COVID-19 EDA

Guok Wei Jie

About the data set

- Large state-level Covid-19 data for countries around the world
- Time series data
- Epidemiological variables such as the cumulative number of confirmed cases, deaths and vaccine doses
- Level of policy measures implemented across different states
- Geographical coordinates of each state

What am I trying to find out?

- How do vaccination efforts affect the case fatality rates across the lower 48 US states?
- Different states may adopt different vaccination policies
- Different states may respond differently to the policies
- People in different states may have different susceptibilities to Covid-19
- How much do vaccines really help in tackling Covid-19, especially as a standalone measure?

Scope of analysis

- lower 48 states (USA has the most entries in the data set)

administrative_area_level_1 <chr>	NumberOfRows <int>
United States	83999
Russia	83996
Thailand	75691
Japan	53815
Ireland	35204
Argentina	34630
Mexico	32768
Italy	31521

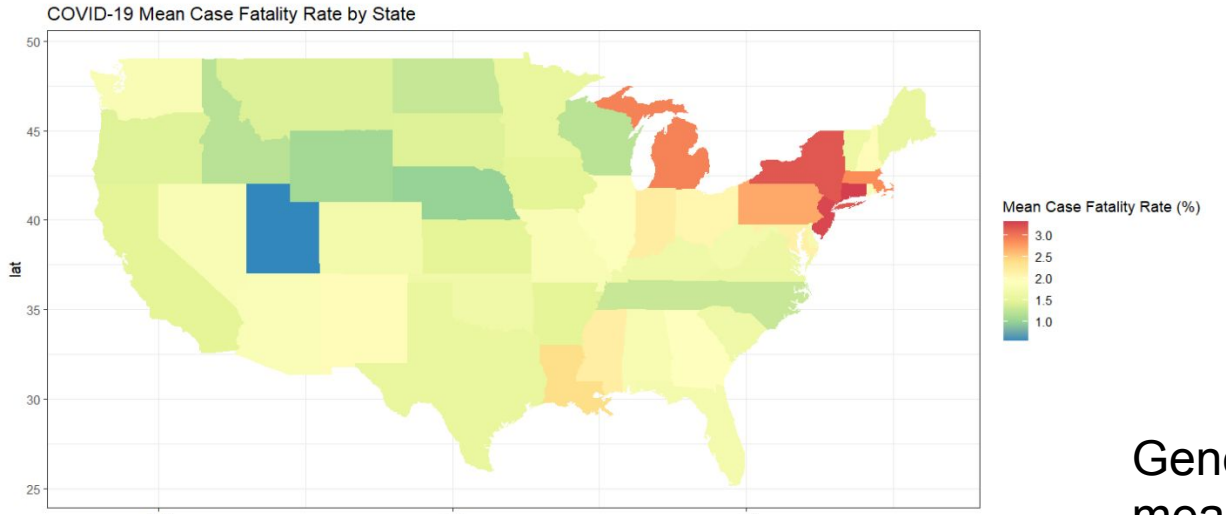
- vaccinations and vaccination policies vs Covid-19 deaths

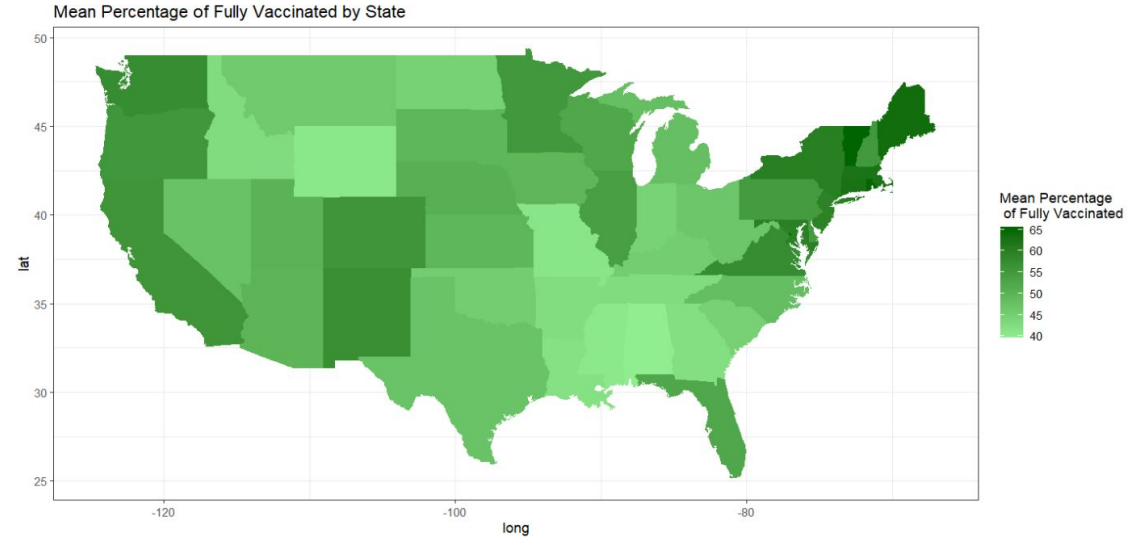
Variables of Interest

- confirmed, deaths
- vaccination_policy
- people_vaccinated, people_fully_vaccinated
- contact_tracing
- Case fatality rate (CFR) = deaths / confirmed * 100
- Percentage Fully Vaccinated (PFV) = people_fully_vaccinated / population * 100
- Vaccination Completion Percentage (VCP) = people_fully_vaccinated / people_vaccinated * 100
- Percentage of days with complete contact tracing (PCT) = number of days where contact_tracing == 2 / total number of days * 100

Data cleaning

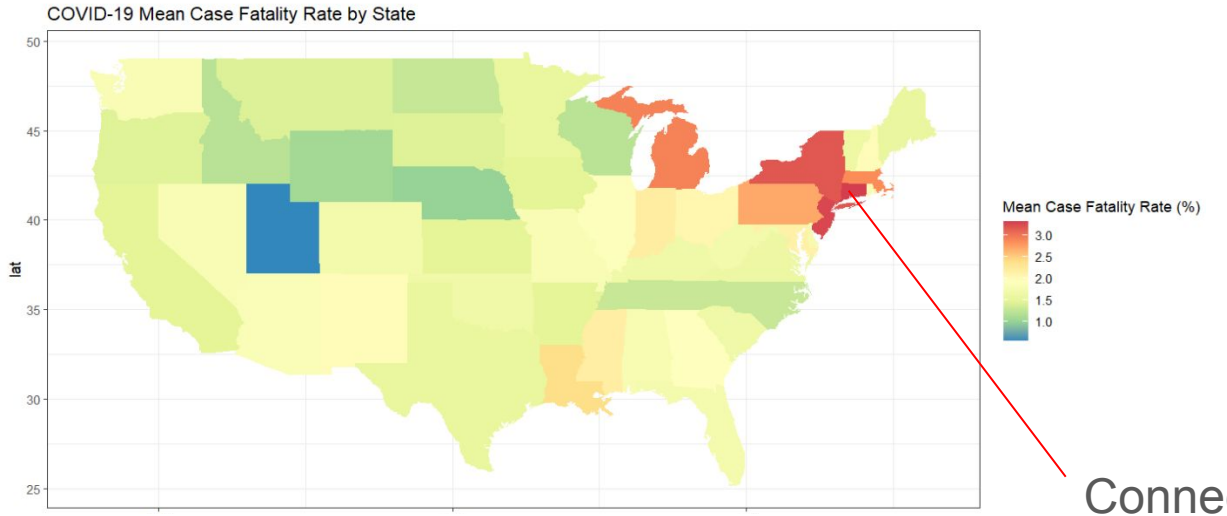
- Removed rows with **NA** values for ‘vaccination_policy’ but non-NA values for ‘vaccines’ => inconsistent
- Removed rows when there was insufficient information to impute NA values with suitable values => “0” does not always work

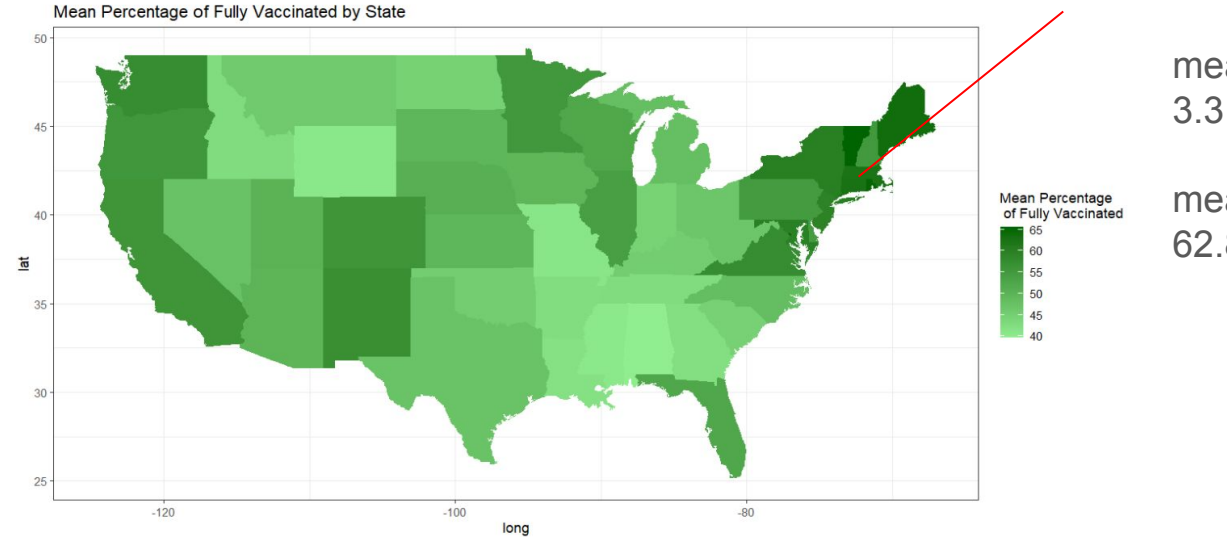




Generally, mean CFR ↓ as mean PFV ↑

BUT there are outliers!

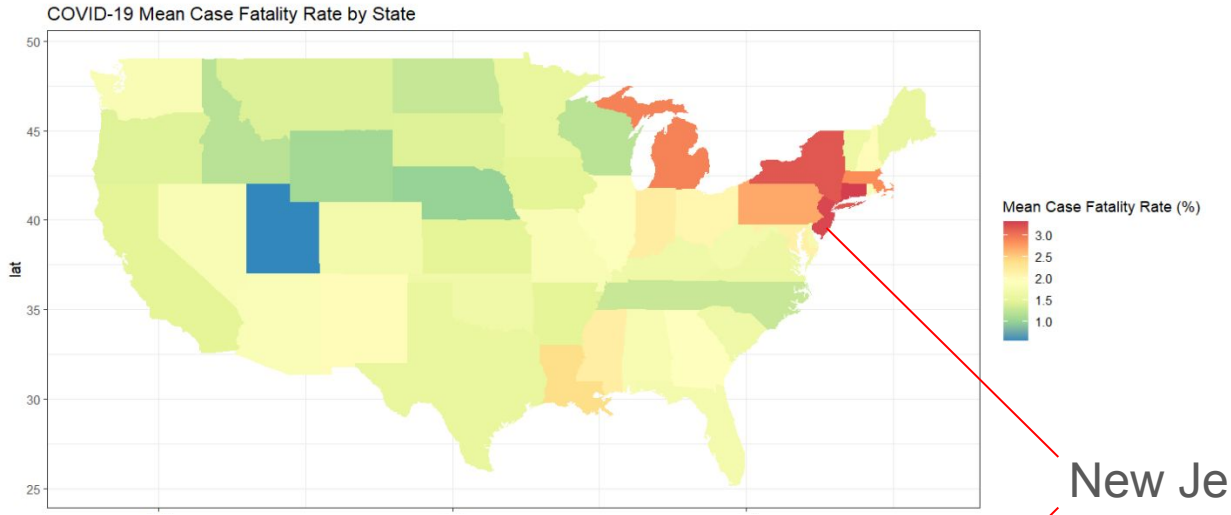


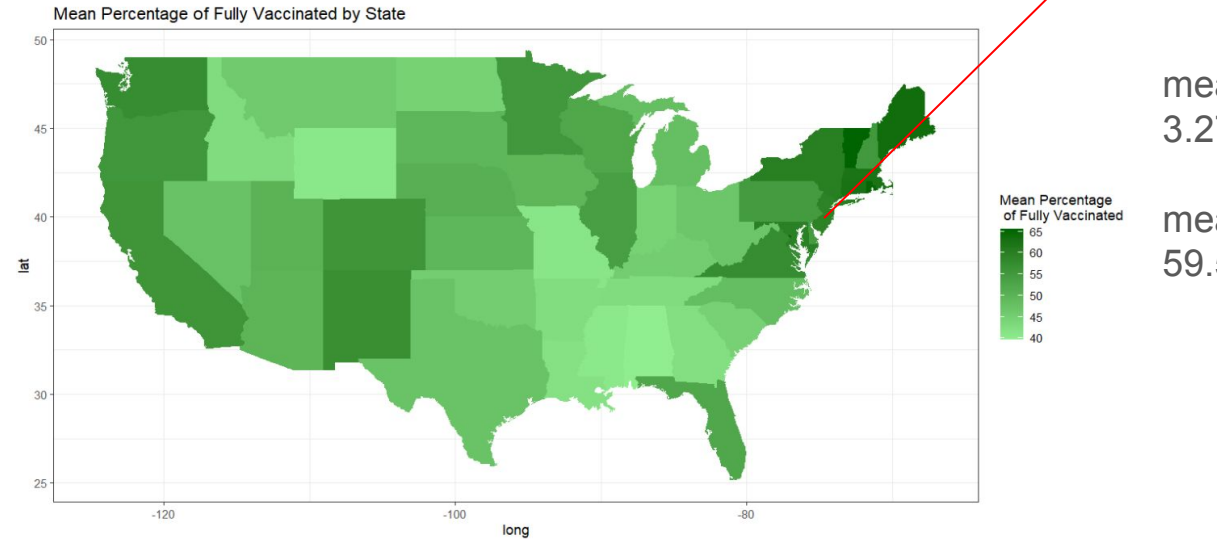


Connecticut

mean case fatality rate:
3.31% (highest)

mean percentage of fully vaccinated:
62.82% (5th highest)

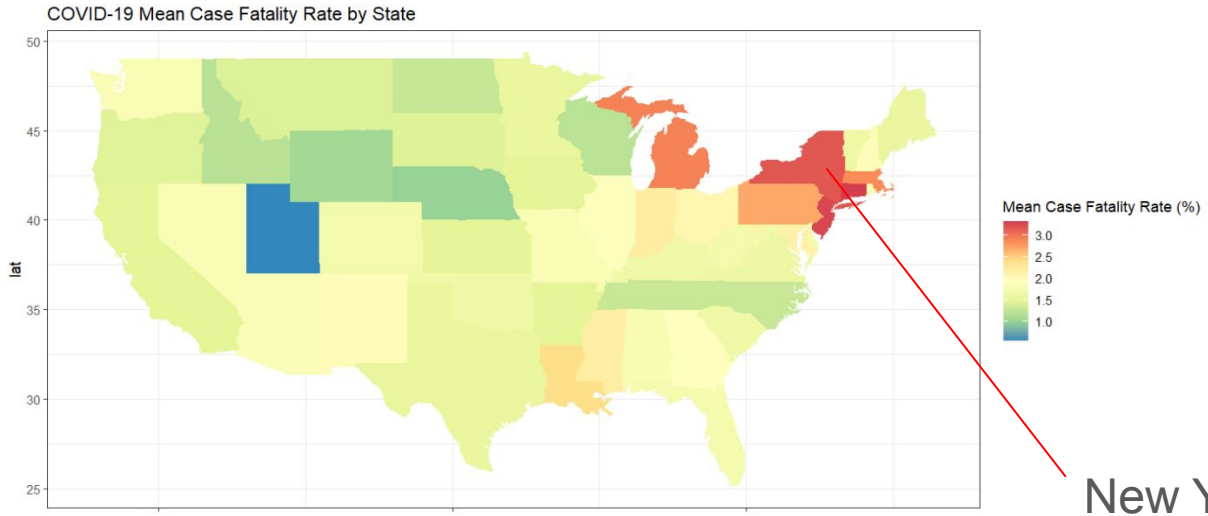


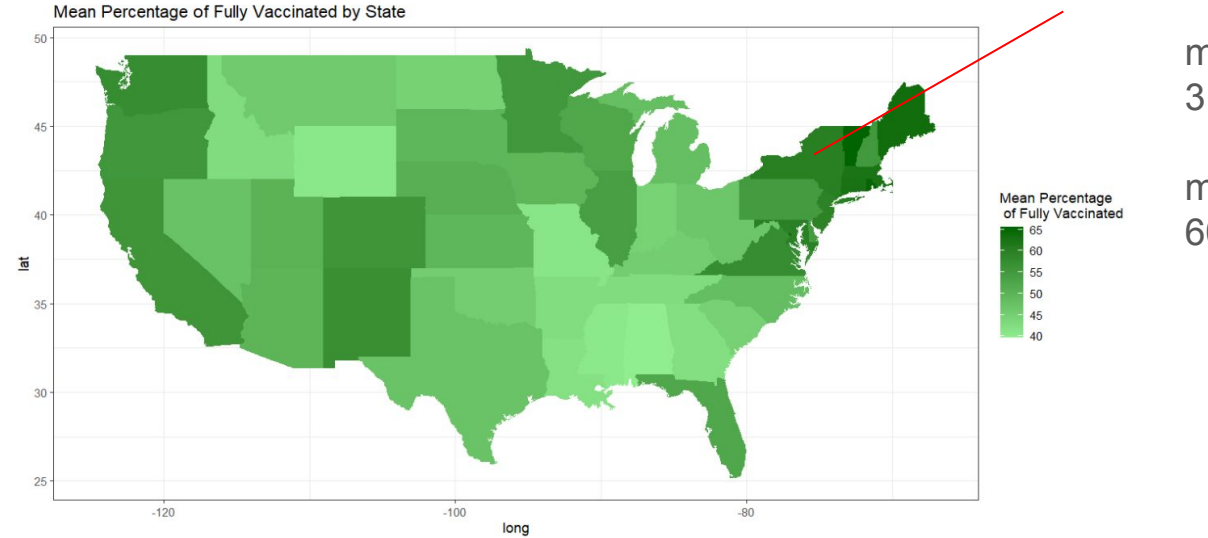


New Jersey

mean case fatality rate:
3.27% (2nd highest)

mean percentage of fully vaccinated:
59.52 % (9th highest)



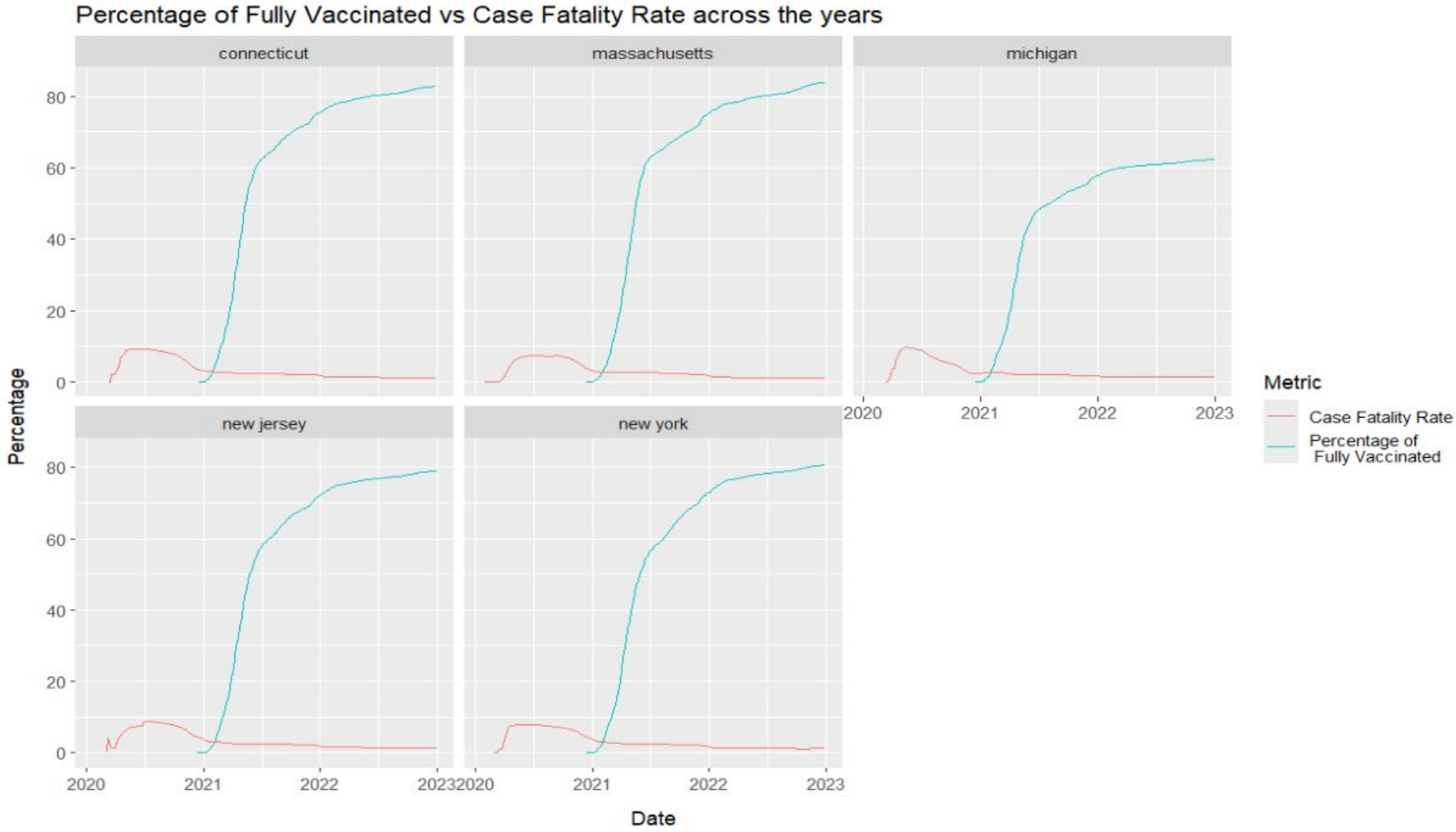


New York

mean case fatality rate:
3.17% (3rd highest)

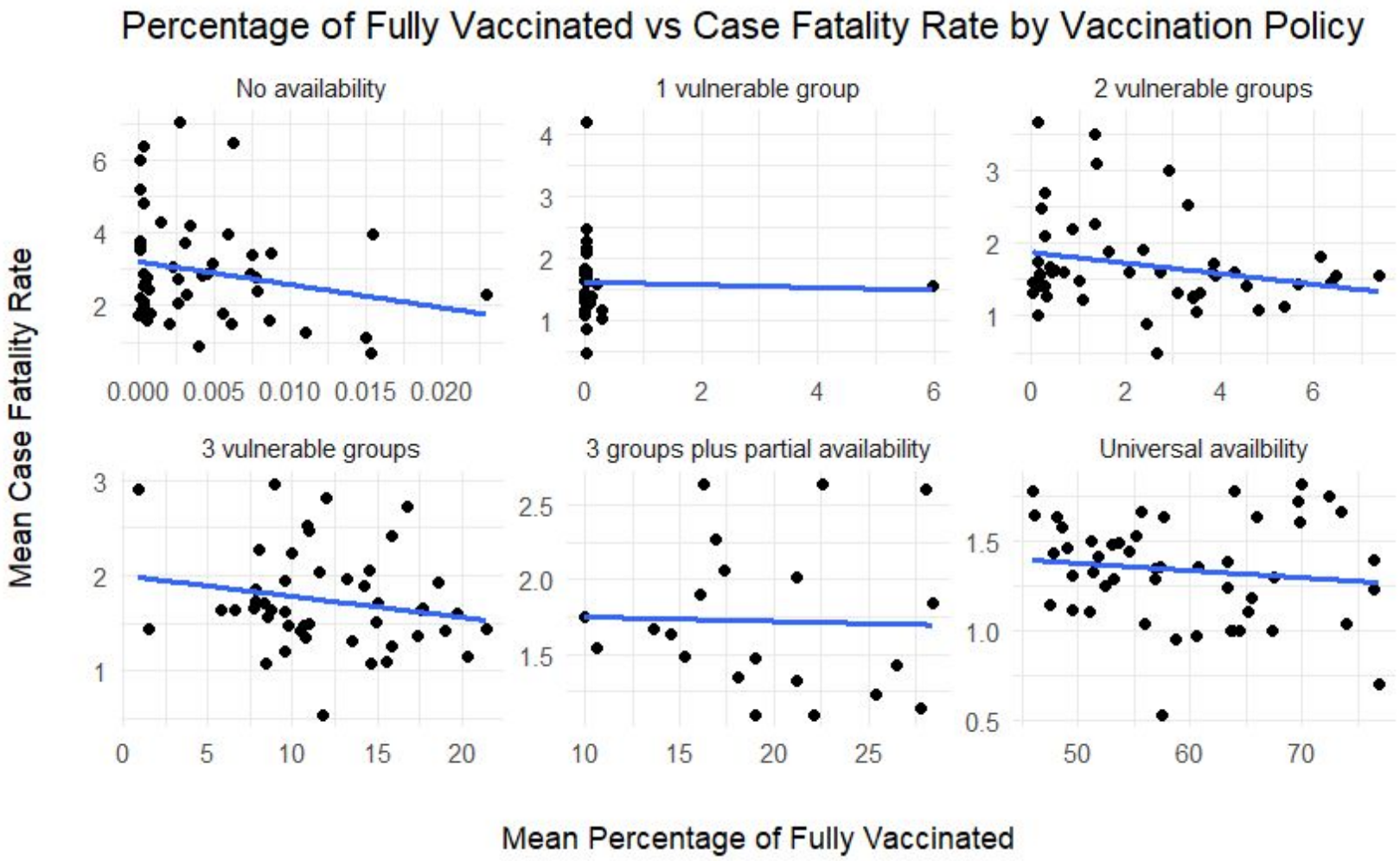
mean percentage of fully vaccinated:
60.1% (7th highest)

Let’s look at the top 5 states with the highest mean case fatality rate!


Percentage of Fully Vaccinated vs Case Fatality Rate across the years
connecticut
massachusetts
michigan
new jersey
new york
Percentage
Date
Metric
Case Fatality Rate
Percentage of Fully Vaccinated
0
20
40
60
80
2020
2021
2022
2023

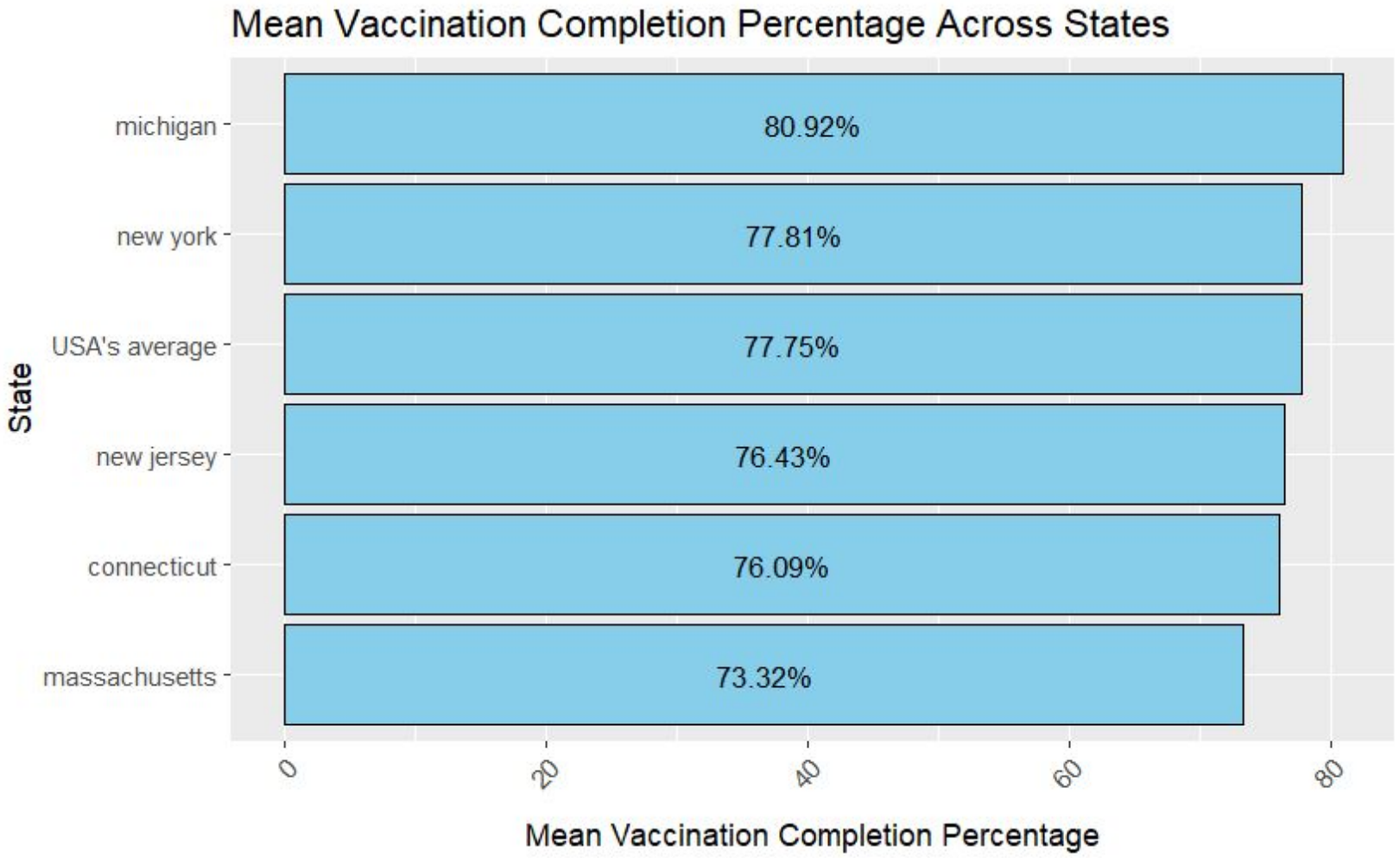
Percentage of Fully Vaccinated vs Case Fatality Rate

- vaccinations only began in December 2020 in the US (Pfizer-BioNTech COVID-19 Vaccine)
- CFR peaked at about 8-10% in mid-2020
- began to drop steadily to around 2-3%, even as PFV increased rapidly in 2021 and slowed down after
- high mean PFVs might be due to the ramping up of vaccinations in 2021
- **BUT** it is also important to consider how **vaccination policies** influenced PFV as it determines **who** the vaccines are available to


Percentage of Fully Vaccinated vs Case Fatality Rate by Vaccination Policy
No availability
1 vulnerable group
2 vulnerable groups
3 vulnerable groups
3 groups plus partial availability
Universal availbility
Mean Case Fatality Rate
Mean Percentage of Fully Vaccinated

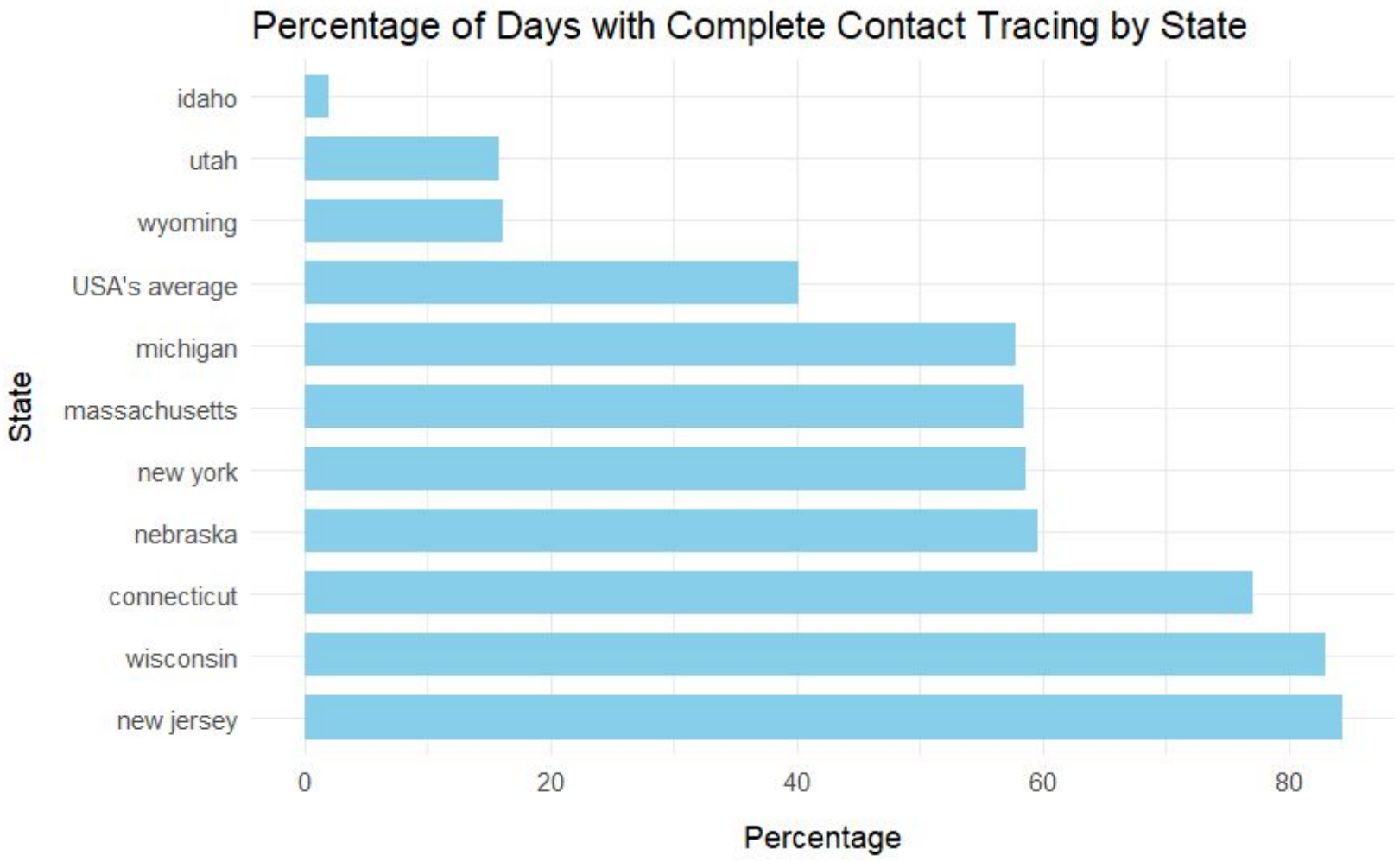
Analyzing by Vaccination Policy

- Generally, mean CFR decreases as mean PFV increases
- interestingly, states with ‘no availability’ of vaccines still do have a small percentage of people who are fully vaccinated => could have gotten their vaccines elsewhere
- sharpest decrease in mean CFR seen in ‘no availability’
- small spread of mean PFV in vaccination policy levels 0 and 1, whereas spread of mean PFV increased significantly in level 2 => level 2 might have been a breakthrough measure


Mean Vaccination Completion Percentage Across States
State
michigan
new york
USA's average
new jersey
connecticut
massachusetts
80.92%
77.81%
77.75%
76.43%
76.09%
73.32%
0
20
40
60
80
Mean Vaccination Completion Percentage

Mean Vaccination Completion Percentage

- shows how many vaccinated individuals have completed all doses => affects the efficacy of the vaccine and hence affects case fatality rate
- filtered out previous 5 states with worst mean CFR, along with USA's average for easier comparison and to not overload data
- Connecticut, New Jersey and Massachusetts have slightly lower mean VCP than USA's average, while New York and Michigan are higher
    - might contribute to the high mean CFR in Connecticut, New Jersey and Massachusetts as people who are not fully vaccinated have lower protection than those who are


Percentage of Days with Complete Contact Tracing by State
idaho
utah
wyoming
USA's average
michigan
massachusetts
new york
nebraska
connecticut
wisconsin
new jersey
State
0
20
40
60
80
Percentage

Percentage of days with Complete Contact Tracing (PCT)

- Research has shown that a 10% increase in proportion of cases identified through contact tracing can lead to COVID-19 mortality reductions of 0.8% to 3.4%
- filtered out the top 5 and bottom 5 mean CFR from the lower 48 states, together with USA's average, for easier comparison
- top 5 states for mean CFR all have higher PCT than USA's average, even though New York, Michigan and Massachusetts have significantly lower PCT than New Jersey and Connecticut
- USA's average PCT (40%) is rather low, increasing PCT may help to reduce mean CFR for the country as a whole

Limitations of analysis

- Acknowledge that there are many factors and policies at play and case fatality rates are not solely influenced by vaccination policy
- Analysis is geographically limited to the lower 48 states of USA
- Missing data for variables like vaccination policy and vaccines make it more challenging to observe accurate trends in the data

Conclusion

- Generally, the higher the percentage of fully vaccinated individuals, the lower the case fatality rate
- Lack of a vaccine in 2020 contributed to higher case fatality rates at the start
- Low vaccination completion rate also contributes to higher case fatality rates
- Level of contact tracing affects the number of confirmed cases identified and also the case fatality rates
- Level of vaccination policy affects who gets the vaccines and hence the percentage of fully vaccinated individuals

Zoom recording link

https://umd.zoom.us/rec/share/FoL6GZXevr3tcm5-vraMX3wpIimgJuJleTIvdRBF5pElGo80SWbA3BXSSc303NS4.mciiLu-LYVqXqYQK?startTime=1712875437000

Passcode: L^Hs6r1H