

# COVID-19 Cases vs Vaccines in the United States

Rachel Goodridge



## Motivation

- To build an informative and interactive website application
- To identify target areas in the United States that could benefit from outreach programs led by public health organizations

# Data

- COVID cases per county
  - Jan 2020 – Feb 2022
  - Johns Hopkins University
  - over 2.5 million rows
- COVID vaccines per state
  - Jan 2021 – Feb 2022
  - Center for Disease Control
  - over 26k rows
- State populations
  - 2020 US Census

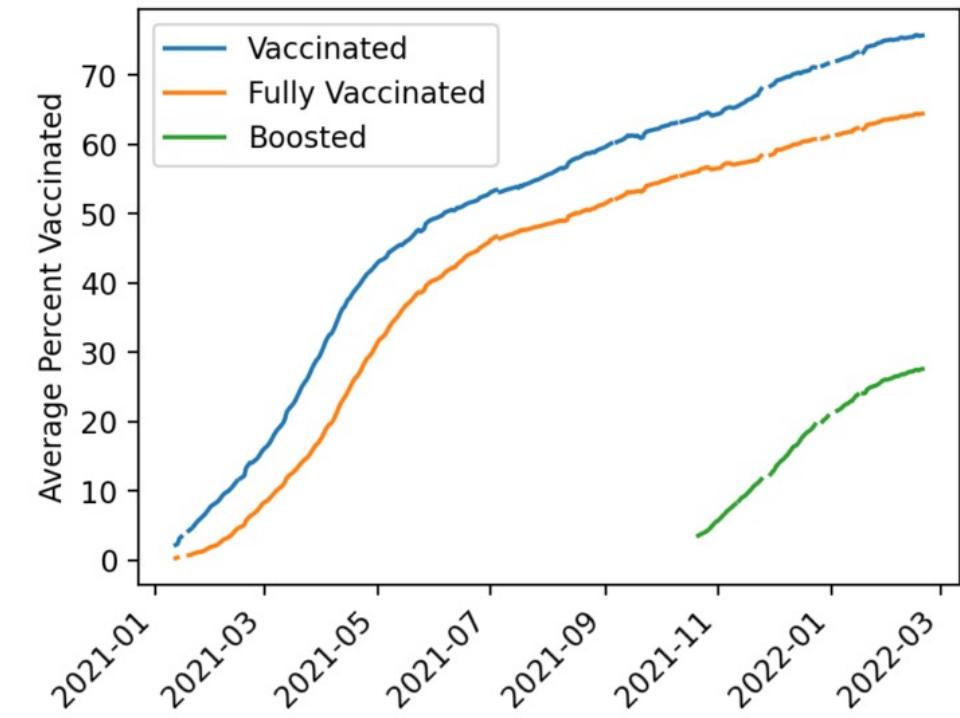
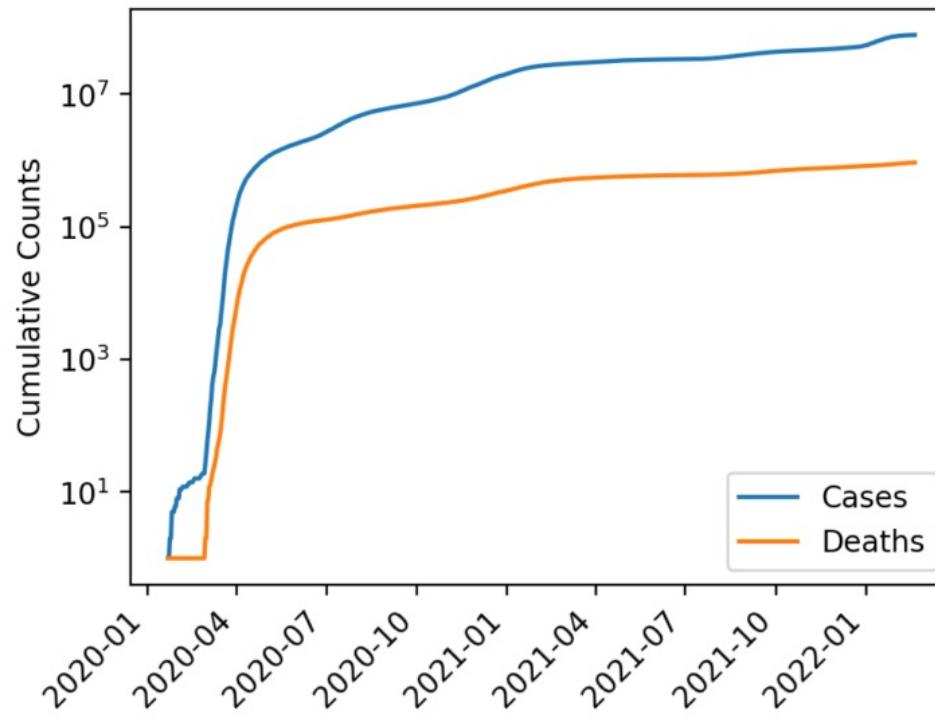
# Data

- COVID cases per state
  - Jan 2020 – Feb 2022
  - Johns Hopkins University
  - over 2.5 million rows
- COVID vaccines per state
  - Jan 2021 – Feb 2022
  - Center for Disease Control
  - over 26k rows
- State populations
  - 2020 US Census

# Tools

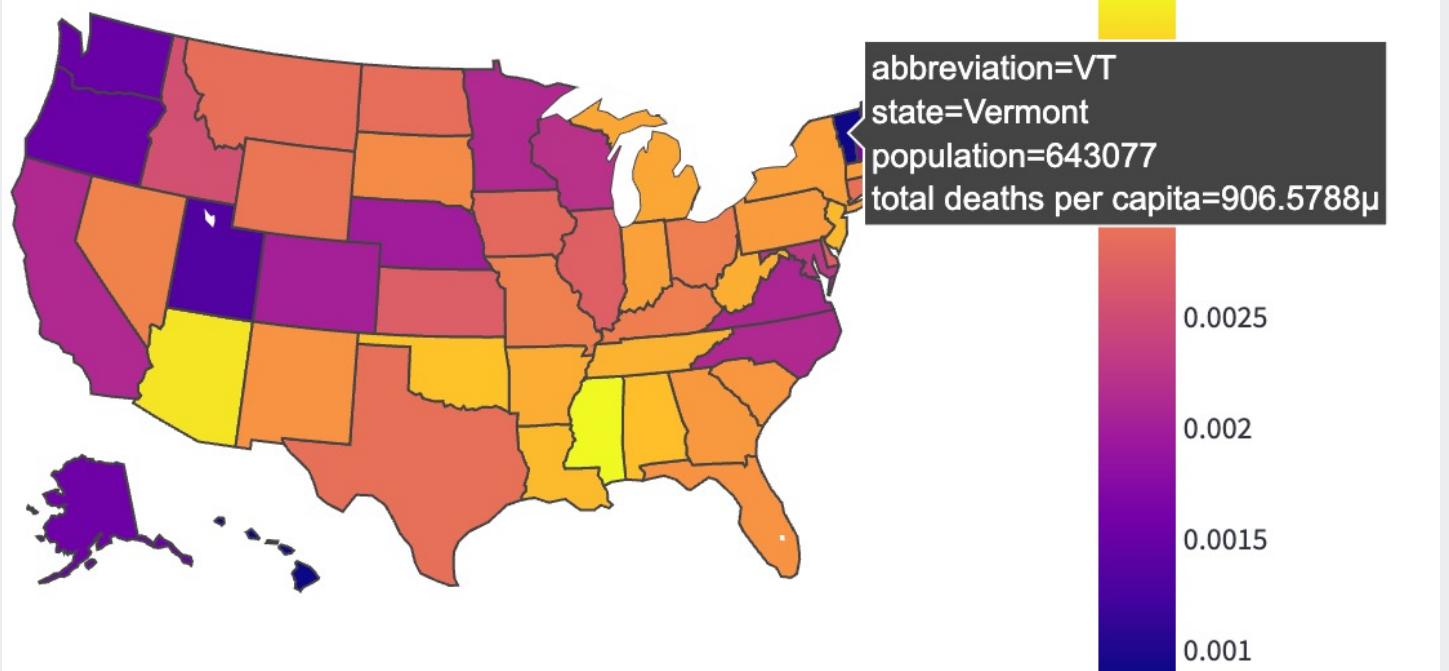
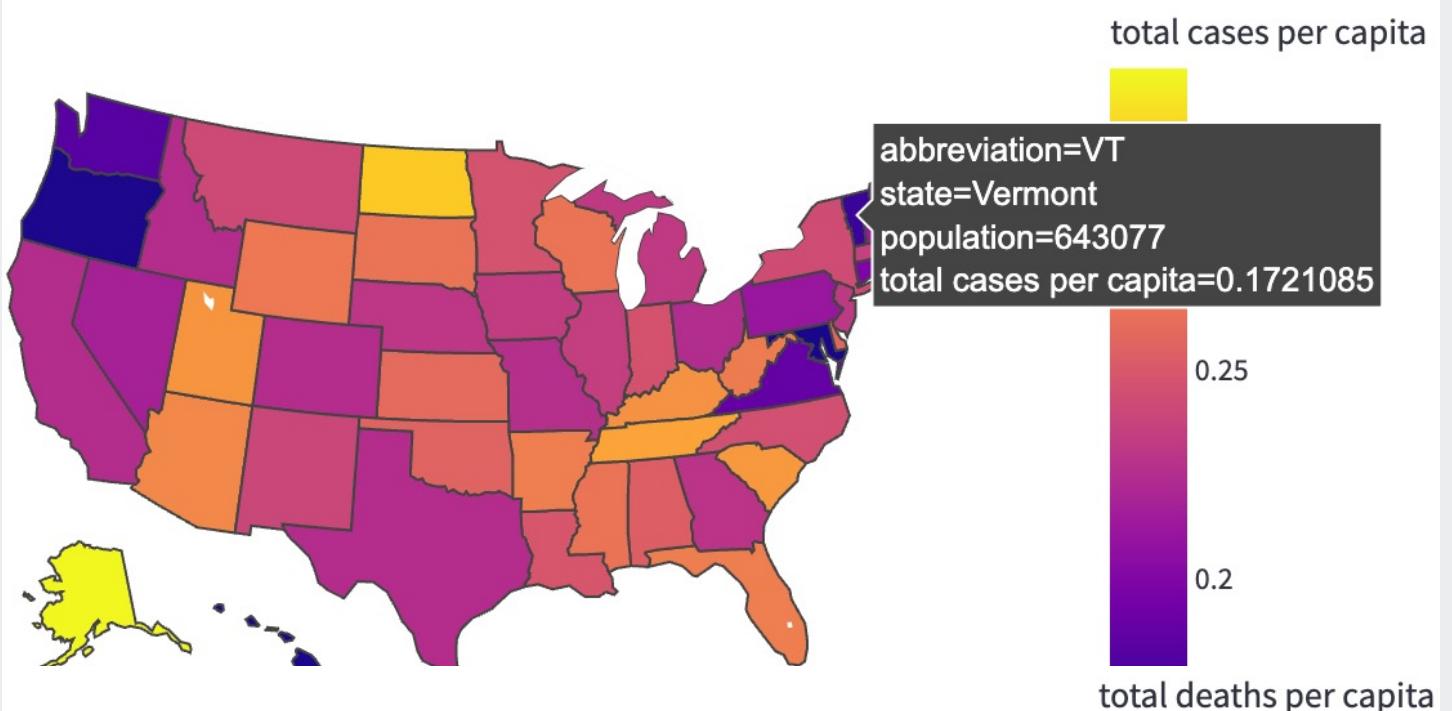
- SQL to store the data
- numpy and pandas for data manipulation
- matplotlib and plotly for visualizations
- streamlit for web app deployment

# COVID Cases Increase Faster than Vaccine Rates



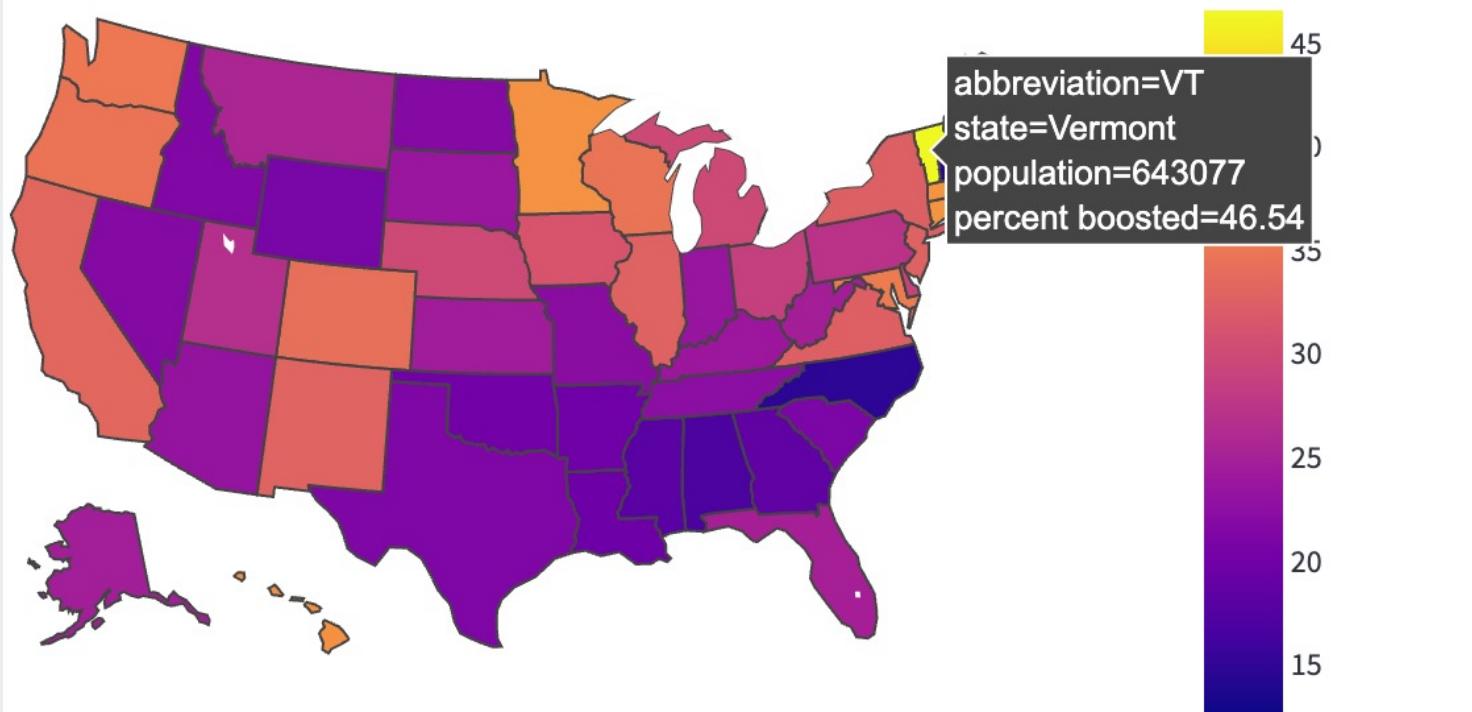
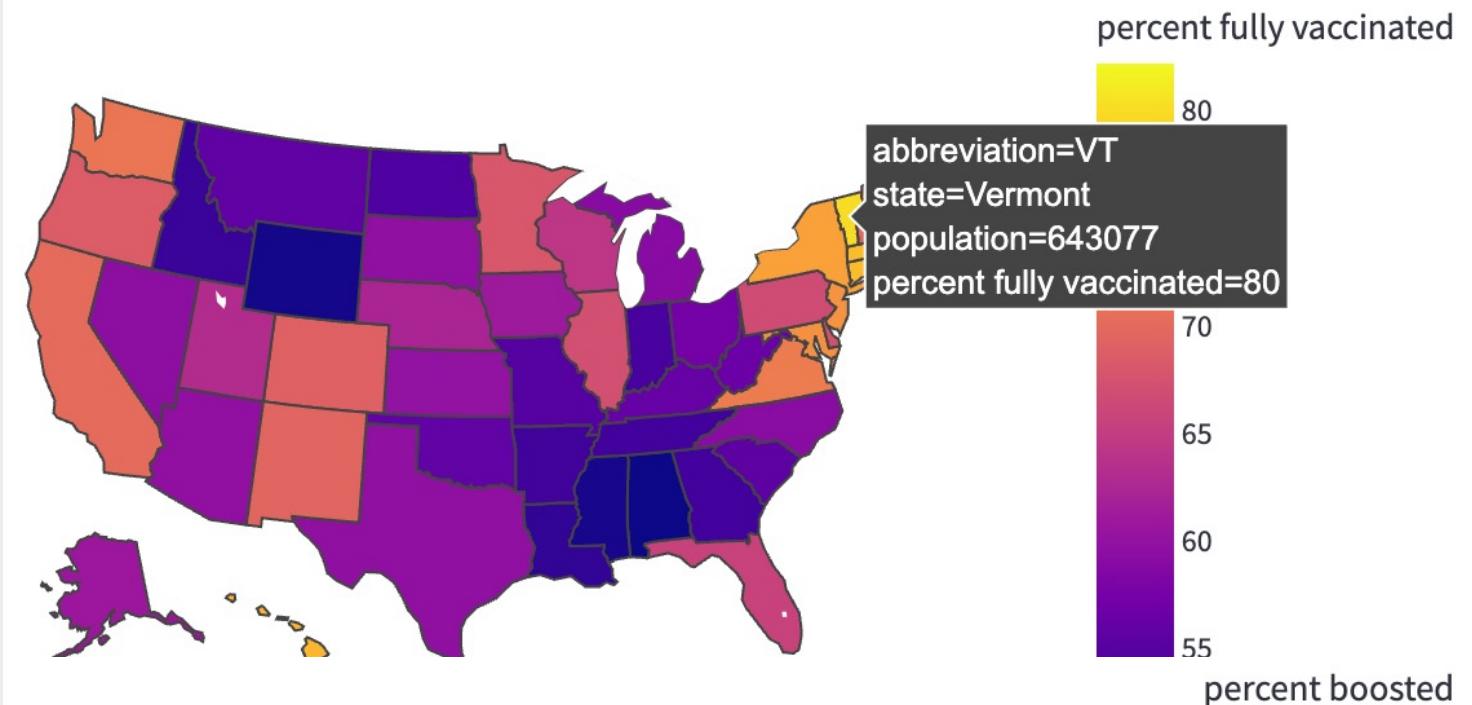
# Case Study 1 : Vermont

- Low cases per capita
- Low deaths per capita



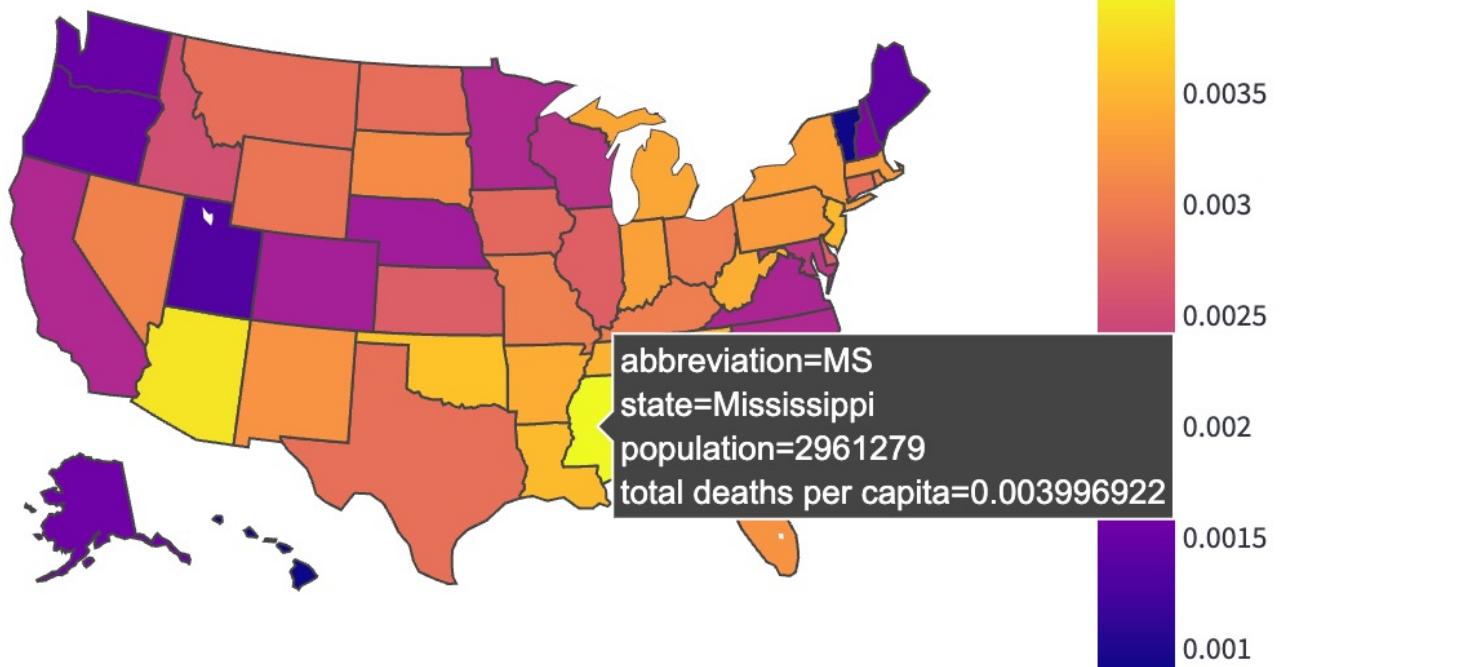
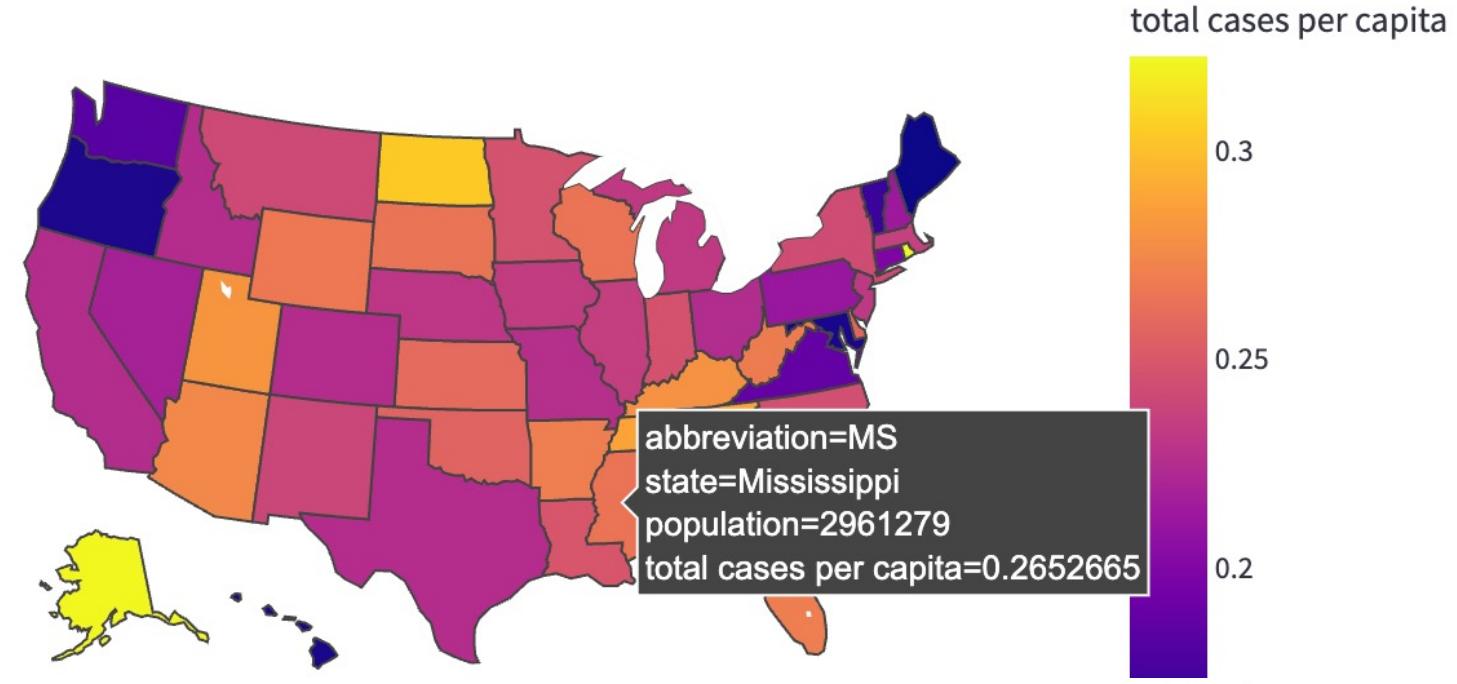
# Case Study 1 : Vermont

- High vaccination rate
- High booster rate



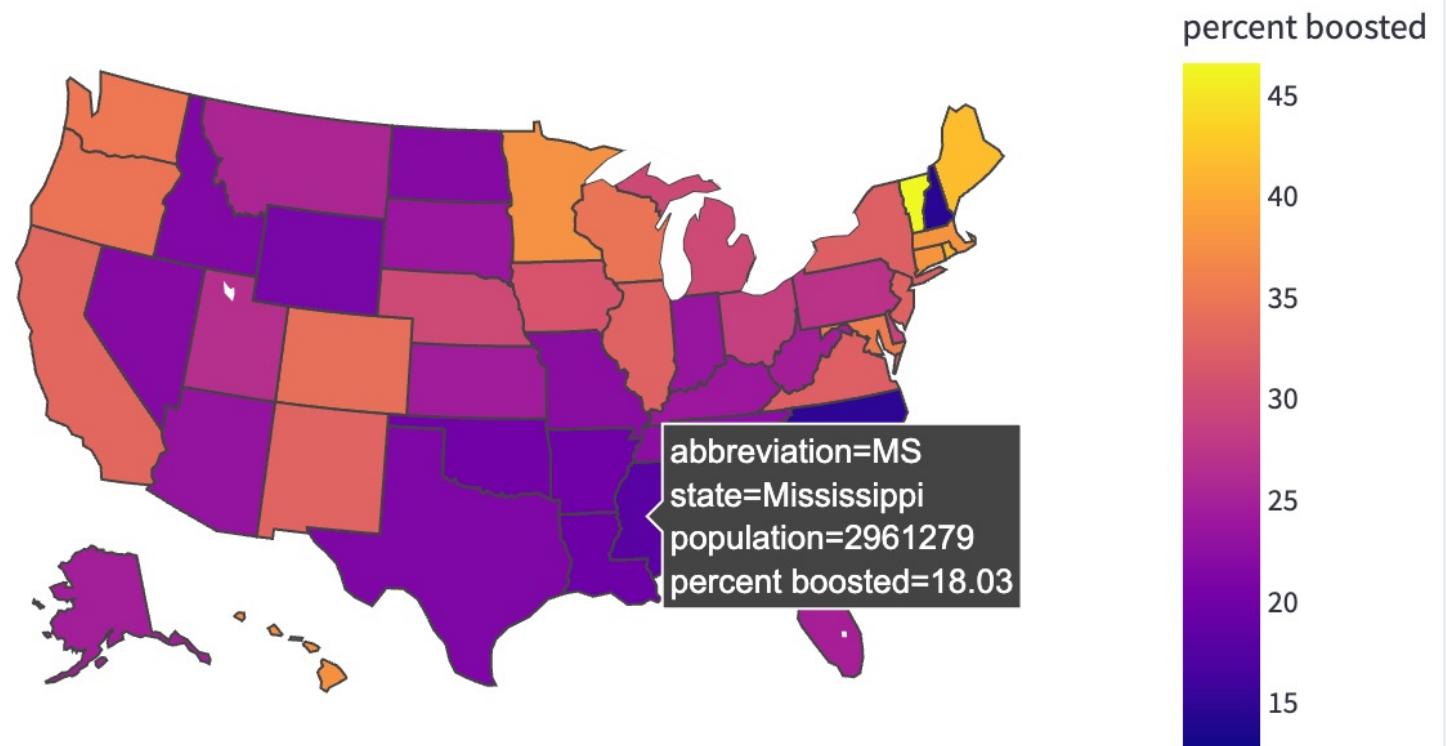
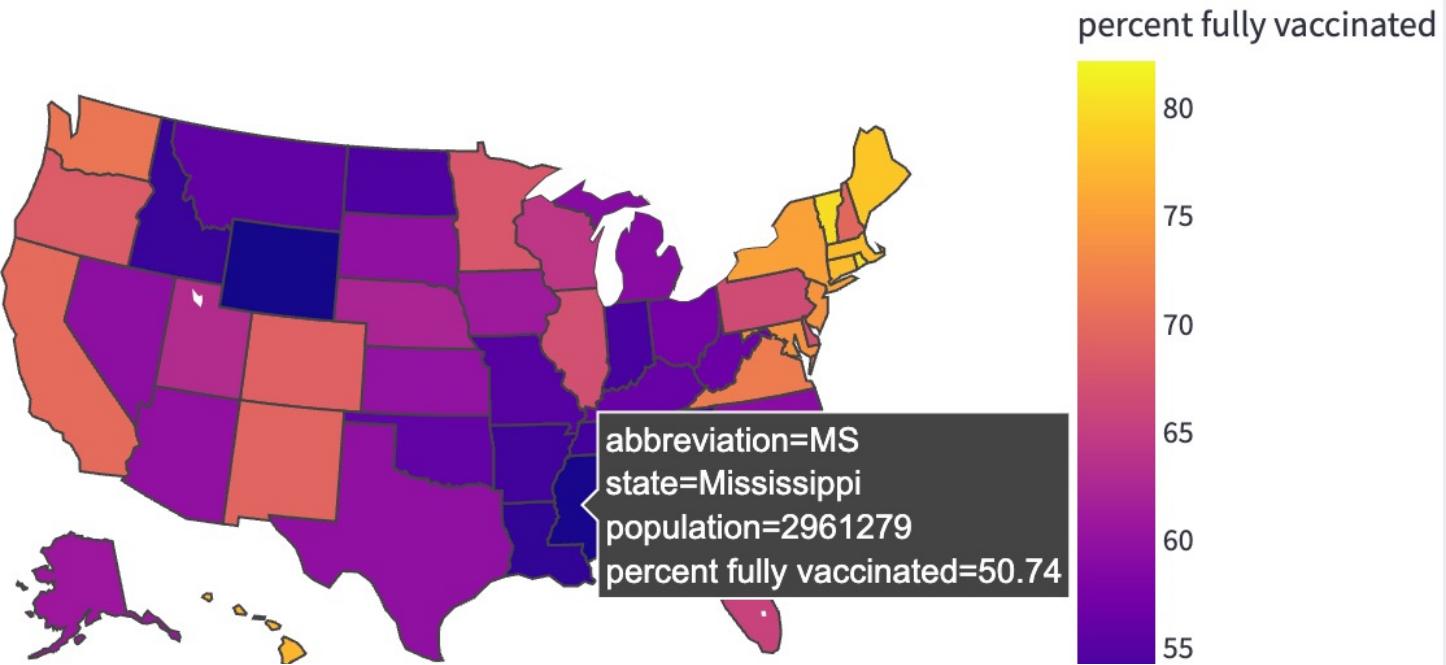
# Case Study 2 : Mississippi

- Medium-high cases per capita
- High deaths per capita



# Case Study 2 : Mississippi

- Low vaccination rate
- Low booster rate



# Website Application Deployment

## COVID-19 Cases vs Vaccines in the U.S.

Distribution of COVID cases and vaccines across the United States

Pick a month and year

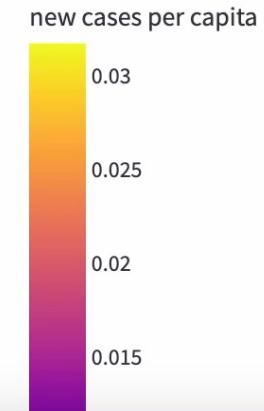
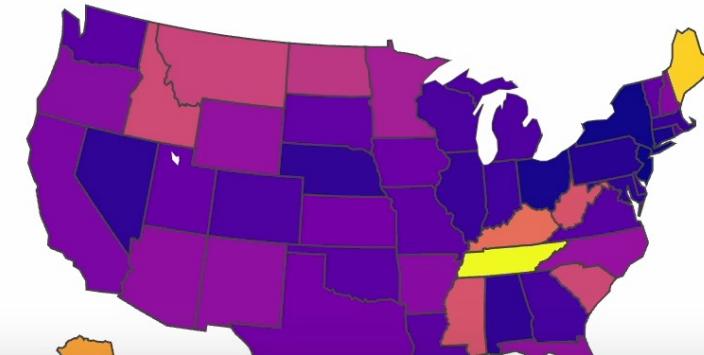
Feb 2022

Pick a topic

new cases

per capita

New COVID Cases Per Capita in the United States in Feb 2022



# Conclusions

- COVID is spreading faster and faster, while the rate of people getting vaccinated has slowed down.
- States with low vaccine rates and high case/death rates (such as Mississippi) would be a good targets for outreach.
- This application is publicly available for anyone who is interested in exploring on their own!
- [https://share.streamlit.io/rachelgoodridge/data\\_engineering/main/streamlit/streamlit\\_app.py](https://share.streamlit.io/rachelgoodridge/data_engineering/main/streamlit/streamlit_app.py)



# Future Directions

- Add another interactive map, broken down into counties.
- Incorporate automatic daily updates to the app.

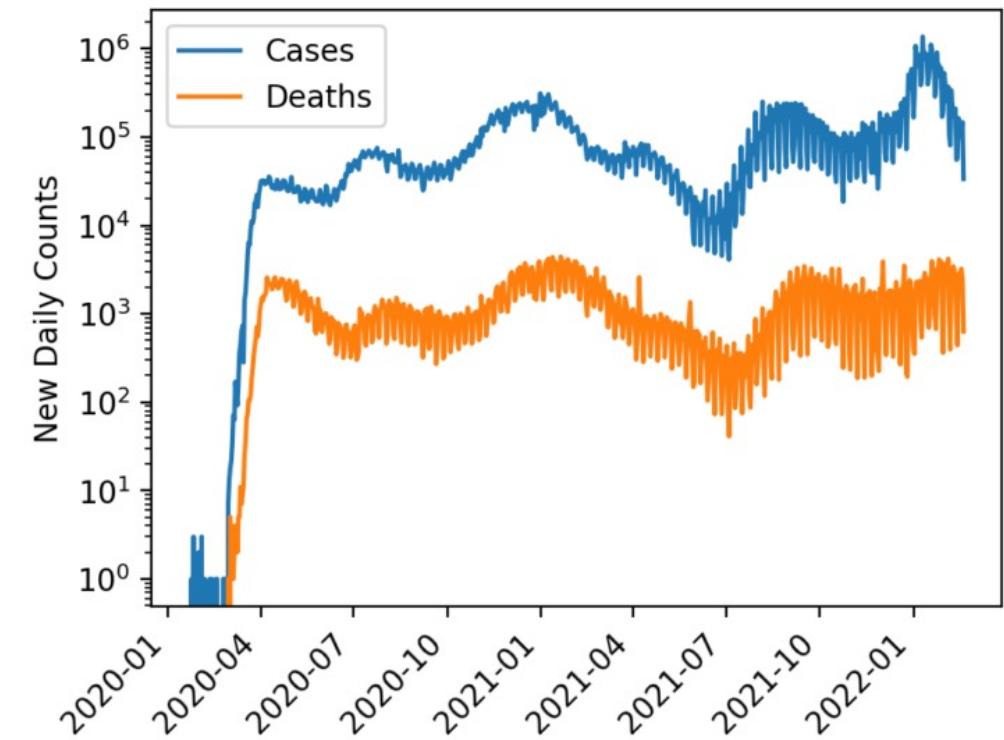
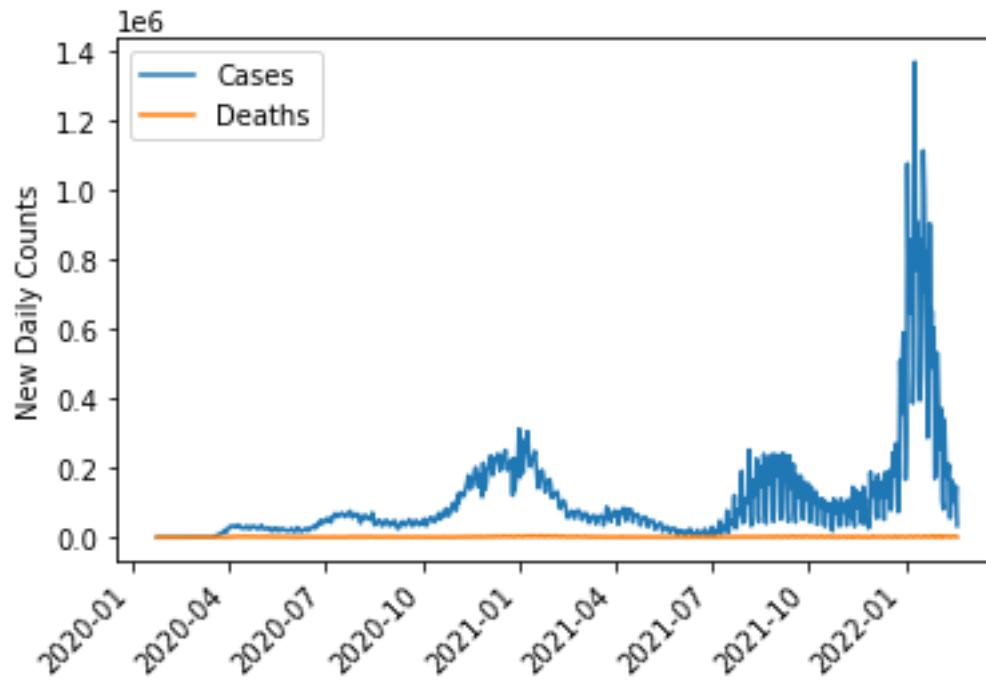


Thank you!

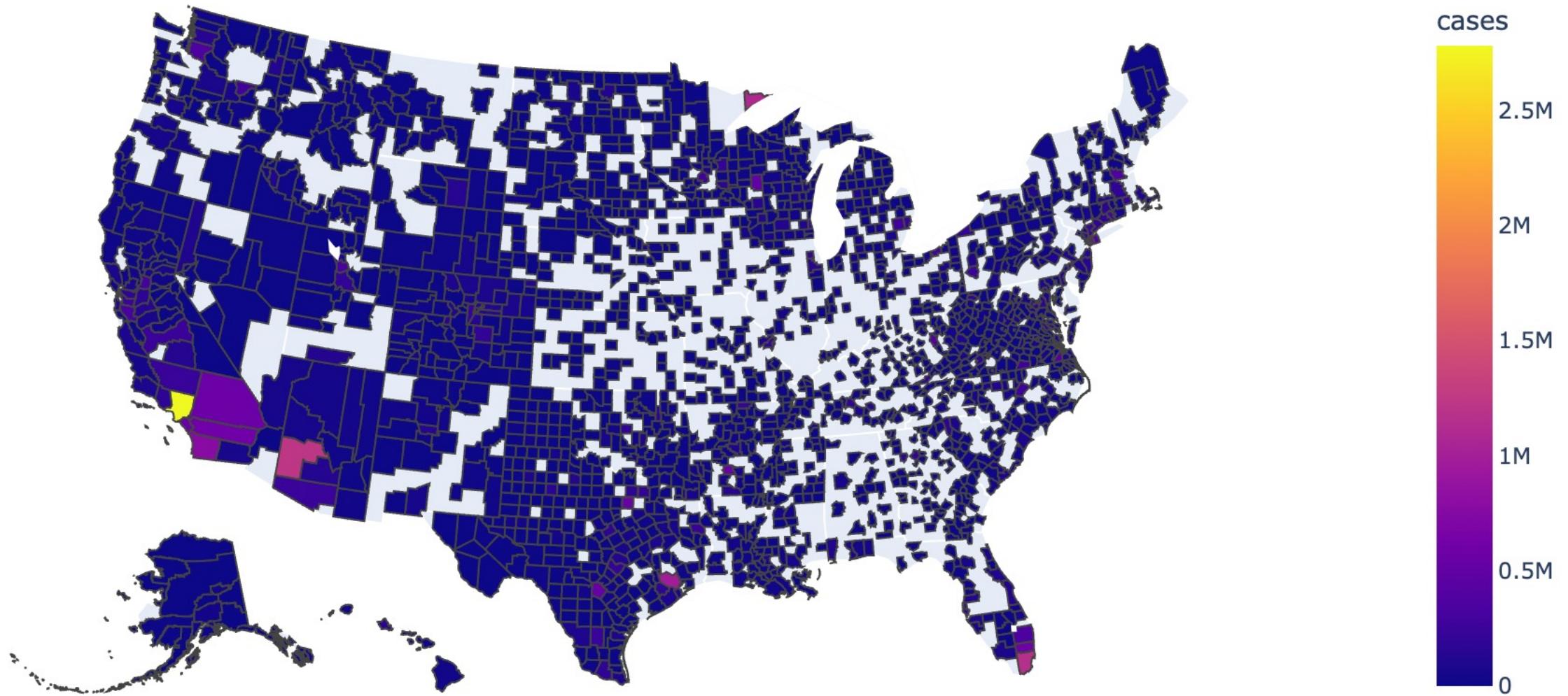
# Appendix



# New Daily COVID Cases and Deaths



# Cumulative COVID Cases by County



# Highest and Lowest Deaths Per Capita

Which 5 states have the highest cumulative number of deaths from COVID per capita?

	state	abbrev	deaths_per_capita
684	Mississippi	MS	0.0040
86	Arizona	AZ	0.0038
1022	Oklahoma	OK	0.0036
8	Alabama	AL	0.0036
528	Louisiana	LA	0.0035

Which 5 states have the lowest cumulative number of deaths from COVID per capita?

	state	abbrev	deaths_per_capita
346	Hawaii	HI	0.0009
1282	Vermont	VT	0.0009
1256	Utah	UT	0.0013
554	Maine	ME	0.0014
1360	Washington	WA	0.0015

# Highest and Lowest Percent Fully Vaccinated

Which 5 states or territories have the highest percentage of fully vaccinated people to date?

	state	abbrev	perc_full_vacc
970	Northern Mariana Islands	MP	82.1600
1100	Puerto Rico	PR	80.8400
1126	Rhode Island	RI	80.3300
60	American Samoa	AS	80.3200
1282	Vermont	VT	80.0000

Which 5 states or territories have the lowest percentage of fully vaccinated people to date?

	state	abbrev	perc_full_vacc
8	Alabama	AL	50.0400
1438	Wyoming	WY	50.4500
684	Mississippi	MS	50.7400
528	Louisiana	LA	52.4100
1308	Virgin Islands	VI	52.5200

# DataFrame Info

```
cases.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2539920 entries, 0 to 2539919
Data columns (total 6 columns):
 #   Column   Dtype  
 --- 
 0   date      object 
 1   fips      object 
 2   county    object 
 3   state     object 
 4   cases     int64  
 5   deaths    int64  
dtypes: int64(2), object(4)
memory usage: 116.3+ MB
```

```
vaccines.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 26148 entries, 0 to 26147
Data columns (total 5 columns):
 #   Column           Non-Null Count  Dtype  
 --- 
 0   date            26148 non-null   object 
 1   state           26148 non-null   object 
 2   perc_vacc       22435 non-null   object 
 3   perc_full_vacc 22360 non-null   object 
 4   perc_boost      6366 non-null   object 
dtypes: object(5)
memory usage: 1021.5+ KB
```