

COVID VACCINE HESITANCY SCIENCE OR HYSTERIA?

Project: MIS-284N Analytics for Unstructured Data



INTRODUCTION

~95% reduction in mortality and severity of covid cases was observed post vaccination with no reported side-effects

How some of the Covid-19 vaccines compare

Company	Туре	Doses	How effective*	Storage	Cost per dose
Oxford Uni- AstraZeneca	Viral vector (genetically modified virus)	x2 /	62-90%	Regular fridge temperature	£3 (\$4)
Moderna	RNA (part of virus genetic code)	x2	95%	-20C up to 6 months	£25 (\$33)
Pfizer- BioNTech	RNA	x2 /	95%	-70C	£15 (\$20)
Gamaleya (Sputnik V)	Viral vector	x2 /	92%	Regular fridge temperature (in dry form)	£7.50 (\$10)

^{*}preliminary phase three results, not yet peer-reviewed

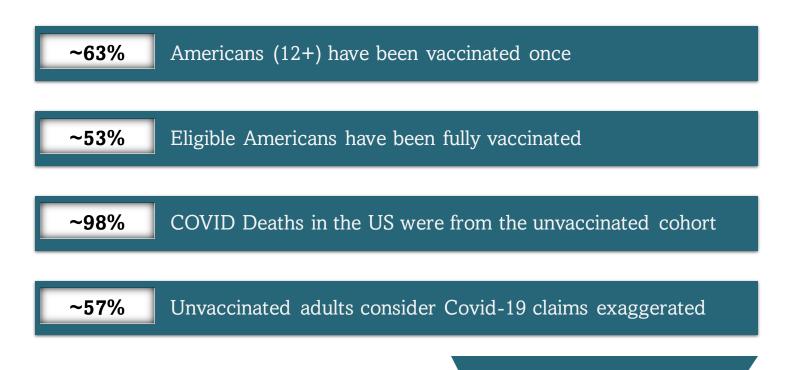
Source: Respective companies, WHO

BBC

- Vaccines prevalent in the US (Pfizer & Moderna) - 95%+ efficacy
- Protection against fatalities from variants in the majority of cases
- Reduction in economic losses by enabling quicker reopening of economies
- Boost the economy by saving \$10 of healthcare costs for every \$1 invested in early-care vaccine programs

INTRODUCTION

Even after targeted government initiatives, more than 40% of the Americans are unvaccinated, which might be due to strong apprehensions regarding the vaccine



Govt Incentives



Problem Statement:

We plan to understand the common perceptions of individuals around COVID-19 vaccination, the key attributes associated with it and any underlying reasons for vaccine hesitancy.

METHODOLOGY

8 topics were identified and the associated sentiment & lift values were analyzed



WEB-SCRAPING

- Scraped tweets
 using Tweepy inluding
 the terms "covid" or
 "vaccine"
- Post extraction, used Selenium to scrape replies to these tweets
- Request Library to extract URL from the shortened version



PRE-PROCESSING

- Removed additional stop words like "RT", "AMP" and other Twitter lingo
- Lemmatized available tweets to simplify word universe
- Removed hashtags, retweets and mentions
- Added keywords from extrcated URLs



SENTIMENT ANALYSIS

- Found sentiments of tweets using Vader
- Word frequency analysis for +ve and -ve sentiment tweet clusters
- Positive Words:
 Effective, Vaccine,
 Health
- Negative Words: Covid,
 Death, Mandates

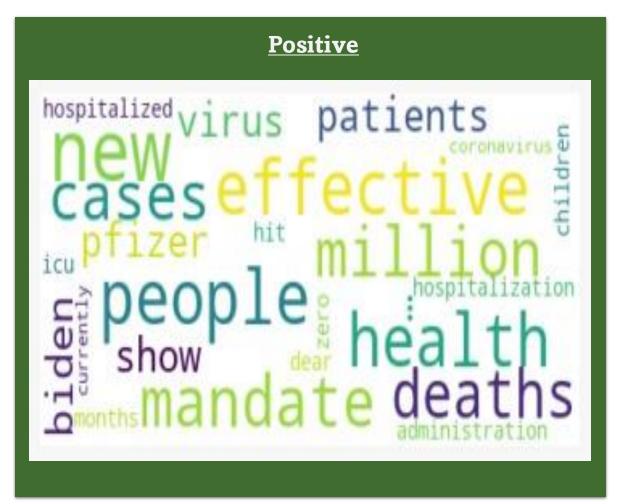


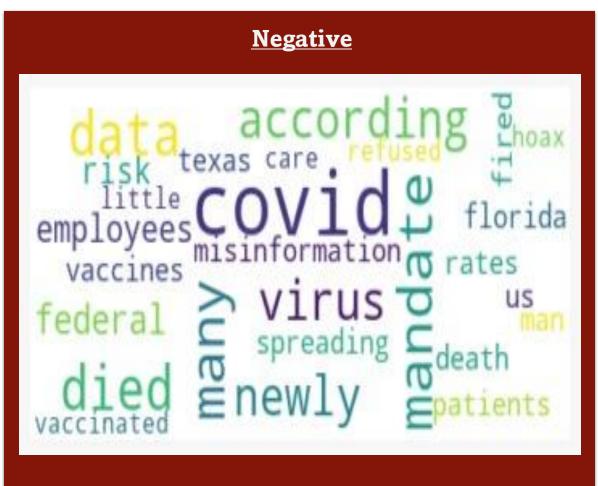
TOPIC MODELLING & LIFT ANALYSIS

- Applied Bag of words & TF-IDF
- Performed LDA to find topics; tuned alpha and eta values to modify the prior Dirichlet distribution
- Found 8 insightful topics; lift analysis and sentiment analysis
- Computed conditional lift to understand reasons for cluster
- Extracted most mentioned accounts for each topic to understand reasons for polarity

RESULTS & INSIGHTS

Frequent words associated with sentiments provided insights into commonly discussed topics





RESULTS & INSIGHTS

CLUSTER1: #VaccinesWork

Getting vaccinated will reduce the spread and deaths from COVID-19	Hospitals are filling up, deaths have risen in Florida	COVID causes more serious illness and more deaths than flu	Getting vaccinated will reduce the spread and deaths from COVID-19
get, 3.5	Florida, 4.3	flu, 4.0	use, 2.9
vaccine	death	illness	vaccine
people, 3.2 vaccine	cases, 5.2 death	flu, 3.7 covid	approve, 1.8 vaccine
vaccine, 0.4	Florida, 8.2	flu, 4.1	effective, 3.4 vaccine
death	hospital	death	
get, 0.5	Florida 2.7	flu, 2.5	booster, 2.1
vaccine	death,	death	vaccine
death	up	less	effective

RESULTS & INSIGHTS

CLUSTER2: #ForcedInjections

Natural immunity works better than vaccines	Vaccine mandates are unconstitutional and anti freedom	Distrust towards government officials is high	Political divide is inhibiting vaccinations
immunity, 9.1	mandate, 6.2	Fauci, 3.3	Biden, 3.4
natural	school	misinform	death
natural, 2.9	mandate, 3.9	America, 2.5	Trump, 0.5 death
good	worker	virus	
immunity, 0.9 vaccine	mandate, 7.7 vaccine	virus, 4.6 made	Biden, 1.7 America
natural, 2.4	school, 2.4	virus, 1.3	death, 2.1 America Biden
immunity	mandate	made	
good	issue	America	

RECOMMENDATIONS & NEXT STEPS

There exists a bias in commonly viewed sources of information, making a neutral source of information the need of the hour

Polarity



There is a visible polarity in public opinion, looking at the most frequent mentions:



Bias



Due to the polarity, the information obtained is biased w.r.t. news sources followed by the individuals

Propoganda • ()



The absence of a trustworthy, fair source of information is further propagating polarizing information:

• Promoting anti-vax behaviors

Spreading falsehood

• Causing unnecessary mass hysteria

THANK YOU!



APPENDIX

DATA SOURCES

S.No.	Data Sources
1	https://usafacts.org/visualizations/covid-vaccine-tracker-states/
2	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2647387/
3	https://www.nationalgeographic.com/science/article/vaccines-are-highly-unlikely-to-cause-side-effects-long-after-getting-the-shot-
4	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4802703/
5	https://www.kff.org/coronavirus-covid-19/poll-finding/kff-covid-19-vaccine-monitor-profile-of-the-unvaccinated/
6	https://apnews.com/article/coronavirus-pandemic-health-941fcf43d9731c76c16e7354f5d5e187
7	https://www.mayoclinic.org/coronavirus-covid-19/vaccine-tracker
8	https://twitter.com/?lang=en