Sentiment analysis of Twitter comments related to COVID-19

Group 5

Dai Dong, Sijia Han, Frank Shi
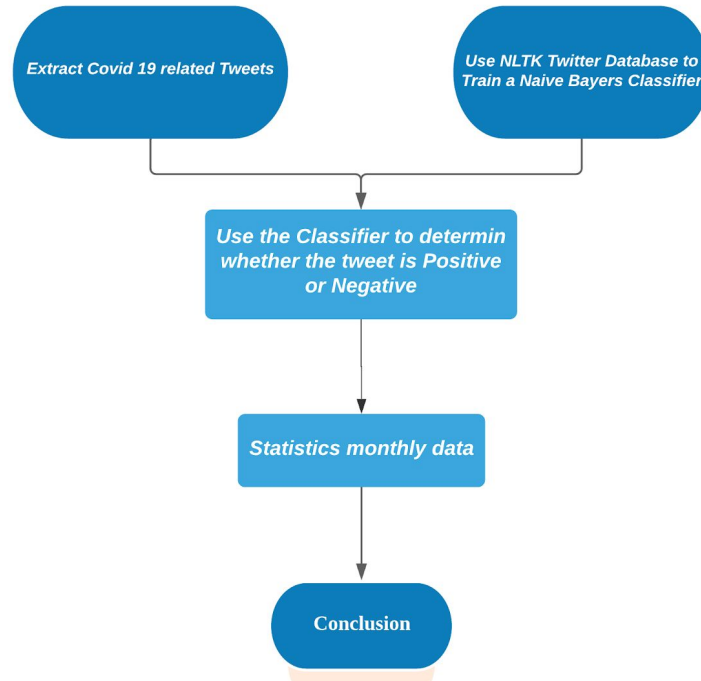
Sentiment Analysis Approaches

machine learning approach

lexicon based approach

hybrid approach

Our Main Approach

Extract Covid 19 related Tweets

Use NLTK Twitter Database to Train a Naive Bayers Classifier

Use the Classifier to determin whether the tweet is Positive or Negative

Statistics monthly data

Conclusion

The N-gram Algorithm Accuracy Comparison

Unigram	0.7133	
Bigram	0.4923	
Trigram	0.5073	
Hybird	0.7133	

Figure 2. The accuracy of the classifier

Average Positive/Negative rate of tweets related Covid-19
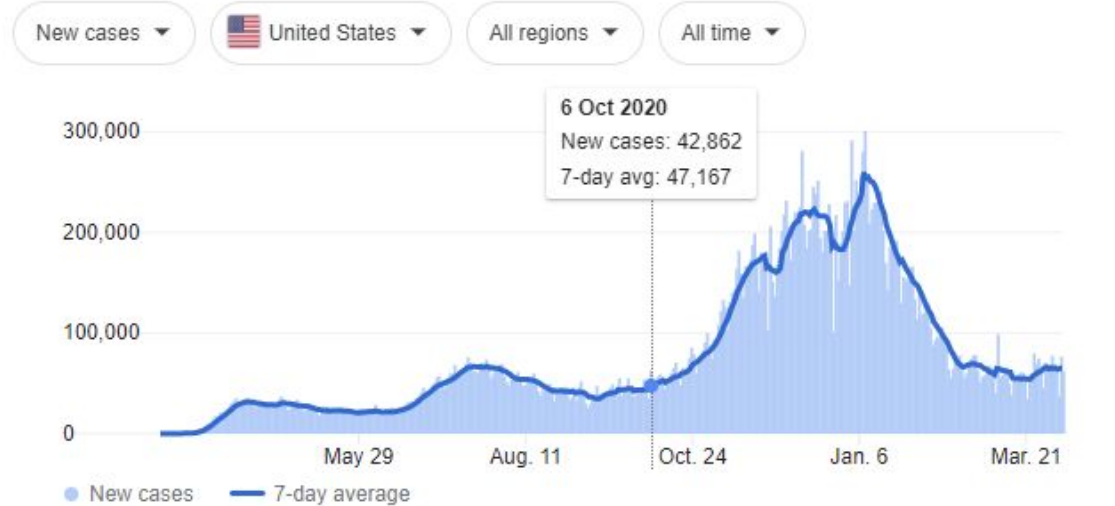
	Positive	Negative
2020-08	0.6755	0.3245
2020-09	0.6738	0.3262
2020-10	0.6787	0.3213
2020-11	0.6625	0.3375
2020-12	0.6548	0.3452
2021-01	0.6495	0.3505
2021-02	0.6685	0.3315
2021-03	0.6804	0.3196
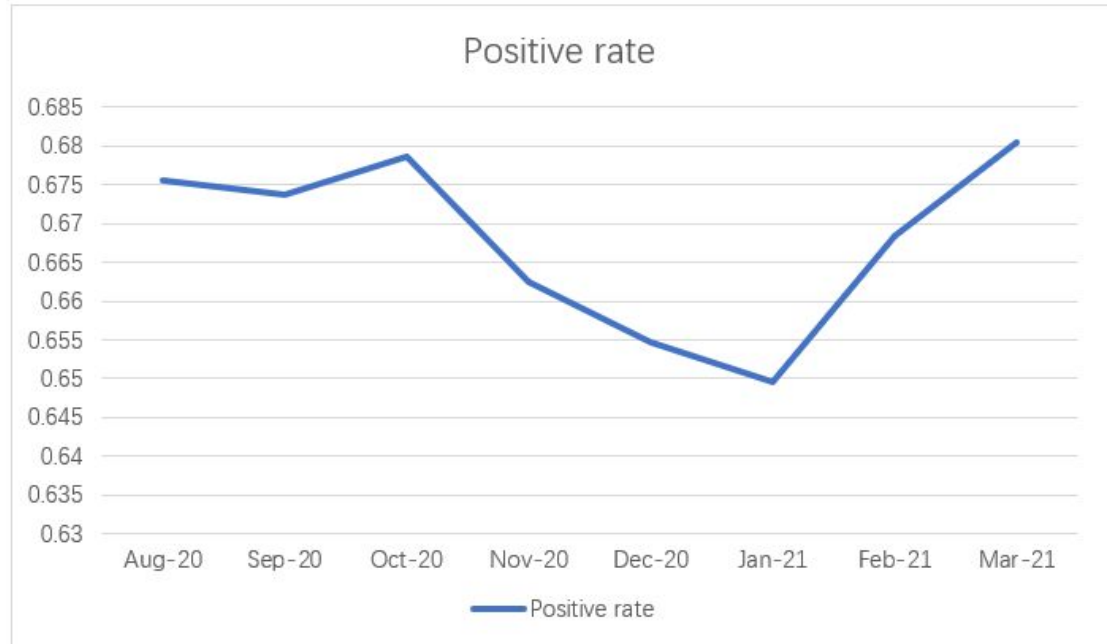
Comparison between positive/negative tweets and covid cases

	Positive	Negative
2020-08	0.6755	0.3245
2020-09	0.6738	0.3262
2020-10	0.6787	0.3213
2020-11	0.6625	0.3375
2020-12	0.6548	0.3452
2021-01	0.6495	0.3505
2021-02	0.6685	0.3315
2021-03	0.6804	0.3196

Daily change

New cases ▼   🇺🇸 United States ▼   All regions ▼   All time ▼

6 Oct 2020
New cases: 42,862
7-day avg: 47,167

300,000

200,000

100,000

0

May 29   Aug. 11   Oct. 24   Jan. 6   Mar. 21

● New cases   — 7-day average

2020 August
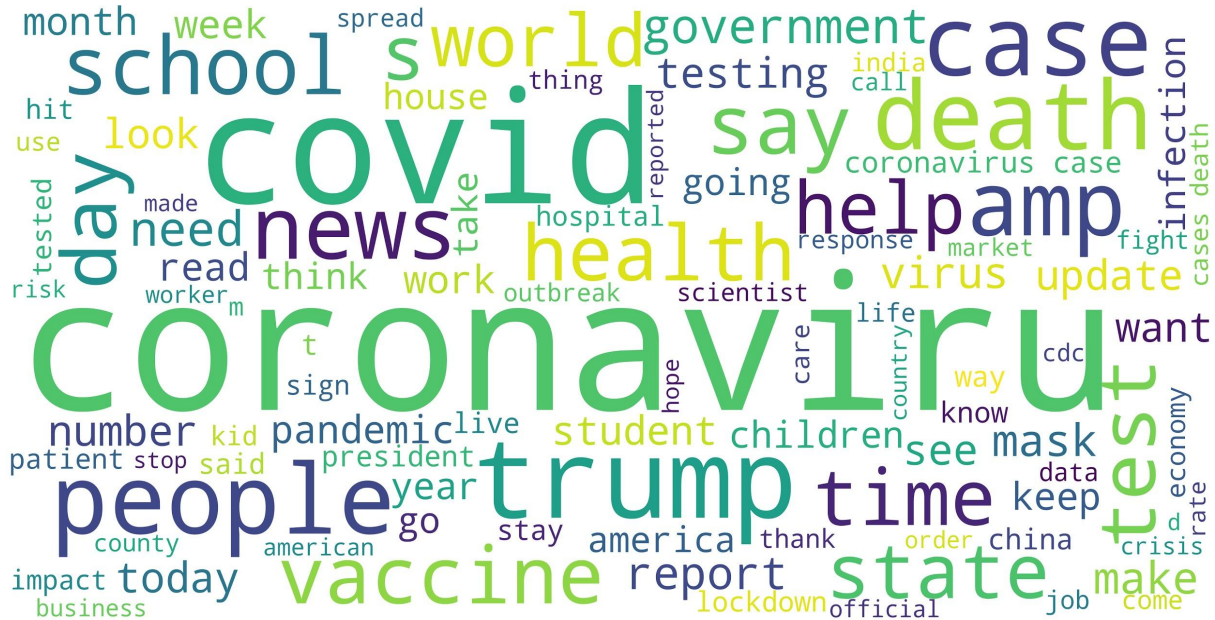
2020 September
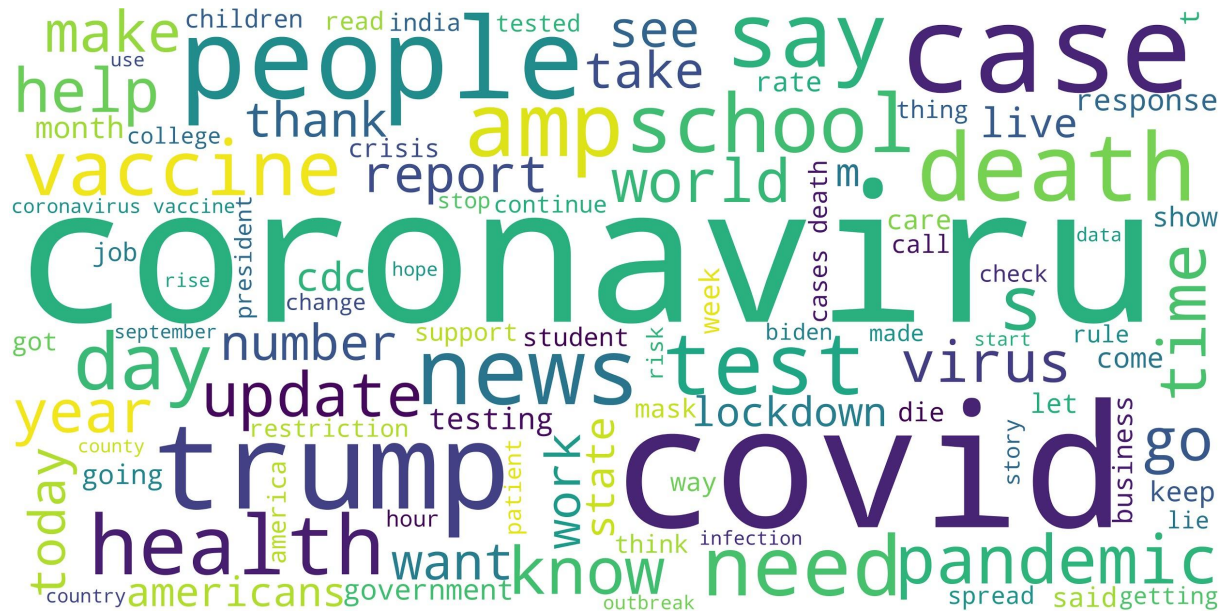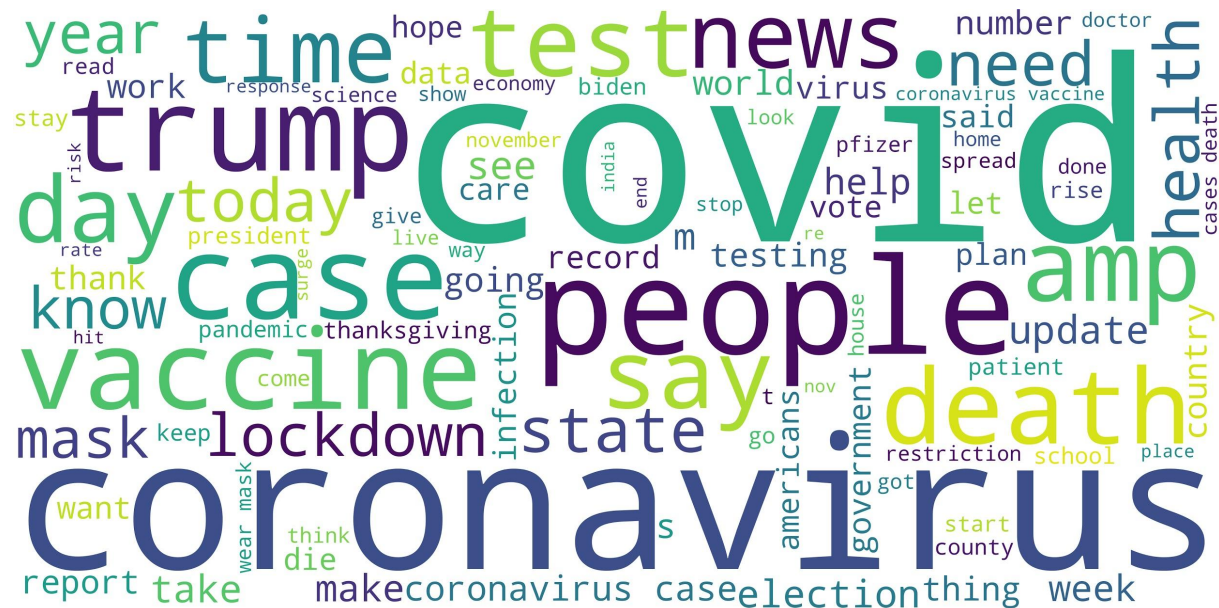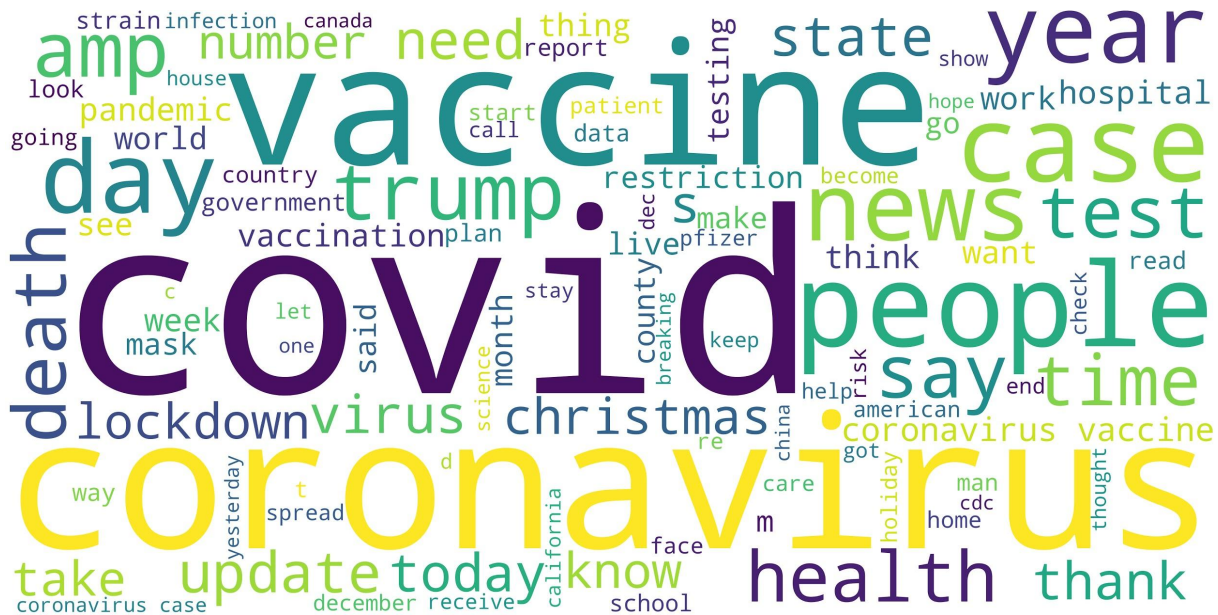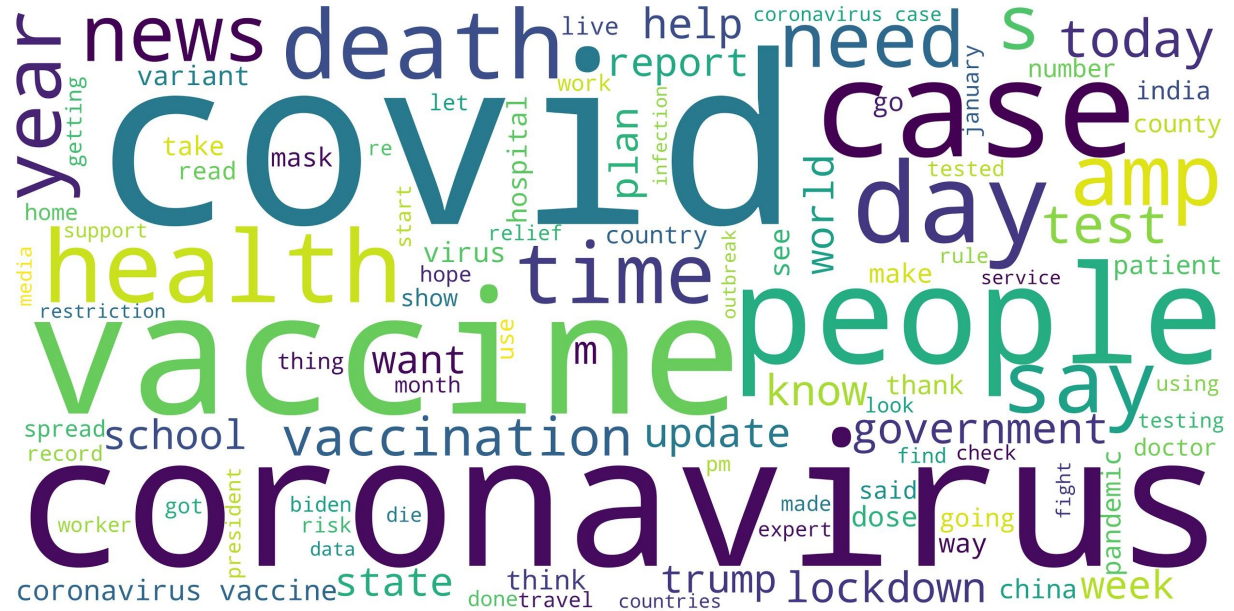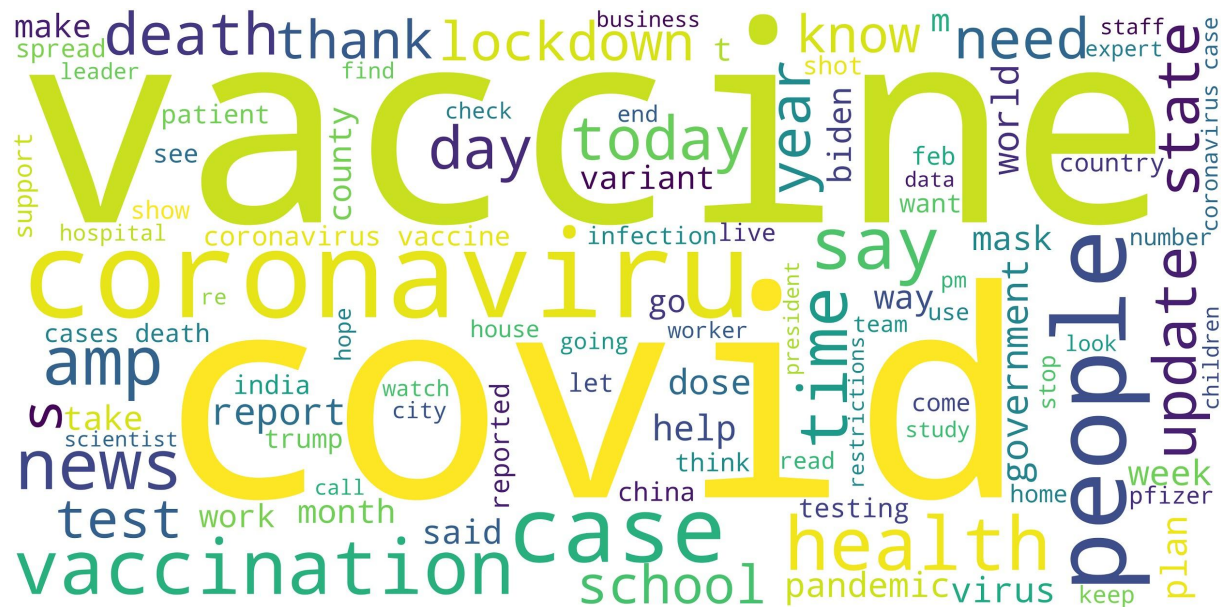
2020 October
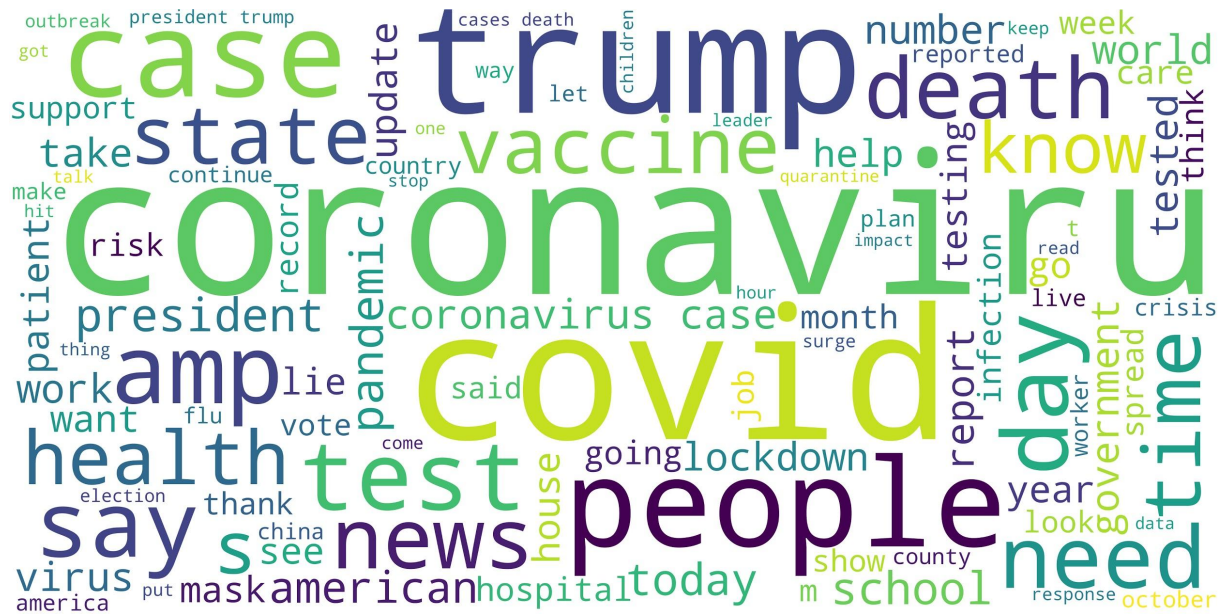
2020 November
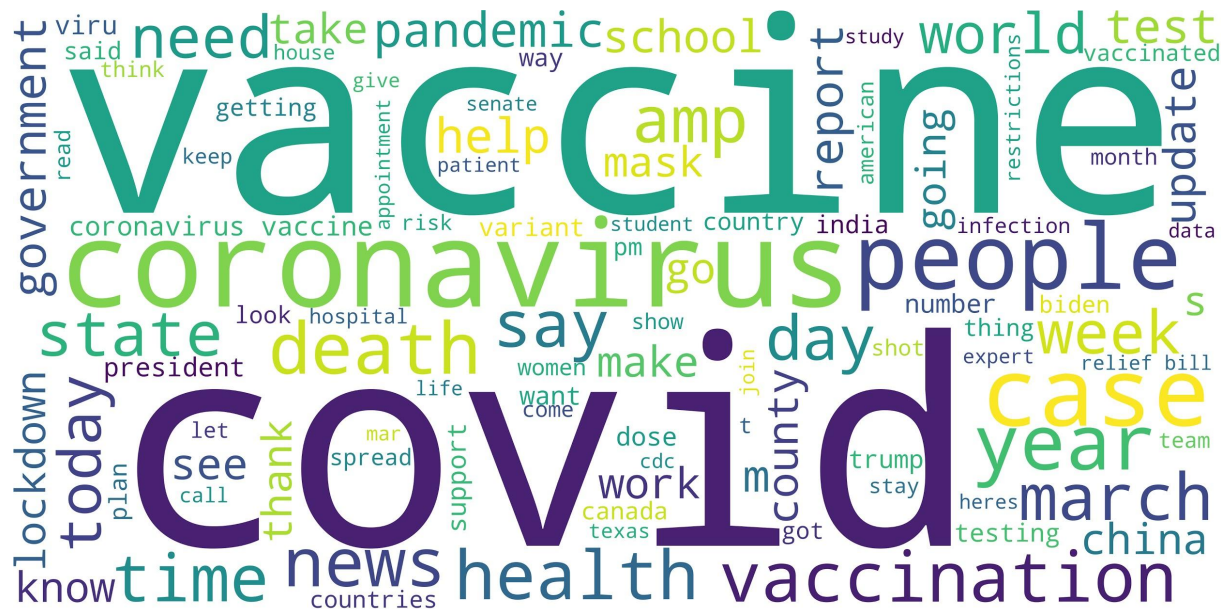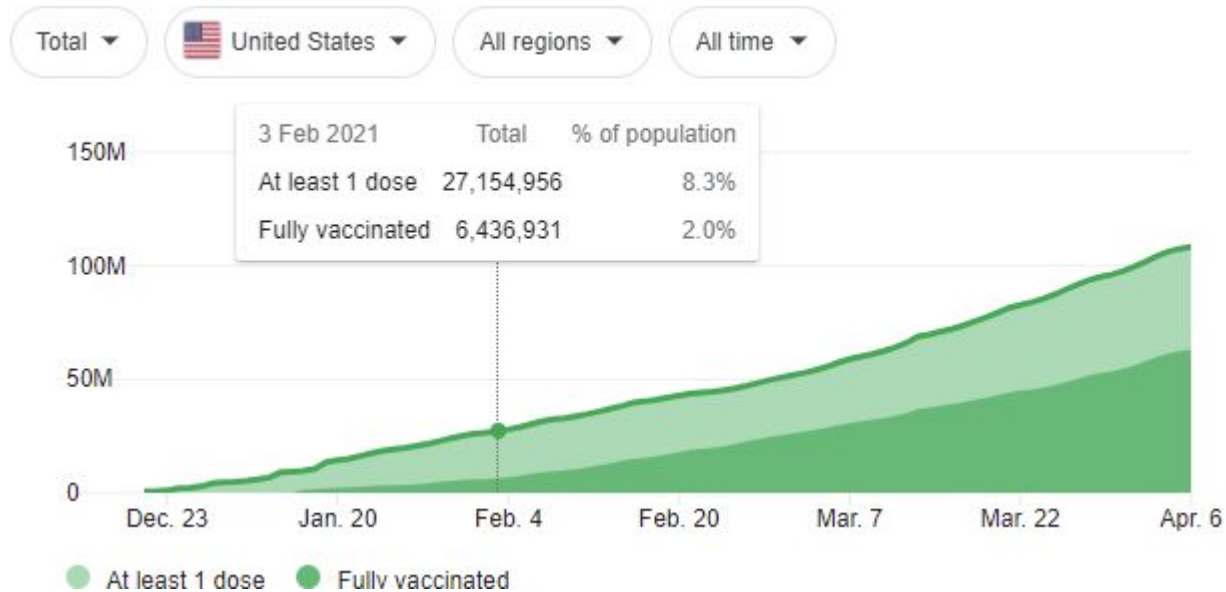
2020 December

2021 January

2021 February

2021 March

Covid Vaccine Dosing History

Thanks for listening !