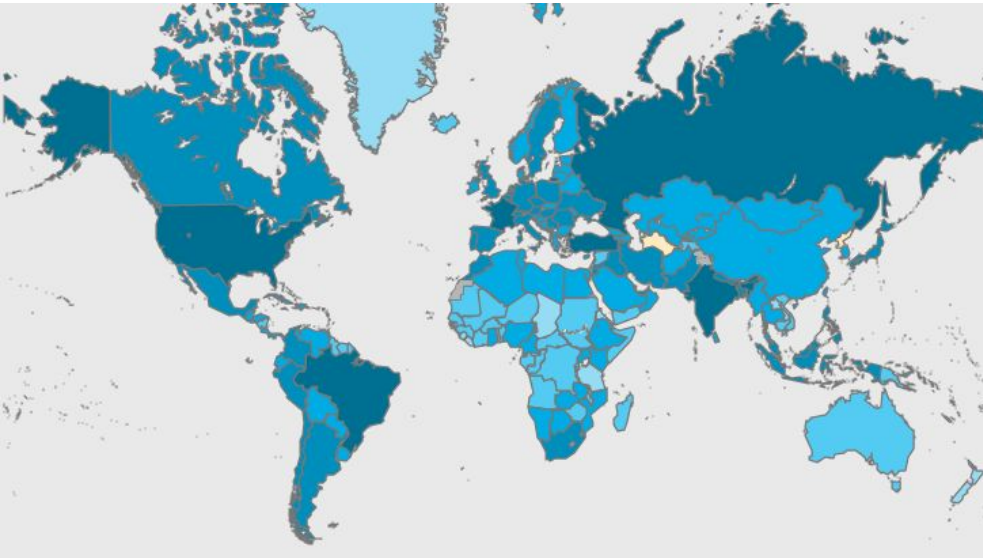Globally, as of 4:00pm CEST, 10 June 2021, there have been 174,061,995 confirmed cases of COVID-19, including 3,758,560 deaths, reported to WHO. As of 10 June 2021, a total of 2,155,894,131 vaccine doses have been administered.

Public Attitudes towards the COVID-19 Vaccine in the U.S. based on Twitter Data

Hangqian Li, Junhong Li, Zijun Cui

11/06/2021

EPFL

Table of Contents

- Introduction
- Goals
- Data
- Methods
- Results & Discussion
- Conclusion

Zijun

Junhong

Hangqian

EPFL

Introduction

Problems:

1) The time cost of vaccine experiments and testing has been greatly compressed.
2) Brand-new technology: mRNA Covid-19 vaccine.
3) Controversial topic: some voices against mass vaccination.
4) Cases of adverse reactions after vaccinating the Covid-19 vaccine.

Public Hesitancy on the Vaccine!

EPFL

How some of the Covid-19 vaccines compare

Company	Type	Doses	Storage
Oxford Uni-AstraZeneca	Viral vector (genetically modified virus)	x2	2 to 8°C (6 months)
Moderna	RNA (part of virus genetic code)	x2	-25 to -15°C (7 months)
Pfizer-BioNTech	RNA	x2	-80 to -60°C (6 months)
Gamaleya (Sputnik V)	Viral vector	x2	-18.5°C (liquid form) 2 to 8°C (dry form)
Sinovac (CoronaVac)	Inactivated virus (weakened virus)	x2	2 to 8°C
Sinopharm	Inactivated virus (weakened virus)	x2	2 to 8°C
Novavax	Protein-based	x2	2 to 8°C
Janssen Johnson & Johnson	Viral vector	x1	2 to 8°C (3 months)

Source: UK government, Reuters

BBC

Introduction

Vaccine hesitancy: "delay in acceptance or refusal of vaccination despite availability of vaccination services"

-- Strategic Advisory Group of Experts on Immunization (SAGE)

3 factors that affect the attitude towards acceptance of vaccination [1]:

1) Complacency: the low perception of the disease risk
2) Convenience: the availability, affordability and delivery of vaccines in a comfortable context
3) Confidence: the trust in vaccination safety and effectiveness

_____

[1] Sallam, M., 2021

EPFL

Introduction

COVID-19 vaccine acceptance rates worldwide (25/12/2020) [2]:

1) the highest: Ecuador (97.0%), Malaysia (94.3%), Indonesia (93.3%), China (91.3%)
2) the lowest: Kuwait (23.6%), Jordan (28.4%), Italy (53.7), Russia (54.9%), Poland (56.3%), US (56.9%), France (58.9%)
3) most studies were from the general public

_____

[2] Sallam, M., 2021

COVID-19 vaccine acceptance rate (%)   0   100

EPFL
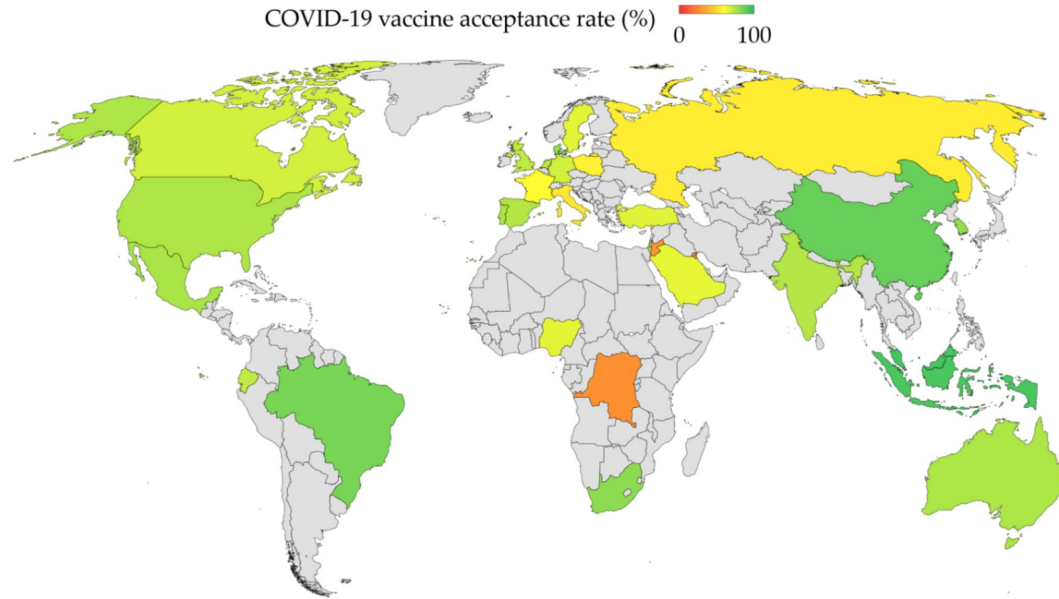
Introduction

COVID-19 vaccine acceptance rates worldwide (25/12/2020):

COVID-19 vaccine acceptance rates in Europe

COVID-19 vaccine acceptance rates in the Middle East

COVID-19 vaccine acceptance rates in Asia

EPFL

Introduction

Vaccine hesitancy in the U.S. (14/04/2021 - 26/04/2021):
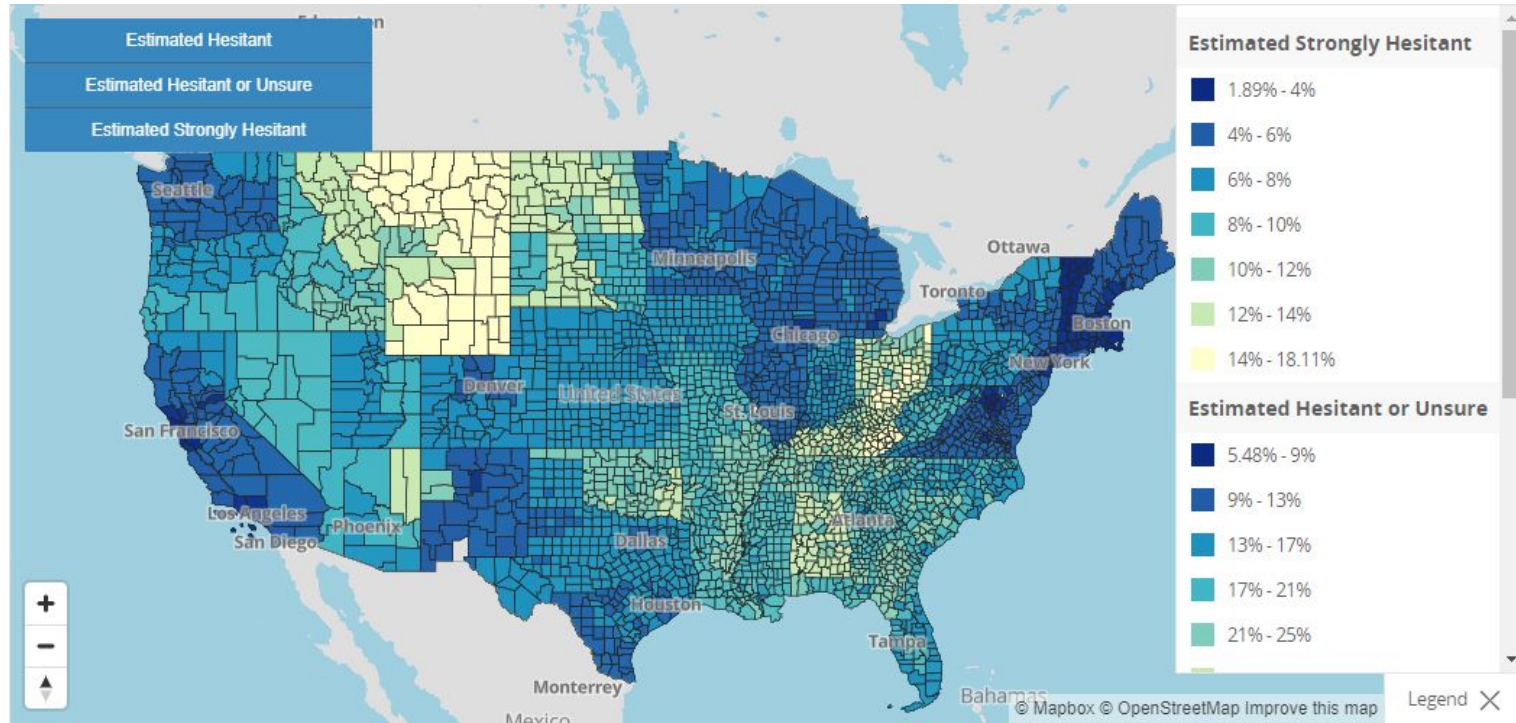
Introduction

Vaccine hesitancy in the U.S.:

"Antivax groups' statements, conspiracy theories, myths and misperceptions, questions about the speed of vaccine development and long-term side effects, and expert opinion on challenges with the COVID-19 vaccine were proliferating in the national media." [3]

religion

political affiliation

place of residence

employment status

family structure

polarized sociopolitical climate

ethnicity

parenthood

income

education

gender

[3] Khubchandani, J., Sharma, S., Price, J.H. et al., 2021

EPFL

Introduction

Situation by WHO Region

Region	Confirmed	
Americas	68,932,684	confirmed
Europe	54,828,356	confirmed
South-East Asia	33,101,517	confirmed
Eastern Mediterranean	10,384,040	confirmed
Africa	3,608,158	confirmed
Western Pacific	3,206,476	confirmed
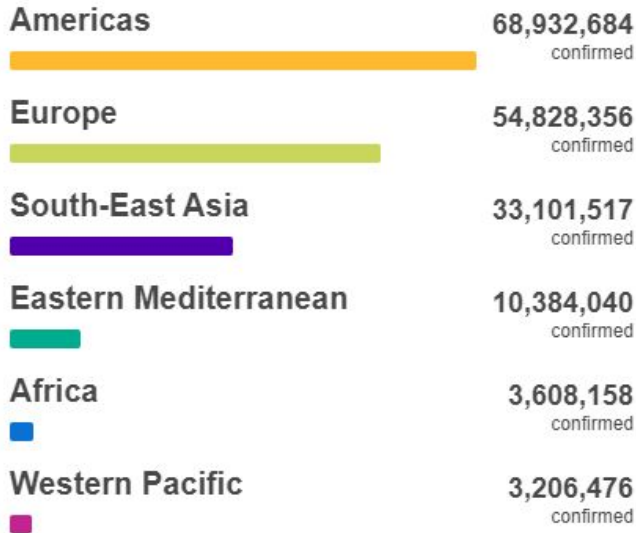
Source: World Health Organization

Questionnaire survey -> Social media data

1) Previous survey methods were mostly based on questionnaire surveys, with limited samples.
2) Social media data can provide a larger range of samples at once.
3) With social media platform, up-to-date data are available for analysis.
4) The health tolls of COVID-19 in the United States have been among the highest in the world.

EPFL

Goals

1. Understand the latest attitudes of the American people in different regions towards the COVID-19 vaccine.
2. Learn the latest vaccine hesitancy in the U.S. for COVID vaccines of different companies (Astrazeneca, Pfizer, Janssen, Sputnik, Moderna, CoronaVac).

EPFL

Data

Data collection for vaccine hesitancy in the U.S.
1. Time: from 2021-May-28 to 2021-June-6
2. Geographical location: 100 miles and 300 miles in radius
     a) Big cities like Los Angeles, New York City
     b) Small counties like Erie County, Moab
3. Hashtag: '#vaccine OR #vaccination OR #vaccineforall OR #vaccinated OR #covidvaccine OR #covid19vaccine OR #getvaccinated OR #astrazeneca OR #oxfordastrazeneca OR #pfizer OR #pfizerbiontech OR #pfizervaccine OR #moderna OR #modernavaccine OR #sputnik OR #sputnikvaccinated OR #janssen OR #johnsonandjohnson

EPFL

Data

Data collection for vaccine hesitancy in the U.S.
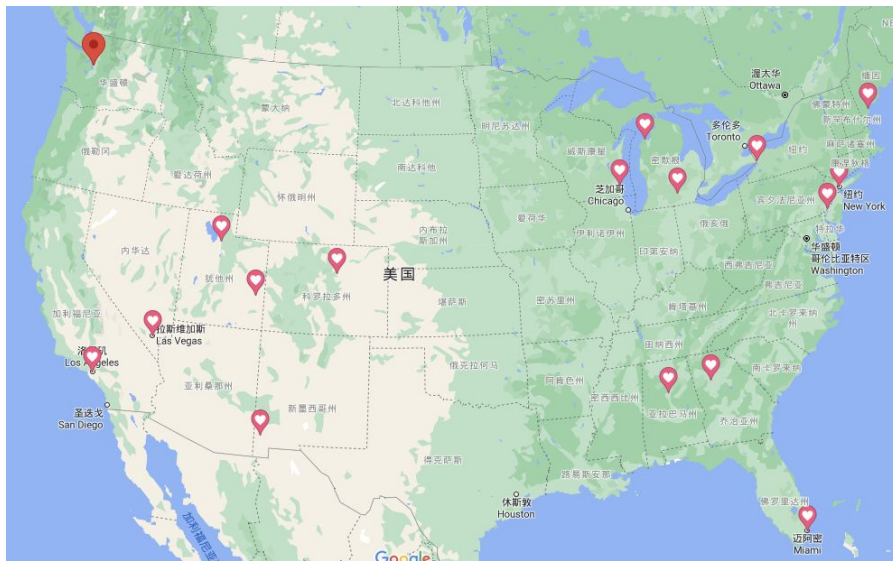
Data1 distribution
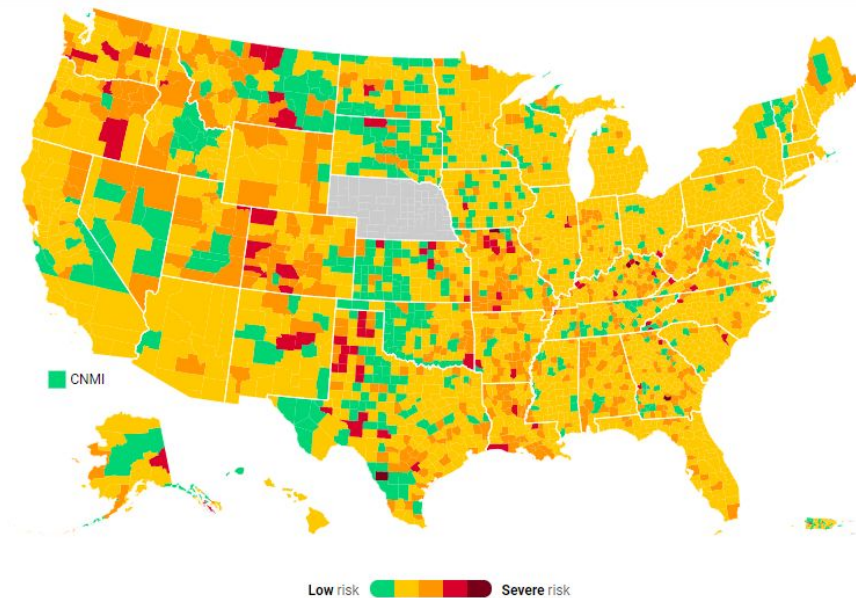
U.S. Covid risk level (until 10/06)

EPFL

Data

Data collection for various vaccines in the U.S.

1) Time: from 2021-May-30 to 2021-June-7
2) Geographical location: 39°20'50.9"N, 100°04'32.2"W, 1500 miles in radius
3) Hashtag: Astrazeneca, Pfizer, Janssen, Moderna, Sputnik, Sinovac as below

```
In [5]:   # hashtag about astrazeneca vaccine:
          text_query_astrazeneca='#astrazeneca OR #oxfordastrazeneca OR #astrazenecavaccine OR #astrazenecacovid OR #astrazenecavaccinat
          ed'
```

```
In [6]:   # hashtag about pfizer vaccine:
          text_query_pfizer='#pfizer OR #pfizervaccine OR #PfizerGang OR #Pfizercovid OR pfizervaccinated OR #pfizerbiontech'
```

```
In [7]:   # hashtag about Janssen vaccine:
          text_query_janssen='#janssen OR #janssenvaccine OR #janssencovid OR #johnsonandjohnson OR #janssenvaccinated OR #JohnsonAndJoh
          nsonCOVIDVaccine'
```

```
In [8]:   # hashtag about moderna vaccine:
          text_query_moderna='#moderna OR #modernavaccine OR #modernacovid OR #modernavaccinated'
```

```
In [9]:   # hashtag about sputnik vaccine:
          text_query_sputnik='#sputnik OR #sputnikvaccine OR #sputnikcovid OR #SputnikV OR #sputnikvaccinated'
```

```
n [10]:   #CoronaVac
          text_query_CoronaVac='#Sinovac OR #CoronaVac OR #SinovacBiotech OR #YinWeidong OR #CoronaVacCovid OR #CoronaVacvaccinated'
```

EPFL

Data

Dataset 1 description:

1) 52748 tweets in total based on area.

2) The numbers of tweets in different areas range from 33 in Lordsburg to 10000 Philidelphia.

3) The total percentage of the collected English Tweets is 89.03%.

Collected Tweet Number in Different Areas
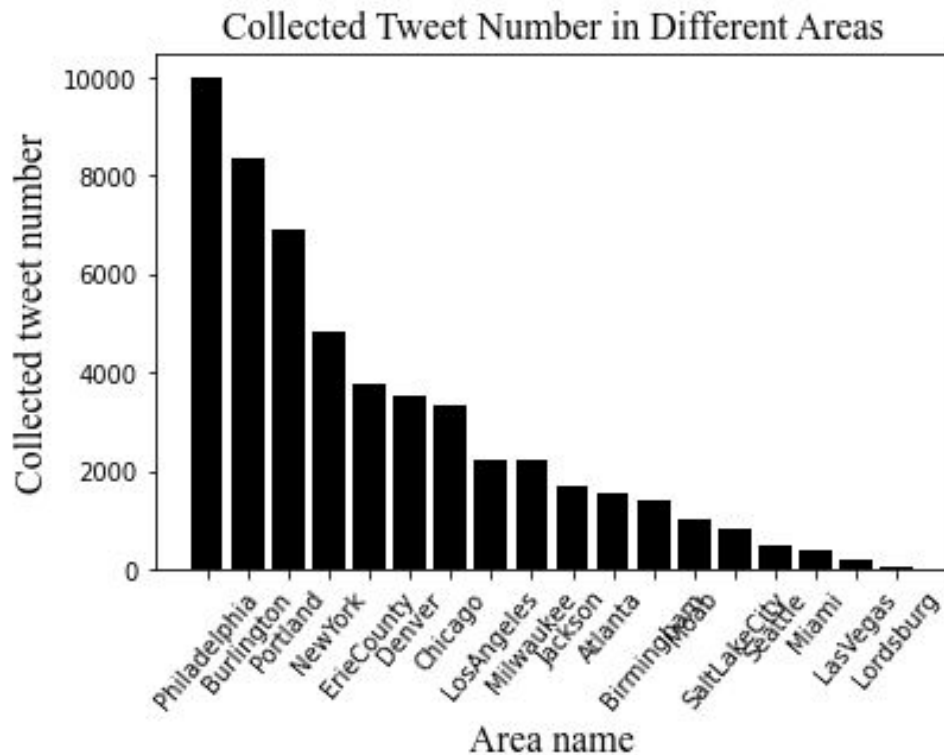
Data

Dataset 1 description:

1) Percentage of tweets that contain URLs is: 38.38%

2) Top five hashtag: #COVID19, #GetVaccinated, #vaccine, #vaccinated, #CovidVaccine,

Retweets Percentage in Different Areas

EPFL
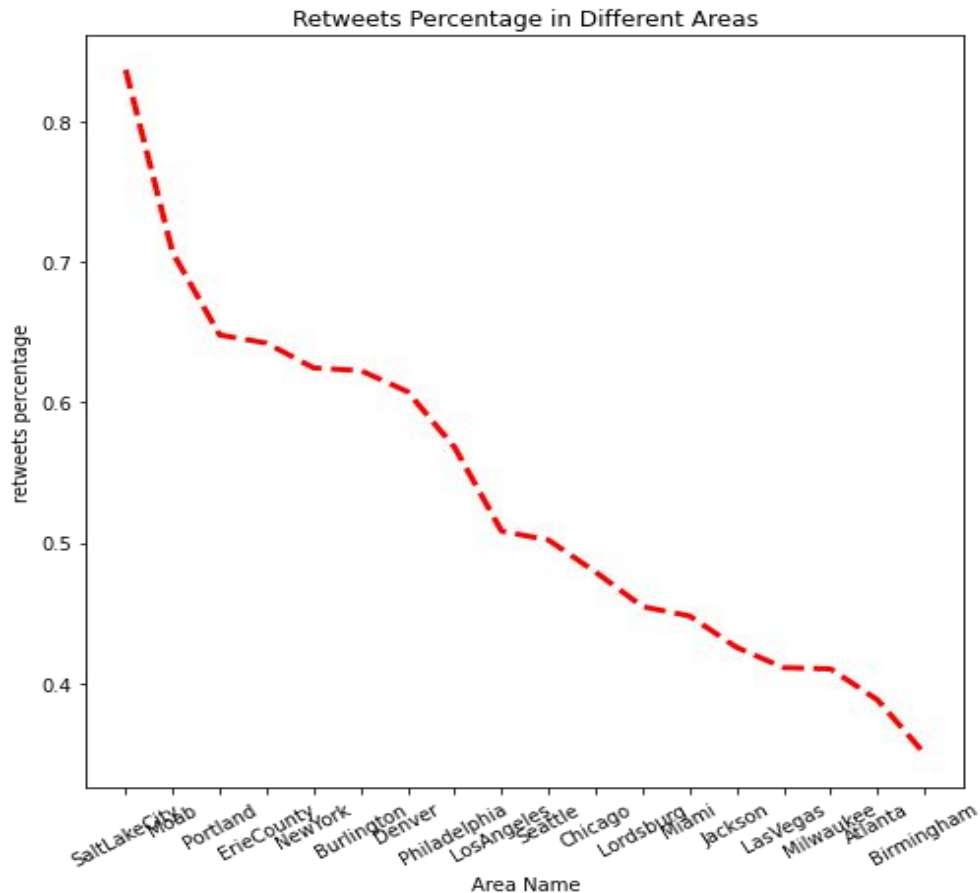
Data

Dataset 2 description:

1) The total percentage of the collected English Tweets is: 45.98%

2) There are 3371 tweets related to pfizer and only 135 tweets related to Corona.

Collected Tweet Number for Different Vaccine

EPFL
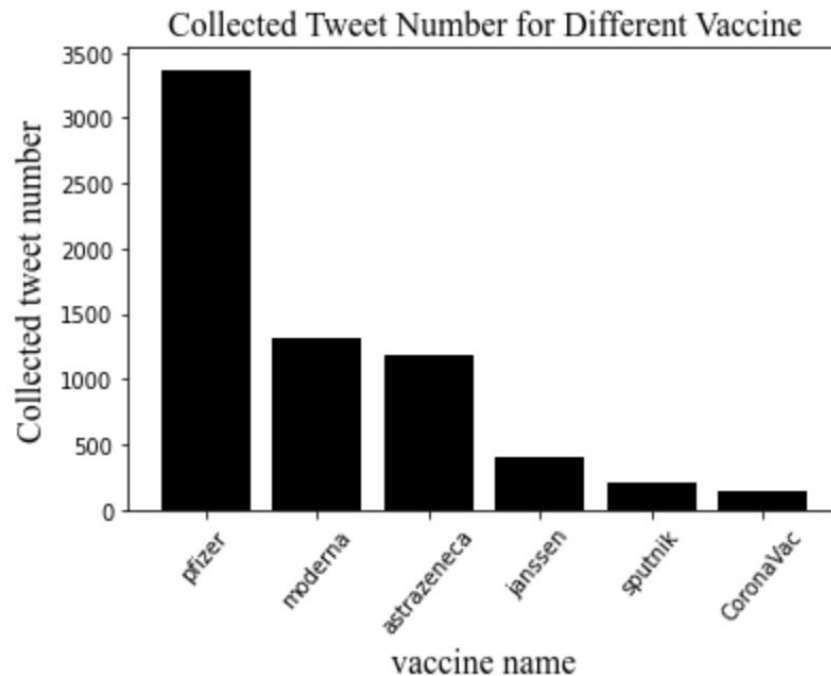
Data

Data pre-processing:

1) Filter out non-English tweets

2) Set the tweets to lowercase

3) Filter URL links

4) Remove username, @, and '#' from tweets

5) Filter emoticons

6) Filter punctuation

7) Lemmatization: converting a word to its base form

8) Remove stopwords

9) Remove short words (length strictly less than 3 characters)

Methods

Work Flow :

1) Collect training data

2) Preprocess training data

3) Vetorize the feature

4) Train the model: Logistics regression & Naive Bayes Model

```
[Labelled dataset] → [Data Preprocess] → [Feature extraction] → [Model training]
```
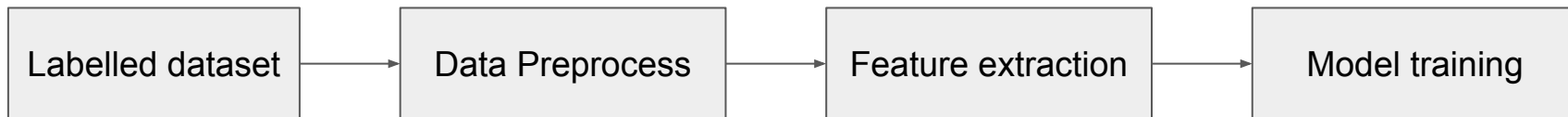
Methods

Labelled dataset: https://www.dropbox.com/s/du1z2m910a68ehk/training.csv?dl=0

Text: original paragraph

Target: Sentiment label

	target	text
0	0	upset cant update facebook texting might cry r...
1	0	dived many times ball managed save 50 rest go ...
2	0	whole body feels itchy like fire
3	0	behaving im mad cant see
4	0	whole crew

Methods

Feature extraction: sklearn.feature_extraction.text.TfidfVectorizer

tf-idf: term frequency–inverse document frequency

$$tf(i,j) = \frac{freq(i,j)}{\max_{k \in T} freq(k,j)} \qquad idf(i) = \log(\frac{n}{n_i}) \in [0, \log(n)]$$

Term weight (tf-idf)

$$w_{ij} = tf(i,j)\, idf(i)$$

EPFL

Methods

Logistics regression(77.74%): sklearn.linear_model.LogisticRegression

$$\hat{y}(\mathbf{x}) = \sigma(\mathbf{w}^T\mathbf{x}) = \frac{1}{1 + \exp(-\mathbf{w}^T\mathbf{x})}$$

$$R(\mathbf{w}) = -\sum_{i \in \text{positive samples}} \ln(\hat{y}_i(\mathbf{w})) - \sum_{i \in \text{negative samples}} \ln(1 - \hat{y}_i(\mathbf{w}))$$

EPFL

Methods

Naive Bayes Model(76.14%):sklearn.naive_bayes.MultinomialNB

$$P(y \mid x_1, \ldots, x_n) = \frac{P(y)P(x_1, \ldots, x_n \mid y)}{P(x_1, \ldots, x_n)}$$

$$P(y \mid x_1, \ldots, x_n) \propto P(y) \prod_{i=1}^{n} P(x_i \mid y)$$

$$\hat{y} = \arg\max_{y} P(y) \prod_{i=1}^{n} P(x_i \mid y)$$

EPFL

Results & Discussion

Attitude positive level in different areas:

1. The highest attitude positive level is in Salt Lake City, which is 79.6%. For big cities like Seattle, Los Angeles and New York, they have over 57% attitude positive levels. And some small towns such as Moab, Lordsburg, they all have high attitude positive level.

2. The lowest is in Miami, which attitude positive level is only 44.0%. The rest four lowest regions are Biemingham, Jackson, denver, and Atlanta, which are 45.4%, 46.9%, 46.9%, 47.9% respectively.

Positive level in Different Areas

EPFL
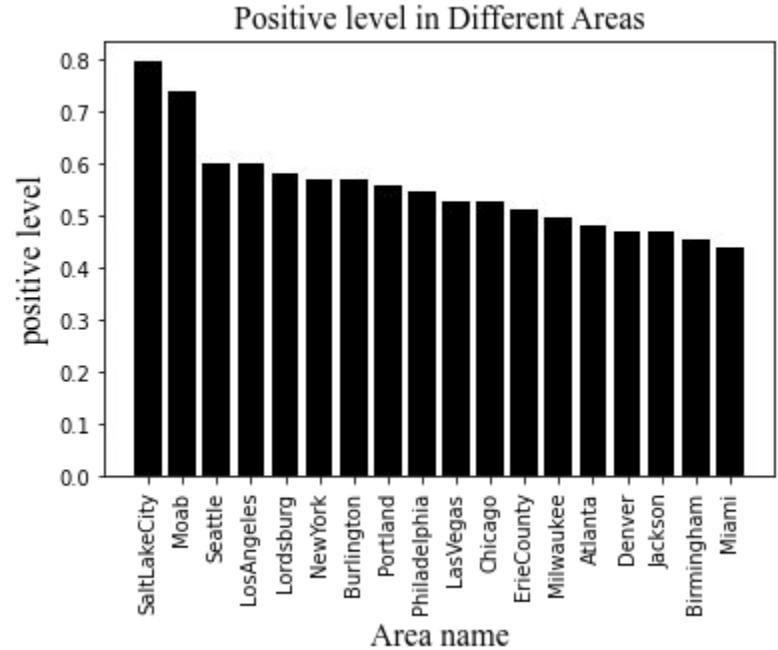
Results & Discussion

Attitude positive level in different areas:

1. Green points means the highest-5 attitude positive level cities, yellow points means the lowest-5 attitude positive levek cities.

2. It seems that people in the West and East has higher attitude positive level, but in the central region, the positive level is relatively low.

The vaccine positive level of the USA
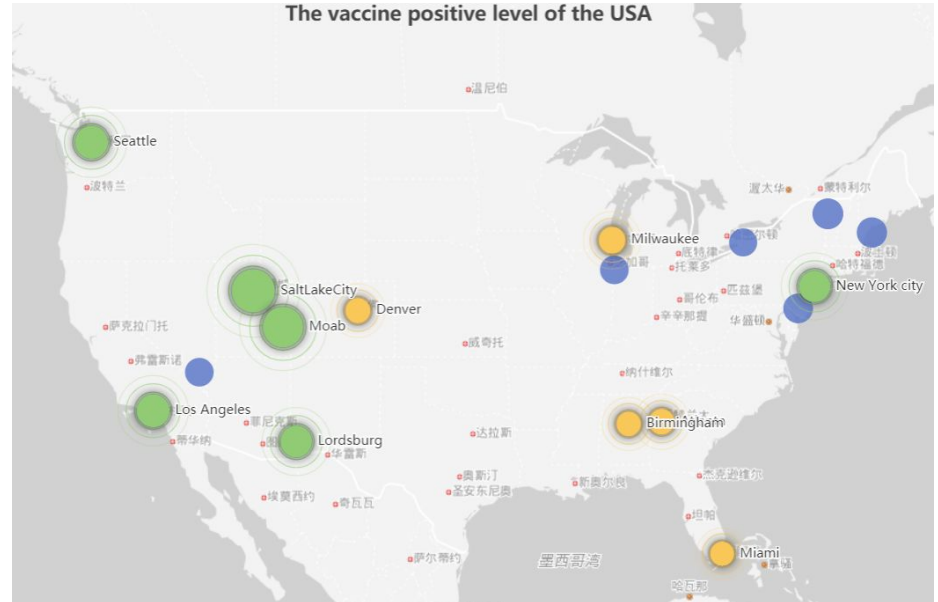
Results & Discussion

Vaccine hesitancy in the U.S. (14/04/2021 - 26/04/2021):
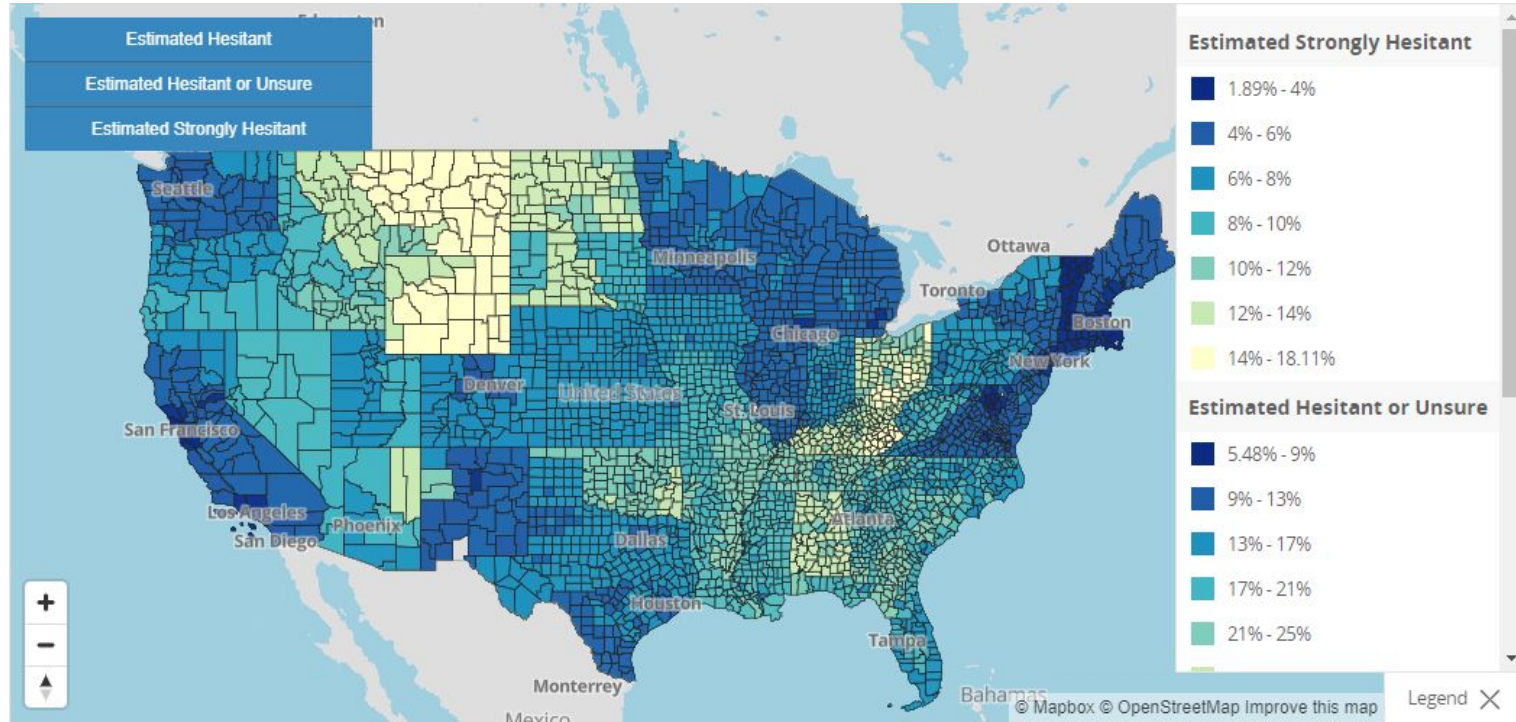
Results & Discussion

Correlation between attitude positive level towards vaccine and vaccinated rate, covid positive rate.

1. By covidactnow.org, we can get the vaccinated and covid positive rate of different regions.

2. The **highest-five covid positive** rate cities have a **mean** altitute positive rate of **0.62**; while the **lowest-five** covid positive rate cities have a **mean** attitude positive of **0.56**.

3. The **highest-five vaccinated rate citie**s have a mean altitute positive level of **0.55**; while the lowest-five covid positive cities have a mean attitude positive level of **0.50**.

4. It seems that the attitude positive level is positively correlated with both covid positive vaccinated rate and covid positive rate.

	area_name	rate	positive_rate	vaccinated
33	SaltLakeCity	0.796053	0.034	0.494
25	Moab	0.740000	0.096	0.526
35	Seattle	0.601643	0.026	0.548
19	LosAngeles	0.600000	0.003	0.552
17	Lordsburg	0.580645	0.011	0.412
27	NewYork	0.570470	0.005	0.645
5	Burlington	0.569606	0.016	0.553
31	Portland	0.557369	0.011	0.654
29	Philadelphia	0.546928	0.025	0.425
15	LasVegas	0.528662	0.037	0.419
7	Chicago	0.525735	0.019	0.581
11	ErieCounty	0.511798	0.016	0.563
23	Milwaukee	0.495055	0.016	0.460
1	Atlanta	0.479473	0.023	0.364
9	Denver	0.469394	0.015	0.447
13	Jackson	0.468707	0.021	0.431
3	Birmingham	0.453774	0.055	0.373
21	Miami	0.440252	0.022	0.378

EPFL

Results & Discussion

Hypothesis testing by Pearson correlation coefficient:

1. Correlation between attitude positive level towards vaccine and vaccinated rate: There is a small (0.38), and not significant (p = 0.12 > 0.05) positive correlation.
2. Correlation between attitude positive level towards vaccine and covid positive rate: There is a small (0.39), and not significant (p = 0.11 > 0.05) positive correlation.
3. The p-value is not low, so the correlation is not statistically significant.
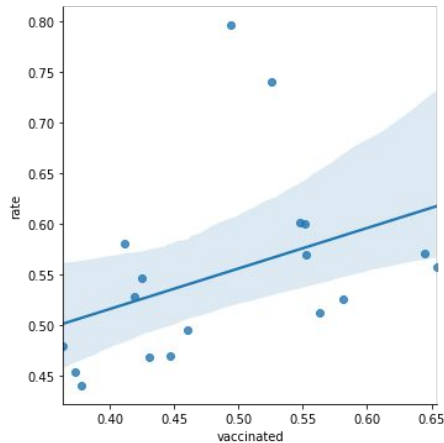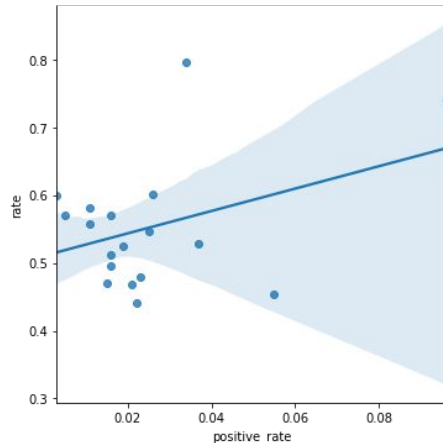
Results & Discussion

1. The three vaccines authorized by the FDA (U.S. Food and Drug Administration)
   a. Pfizer-BioNTech: **95%** efficacy in preventing COVID-19 in those without prior infection. In clinical trials, the vaccine was 100% effective at preventing severe disease.
   b. Moderna: **94.1%** effective at preventing symptomatic infection in people with no evidence of previous COVID-19 infection, and the efficacy rate drops to 86.4% for people ages 65 and older.
   c. Johnson & Johnson: **72%** overall efficacy and 86% efficacy against severe disease in the U.S.
2. Vaccines not (yet) available in the U.S.
   a. Oxford-AstraZeneca:
      i. This vaccine, which is currently being distributed in the United Kingdom and other countries, it's **cheaper** to make per dose, and it can be **stored**, transported, and handled in normal refrigeration for at least six months.
      ii. showing its vaccine to be **76%** effective at reducing the risk of symptomatic disease after receiving the two doses, and 100% against severe disease.
   b. Russia's Sputnik: According to an interim analysis of phase 3 clinical trial data published in The Lancet, the vaccine's efficacy is **91.6%**Trusted Source.
   c. China's CoronaVac: shows an effectiveness rate of just **56%-65%**.

   Data from Yale Medicine, June 2, 2021, and healthline, May 14, 2021

EPFL

Results & Discussion

Some covid vaccine news recently

AstraZeneca covid-19 vaccine may hinder blood clotting in rare cases

HEALTH 9 June 2021

SWI swissinfo.ch   Swiss perspectives in 10 languages

Malaysia says delivery of Thai-made AstraZeneca vaccines delayed

Moderna seeks COVID vaccine approval for kids 12-17

By CNN staff
Published: Jun. 10, 2021 at 5:53 PM GMT+2 | Updated: 18 minutes ago

The New York Times

AstraZeneca Vaccine and Blood Clots: What Is Known So Far

Healthcare & Pharma

WHO approves Sinovac COVID shot in second Chinese milestone

Stephanie Nebehay

BIONTECH

Pfizer and BioNTech to Provide 500 Million Doses of COVID-19 Vaccine to U.S. Government for Donation to Poorest Nations

SWI swissinfo.ch   Swiss perspectives in 10 languages

Biden to donate 500 million Pfizer doses, urge others to join in

AM Akron Beacon Journal
Nearly 200,000 COVID-19 vaccine doses could go to waste in Ohio
"The Johnson & Johnson vaccine is one of those options. It is safe and effective and only requires one shot." His office said that there are no legal options to give ...
2 天前

Global Acceptance for Sinovac's COVID-19 Vaccine Grows Despite Disparate Data

Published: Jun 04, 2021 | By Brandon May

COVID-19

Unused Johnson & Johnson Covid Doses Are Piling Up as FDA Waits to See if Shelf Life Can Be Extended

EPFL

Results & Discussion

Attitute positive level of Twitter users in the U.S.:

1. AstraZeneca ranks highest among those six vaccines, with around 0.673 attitude positive level;

2. Moderna ranks second highest among those six vaccines, with around 0.671 attitude positive level;

3. Janssen, and Sputnik rank third and fourth with 0.660 and 0.595 seperately;

4. Pflizer ranks fifth with 0.528 attitude positive level;

5. CoronaVac ranks lowest with 0.481 attitude positive level

Positive level in Different Vaccine

EPFL
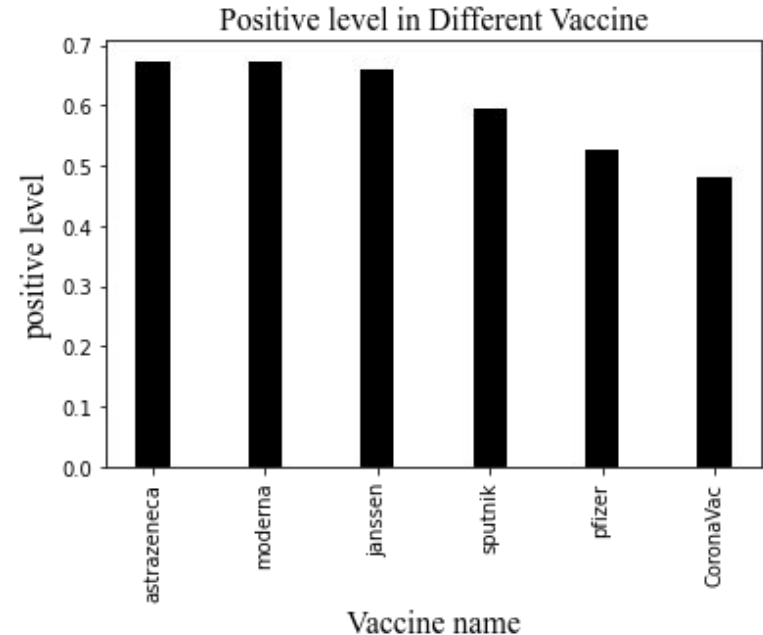
Results & Discussion

Pfizer-BioNTech:

1. Some negative remarks about pfizer:
    a. Another tone-deaf tweet from a self-entitled politician. According to the Australian Government's vaccine roll-out programme, #AnnastaciaPalaszczuk SHOULD have received the #AstraZeneca vaccine. But somehow, she qualified for the less problematic #Pfizer. Lucky her.
    b. I've got nothing nice to say either. #AnnastaciaPalaszczuk is over 50, just like me, but I only qualify for the AstraZeneca vaccine, which carries more risk. Don't tell this is blatant discrimination because it is.
    c. My reluctance to be vaccinated is simple...lack of choice...at 54 I would like the choice to receive the #Pfizervaccine...why…
    d. Has anyone experience leg pain following the 2nd dose of #Pfizer? My 13 year old is having increasing pain in his left thigh. Seems weird.
    e. Surprise the virus, mutate first!
    f. Bloody hell! Most of the people I've spoken with have had the #Pfizer jab
2. Some negative comments are not about vaccine itself.

EPFL

Results & Discussion

China's Sinovac

1. Some negative remarks about Sinovac:
    a. Any Doctor ? Yesterday my cousin got sinovac shot now today he have some allergic rashes on his face and no other reaction…'
    b. Are you getting all the info on #COVID19? If you read Reuters, don't miss out on what 15 other sources have to say #Sinovac #Coronavirus #Vaccines #WorldHealthOrganization
    c. WHO Requests that Sinovac Share more Vaccine Data to Consider EUA.
2. The data sample is only 135, too small comparing to pfizer 3371 and moderna 1307.
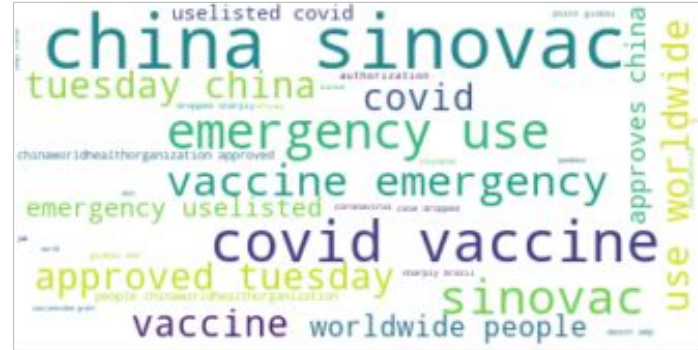
Results & Discussion

Moderna:

1. Some positive remarks about Moderna:
   a. #Moderna has recently announced their #COVID19 vaccine is safe and effective for patients aged 12-17. So relieved Mom's gotten her second shot. Doing much better with #Moderna this round. We came ready with Tylenol, Gravol, Coffee and some snacks. Just a bit of soreness.
   b. #Pfizer and #Moderna to launch new mRNA product to vaccinate against lightning strikes.'
   c. The @CDCgov recently confirmed the safety of the #Pfizer and #Moderna vaccines for expectant mothers and babies.

EPFL

Results & Discussion

Oxford-AstraZeneca:

1. Some positive remarks about Oxford-AstraZeneca:
   a. Government regulators: "Be rest assured that we have done rigorous testing for the #vaccine's safety.
   b. Thank you Denmark for pledging the direct donation of these #AstraZeneca vaccines! #Nepal will be forever grateful for this kind gesture.
   c. 'Health Canada just extended the expiration of #astrazenecavaccine by a month. Sensible. #cdnpoli #Covid19MB
   d. 'Just got my 2nd #AstraZeneca shot tonight (at a Shoppers near home). But this is good info for those in the hunt!!!! https://t.co/fLiZ8VNU0H'
   e. Sure glad I got my second doze of #astrazenecavaccine yesterday at Shoppers Osborne Village! #GetVaccinated 🤟
   f. Great joy! Today at 12:24pm I had my 2nd #AstraZeneca Covid vaccination jab! I was due to have it on 12th June but yesterday I was summoned to have an earlier appointment. Congratulations to our Glorious Leader Boris Johnson and unassailable Health Secretary Matt Hancock! 👋
2. Many positive remarks are not about the vaccine itself, but about donation.

Conclusion

1. The public hecitancy towards to COVID-19 vaccine due to various internal and external reasons, and their different attitudes towards the vaccine vary from person to person. This requires government policy makers and medias to adopt more appropriate measures to build COVID-19 vaccination trust amont the general public through public messages.
2. Limitation:
    a. Due to the limitation of Tweet API and computing power, we only selected a short period (2021-May-29 to 2021-June-6) of time, but did not choose a longer period of time.
    b. We only sampled 18 regions in the U.S.A., rather than the whole U.S.A..
    c. As for some regions and vaccines, there are only a few hundred tweets. Many of these tweets might be retweets. A large amount of similar tweet might affect the accuracy of the results.
    d. Some remarks are not about the attitude towards the vaccines themselves, might be affected by some specific events about the COVID-19 vaccine.

EPFL

References

1. Sallam, M. (2021). COVID-19 Vaccine Hesitancy Worldwide: A Concise Systematic Review of Vaccine Acceptance Rates. Vaccines, 9(2), 160. https://doi.org/10.3390/vaccines9020160
2. Khubchandani, J., Sharma, S., Price, J.H. et al. COVID-19 Vaccination Hesitancy in the United States: A Rapid National Assessment. J Community Health 46, 270–277 (2021). https://doi.org/10.1007/s10900-020-00958-x
3. Analyzing Twitter Data to Evaluate the People's Attitudes to Public Health Policies and Events in the Era of COVID-19. https://assets.researchsquare.com/files/rs-48957/v1_stamped.pdf
4. Vaccine Hesitancy on Social Media: Sentiment Analysis from June 2011 to April 2019. https://www.mdpi.com/2076-393X/9/1/28/pdf
5. Twitter text analytics reveals covid-19 vacine hesitancy tweets have crazy traction. https://dataconomy.com/2020/12/twitter-text-analytics-reveals-covid-19-vaccine-hesitancy-tweets-have-crazy-traction/
6. Global attitudes : COVID-19 vaccines. https://www.ipsos.com/en/global-attitudes-covid-19-vaccine-january-2021
7. Growing Share of Americans Say They Plan To Get a COVID-19 Vaccine – or Already Have. https://www.pewresearch.org/science/2021/03/05/growing-share-of-americans-say-they-plan-to-get-a-covid-19-vaccine-or-already-have/
8. WHO Coronavirus (COVID-19) Dashboard https://covid19.who.int/
9. Estimates of vaccine hesitancy for COVID-19 https://data.cdc.gov/stories/s/Vaccine-Hesitancy-for-COVID-19/cnd2-a6zw/
10. U.S. COVID Risk & Vaccine Tracker https://covidactnow.org/?s=1918836

EPFL

Q&A