Text Analysis of COVID-19 vaccine tweets

Sydney Joseph (smj), Satvika Neti (sneti), Pooja Puvvadi (ppuvvadi), & Sinduja Sriskanda (ssriskan)

Unstructured Data Analytics for Policy
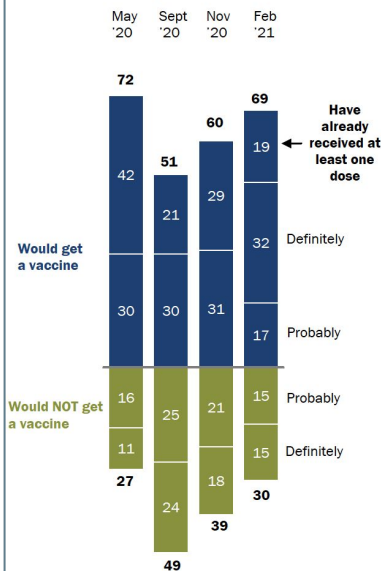Heinz College, Carnegie Mellon University

Background: Public Perception of COVID-19 Vaccines

Despite evidence that fully vaccinated people are less likely to be infected by and spread COVID-19, public perception of COVID-19 vaccines has been varied which is largely due to misinformation.

Half of Americans intend to get a COVID-19 vaccine; 19% already have

% of U.S. adults who say, thinking about vaccines to prevent COVID-19, they ...

	May '20	Sept '20	Nov '20	Feb '21	
	72	51	60	69	
	42	21	29	19	← Have already received at least one dose
				32	Definitely
	30	30	31	17	Probably
	16	25	21	15	Probably
	11			15	Definitely
	27	24	18	30	
		49	39		

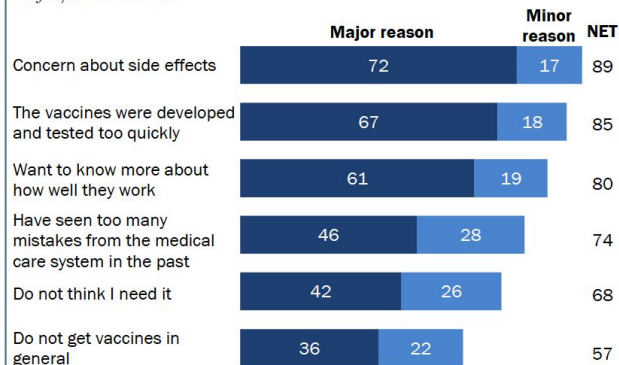**Would get a vaccine**

Would NOT get a vaccine

Note: Respondents who did not give an answer are not shown.
Source: Survey conducted Feb. 16-21, 2021.
"Growing Share of Americans Say They Plan To Get a COVID-19 Vaccine – or Already Have"

PEW RESEARCH CENTER

Those disinclined to be vaccinated cite concerns about side effects, pace of vaccine development and desire for more information as top reasons why

Among the U.S. adults who say they probably/definitely will NOT get a vaccine to prevent COVID-19, % who say each of the following is a major/minor reason

	Major reason	Minor reason	NET
Concern about side effects	72	17	89
The vaccines were developed and tested too quickly	67	18	85
Want to know more about how well they work	61	19	80
Have seen too many mistakes from the medical care system in the past	46	28	74
Do not think I need it	42	26	68
Do not get vaccines in general	36	22	57

Note: Based on those who say they definitely/probably will NOT get a vaccine to prevent COVID-19. Respondents who gave other responses or did not give an answer are not shown.
Source: Survey conducted Feb. 16-21, 2021.
"Growing Share of Americans Say They Plan To Get a COVID-19 Vaccine – or Already Have"

PEW RESEARCH CENTER

2

💡 Policy Questions

1. How does trust of the COVID-19 vaccine change over time around the world (specifically US, UK, and India) ?

    a. Is the trust changing is related to the number of cases in a specific country?

    b. Does trust vary across brands/companies?

2. How does being verified on twitter or number of followers relate to the vaccine sentiment?

Data: Selection and Cleaning

COVID-19 Vaccine Tweets

All tweets worldwide about each COVID vaccine

Language Filtering

Selected a subset of English-language tweets

Location Filtering

Selected tweets from the US, the UK, and India

Location Processing

Filtered out fake locations and geocode the country name of real locations

Date Filtering

Created a subset of tweet data from 12/8/2020 - 4/22/2021

COVID-19 Daily Cases

JHU repository of aggregated worldwide COVID cases

Location Filtering

Selected case data from the US, the UK, and India

Date Filtering

Created a subset of case data from 12/8/2020 - 4/22/2021

Data Transformation

Converted cumulative totals to daily new cases

4

Topic Modeling & Clustering

- TSNE
  - Vaccine label
  - Geography
- Topic Modeling
  - Geography
  - Efficacy
  - Getting vaccinated

TSNE: Vaccine Types and Countries
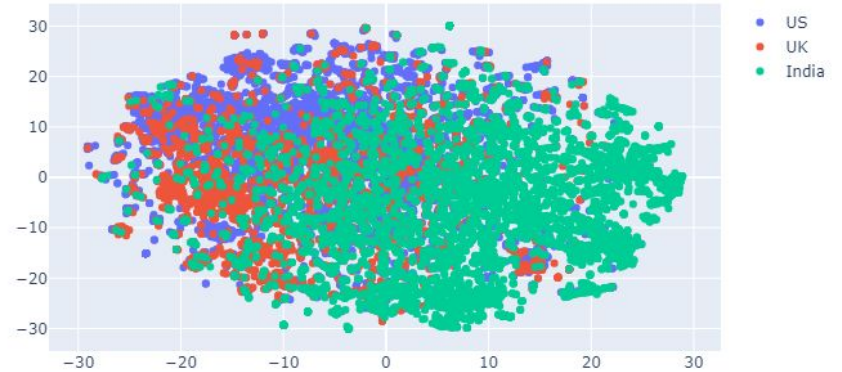
Topic Modeling

Geography

◎ Topic 6 (UK/EU): oxfordastrazeneca, oxford, eu, pfizer, safe
◎ Topic 8 (China): sinovac, sinopharm, china, positive, new
◎ Topic 7 (India): narendramodi, bharatbiotech, modi, bharat, ji

Efficacy/Trials

◎ Topic 1: efficacy, phase, trial, clinical, trials, results

Getting vaccinated

◎ Topic 5 : got, dose, vaccinated, shot, today, second, thank, fully
◎ Topic 9: arm, effects, got, feel, dose, second, sore, little, feeling
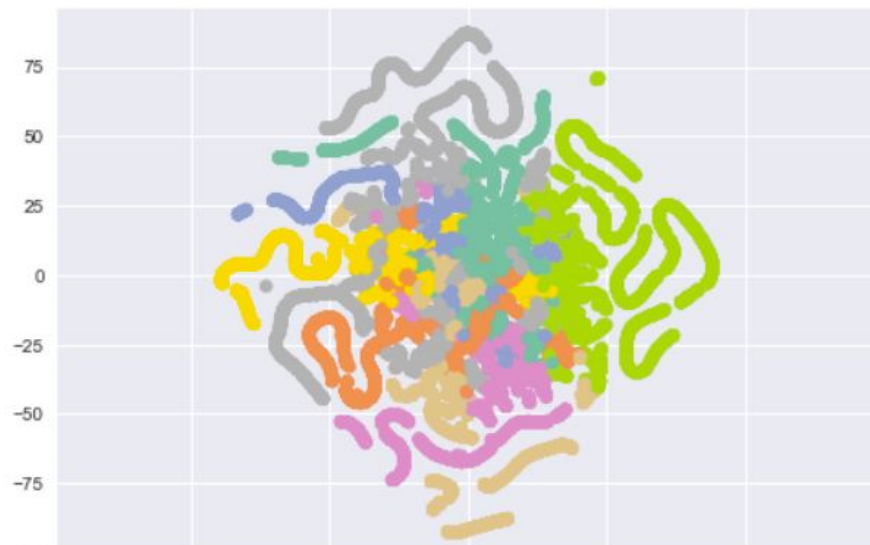
Sentiment Analysis

- Geography
- Verification Status
- Number of Followers
- Number of Retweets/Favorites
- Vaccines
- Vaccines over Time

👍 Subjectivity and Polarity
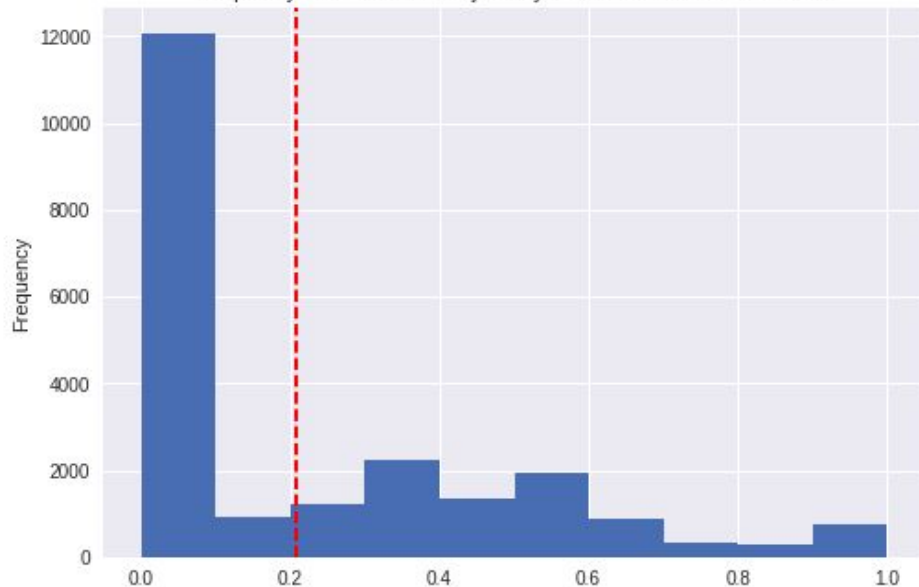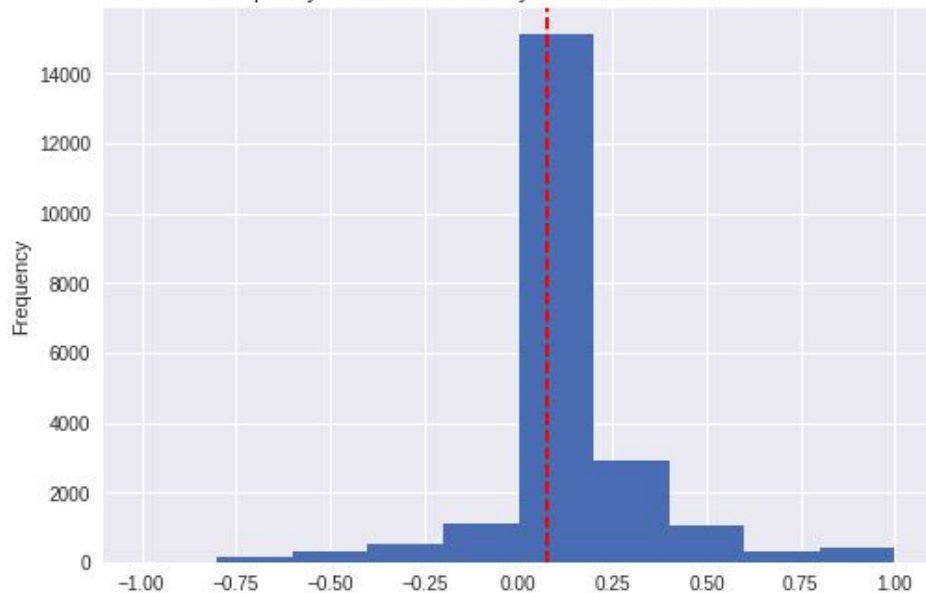
# 👍 Sentiment Analysis: Geography and Verification

No significant difference in polarity or subjectivity based on geographic difference or verification status

Subjectivity by Country

Polarity by Country

Subjectivity Score by Verification Status

Polarity Score by Verification Status
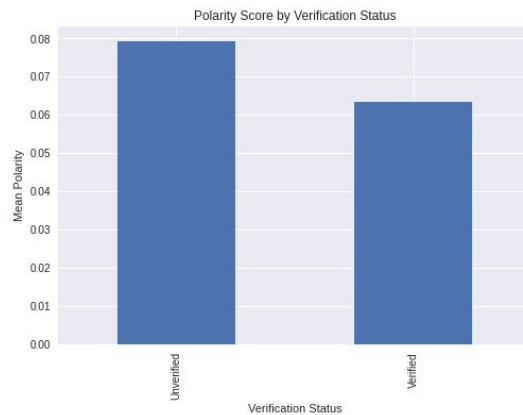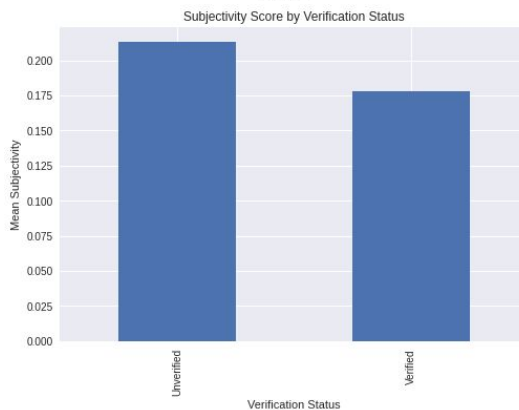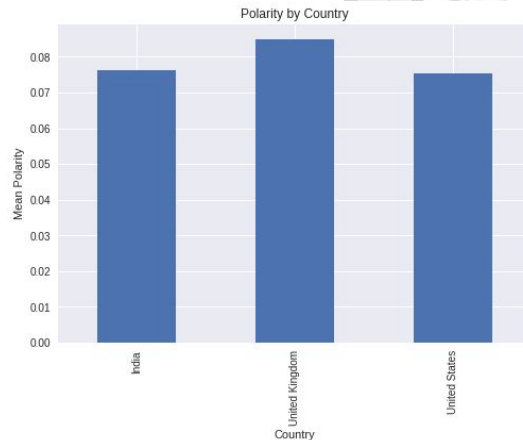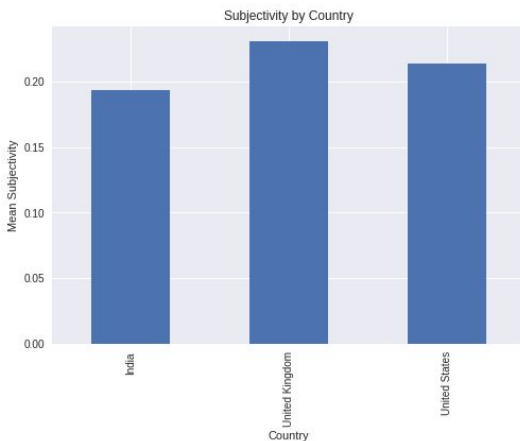
# Sentiment Analysis: Number of Followers and Tweet Interactions
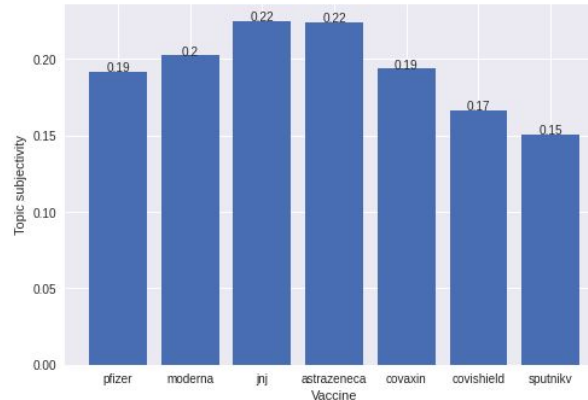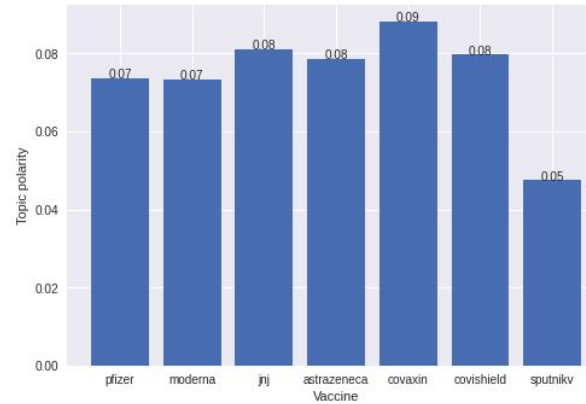
There does not appear to be a relationship between:

- Subjectivity and number of followers
- Polarity and number of followers
- Subjectivity and number of tweet interactions
- Polarity & number of tweet interactions
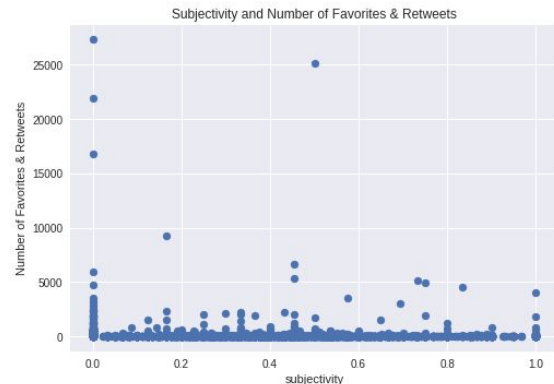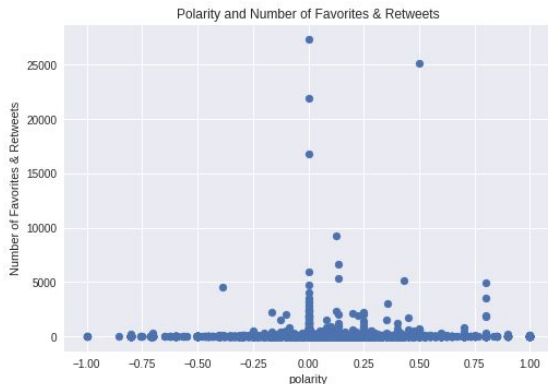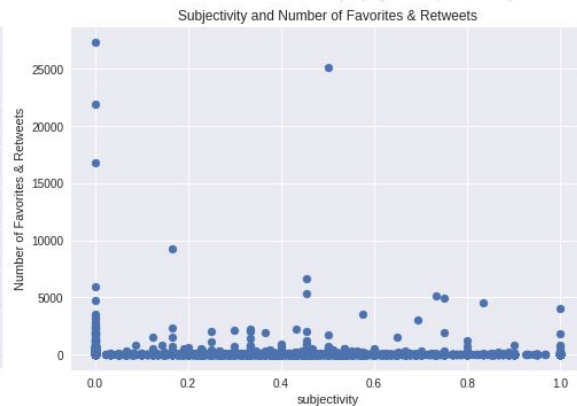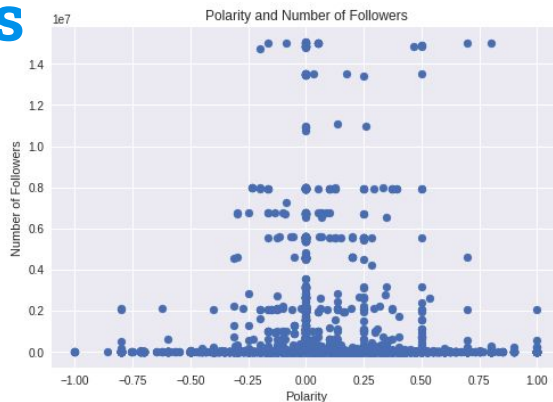
👍 Sentiment Analysis - Vaccines

Discussion started off more positive and subjective in December 2020 and got more neutral and objective over time
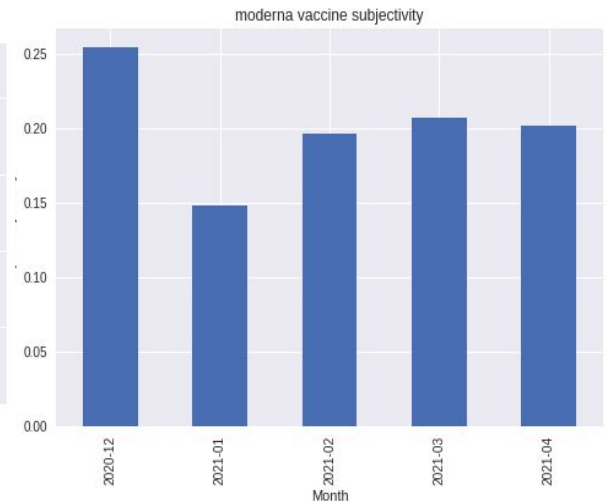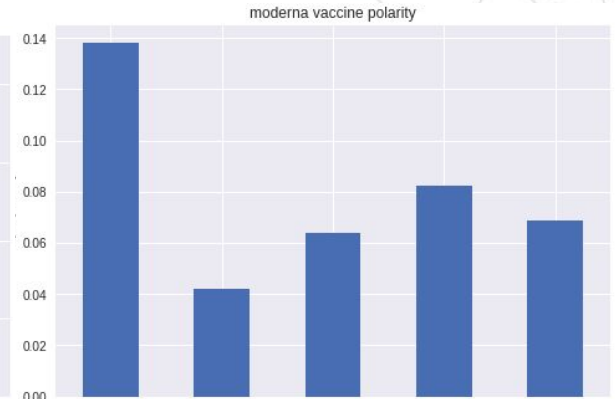
👍 Bag of Words for Positive and Negative Tweets

Positive Tweets

Count	Word Groupings
912	first dose
647	covid19 vaccin
582	got first
500	moderna vaccin
334	dose moderna
298	covid vaccin
272	first shot
247	receiv first
236	first dose moderna
222	got first dose
210	side effect
202	first moderna
201	vaccin today
175	moderna shot
162	get first
158	second dose
152	first vaccin
140	receiv first dose
137	dose covid19
135	pfizer moderna
130	vaccin moderna
128	oxfordastrazeneca vaccin
126	dose covid19 vaccin
123	took first

Negative Tweets

Count	Word Groupings
107	moderna vaccin
100	covid19 vaccin
63	side effect
62	covid vaccin
52	moderna shot
49	subject expert
47	subject expert committe
47	expert committe
46	second dose
43	dose moderna
40	emerg use
37	pfizer moderna
36	view odd
36	second moderna
35	feel like
34	get vaccin
33	sputnikv vaccin
32	vaccin moderna
30	first dose
30	2nd moderna
28	2nd dose
26	harsh vardhan
24	got second
24	fulli vaccin
23	vaccin shot
23	sore arm

Conclusion

- Limitations
- Policy Recommendations

Limitations

◎ Locations - assuming people are being truthful, spaCy's entity recognition limitation

◎ Positive/negative sentiments might not mention specific brands but more general ("vaccine")

◎ Character limit: not many words per tweet, word shortening and abbreviations

Policy Recommendations

◎ Twitter is a good place for public officials to focus on to spread more information about the vaccines and vaccine updates

◎ Leaders should focus on transparency about the vaccine process

  ○ This includes vaccine formulation and funding