times. Furthermore, two unique IDs have 319 and 489 sentiment changes that point to celebrities and Covid-Related Companies. The pattern of positive sentiments growth and the decay in the number of Anti-mask trends express that people are willing to use protective tools against COVID-19 and tend to quite the virus transmission.

  Various factors may influence shifting sentiment from negative to positive. We recognize positive sentiment towards the COVID-19 mask. Most tweets with positive sentiments contain the fear of getting the Coronavirus with its negative sides and many viewpoints about who supports Anti-mask. For example:

  "I had three patients this morning who said they were not going to wear masks " , "please wear your mask when ordering/interacting with drive thru workers. They are frontline and deserve protection, as do you. " "Face mask are no longer required in Texas as a main Anti-mask state. Now we're all gonna know how ugly people are. I'll still be wearing mine..." , "Unless you are a service worker and have to deal with people who have no consideration for your safety... their rights are more important than your health .They are really stupid" , "Anti maskers act like people who want to wear mask wanted to wear one. Like no, we just donâ€™t want to be sick or be responsible for getting someone else sick ."Trump supporters that were against wearing masks and went to many Trump rallies without respecting protective manners are the main source of spreading coronavirus. All of sudden he is taking the virus seriously because it is affecting him." , "Those anti-maskers will kill over someone impeding on their freedom to spread Covid. BackTheBlue anti-maskers CopKilled Man who killed New Orleans officer was denied entry into game for not wearing mask" , "Most idiots I see think it's all over...no masks, chatting in the street, no social distancing... Sleepwalking into another lockdown and thousands more deaths". Another reason that changed many beliefs about wearing mask into positive sides are the reaction of celebrities and academic researchers about wearing masks, for example, Bernie Sanders: "I am once again asking you to wear a mask." JOHN W. LUNDIN(Lawyer, Historian, Author) : "Wearing masks is not a political statement, it is an IQ test".

## 4.10 Limitation

There are some limitations in using the Twitter dataset. Twitter expresses public interaction and does not represent all populations of the United States. Active users in Twitter are only 15% of adults between 18 to 29 years old that tend to participate in discussions in this online social media than the overall population. About 16.5% of the American population was 65 years old or over 2021 and would not use social media[10]. In addition, active and passive users are more widespread than users who post randomly [70]. So the specific sentiments can be biased[71] thus, careful analysis and predictions are needed. Moreover, Twitter bots are programmed and employed to copy, share and spread the content of specific topics daily. These kinds of tweets could be tricks and fakes and can have significant effects on the perceptions of the real Twitter users[72].

[10]https://www.statista.com/

# CHAPTER 5

# CONCLUSION

The Coronavirus causes many difficulties to the public and has significant effects on every individual daily life. It is essential to work on public attitudes in reacting to all measures against the COVID-19 pandemic and assess its impacts on mental health. In this study, we conduct sentiment analysis of the tweets in all 50 states of the united state on the most frequent topics related to COVID-19 including, face masks and vaccines. We apply topic modeling and Vader methods to obtain the most frequent words and the sentiments of tweets relevant tweets.

Our visualization of sentiment scores presents information about the users' reactions to the virus ,vaccination and wear masks . The statistics and figures illustrate the trend of public and individual sentiments and their changes in different periods from December 2020 to March 2021. Our results show the difference in the frequency of positive, negative, and neutral stances in both vaccination and wearing masks, with the dominant polarity of positive sentiments. Regardless of the overall strength of positive opinions, negative feelings about COVID-19 vaccines and wearing masks remain among people besides vaccination hesitancy.

By reading the text of tweets and another papers related to the reason of perceptions and sentiments, The most probable the main reasons for sentiment changes appears to be the belief that such tweets are providing accurate information and also because of the COVID-19 as discussed by well-regarded organizations such as CDC increased the user's knowledge about negative insights of COVID-19 and led them to be more careful about their health. At the same time, however, inaccurate information, mistrust of well-regarded organizations or media, and the impact of celebrities on their followers does push a segment of users into hesitancy and negative viewpoints about masks and vaccination.

# BIBLIOGRAPHY

[1] Domenico Cucinotta and Maurizio Vanelli. Who declares covid-19 a pandemic. *Acta Bio Medica: Atenei Parmensis*, 91(1):157, 2020.

[2] Thushara Galbadage, Brent M Peterson, and Richard S Gunasekera. Does covid-19 spread through droplets alone? *Frontiers in public health*, 8:163, 2020.

[3] Ah-Hwee Tan et al. Text mining: The state of the art and the challenges. In *Proceedings of the pakdd 1999 workshop on knowledge disocovery from advanced databases*, volume 8, pages 65–70. Citeseer, 1999.

[4] Patrick Peretti-Watel, Valérie Seror, Sébastien Cortaredona, Odile Launay, Jocelyn Raude, Pierrea Verger, Lisa Fressard, François Beck, Stéphane Legleye, Olivier l'Haridon, et al. A future vaccination campaign against covid-19 at risk of vaccine hesitancy and politicisation. *The Lancet Infectious Diseases*, 20(7):769–770, 2020.

[5] Andrew M Guess, Brendan Nyhan, Zachary O'Keeffe, and Jason Reifler. The sources and correlates of exposure to vaccine-related (mis) information online. *Vaccine*, 38 (49):7799–7805, 2020.

[6] Edward Velasco, Tumacha Agheneza, Kerstin Denecke, Goeran Kirchner, and Tim Eckmanns. Social media and internet-based data in global systems for public health surveillance: a systematic review. *The Milbank Quarterly*, 92(1):7–33, 2014.

[7] Junling Gao, Pinpin Zheng, Yingnan Jia, Hao Chen, Yimeng Mao, Suhong Chen, Yi Wang, Hua Fu, and Junming Dai. Mental health problems and social media exposure during covid-19 outbreak. *Plos one*, 15(4):e0231924, 2020.

[8] Zahra Fatemi and Elena Zheleva. Network experiment design for estimating direct treatment effects. In *KDD Workshop on Mining and Learning with Graphs(MLG)*, 08 2020.

[9] Hilary Piedrahita-Valdés, Diego Piedrahita-Castillo, Javier Bermejo-Higuera, Patricia Guillem-Saiz, Juan Ramón Bermejo-Higuera, Javier Guillem-Saiz, Juan Antonio Sicilia-Montalvo, and Francisco Machío-Regidor. Vaccine hesitancy on social media: Sentiment analysis from june 2011 to april 2019. *Vaccines*, 9(1):28, 2021.

[10] Florian Kunneman, Mattijs Lambooij, Albert Wong, Antal Van Den Bosch, and Liesbeth Mollema. Monitoring stance towards vaccination in twitter messages. *BMC medical informatics and decision making*, 20(1):1–14, 2020.

[11] Jeremy Howard, Austin Huang, Zhiyuan Li, Zeynep Tufekci, Vladimir Zdimal, Helene-Mari van der Westhuizen, Arne von Delft, Amy Price, Lex Fridman, Lei-Han Tang, et al. Face masks against covid-19: an evidence review. 2020.

[12] Maryke S Steffens, Adam G Dunn, Kerrie E Wiley, and Julie Leask. How organisations promoting vaccination respond to misinformation on social media: a qualitative investigation. *BMC public health*, 19(1):1–12, 2019.

[13] Peter J Hotez, Tasmiah Nuzhath, and Brian Colwell. Combating vaccine hesitancy and other 21st century social determinants in the global fight against measles. *Current opinion in virology*, 41:1–7, 2020.

[14] Hanjia Lyu, Junda Wang, Wei Wu, Viet Duong, Xiyang Zhang, Timothy D Dye, and Jiebo Luo. Social media study of public opinions on potential covid-19 vaccines: Informing dissent, disparities, and dissemination. *arXiv preprint arXiv:2012.02165*, 2020.

[15] Zahra Fatemi and Elena Zheleva. Minimizing interference and selection bias in network experiment design. In Munmun De Choudhury, Rumi Chunara, Aron Culotta, and Brooke Foucault Welles, editors, *Proceedings of the Fourteenth International AAAI Conference on Web and Social Media, ICWSM 2020, Held Virtually, Original Venue: Atlanta, Georgia, USA, June 8-11, 2020*, pages 176–186. AAAI Press, 2020. URL https://aaai.org/ojs/index.php/ICWSM/article/view/7289.

[16] Martin Saveski, Jean Pouget-Abadie, Guillaume Saint-Jacques, Weitao Duan, Souvik Ghosh, Ya Xu, and Edoardo M. Airoldi. Detecting network effects: Randomizing over randomized experiments. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '17, page 1027–1035, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450348874. doi: 10.1145/3097983.3098192. URL https://doi.org/10.1145/3097983.3098192.

[17] Skye Bender-deMoll and Daniel A McFarland. The art and science of dynamic network visualization. *Journal of Social Structure*, 7(2):1–38, 2006.

[18] Zahra Fatemi, Mostafa Salehi, and Matteo Magnani. A generalized force-directed layout for multiplex sociograms. In Steffen Staab, Olessia Koltsova, and Dmitry I. Ignatov, editors, *Social Informatics - 10th International Conference, SocInfo 2018, St. Petersburg, Russia, September 25-28, 2018, Proceedings, Part I*, volume 11185 of *Lecture Notes in Computer Science*, pages 212–227. Springer, 2018. doi: 10.1007/978-3-030-01129-1_13. URL https://doi.org/10.1007/978-3-030-01129-1_13.

[19] Md Shahriare Satu, Md Imran Khan, Mufti Mahmud, Shahadat Uddin, Matthew A Summers, Julian MW Quinn, and Mohammad Ali Moni. Tclustvid: a novel machine learning classification model to investigate topics and sentiment in covid-19 tweets. *Knowledge-Based Systems*, page 107126, 2021.

[20] Muzafar Bhat, Monisa Qadri, Majid Kundroo Noor-ul Asrar Beg, Naffi Ahanger, and Basant Agarwal. Sentiment analysis of social media response on the covid19 outbreak. *Brain, Behavior, and Immunity*, 2020.

[21] László Nemes and Attila Kiss. Social media sentiment analysis based on covid-19. *Journal of Information and Telecommunication*, pages 1–15, 2020.

[22] Hamed Jelodar, Yongli Wang, Rita Orji, and Shucheng Huang. Deep sentiment classification and topic discovery on novel coronavirus or covid-19 online discussions: Nlp using lstm recurrent neural network approach. *IEEE Journal of Biomedical and Health Informatics*, 24(10):2733–2742, 2020.

[23] Abraham C Sanders, Rachael C White, Lauren S Severson, Rufeng Ma, Richard McQueen, Haniel C Alcântara Paulo, Yucheng Zhang, John S Erickson, and Kristin P Bennett. Unmasking the conversation on masks: Natural language processing for topical sentiment analysis of covid-19 twitter discourse. *medRxiv*, pages 2020–08, 2021.

[24] Danny Valdez, Marijn Ten Thij, Krishna Bathina, Lauren A Rutter, and Johan Bollen. Social media insights into us mental health during the covid-19 pandemic: Longitudinal analysis of twitter data. *Journal of medical Internet research*, 22(12): e21418, 2020.

[25] Abdullah Wahbeh, Tareq Nasralah, Mohammad Al-Ramahi, and Omar El-Gayar. Mining physicians' opinions on social media to obtain insights into covid-19: mixed methods analysis. *JMIR public health and surveillance*, 6(2):e19276, 2020.

[26] Amir Hussain, Ahsen Tahir, Zain Hussain, Zakariya Sheikh, Mandar Gogate, Kia Dashtipour, Azhar Ali, and Aziz Sheikh. Artificial intelligenceâ€" enabled analysis of public attitudes on facebook and twitter toward covid-19 vaccines in the united kingdom and the united states: Observational study. *Journal of Medical Internet Research*, 23(4), 2021.

[27] Francis E Andre, Robert Booy, Hans L Bock, John Clemens, Sibnarayan K Datta, Thekkekara J John, Bee W Lee, S Lolekha, Heikki Peltola, TA Ruff, et al. Vaccination greatly reduces disease, disability, death and inequity worldwide. *Bulletin of the World health organization*, 86:140–146, 2008.

[28] Malak Alsabban. Comparing two sentiment analysis approaches by understand the hesitancy to covid-19 vaccine based on twitter data in two cultures. In *13th ACM Web Science Conference 2021*, pages 143–144, 2021.

[29] Eugene J Gangarosa, Artur M Galazka, Cameron R Wolfe, Lynelle M Phillips, Elizabeth Miller, Robert T Chen, and RE Gangarosa. Impact of anti-vaccine movements on pertussis control: the untold story. *The Lancet*, 351(9099):356–361, 1998.

[30] LP Wong, PF Wong, and S AbuBakar. Vaccine hesitancy and the resurgence of vaccine preventable diseases: the way forward for malaysia, a southeast asian country. *Human vaccines & immunotherapeutics*, 16(7):1511–1520, 2020.

[31] Mike Thelwall, Kayvan Kousha, and Saheeda Thelwall. Covid-19 vaccine hesitancy on english-language twitter. *Profesional de la información (EPI)*, 30(2), 2021.

[32] Bert Baumgaertner, Juliet E Carlisle, and Florian Justwan. The influence of political ideology and trust on willingness to vaccinate. *PloS one*, 13(1):e0191728, 2018.

[33] Matthew J Hornsey, Matthew Finlayson, Gabrielle Chatwood, and Christopher T Begeny. Donald trump and vaccination: The effect of political identity, conspiracist ideation and presidential tweets on vaccine hesitancy. *Journal of Experimental Social Psychology*, 88:103947, 2020.

[34] Erika Bonnevie, Allison Gallegos-Jeffrey, Jaclyn Goldbarg, Brian Byrd, and Joseph Smyser. Quantifying the rise of vaccine opposition on twitter during the covid-19 pandemic. *Journal of Communication in Healthcare*, pages 1–8, 2020.

[35] Larissa G Malagoli, Julia Stancioli, Carlos HG Ferreira, Marisa Vasconcelos, Ana Paula Couto da Silva, and Jussara M Almeida. A look into covid-19 vaccination debate on twitter. In *13th ACM Web Science Conference 2021*, pages 225–233, 2021.

[36] Rachael H Dodd, Kristen Pickles, Brooke Nickel, Erin Cvejic, Julie Ayre, Carys Batcup, Carissa Bonner, Tessa Copp, Samuel Cornell, Thomas Dakin, et al. Concerns and motivations about covid-19 vaccination. *The Lancet. Infectious Diseases*, 21(2):161, 2021.

[37] Jeffrey V Lazarus, Scott C Ratzan, Adam Palayew, Lawrence O Gostin, Heidi J Larson, Kenneth Rabin, Spencer Kimball, and Ayman El-Mohandes. A global survey of potential acceptance of a covid-19 vaccine. *Nature medicine*, 27(2):225–228, 2021.

[38] Malik Sallam. Covid-19 vaccine hesitancy worldwide: a concise systematic review of vaccine acceptance rates. *Vaccines*, 9(2):160, 2021.

[39] Quyen G To, Kien G To, Van-Anh N Huynh, Nhung TQ Nguyen, Diep TN Ngo, Stephanie J Alley, Anh NQ Tran, Anh NP Tran, Ngan TT Pham, Thanh X Bui, et al. Applying machine learning to identify anti-vaccination tweets during the covid-19 pandemic. *International journal of environmental research and public health*, 18 (8):4069, 2021.

[40] J Garay, R Yap, and MJ Sabellano. An analysis on the insights of the anti-vaccine movement from social media posts using k-means clustering algorithm and vader sentiment analyzer. In *IOP Conference Series: Materials Science and Engineering*, volume 482, page 012043. IOP Publishing, 2019.

[41] Matthew DeVerna, Francesco Pierri, Bao Truong, John Bollenbacher, David Axelrod, Niklas Loynes, Cristopher Torres-Lugo, Kai-Cheng Yang, Fil Menczer, and John Bryden. Covaxxy: A global collection of english twitter posts about covid-19 vaccines. *arXiv e-prints*, pages arXiv–2101, 2021.

[42] Didi Surian, Dat Quoc Nguyen, Georgina Kennedy, Mark Johnson, Enrico Coiera, and Adam G Dunn. Characterizing twitter discussions about hpv vaccines using topic modeling and community detection. *Journal of medical Internet research*, 18 (8):e6045, 2016.

[43] Gloria J Kang, Sinclair R Ewing-Nelson, Lauren Mackey, James T Schlitt, Achla Marathe, Kaja M Abbas, and Samarth Swarup. Semantic network analysis of vaccine sentiment in online social media. *Vaccine*, 35(29):3621–3638, 2017.

[44] Steffen E Eikenberry, Marina Mancuso, Enahoro Iboi, Tin Phan, Keenan Eikenberry, Yang Kuang, Eric Kostelich, and Abba B Gumel. To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the covid-19 pandemic. *Infectious Disease Modelling*, 5:293–308, 2020.

[45] Anna Davies, Katy-Anne Thompson, Karthika Giri, George Kafatos, Jimmy Walker, and Allan Bennett. Testing the efficacy of homemade masks: would they protect in an influenza pandemic? *Disaster medicine and public health preparedness*, 7(4): 413–418, 2013.

[46] Abraham Sanders, Rachael White, Lauren Severson, Rufeng Ma, Richard McQueen, Haniel Campos Alcanatara Paulo, Yucheng Zhang, John S Erickson, and Kristin P Bennett. Unmasking the conversation on masks: Natural language processing for topical sentiment analysis of covid-19 twitter discourse. *medRxiv*, 2020.

[47] Jiang Wu, Fujie Xu, Weigong Zhou, Daniel R Feikin, Chang-Ying Lin, Xiong He, Zonghan Zhu, Wannian Liang, Daniel P Chin, and Anne Schuchat. Risk factors for sars among persons without known contact with sars patients, beijing, china. *Emerging infectious diseases*, 10(2):210, 2004.

[48] Lu He, Changyang He, Tera L Reynolds, Qiushi Bai, Yicong Huang, Chen Li, Kai Zheng, and Yunan Chen. Why do people oppose mask wearing? a comprehensive analysis of us tweets during the covid-19 pandemic. 2021.

[49] Qihuang Zhang, Grace Y Yi, Li-Pang Chen, and Wenqing He. Text mining and sentiment analysis of covid-19 tweets. *arXiv preprint arXiv:2106.15354*, 2021.

[50] Pew Research Center. Republicans, democrats move even further apart in coronavirus concerns, 2020.

[51] Xueting Wang, Canruo Zou, Zidian Xie, and Dongmei Li. Public opinions towards covid-19 in california and new york on twitter. *medRxiv*, 2020.

[52] Jun Lang, Wesley W Erickson, and Zhuo Jing-Schmidt. # maskon!# maskoff! digital polarization of mask-wearing in the united states during covid-19. *PloS one*, 16(4):e0250817, 2021.

[53] Nasser M Nasrabadi. Pattern recognition and machine learning. *Journal of electronic imaging*, 16(4):049901, 2007.

[54] Jan Van den Broeck, Solveig Argeseanu Cunningham, Roger Eeckels, and Kobus Herbst. Data cleaning: detecting, diagnosing, and editing data abnormalities. *PLoS Med*, 2(10):e267, 2005.

[55] Emily Chen, Kristina Lerman, and Emilio Ferrara. Tracking social media discourse about the covid-19 pandemic: Development of a public coronavirus twitter data set. *JMIR Public Health and Surveillance*, 6(2):e19273, 2020.

[56] Fred Morstatter, Jürgen Pfeffer, Huan Liu, and Kathleen Carley. Is the sample good enough? comparing data from twitter's streaming api with twitter's firehose. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 7, 2013.

[57] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *the Journal of machine Learning research*, 3:993–1022, 2003.

[58] Clayton Hutto and Eric Gilbert. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 8, 2014.

[59] Shihab Elbagir and Jing Yang. Twitter sentiment analysis using natural language toolkit and vader sentiment. In *Proceedings of the International MultiConference of Engineers and Computer Scientists*, volume 122, page 16, 2019.

[60] Michelle L Holshue, Chas DeBolt, Scott Lindquist, Kathy H Lofy, John Wiesman, Hollianne Bruce, Christopher Spitters, Keith Ericson, Sara Wilkerson, Ahmet Tural, et al. First case of 2019 novel coronavirus in the united states. *New England Journal of Medicine*, 2020.

[61] M Weiss, A Schwarzenberg, R Nelson, Karen M Sutter, and Michael D Sutherland. Global economic effects of covid-19. *Congressional Research Service*, 2020.

[62] Francesco Pierri, Brea Perry, Matthew R DeVerna, Kai-Cheng Yang, Alessandro Flammini, Filippo Menczer, and John Bryden. The impact of online misinformation on us covid-19 vaccinations. *arXiv preprint arXiv:2104.10635*, 2021.

[63] Nig J Pure. Appraisal of public opinions towards potential covid-19 vaccination in fct-abuja nigeria.

[64] Steven Taylor, Caeleigh A Landry, Michelle M Paluszek, Rosalind Groenewoud, Geoffrey S Rachor, and Gordon JG Asmundson. A proactive approach for managing covid-19: the importance of understanding the motivational roots of vaccination hesitancy for sars-cov2. *Frontiers in psychology*, 11:2890, 2020.

[65] Anthony Paulo Sunjaya and Christine Jenkins. Rationale for universal face masks in public against covid-19. *Respirology (Carlton, Vic.)*, 2020.

[66] Josh Katz, Margot Sanger-Katz, and Kevin Quealy. A detailed map of who is wearing masks in the us. *The New York Times*, (July 17), 2020.

[67] Jiao Wang, Lijun Pan, Song Tang, John S Ji, and Xiaoming Shi. Mask use during covid-19: A risk adjusted strategy. *Environmental Pollution*, 266:115099, 2020.

[68] Mark Egan, Amish Acharya, Viknesh Sounderajah, Yihan Xu, Abigail Mottershaw, Rosie Phillips, Hutan Ashrafian, and Ara Darzi. Evaluating the effect of infographics on public recall, sentiment and willingness to use face masks during the covid-19 pandemic: a randomised internet-based questionnaire study. *BMC public health*, 21 (1):1–10, 2021.

[69] Yu Wang, Huaiyu Tian, Li Zhang, Man Zhang, Dandan Guo, Wenting Wu, Xingxing Zhang, Ge Lin Kan, Lei Jia, Da Huo, et al. Reduction of secondary transmission of sars-cov-2 in households by face mask use, disinfection and social distancing: a cohort study in beijing, china. *BMJ global health*, 5(5):e002794, 2020.

[70] Man Hung, Evelyn Lauren, Eric S Hon, Wendy C Birmingham, Julie Xu, Sharon Su, Shirley D Hon, Jungweon Park, Peter Dang, and Martin S Lipsky. Social network analysis of covid-19 sentiments: Application of artificial intelligence. *Journal of medical Internet research*, 22(8):e22590, 2020.

[71] Ross Joseph Gore, Saikou Diallo, and Jose Padilla. You are what you tweet: connecting the geographic variation in america's obesity rate to twitter content. *PloS one*, 10(9):e0133505, 2015.

[72] Yusuf Mücahit Çetinkaya, İsmail Hakkı Toroslu, and Hasan Davulcu. Developing a twitter bot that can join a discussion using state-of-the-art architectures. *Social network analysis and mining*, 10(1):1–21, 2020.