HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

**English Studies Master’s Program**

**Module: Computational Literacy**

**Learner: Jacqueline Guadalupe Armijos Monar**

**Student number: 015396398**

**Period: II**

# Evaluative language attitudes of Covid-19 vaccine-related tweets from January 2020 to September 2022: A sentiment analysis and Topic modelling

## 1. Introduction

My class research topic is Evaluative language attitudes of Covid-19 vaccine-related tweets from January 2020 to September 2022: A sentiment analysis and Topic modelling. Although there are many studies about Covid-19 vaccines, there are still many concerns to solve and learn from, such as the disease itself and people's attitudes to the disease and to the vaccination program. Mellor (2022), for example, argues that we are still only in the first of five stages of the pandemic. Moreover, the World Health Organization, WHO (2021) states that the end of the coronavirus pandemic will only be when its symptoms are generally seen as similar to seasonal influenza. Therefore, it is an academic contribution from the lens of Digital humanities to generate and update new knowledge not only in the Computational Literacy field but also to influence governments on the creation of more effective policies and strategies when dealing with Covid-19 vaccine and its impact on public sentiment. The final results show to what extent people's attitudes are for or against Covid-19 vaccines by using a Sentiment analysis and topic modeling in a corpus from Covid vaccine Tweets as the Kaggle dataset. In this present analysis, I analyzed only one variable: text descriptions; meanwhile, dates allowed me to limit the scope of the study. Consequently, I excluded names of any geographical locations or unnecessary data to detect public sentiment.

## 2. Background

Two pressing concerns are apparent from previous studies on the covid-19 vaccine and public sentiment toward it on social media. One is infodemic/misinformation, and the other is the vague presence of official voices either or debunk or ratify information about the Covid-19 vaccine. The latter, the official voice, may increase vaccine campaigns that speed up herd immunity worldwide. Most researchers have used data from social networks such as Twitter, Parler, and Facebook mainly and employed a variety of social theories, and approaches (Social judgment theory, Extended parallel process model, Systematic functional linguistics or Appraisal analysis, and Semantic network analysis among others), corpora, and statistical measuring tools (Likelihood model, Data mining, Natural language processing, Weighted VADER, LDA model, cross-sectional studies, sentiment analysis, and so on) analyze and draw conclusions.

In a broad sense, interpersonal meaning or attitude analysis of comments on social media are divided into three categories: 1) The anti-vaccine movement and political/conspiracy theories against Covid-19 vaccination, 2) The pro-vaccine movement, and 3) Topics of concern waiting for scientific responses.

**2.1** Scannell et al. (2021) have studied how tweeters promoted anti-vaccine movements by employing persuasion techniques. The message contents focussed on safety, and the value of personal choice. The main purpose is to promote fear behavior; the anti-vaccine movement gets stronger by the presence of cyberattacks (social bots, trolls, and sock puppets), which creates an environment of perceived severity and perceived susceptibility among Twitter users, threatening human lives.

Culture also may determine behavioral norms. Another research paper by Luo, Chen  Cui, & Liao (2021) done on two social media platforms, Weibo in China and Twitter in the United States, clearly shows how users express their attitudes towards the Covid-19 vaccine. Particularly, Weibo users evidenced appreciation of authorities and official institutions for vaccinating people; their positive feelings towards the vaccine campaign are closely related to their collectivist culture. In other words, they concentrate on the aims of the group. Moreover, comments on Weibo are based on what national and international organizations WHO are mainly communicating to the public. Apart from what has already been reported, Chinese authorities have imposed strict control over Weibo content. On the other hand, Twitter users are mostly Americans, representing an individualist culture that prioritizes personal goals, self-reliability, and personal control over physical condition consequences. Therefore, Twitter users openly express their reluctant attitudes toward the Covid-19 vaccine. Another of her study carried out by Na, Cheng, Li, Lu, & Li (2021) reveals how celebrities influence public opinion and opinion fluctuations among Twitter users from the United States and the United Kingdom. 40% of negative public opinion in these sites is against Covid-19 vaccination in these countries. In many cases, Americans express concerns about several people dying after receiving their vaccination, most likely as side effects. The rest of the percentage of Facebook posts is divided among the health damage produced by the virus, vaccine distribution, vaccination to schooling communities, and manufacturing vaccine brand concerns.

It is still evident that Twitter users are highly unmotivated about the Covid-19 vaccine in general. Other research conducted by Cascini et al. (2022) reached almost the same conclusions. They collected 156 articles to assess the relationship between social media use and Covid-19 vaccine hesitancy through a cross-sectional method. The results concluded that specific dominant topics against Covid-19 vaccines occur. The study recognizes that social media specifies significant events as negative co-factors against vaccinations. For example, the lack of vaccination across the globe vary due to geographical, social, and cultural contexts, including diverse political and ideological orientations, and weak medical intervention (medical organizations) on social media regarding a pro-vaccine campaign.

Not only do co-factors such as ideological, geographical, and limited medical information campaigns delay intensive and massive Covid-19 vaccination progress, but Facebook posts also revealed that there was significant fear and sadness when Facebook users discussed vaccine emergency approval and trial results (Zhang et al., 2022). Thus, Facebook likes were interpreted as a call for action when having to do with vaccines e.g., the usefulness of vaccines, shipment, new cases of Corona infections, and Coronavirus deaths, and testing vulnerable groups (healthcare staff and elders).

**2.2** Analyzing these comments on social media shows that people’s preference is for the Pfizer vaccine (Na et al., 2021). The same finding is confirmed by (Zhang, Mukerjee , & Qin, 2022) in their study “Topics and Sentiments Influence Likes: A Study of Facebook Public Pages’ Posts about Covid-19 Vaccination.” The discoveries affirmed that Facebook posts liked the rollout of the first opening dosage of the vaccine Pfizer-BioNTech among Americans; there was a positive appraisal of the trial results and effectiveness, and testing approaches. Another recent study by Yin, Xiangyu, Shuiqiao, & Jianxin (2022) reveals general sentiment polarity was positive for the Covid-19 vaccine from Twitter tweets. People’s positive tweets expressed their appreciation or gratitude for receiving the vaccine. People hope the pandemic can be under control, they can resume their everyday lives as soon as possible.

**2.3** There here are more questions than answers on the topic of Covid-19 vaccines. Many of those concerns are Covid-19 vaccine side effects, population reduction through the Covid-19 vaccine, children's vaccination without parental approval, and health issues concerning Covid-19. Although, for example, Yin, Xiangyu, Shuiqiao, & Jianxin (2022) concluded that there is positive polarity sentiment in people's tweets towards a covid-19 vaccine. The researchers also found that negative tweets-text complaints about fever, sore arms, and so on as part of the vaccination side-effects. Moreover, those concerns are nourished and reinforced by like-minded people's beliefs on social media platforms, particularly on Parler, which is a newer social media network (see Baines, Ittefaq, & Abwao, 2021). In the same study, these researchers suggested that medical authorities, public health experts, and organizations should play a proactive role in denying misinformation and stopping like-minded people from spreading conspiracy theories that may have deadly consequences.

These findings suggest that an infodemic on social media can create a climate of fear. Social media seems to be a hotbed where users can be easily misinformed about the true nature of coronavirus and its current situation. One raised question is why have not WHO and scientific communities, as official voices in the scientific field, had a more proactive role on social media to deny these conspiracy theories. Another question is, what scientific alternatives have been implemented by WHO? The anti-vaccine movement is justifiable among social media users since they have not heard medical and persuasive official voices to challenge questionable Covid-19 development procedures. Another critical point to highlight is the necessity of creating solid policies to regulate and stop hate speech and misinformation generation in all existing social media platforms. It is expected that ordinary social media users' fears, mistrust, skepticism, and hesitancy about Covid-19 vaccines will decrease due to exposure to valid scientific information.

To conduct the analysis research the following research questions and objectives are employed:

**Search questions:**

- What is the mass opinion statistically found in Covid-19 vaccine-related tweets collected from January 2020 and September 2022 by executing a Sentiment analysis?
- What are the most salient semantic features detected in a polarity Covid-19 vaccine-related tweets corpus from January 2020 and September 2022 by executing LDA as a Topic modeling?

**General objective:**

To evaluate language attitudes of Covid-19 vaccine-related tweets from January 2020 to September 2022 through a Sentiment analysis and Topic modeling

**Specific objectives:**

1) To clean data from the Covid-19 vaccine-related tweets
2) To evaluate language attitudes to the Covid-19 vaccine-related Tweets by Sentiment analysis.
3) To visualize the evaluative language features of Covid-19 vaccine-related Tweets by Topic modeling.

## 3. Literature review of the main components of the research topic

In my literature review, I discuss relevant concepts and ideas according to my research questions and the research methodology so that I could achieve the research aim. In this regard, I display in a synthesized way what  Kaggle is as a data repository, Covid vaccine Tweets-Kaggle, Sentiment analysis, and Topic modeling mean as the main components of this study.

### 3.1 Kaggle

It is a platform that hosts quantitative web-based data[1] mostly as a Google LLC subsidiary. Additionally, it has turned into a digital and scientific community to promote competitions, finding, exploring, and publishing datasets by employing data science and machine learning. Within Kaggle's web-based data, there are

[1] https://www.kaggle.com/

many available datasets for exploring, analyzing, and sharing quality data. Covid vaccine tweets[2] are one accessible dataset; its tabular data is supported with a Comma-Separated List (CSV) file type.

## 3.2 Sentiment analysis (SA)

It is defined as an artificial intelligence tool to decode people's emotional opinions or feelings from a digital text, usually known as opinion mining. All words can be classified as positive, negative, and neutral as polar sentiments of a particular written text. The polarity detection of the data is an opinion-mining approach to Natural Language Processing (NLP). The SA operates together with VADER (Valence Aware Dictionary for Sentiment Reasoning (Yin, Xiangyu, Shuiqiao, & Jianxin, 2022).

However, recent studies have expressed doubts. When using computational text analysis, Nguyen et al. (2020) warn about dealing with born-digital data, for example, Twitter. They claimed that Twitter is not designed to assess public sentiment because the born-digital data is somewhat restricted for research purposes. Hence, the results may show some possible bias and weak conclusions. In the framework of computational text analysis, big data is a challenge. Boyd & Crawford (2012) claim that social networking platforms such as Facebook and Twitter databases are unreliable sources for conducting social research that uses a quantitative approach. These social media companies often have reservations about sharing their data with people outside their companies. Other factors that may alter raw analytics are the presence of trolls or hackers such as those who post hate comments on social media. In many cases, social media companies like Facebook or Twitter do not always debug those derogatory comments. The hate comments alter the statistical analysis and interpretation of the original data set. Big data is a powerful tool to address social problems using quantitative analysis, but one must continuously be aware of the problem of subjectivity.

[2] https://www.kaggle.com/datasets/kaushiksuresh147/covidvaccine-tweets

### 3.3 Topic modeling

Latent Dirichlet Allocation (LDA) is a popular topic model to extract high-level semantic subjects across an unstructured and large corpus. The LDA finds a distribution of words to form hot-topic discussions separately under the logic of discrete probability semantic distributions. The LDA distributes the topics in positive and negative sentiments to form word clouds that human logic interprets (Yin, Xiangyu, Shuiqiao, & Jianxin, 2022).

## 4. Data and Methods

### 4.1 Data

Initial data: 399645 from Covid Vaccine Tweets

Clean data: 3525 in Microsoft Excel

**Methods**

R-studio

### 4.2 Methodology

In this study, I analyzed attitudinal reactions to Tweeter-tweets published by Twitter's users. All the tweets, as raw information, were retrieved from the Kaggle repository, specifically from the Covid-19 vaccine data. The published Twitter tweets in English went from January 2020 to September 2022 and are easily accessible to download in a CSV format. After having the needed data. I selected R studio which is software to process data on large-scale data. I divided the quantitative analysis into three stages, namely 1) Pre-processing of the data, 2) Sentiment analysis, 3) Topic modelling.

#### 4.2.1 Pre-processing of the data

At this stage, I verified the number of variables. They also needed to be characterized or sorted into my research criteria. For example, on the raw data, I verified the dates, text descriptions, and other special characters: e.g @, hangtags, etc. Additionally, I detected and deleted empty data which meant that rows did not contain the needed information under the following abbreviation N/A (Not available), unwanted columns.

After having done the cleaning of the data on each variable. I continued to process or clean texts published on the Twitter platform. Mainly, I changed letters from uppercase to lowercase, and remove whitespaces, blank cells, symbols or abbreviations (RT) (retweet), duplicate cells numbers, punctuation characters, a mixture of numbers and letters, and common words known as stopwords, meaningless words. The pre-processing data began with 399645 observations of the variables, and cleaning the whole amount of data, I obtained 3525 text containing data employed in the other type of analysis, namely the Sentiment analysis and the Topic modeling.

### 4.2.2 Sentiment analysis

I conducted the Sentiment analysis to know the different types of attitudinal expressions registered in each tweet-texts. For doing so, I employed the VADER dictionary (Valence Aware Dictionary and Sentiment Reasoner), which contains word classifications expressing positive, negative, or neutral sentiments on a scale from -4 to 4+. As a result, it expresses -1 as the most extreme negative value and 1+ as the most extreme positive value, which is part of the VADER's s metrics or Compound score. My threshold values of the compound scores were (positive sentiment >=0.05, neutral sentiment > -0.05, and negative sentiment <=-0.05. I activated the VADER dictionary library within the R studio software its function is vader_df. VADER library executes the word analysis within a text.

The following graphs show representations of the compound scores via Standard Distribution (Figure 1) and Doughnut chart compound score (Figure 2).

1) **The Standard curve distribution of the compound score**

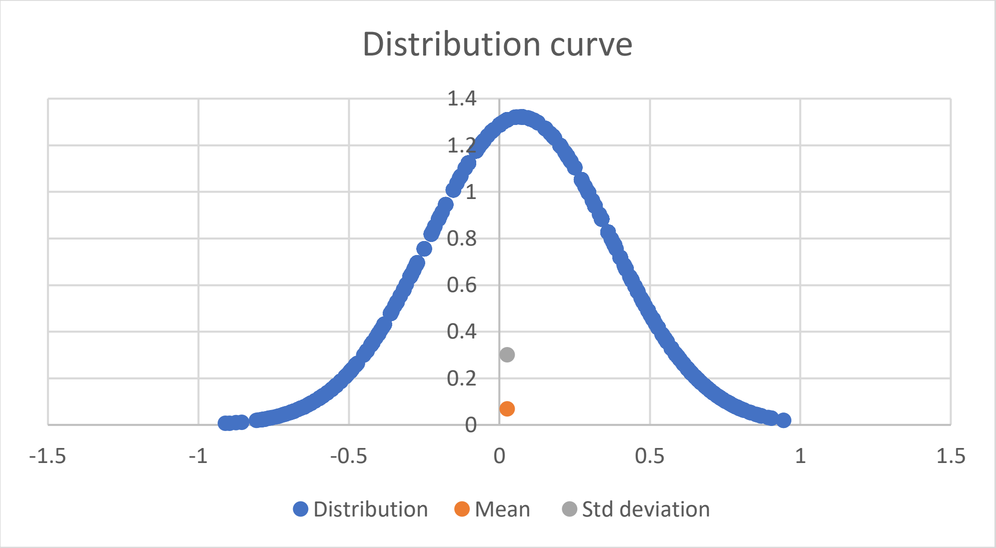


**Figure 1**

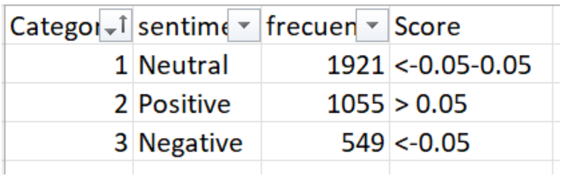
**Source:** The author

This represents an output of the data per each Twitter tweet. I evaluated the mean and the Standard deviation (SD) scores. The mean score represents 0.069032, and the SD reached 0.301662. Thus, to calculate the distribution, I used an excel sheet that allowed plotting the distribution curve.

As can be seen, the compound scores present an almost normal distribution pattern. It shows us that the mean is more than zero. The increase of the curve goes more than -0.05 and 0.05 in the SD. It is interpreted that most of the tweets are neutral with 54%.

### 2) A doughnut chart of the compounded score

The compound scores are represented into three divisions or categories as part of the sentiments: Neutral, Positive, and Negative.



| Categor | sentime | frecuen | Score |
|---|---|---|---|
| 1 | Neutral | 1921 | <-0.05-0.05 |
| 2 | Positive | 1055 | > 0.05 |
| 3 | Negative | 549 | <-0.05 |

**Table 1**

**Source:** The author

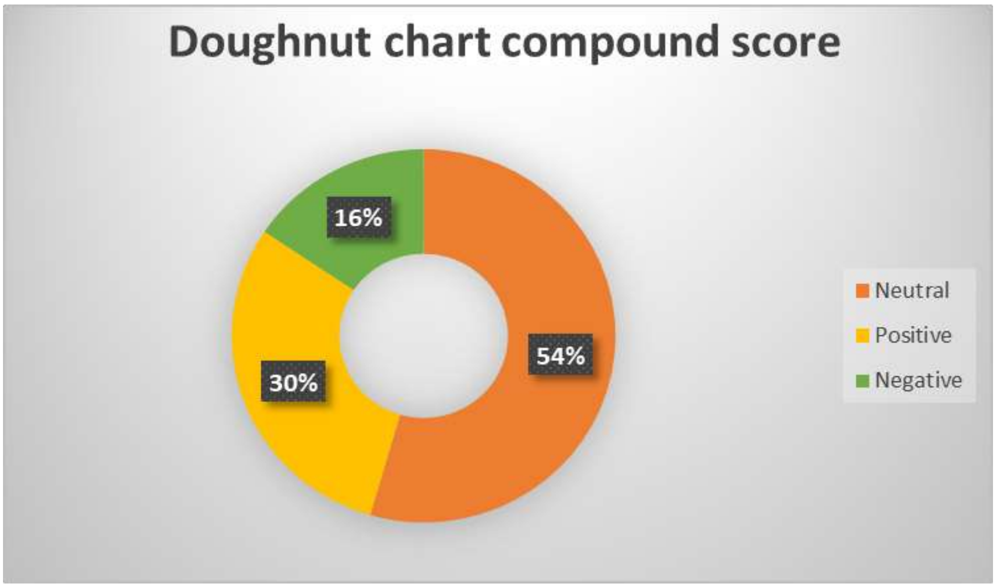


**Figure 2**

**Source:** The author

Table 1 provides information about each sentiment with each one of their frequencies. For example, based on the formula compound score, the neutral sentiment frequency lies on the score -0.05 -0.05.

Figure 2 projects through a doughnut chart visualization the division or polarity of Twitter tweets. 54% of the evaluated tweets reached a neutral sentiment. It is followed by a 30% frequency score of positive sentiment. The lowest percentage goes to the people who tweeted negative phrases. The negative sentiment indicates 16%. However, if I compare the positive and negative sentiment percentages, the positive sentiment is greater than the negative one.

### 4.2.3 Topic modeling

For the next study stage, I carried out the topic modeling to visualize expressions or words based on Sentiment classification. I divided the dataset into three matrixes according to sentiments or attitudes such as positive, negative, or neutral. After dividing the dataset, I included the library (topic models) and the library (quanteda) using the LDA function. It consists of sorting out words from the whole text to determine frequencies and usage percentages in the tweets. Afterward, I activated the word cloud function to visualize the most frequently used words in the shape of a cloud per each sentiment or attitude.

After applying both the Sentiment analysis and the Topic modeling, the results were recorded and exported to Excel files. Then, I can evidence and visualize the classification of the texts through the expressed sentiments into three groups, positive, neutral, and negative.

**Neutral sentiment word cloud**

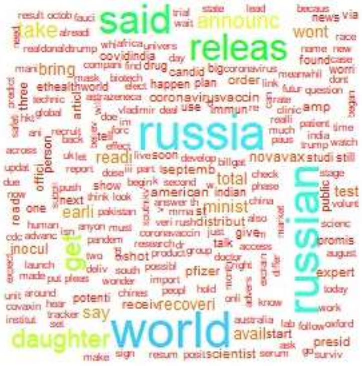


**Figure 3**

**Description e interpretation**

Figure 3 shows the high-frequent words in the studied dataset. The most popular tokened words are visualized in a word cloud. There is not a detectable sentiment in terms of polarity, even though there are salient words like *Russia, Russian world get, daughter, recovery*. They do not make a coherent connection under the hot-discussed topic: covid-19 vaccine.

**Positive sentiment word cloud**

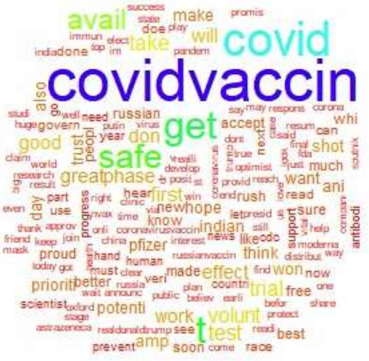


**Figure 4**

**Description e interpretation**

The dataset shows prevalent positive words from tweets in figure 4. According to this word cloud, the most dominant words in their internal frequency scale have become *covid vaccine get-safe*, including the words *available, take, good, will, effect, trail, work*. It clearly shows that people link covid vaccine with a positive attitude. Those words represent a potential sentiment spread and replicated to protect themselves from the infection.

**Negative sentiment word cloud**

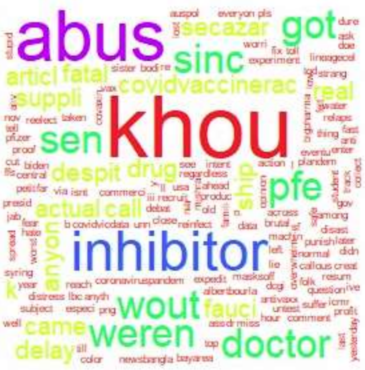


**Figure 5**

**Description e interpretation**

The results in figure 5 display the most frequent numbers in negative polarity. The number of negative words is equal to in the positive word cloud. In the group of the most salient words appear for instance: *inhibitor, abuse, despite, drug, out, delay, doctor, fauci, ship* Nonetheless, they express some emotional concerns towards covid-19 vaccine. The negative polarity shows a delay in the vaccination program worldwide due to vaccine shortages transported on ships. Another hot-discussion topic was the imposition of a covid vaccine meaning an inhibitor of every person's will.

## 5. Analysis and discussion

The **Sentiment analysis results** generated by VADER tool reveal that there is a strong *neutral polarity sentiment* toward the covid-19 vaccine in Twitter tweets. Its 54% represents more than half of the studied data, fig.2, and it had a standard deviation of more than 2.5 regarding the mean, fig. 1, based on 1921 token features, tabl.1, over the studied period. The second sentiment polarity is *positive*, showing that the number of positive people's tweets doubles its percentages, 30%, if compared to the *negative* sentiment, 16%. In other words, the percentage of positive tweets is more significant than the negative reactions on Twitter.

According to some random observations done in an excel file[3], about polarity sentiment, particularly the neutral sentiment observed in the studied texts display neutral phrases. Example: *mom said corona dad heard karona and her it come coroni covidvaccin coronapandem.* On the other hand, positive tweets show aspects such as safety, effectiveness. Exampple: *until a safe amp* ***effect covidvaccin*** *can provid immun mask increas sanit amp social distanc pra data data and more data will make a coronavirus vaccin* ***safe*** *usa today vaccin panel say*. These findings portray some similarities with Yin, Xiangyu, Shuiqiao, & Jianxin (2022). Their sentiment analysis results present that people's tweets consider covid-19 vaccine as an effective protection, and bring them a feeling of gratitude towards the vaccination. Almost similar results were found in (Na et al., 2021) and (Zhang,

[3] sentiment-analysis-visualization-standard-deviation

Mukerjee , & Qin, 2022), their studies showed that people on mass media had a high preference for the Pfizer vaccine because of their effectiveness and trial-testing results.

In contracts, negative sentiment polarity in my current findings focused on covid vaccine shipping delay, vaccine shortage, and lack of trust. Examples from the excel file: *is central vaccin polici of icmrdelhi* ***delay covidvaccin*** *slow clinic trial amp test delay. still* ***sceptic*** *about the vaccin covid covidvaccin. what if we are just part of their* ***pilot project*** *and then we suffer more covidvaccin. blackamerican of us popul of covid* ***death would refus a covidvaccin*** *if offer today.* can there ani doubt whi in american who are not antivaxx wont **trust a covidvaccin** that has been polit. My findings also share similarities with other previous research studies, where Yin, Xiangyu, Shuiqiao, & Jianxin (2022) detected topics like vaccine scarcity, side effects, and deaths due to vaccinations. The same observations appeared in other studies (Baines, Ittefaq, & Abwao, 2021).

Even though there are many similarities found between my results compared with other previous studies, there is a big difference in sentiment polarity percentage. Most of the tweets reported an overall emotion of neutrality. It means most people preferred not to express their attitudes toward vaccines via Tweeter-tweets.

The **Topic modelling of the covid-19 vaccine results**

The employment of LDA generated topics to observe which aspects or concerns tweets point out in their texts. For example, each word cloud's big and distinguishable colors were interpreted as neutral sentiment, fig. 3, positive sentiment, fig.4, and negative fig, 5. In a broad, the salient semantic features do not provide any clear message from the cloud; meanwhile in the positive Twitter tweets clouds, people feel safe getting the covid vaccine. The same topics occurred in (Yin, Xiangyu, Shuiqiao, & Jianxin, 2022).

In the negative word cloud, most tweets see the covid vaccine as an inhibitor. That statement cannot be interpreted because of a lack of clarity. Possibly, the semantic feature refers to covid vaccine shipping delays, as it was also investigated in (Yin, Xiangyu, Shuiqiao, & Jianxin, 2022). Vaccination program as abuse which does not appear in other previous studies

## 6. Conclusion

This class project evaluated language attitudes of Covid-19 vaccine-related tweets from January 2020 to September 2022 through a Sentiment analysis and Topic modeling. A total number of 3525 tweets as cleaned or processed data was obtained. The data display only written texts about people's covid-19 vaccine opinions, posted on their Twitter accounts; information such as geographical locations, dates, and users' names was eliminated due to the nature of the study. According to the Sentiment analysis results, the general sentiment polarity is neutral, and the number of neutral tweets represents approximately more than half of the total tweets, showing an absence of interest in commenting and reacting on the covid-19 vaccine topic on Twitter.

On the other hand, when comparing the positive and negative sentiment of the tweets, the positive sentiment doubles its percentage since the people project the idea of feeling safe. Furthermore, they see the covid vaccine as effective too; both safety and effectiveness are seen as the mass opinion among the tweets posted. Additionally, safety and effectiveness are determined as the most salient semantic features when executing the topic modeling. In contrast, the negative sentiment polarity focuses on covid vaccine delays, vaccine shortages, and lack of trust. Nonetheless, those mentioned factors are not clearly seen in the SA and LDA analysis, which does not mean that SA and LDA lack accuracy, they show a promising level of reliability instead.

In the future, this negative sentiment polarity deserves more attention to identify the main concerns of the people's Twitter tweets. Then, local governments, lawmakers, health administrators, and people in charge of vaccination programs, could interpret those concerns into possible solutions. Finally, the remaining deviation towards the negative sentiment in the distribution curve may decrease until people's opinions consider the covid-19 disease symptoms to have become not deadly.

## References

Boyd, D., & Crawford, K. (2012). CRITICAL QUESTIONS FOR BIG DATA. *Taylor & Francis Online*, 662-679.

Baines, A., Ittefaq, M., & Abwao, M. (2021). #Scamdemic, #Plandemic, or #Scaredemic: What Parler Social Media Platform Tells Us about COVID-19 Vaccine. (R. J. DiClemente, Ed.) *MDPI, 9*(5), 1-15. doi: 10.3390/vaccines9050421

Cascini, F., Pantovic, A., Al-Ajlouni, Y., Failla, G., Puleo , V., Melynk , A., . . . Ricciardi , W. (2022, June 01). Social media and attitudes towards a Covid-19 vaccination: A systematic review of the literature. *eClinicalMedicine, 48*, 1-43. doi:10,1016/j.eclinm.2022.101454

Luo, C., Chen , A., Cui , B., & Liao, W. (2021). Exploring public perceptions of the COVID-19 vaccine online from a cultural perspective: Semantic network analysis of two social media platforms in the United States and China. *Telematics and Informatics*, 2-13. doi:10,1016/j.tele.2021,101712

Mellor, S. (2022, January 18). *FORTUNE*. Retrieved from Fauci says there are 5 stages of the COVID pandemic—and we are still in phase 1: https://fortune.com/2022/01/18/fauci-covid-pandemic-five-stages/

Na, T., Cheng , W., Li , D., Lu, W., & Li, H. (2021). Insight from NLP Analysis: COVID-19 Vaccines Sentiments on Social Media. *Cornell University*, 1-10. doi:https://doi.org/10.48550/arXiv.2106.04081

Nguyen, D., Liakata , M., DeDeo , S., Eisenstein , J., Mimno, D., Tromble , R., & Winters, J. (2020, August 25). How we do things with words: Analyzing text as. *Frontiers in Artificial Intelligence*, 1-24. doi:https://doi.org/10.3389/frai.2020.00062

Scannell, D., Desens, L., Guadagno, M., Tra, Y., Acker, E., Sheridan , K., . . . Fulk , M. (2021). COVID-19 Vaccine Discourse on Twitter: A Content Analysis of Persuasion Techniques, Sentiment, and Mis/Disinformation. *Journal of Health Communication,* 443-459. doi:10.1080/10810730.2021.1955050

World Health Organization. (2021, January 29). *World Health Organization*. Retrieved from Listing of WHO's response to COVID-19: https://www.who.int/news/item/29-06-2020-covidtimeline

Yin, H., Xiangyu, S., Shuiqiao, Y., & Jianxin, L. (2022). Sentiment analysis and Topic modeling for Covid-19 vaccine discussions. *World Wide Web,* 1067–1083. doi:https://doi.org/10.1007/s11280-022-01029-y

Zhang, W., Mukerjee, S., & Qin, H. (2022, September 14). Topics and Sentiments Influence likes: A study of Facebook Public Pages' Posts About COVID-19 Vaccination. *Cyberpsychology, Behavior, and Social Networking, 25*(9), 552-560. doi:https://doi.org/10.1089/cyber.2022.0063