COVID-19 in the United States

By:
Elizabeth Axworthy
Cherngywh Lee
Christopher Velasquez Vonderach
Michael Wright

Scenario

Client:

- Our client is a United States federal entity concerned with disease control and prevention looking for ways to limit Covid-19 transmission and reduce future cases.

Us:

- An analysis firm hired to review data concerning COVID-19 cases, vaccination rates, and additional data such as public perception and sentiment.

Project Specifications

Problem:

- What are the factors leading to higher amounts of confirmed COVID-19 cases?

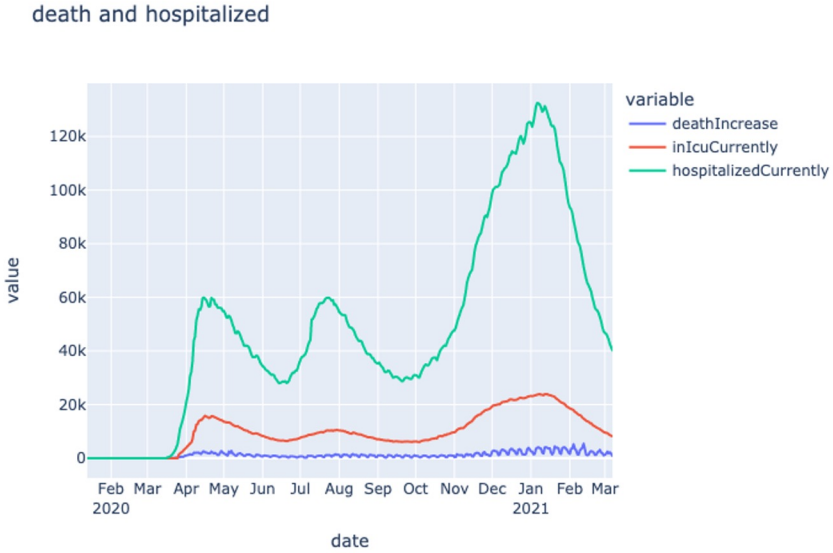
Hypotheses:

- higher vaccination rates = fewer confirmed cases
- more positive vaccine sentiment = higher vaccination rates
- more positive vaccine sentiment = fewer confirmed cases

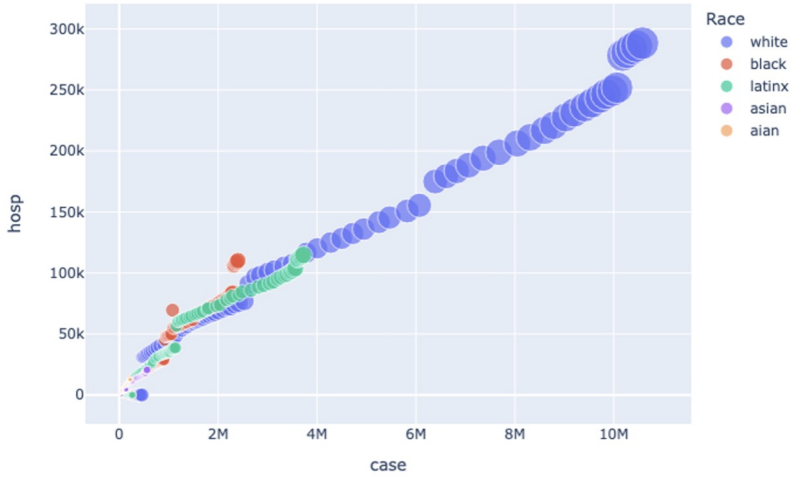
Data Observations

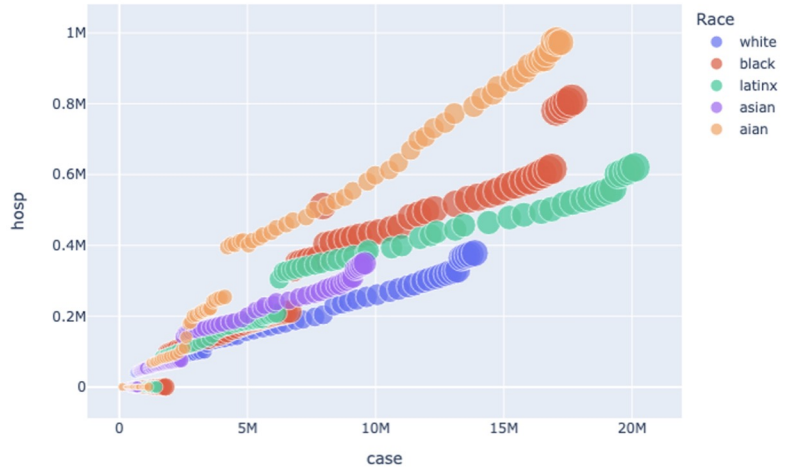
- No correlation between cases per capita and state population.
- Vaccine sentiment in U.S. leans more positive.
- More likely for a negative sentiment Tweet to be re-tweeted than positive sentiment Tweet.
- Higher raw numbers of vaccinations and COVID-19 cases correspond with population numbers.
- Higher amounts of vaccinations per capita follow political borders.

1st YR of Covid '20 data



- Death, hospitalized and ICU rate kept increasing before vaccine was available. Cases surged in Oct while Delta variant outbreak
- Most white people got infected, but higher proportion of black and Latino got infected



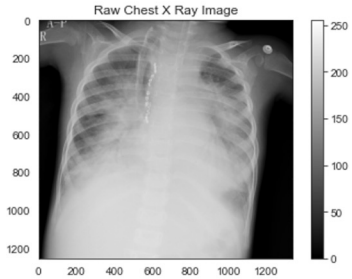


1st YR of Covid - CNN

Create a Deep Learning model to classify X-Ray image to quickly determine it's Normal or Pneumonia.

Normal

Pneumonia



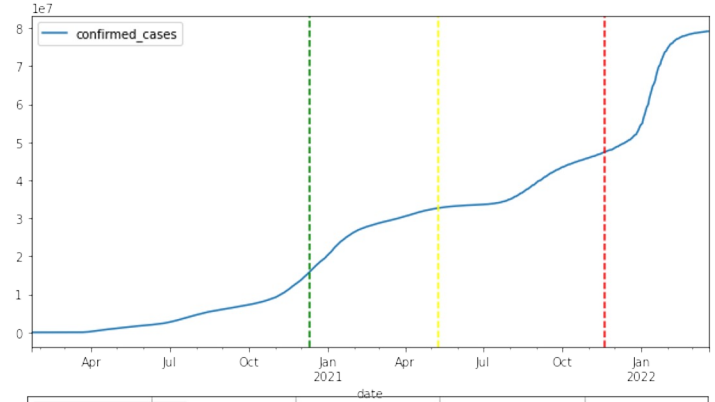
Train set: PNEUMONIA = 3875, NORMAL = 1341
Test set: PNEUMONIA = 390, NORMAL = 234
Validation set: PNEUMONIA = 8, NORMAL = 8

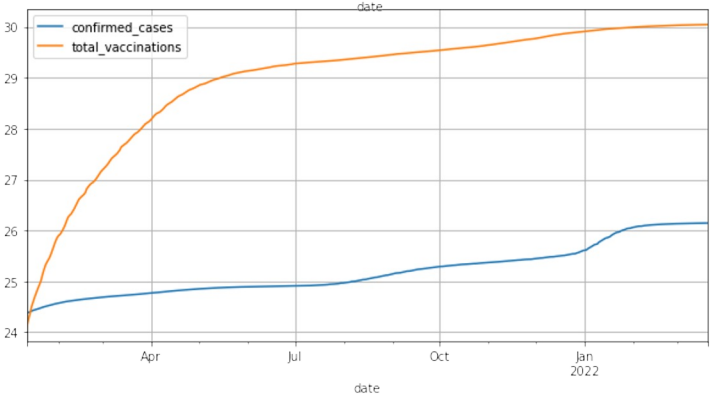
We also applied transfer learning by using champion models:

- **DenseNet:** utilizes dense connections between layers, through Dense Blocks, where we connect all layers (with matching feature-map sizes) directly with each other.
- **Inception-V3:** makes several improvements including using Label Smoothing, Factorized 7 x 7 convolutions, and the use of an auxiliary classifier to propagate label information lower down the network
- **EfficientNetB4:** a novel model scaling method that uses a simple yet highly effective *compound coefficient* to scale up CNNs in a more structured manner.

Model	Precision	Recall	F1	Accuracy
Sequential	76.5%	67.3%	67.9%	74.6%
DenseNet	81.6%	51.5%	41.6%	63.6%
Inception-V3	86.3%	84.6%	85.3%	86.5%
EfficientNetB4	67%	55.2%	39%	44.4%

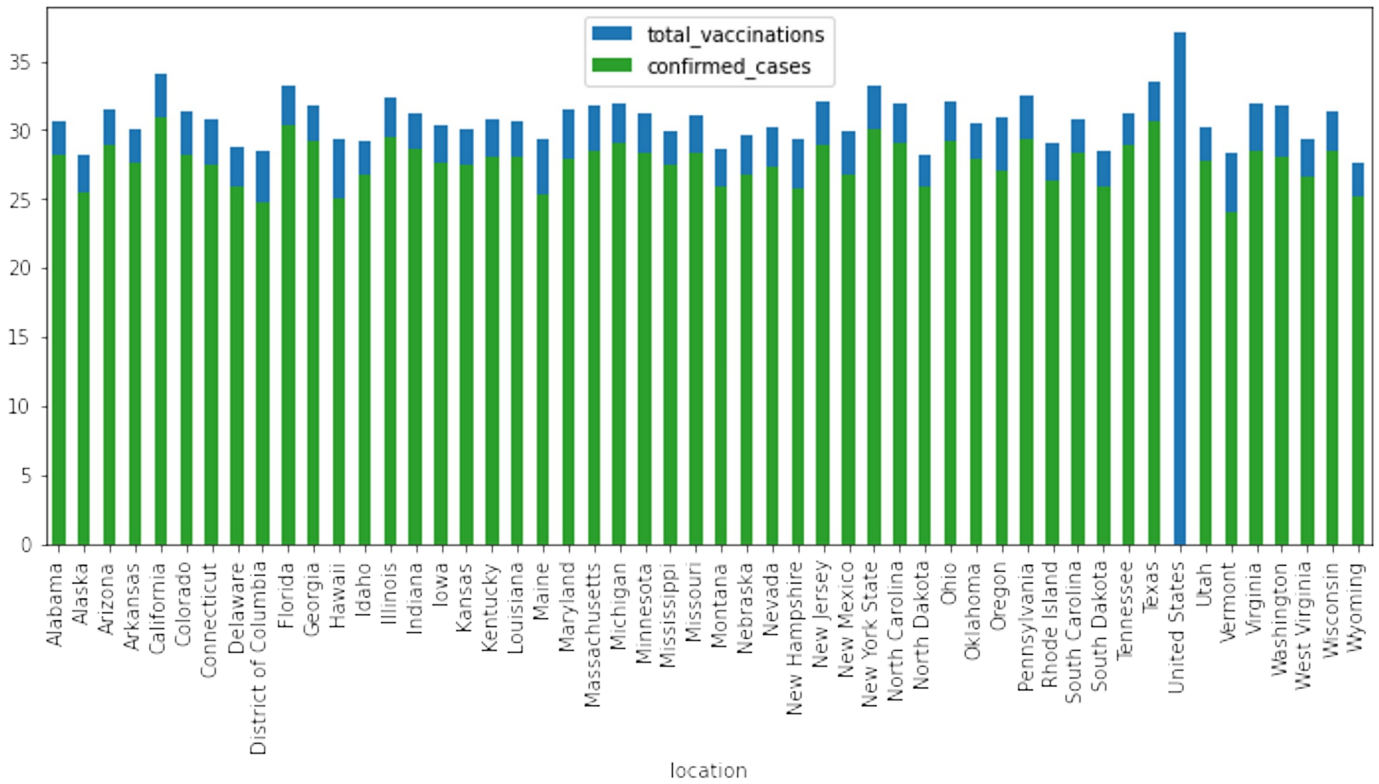
Cases and Vaccination Rates



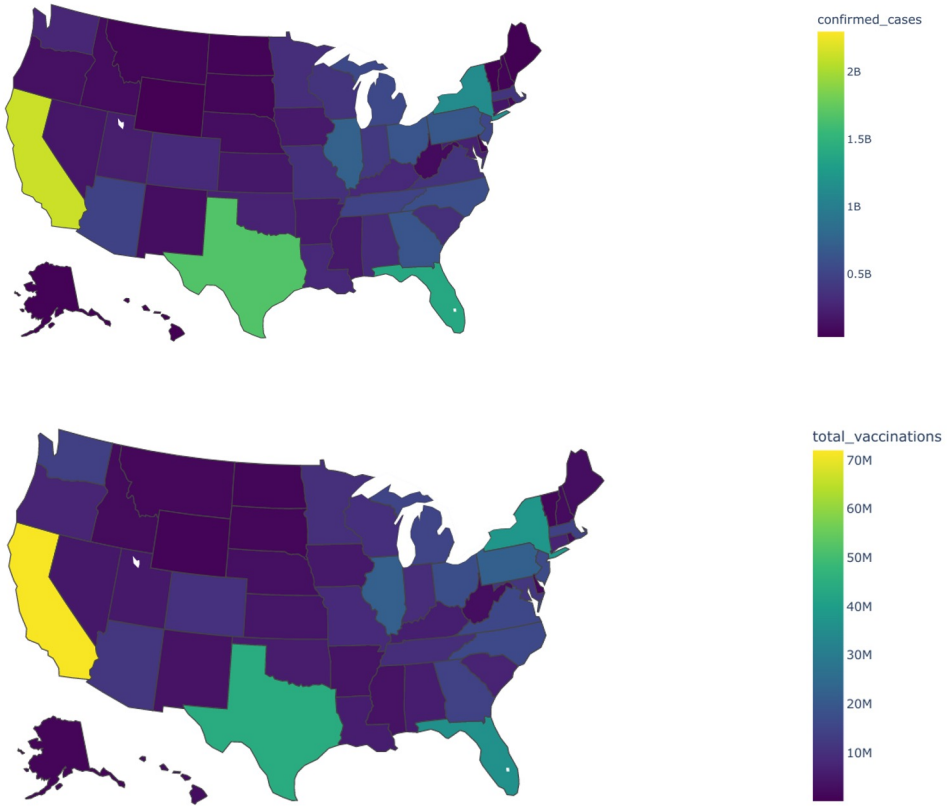


- The introduction of vaccinations lowered the rate of increase in total accumulated US COVID cases.
- Availability of boosters didn't have as strong an effect.
- Noticeable jump in cases January.
- Total vaccination count far exceeds confirmed case count, even logged.

Cases and Vaccination Rates

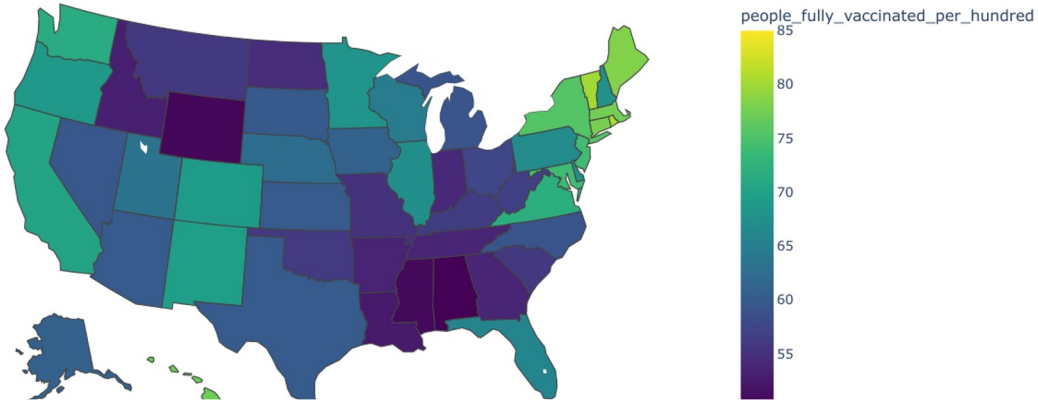


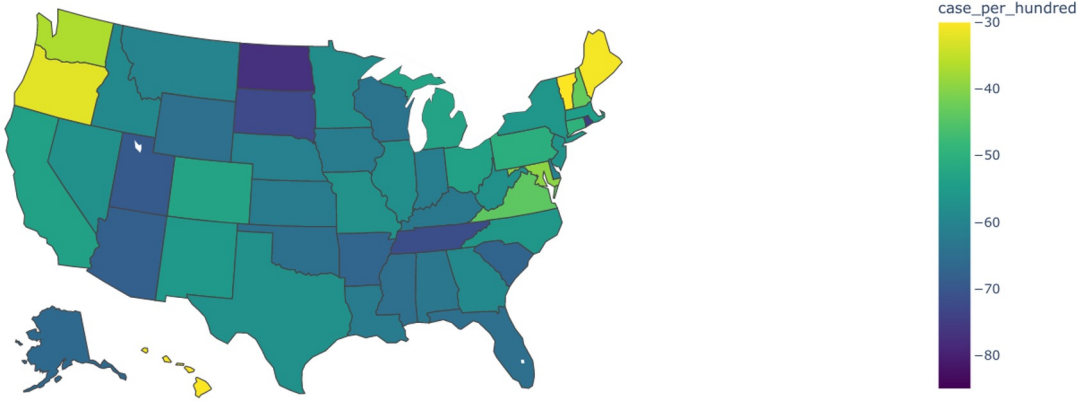
Cases and Vaccination Rates



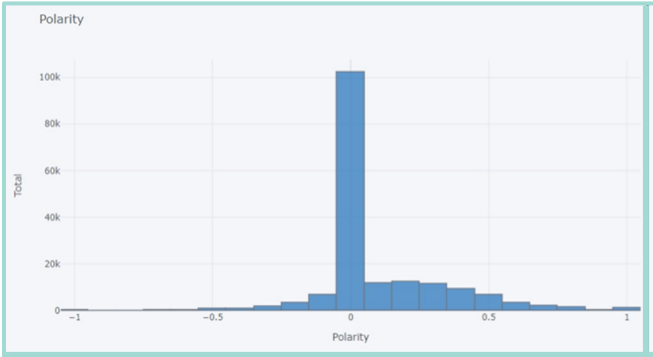
- The map of confirmed cases is visually similar to total vaccinations, though in confirmed cases Texas and Florida are closer to California than they are when looking at vaccination counts.

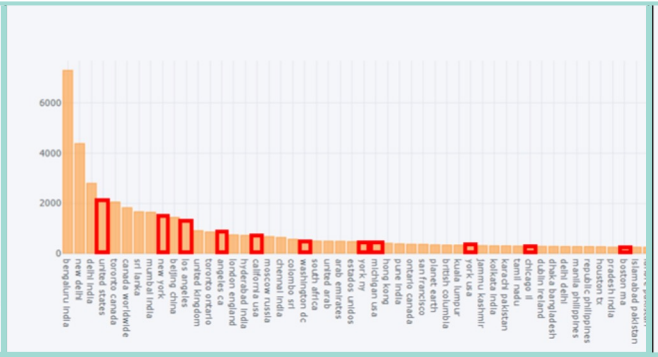
Cases and Vaccination Rates

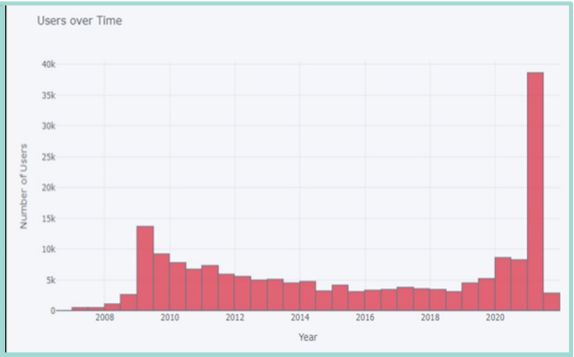




Sentiment Analysis Findings





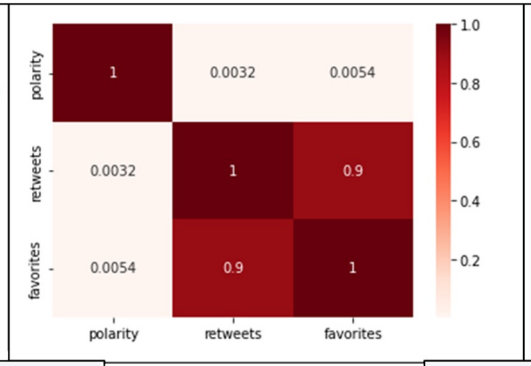


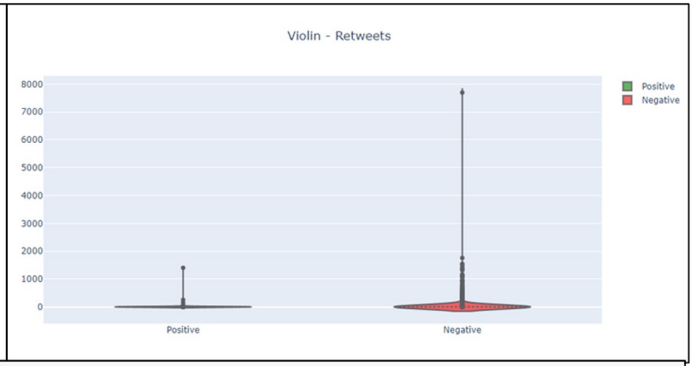
- Polarity from TextBlob Package
- ~45% of tweets as 0.
- Majority of other tweets are positive sentiment.
- Cluster from 0.05 - 0.5

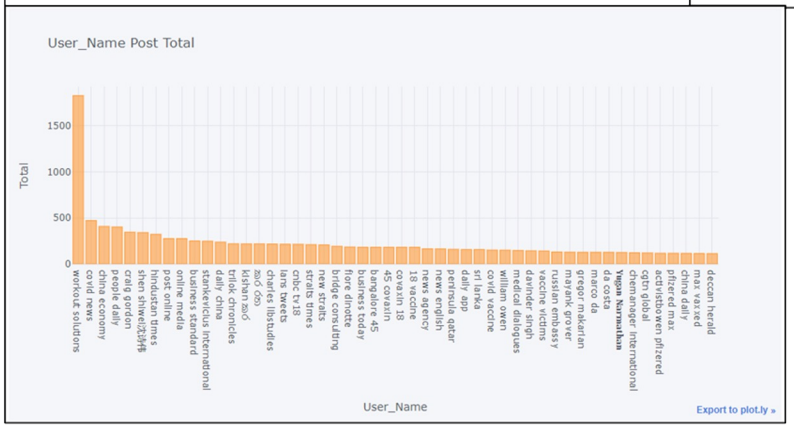
- Data set is not limited to US
- Data did not provide uniformity for location
- ~8,100 tweets are from US that are depicted.

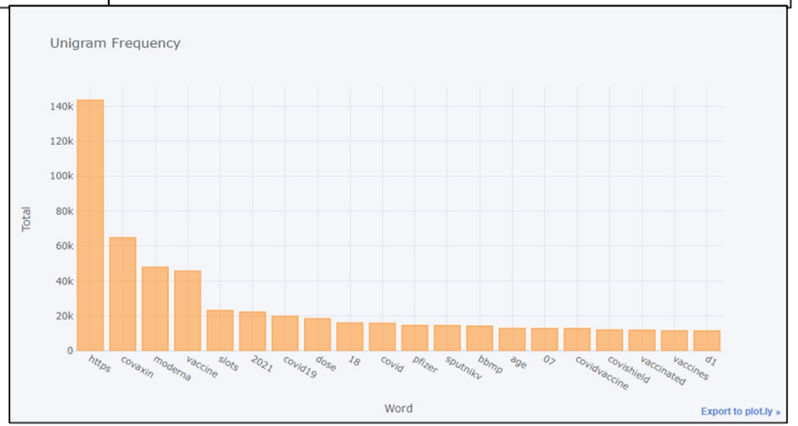
- Downward trend of users created from 2009-2020.
- Highest spike of ~40K accounts from Jan-Jul 2020.
- That spike of accounts makes up ~17% of the dialogue

Sentiment Analysis Findings - Continued









Recommendations to Decrease Risk of Infection

- Be wary of Covid-19 info on social media; use reliable sources for information pertaining to Covid-19
- Get vaccinated
- Wear a well-fitting mask
- Social distance when possible
- Practice good hygiene

Thank you!