

# Fear of COVID-19 Vaccine Hesitancy using Sentiment Analysis and World Population Data

Oluwatimilehin Ajayi<sup>1</sup>, Gurmol Sohi<sup>1</sup>, and Joshua George<sup>2</sup>

<sup>1</sup>University of Calgary

<sup>2</sup>University of Alberta

June 1, 2021

## Abstract

Vaccines are touted to be the most successful public health measure and they have been around since the 1800s and vaccine hesitancy likewise. COVID-19 (severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)) is a viral infection that likely originated in bats and was transmitted to humans, this disease first appeared in Wuhan, China and has spread across the world resulting in the world health organization declaring the virus a pandemic. With many governments, and organizations invested in returning to life pre-COVID-19 (without masks, governmental restrictions, lockdowns etc.), vaccines were developed namely the Moderna, Pfizer-BioNTech (Pfizer), Oxford-AstraZeneca (AstraZeneca) and Johnson Johnson vaccines. Just like the groups who are invested in slowing the spread of COVID-19, there are other groups who look to benefit from COVID-19 by spreading information and misinformation about the virus, its pathology and the new vaccines. Through infodemiology, this report aims to address myths and provide answers relating to the COVID-19 vaccines. Data was collected from research papers, websites, and kaggle generated twitter datasets. They were then analyzed using Jupyter, and Matlab, Numoy, Textblob, Scipy, Pandas and Seaborn to determine polarity and subjectivity of the datasets. themes and extract patterns of knowledge reflecting ideologies behind the beliefs of those content creators. Although the extraction of twitter datatweets had a neutral polarity and marginally positive subjectivity showed a neutral data set, there however was however a strong positive relationship between vaccine fear and vaccine hesitancy. The neutrality of the twitter dataset may conceal the polarity of

beliefs both extremely for and extremely against the vaccine. Fear of vaccinations is proving to be a huge public health concern because it signals that the public has low confidence in science and modern medicine leading to a growing population that is vaccine hesitant. We provide a possible reason a growing number of the population is fearful about vaccines, misinformation (infodemic), and conclude that scientists, corporations and governments must do more to eliminate fear of vaccines and that will in turn decrease the amount of people who are vaccine hesitant. We recommend future studies to take stock of different cultural, socioeconomic, and religious populations and their attitudes towards vaccinations and take steps to restore the confidence in the safety, efficacy and utility of vaccines against diseases and viruses like COVID-19.

## Keywords

COVID-19, Vaccine, Immunization, Vaccine Hesitancy

## 1 Introduction

On March 11, 2020, the world health organization (WHO) classified severe acute respiratory syndrome coronavirus (SARS-CoV-2) or more commonly COVID-19 as a global pandemic (1). This virus causes respiratory tract infection when transmitted human-to-human via respiratory droplets and has an incubation period of 1-14-days before the hallmark symptoms (fever, dry cough and fatigue) appear [2]. While most cases are mild, some result in acute respiratory distress syndrome, coagulation, dysfunction, pneumonia, and septic shock, these adverse events occur predominantly in the elderly and those who have chronic and/or pre-existing

diseases [2]. This is not the first time a coronavirus has been introduced into the human species as severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV) have been identified long before the global outbreak of COVID-19, infact COVID-19 is the seventh known member of the coronavirus family that can infect humans [2, 3].

As of May 29, 2021 according to WHO, there have been over 169 million confirmed positive cases of COVID-19 worldwide and over 3.5 million resulting deaths [4]. Researchers called for a comprehensive strategy including surveillance, development of vaccines and drugs and more to prevail against COVID-19 [2]. As this virus has impacted medical and public health facilities, economists, politicians and the general public, it is no surprise that many groups have a vested interest in the virus and creating an effective vaccine against this virus. The debate is not about vaccines and their utility in preventing diseases as they have been credited with the eradication of poliomyelitis and smallpox, instead it is about the confidence people have in vaccines and their developers [5].

Below are background information on some of the companies who have successfully created a COVID-19 vaccine:

Moderna is an American pharmaceutical and biotechnology company that focuses on vaccine technologies based on messenger ribonucleotide (mRNA) delivery systems who partnered with the National Institutes of Health (NIH) to create a COVID-19 vaccine [2].

Pfizer is an American multinational biopharmaceutical company that aims to impact healthcare through the discovery, development and delivery of medicines and vaccines [6].

AstraZeneca is a British-Swedish multinational company that focuses on the discovery, development and commercialisation of prescription medicines, primarily for the treatment of diseases in three therapy areas: oncology, cardiovascular metabolic diseases and respiratory diseases [7].

Johnson Johnson, another American multinational corporation, tout of being the largest and most-broadly based healthcare company, committed to changing the trajectory of health for all humanity through their involvement in consumer health, medical devices and pharmaceuticals [8].

Another aspect of biotechnology and pharmaceuticals that does not relate to the companies and their technologies is the media and their coverage of the news. With respect to AstraZeneca for example, the issue of blood clots following

the COVID-19 vaccine was unfairly perpetuated as common but in reality was extremely rare [9,10]. Information or more specifically misinformation compounds on the growing amounts of vaccine hesitancy. 1 Researchers postulated that 5-10

This paper examines - at least on a surface level - virology, immunology and the social sciences as we aim to address why some populations are hesitant and all people in general are polarized either for or against the COVID-19 vaccine. In this paper, we present data generated from different data sources compiled together and use them to address misconceptions and dispel myths regarding the COVID-19 vaccine.

## 2 Materials and Methods

A variety of datasets were used from kaggle as well as data from cited research papers. Data analysis was performed using numpy, Textblob, scipy, pandas and seaborn. The following are the datasets used:

- Covid Vaccine Tweets.[21]
- COVID-19 World Vaccination Progress.[22]

Sentiment analysis - the interpretation and classification of emotions, opinions, expression or attitude towards a topic - was performed. Polarity in sentiment analysis is defined as the float value which lies in the range of  $[-1, 1]$  where 1 means positive statement and  $-1$  means a negative statement and 0 being a neutral statement, while subjectivity in sentiment analysis generally refers to personal opinion, emotion or judgment whereas objectivity refers to factual information. Subjectivity is also a float which lies in the range of  $[0, 1]$ .

Linear Regression - modelling the relationship between an independent and dependent variable so as to predict one value based on the other - was also performed using fear of vaccines and vaccine hesitancy as the variables of interest. The dependent variable is displayed on the y-axis and independent on the x-axis.

## 3 Results

Figure 1: Extraction of twitter data about COVID-19 vaccines. Number of tweets, in intervals of 10, 000. The graph represents the number of tweets (from 46,048 tweets, in intervals of 10, 000s) and their polarity distribution. The value of polarity ranges from  $-1$  to  $1$  and the mean is  $0.1$  with a standard deviation of  $0.23$

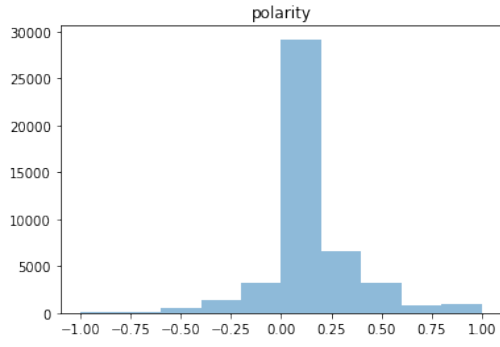


Figure 1: Polarity of COVID-19 vaccine related tweets

which indicates that a majority of people make fairly neutral statements about COVID-19 vaccines.[21]

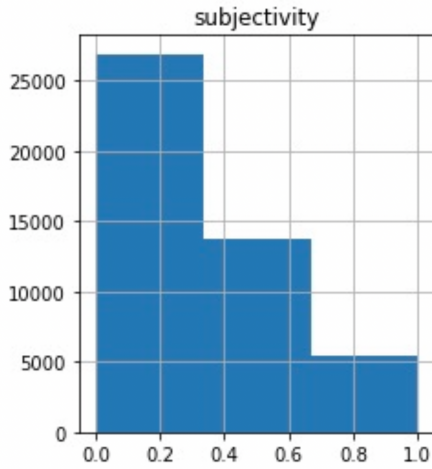


Figure 2: Subjectivity of COVID-19-vaccine related tweets

Figure 2: The graph represents the number of tweets (the same number as in figure 1) and the value of their subjectivity. The value of polarity ranges from  $-1$  to  $1$  and the mean is  $0.26$  with a standard deviation of  $0.29$  which indicates that people have fairly positive opinions on the vaccine. [21]

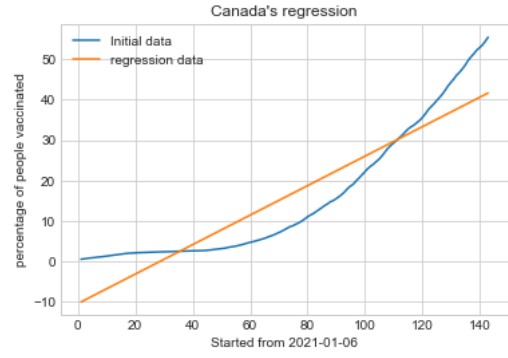


Figure 3: Canada's vaccination regression

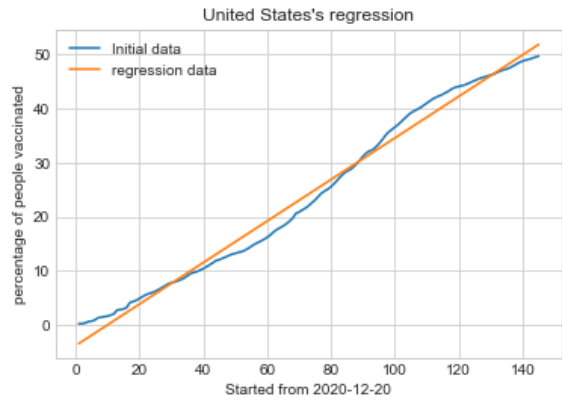


Figure 4: The United States of America's vaccination regression

Figure 3 & 4: Represents the percentage of people living in Canada and the US vaccinated with a regression model.

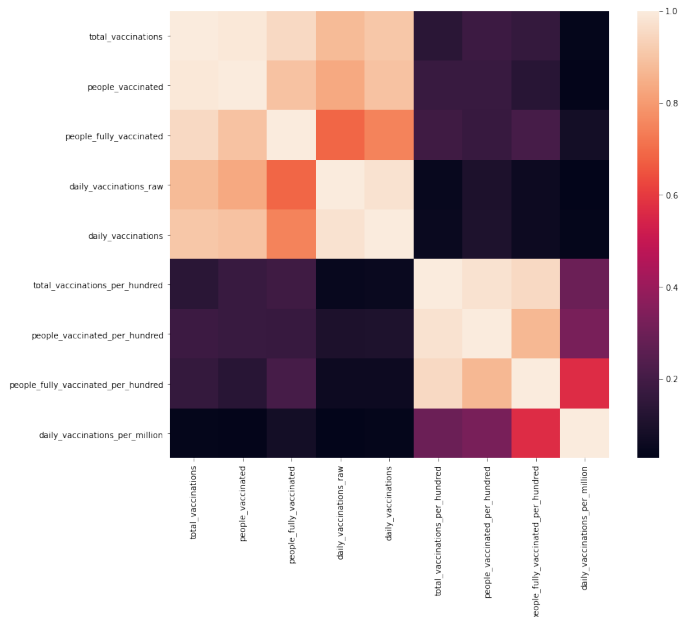


Figure 5: Vaccination summary

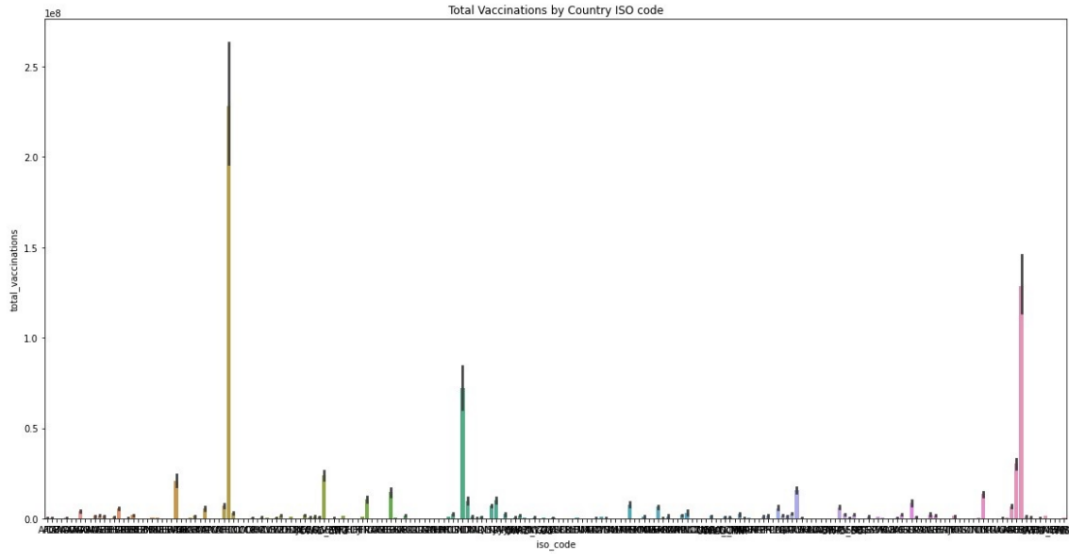


Figure 6: Total vaccinations by country ISO code

Figure 6: Country ISO codes are plotted against the number of total vaccinations. The y-axis values are in 100,000 intervals. [22]

## 4 Discussion

COVID-19 has hampered all world systems and left a severe adverse impact on the daily lives of people worldwide. Lockdowns, mask wearing and social distancing were once strange terms that have now become normal following or even preceding the first confirmed case of COVID-19 in cities and towns worldwide [11]. On December 9, 2020, Health Canada authorized the first COVID-19 vaccine (Pfizer-BioNTech) [12]. This was a turnpoint in the lives of Canadians as it signaled the potential end of COVID-19 and the return to normal life as it were. This breakthrough in modern medicine was not without its detractors as fear about the vaccine was widespread, generated by politicians, media and confusion in the general public. Confusion about the method which mRNA vaccines work, fear of the vaccine altering one’s DNA, the prevalence of blood clots following the AstraZeneca vaccine, etc.

As demonstrated by the sentiment analysis, previous studies and popular websites, there is a strong belief that vaccines and more specifically the COVID-19 vaccine should be avoided [13, 14, 15, 16, 17]. Figure 1 shows the extraction of twitter data and revealed a net neutral dataset in intervals of 10, 000 tweets. There was a high number of slightly positive tweets, as indicated by the polarity in figure 1, tweets showing that over 25,000 tweets spoke positively about

the vaccine. Figure 2 shows the subjectivity of the same data, indicating that despite the positive statements, people are not really convicted and have marginally positive opinions about the vaccine. Figure 5 summarizes vaccinations per different indices.

Figure 6 shows the vaccination status of different countries via ISO code. Gibraltar and the United Arab Emirates have the highest ISO code, Zambia and Nigeria have the lowest code meaning they respectively have the highest and lowest vaccination rates. Comparing the differences in how many people are vaccinated in some countries versus others, there could be a link to vaccine hesitancy present or it could be access to the vaccine that confounds the data. Therefore, a limitation is the uncertainty that underlies the differences in vaccination status in the countries analyzed. Another limitation of the present study is the lack of data about the true reasons why people are polarized regarding the COVID-19 vaccine and why such populations are hesitant to receive the vaccine. With more knowledge about why, scientists, governments and corporations can begin to answer those questions and alleviate the concerns about vaccines and their safety and efficacy.

Figure 3 and 4 shows a positive relationship between fear of vaccines and vaccine hesitancy. This perhaps the most significant finding as it represents true convictions of the general public. One of the benefits of science and medicine is their ability to make the world a better place to live in and prevent mortality. Technology bolsters this ability. Our analysis, taken together, suggests a growing number of people are losing

confidence in science and medicine. There is a growing population of vaccine hesitant people and people who are fearful and confused about vaccines. This represents a disconnect between science and the general public. There are many sources of misinformation (websites, politicians, social media, and friends) places for people to go to acquire false information and not enough resources to correct misinformation [18]. Responding and eliminating misinformation is crucial to restoring the confidence the general public has with scientists and vaccination developers [19, 20]. Politicians are not exempt from helping restore this relationship as their words have a tremendous impact on their constituents. More attention ought to be paid to the social side of science, greater efforts from scientists, corporations and governments need to be invested in disseminating research findings for example using infographics which are easy to understand and to create.

## Conclusion

We primarily relied on twitter datasets as there is a paucity of published research regarding people's thoughts and feelings about COVID-19 vaccines. Our analysis did show that people made neutral tweets about COVID-19 vaccines and feel marginally positive about them. There was also a positive relationship between fear of vaccines and vaccine hesitancy which may prove to be the next public health disaster. The infodemic or misinformation concurrent with COVID-19 also contributed to fear of COVID-19. We provided some suggestions for corporations, scientists and governments to take to ameliorate the relationship between the public and them. Future studies should investigate those who have and have not received the COVID-19 vaccine and their intentions for doing so.

## Acknowledgements

We would like to express our special thanks of gratitude to the Canadian STEM fellowship for organizing the Big Data Challenge and allowing us to learn, collaborate and enhance our skills in data science.

We would also like to thank all faculty members who held workshops for guiding us throughout the challenge.

## References

- [1] Cucinotta, D., Vanelli, M. (2020). WHO Declares COVID-19 a Pandemic. *Acta bio-medica : Atenei Parmensis*, 91(1), 157–160. <https://doi.org/10.23750/abm.v91i1.9397>
- [2] Yang P., Wang X. (2020). COVID-19: a new challenge for human beings. *Cellular Molecular Immunology* 17(5), 555–557. <https://dx.doi.org/10.1038%2Fs41423-020-0407-x>
- [3] Ramshaw, R. E., Letourneau, I. D., Hong, A. Y., Hon, J., Morgan, J. D., Osborne, J., Shirude, S., Van Kerkhove, M. D., Hay, S. I., Pigott, D. M. (2019). A database of geopositioned Middle East Respiratory Syndrome Coronavirus occurrences. *Scientific data*, 6(1), 318. <https://doi.org/10.1038/s41597-019-0330-0>
- [4] WHO COVID-19 Explorer. (n.d.). <https://worldhealthorg.shinyapps.io/covid/>.
- [5] Dubé, E., Laberge, C., Guay, M., Bramadat, P., Roy, R., Bettinger, J. (2013). Vaccine hesitancy: an overview. *Human vaccines immunotherapeutics*, 9(8), 1763–1773. <https://doi.org/10.4161/hv.24657>
- [6] About Us. Pfizer Canada. (2019, August 8). <https://www.pfizer.ca/about-us>.
- [7] AstraZeneca. (2020). AstraZeneca Modern Slavery Act. [https://www.astrazeneca.com/content/dam/az/PDF/Sustainability/Modern\\_Slavery\\_Act\\_Statement.pdf](https://www.astrazeneca.com/content/dam/az/PDF/Sustainability/Modern_Slavery_Act_Statement.pdf).
- [8] ABOUT US. Johnson Johnson Canada. (2020, February 4). <https://www.jnjcanada.com/about-us>.
- [9] Mahase, E. (2021, March 19). Covid-19: AstraZeneca vaccine is not linked to increased risk of blood clots, finds European Medicine Agency. The BMJ. <https://www.bmj.com/content/372/bmj.n774.long>.
- [10] Wise, J. (2021, April 14). Covid-19: How AstraZeneca lost the vaccine PR war. The BMJ. <https://www.bmj.com/content/373/bmj.n921>.
- [11] BBC. (2020, April 6). Coronavirus: The world in lockdown in maps and charts. BBC News. <https://www.bbc.com/news/world-52103747>.

- [12] Canada, H. (2021, January 21). *Health Canada authorizes first COVID-19 vaccine*. Canada.ca. <https://www.canada.ca/en/health-canada/news/2020/12/health-canada-authorizes-first-covid-19-vaccine0.html>.
- [13] Pandolfi, F., Franza, L., Todi, L., Carusi, V., Centrone, M., Buonomo, A., Chini, R., Newton, E. E., Schiavino, D., Nucera, E. (2018). The Importance of Complying with Vaccination Protocols in Developed Countries: "Anti-Vax" Hysteria and the Spread of Severe Preventable Diseases. *Current medicinal chemistry*, 25(42), 6070–6081. <https://doi.org/10.2174/0929867325666180518072730>
- [14] Dror, A. A., Eisenbach, N., Taiber, S., Morozov, N. G., Mizrachi, M., Zigron, A., Srouji, S., Sela, E. (2020). Vaccine hesitancy: the next challenge in the fight against COVID-19. *European journal of epidemiology*, 35(8), 775–779. <https://doi.org/10.1007/s10654-020-00671-y>
- [15] ProCon.org. (2021, May 12). Should Vaccines Be Required for Children? Vaccines. <https://vaccines.procon.org/>.
- [16] Saldana, G. (2020, December 16). Why You Should Consider NOT Getting The COVID Vaccine. Freedom Wire. <https://freedomwire.com/argument-against-covid-vaccine/>.
- [17] Browder, K. (2021, February 1). Opinion: Americans Should Avoid Getting The COVID-19 Vaccine. The Liberty Champion. <https://www.liberty.edu/champion/2021/02/opinion-americans-should-stay-avoid-getting-the-covid-19-vaccine/>.
- [18] Kulke, S. (2020, September 23). Social media contributes to misinformation about COVID-19. <https://news.northwestern.edu/stories/2020/09/social-media-contributes-to-misinformation-about-covid-19/>.
- [19] Vraga, E. K., Bode, L. (2021). Addressing COVID-19 Misinformation on Social Media Preemptively and Responsively. *Emerging Infectious Diseases*, 27(2), 396–403. <https://doi.org/10.3201/eid2702.203139>.
- [20] Zarocostas, J. (2020). How to fight an infodemic. *The Lancet*, 395(10225):676, [https://doi.org/10.1016/S0140-6736\(20\)30461-X](https://doi.org/10.1016/S0140-6736(20)30461-X)
- [21] Limeng Cui and Dongwon Lee (2020). *CoAID: COVID-19 Healthcare Misinformation Dataset* [Covid all Tweets]
- [22] Gabriel Preda (2021). *COVID-19 World Vaccination Progress* (Version 120) [Daily and Total Vaccination for COVID-19 in the World from Our World in Data]