THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC

Data Management & Analytics
DNSC 6305

Final Project Presentation

Group 7

Himanvesh Maddina
Elias Moreno
Chanmi Shin
Andrea Ho
Hyemin Park
Abid Shafiullah
Virat Pansuriya

Project Description

- Descriptive analyses of COVID-19 issues in the United States
- Cause and effect analyses of certain policies, testing, and vaccination in the COVID-19 status by location and time
- All analyses based on a database of actual COVID-19 data and designed to answer specific questions of the project

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC

Data Sources

Main Data (source: Johns Hopkins University)

- Cases: Data of COVID-19 cases information
- Vaccine: Data of COVID-19 vaccine information
- Times: Data of COVID-19 test information

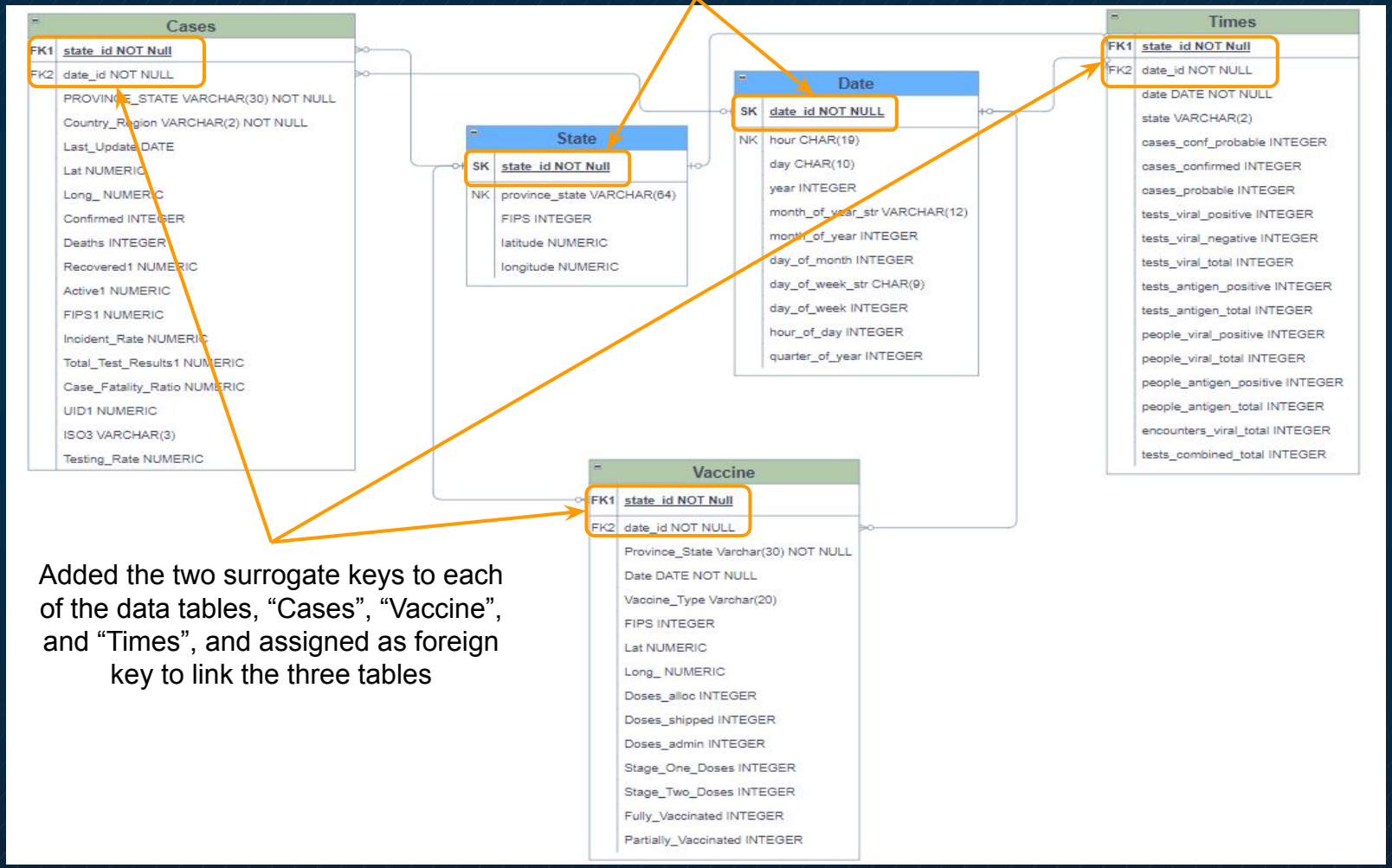
Supplemental Data

- Population: Population data by state
  - https://www.census.gov/quickfacts/fact/table/US/PST045219
- Restriction: Restriction period by state
  - https://ballotpedia.org/Travel_restrictions_issued_by_states_in_response_to_the_coronavirus_(COVID-19)_pandemic,_2020-2021

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC

Dimensional Model

Created surrogate keys with unique values to “State” and “Date” tables based on their natural keys



Added the two surrogate keys to each of the data tables, “Cases”, “Vaccine”, and “Times”, and assigned as foreign key to link the three tables

Analytical Questions

1. What's the percentage of people who are affected by COVID ?
2. What is the confirmed vs recovery vs death cases ?
3. Do vaccines have an actual effect on covid cases ?
4. Do testing rates affect confirmed cases ?
5. Were there any changes in confirmed cases before/after restrictions ?

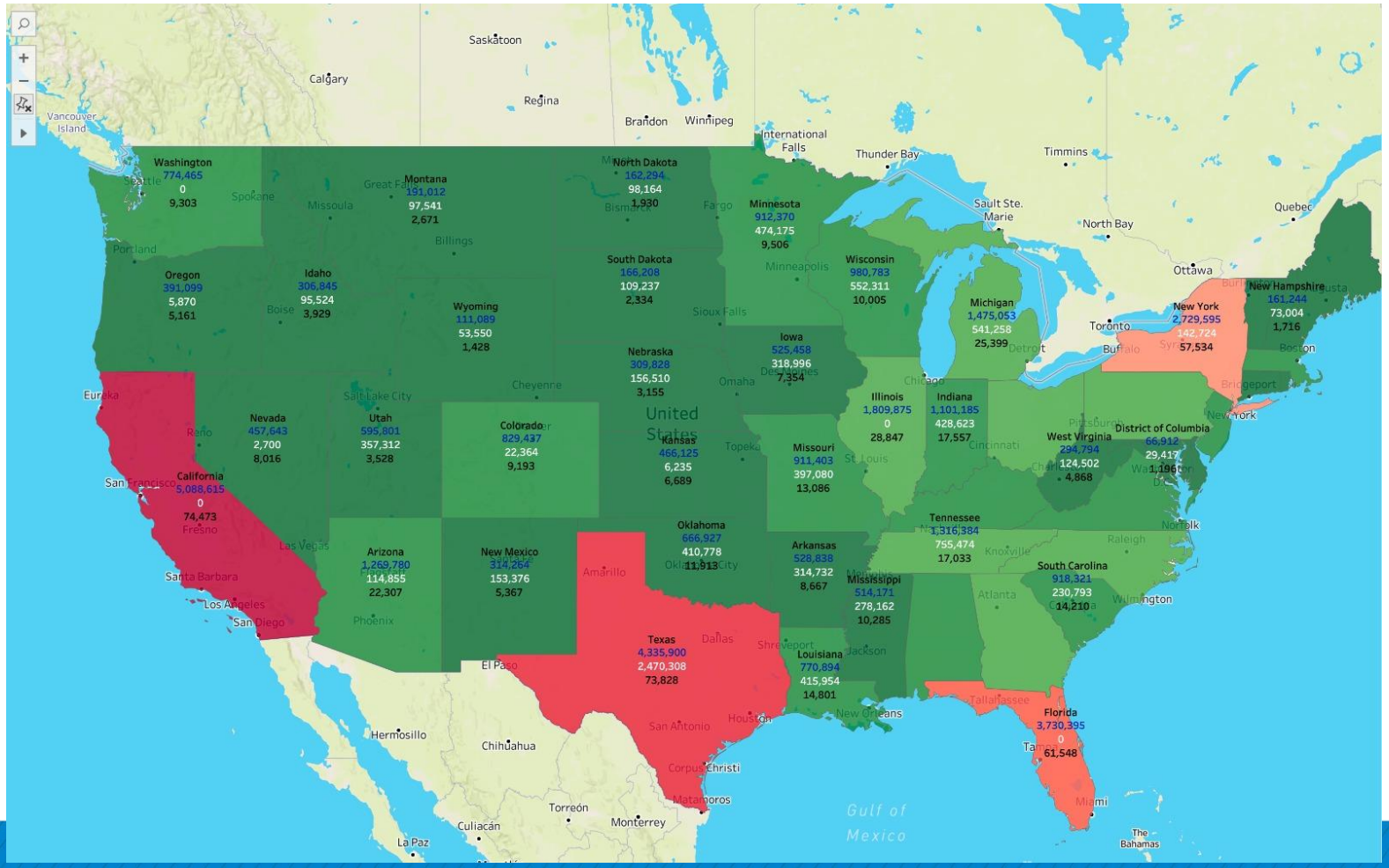
THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC

1. What is the % of people infected with COVID?

As we can see from the above table North Dakota and Alaska have higher covid affected people since they are states with lower population under 1 million people. All top 10 states with highest covid percentage of covid confirmed cases besides Tennessee are all under 5 million population.

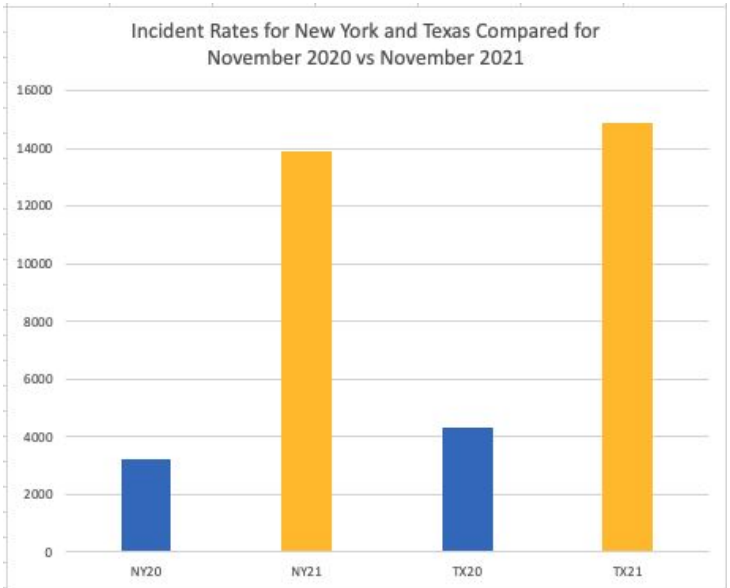
fips1	province_state	population	total_cases_confirmed	percentage
38	North Dakota	762062	162294	21.2967
2	Alaska	731545	150996	20.6407
47	Tennessee	6829174	1316384	19.2759
56	Wyoming	578759	111089	19.1943
46	South Dakota	884659	166208	18.7878
49	Utah	3205958	595801	18.5842
44	Rhode Island	1059361	191763	18.1018
30	Montana	1068778	191012	17.8720
45	South Carolina	5148714	918321	17.8359
21	Kentucky	4467673	785926	17.5914

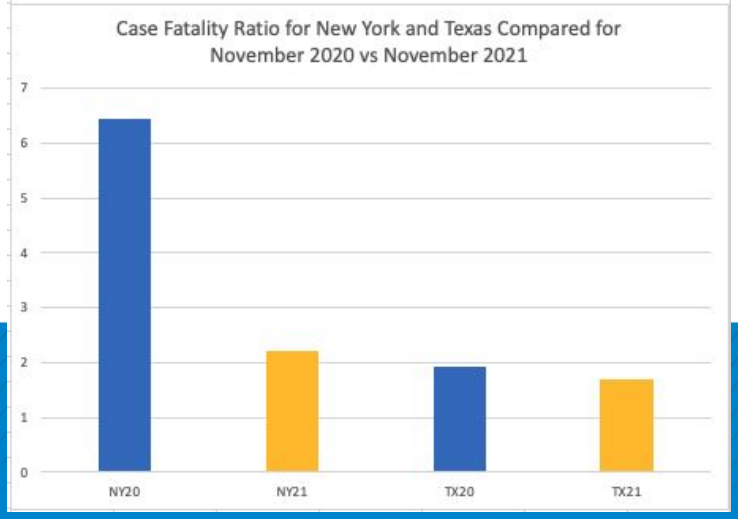
2. What is the Confirmed vs Recovery vs Death cases?

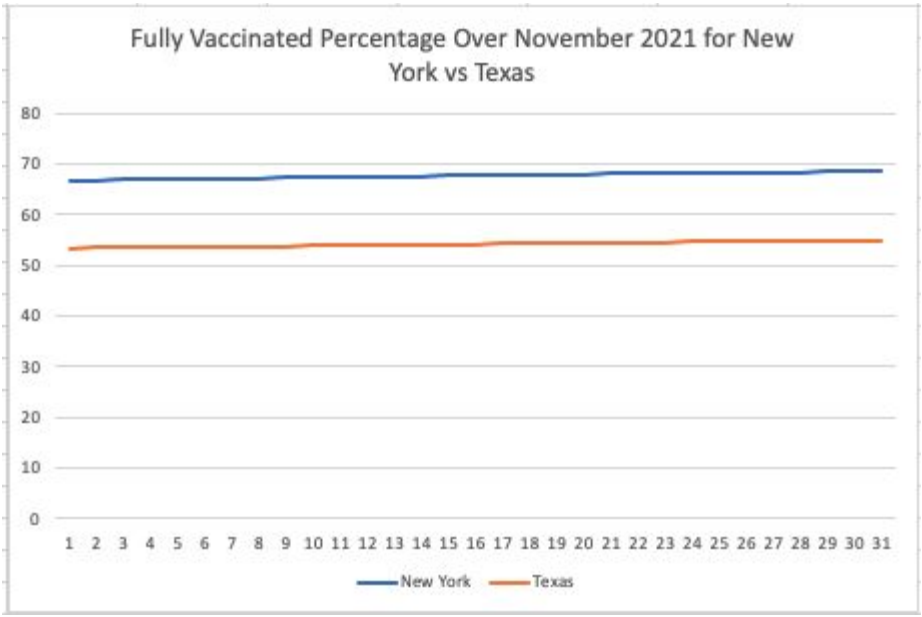


THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC

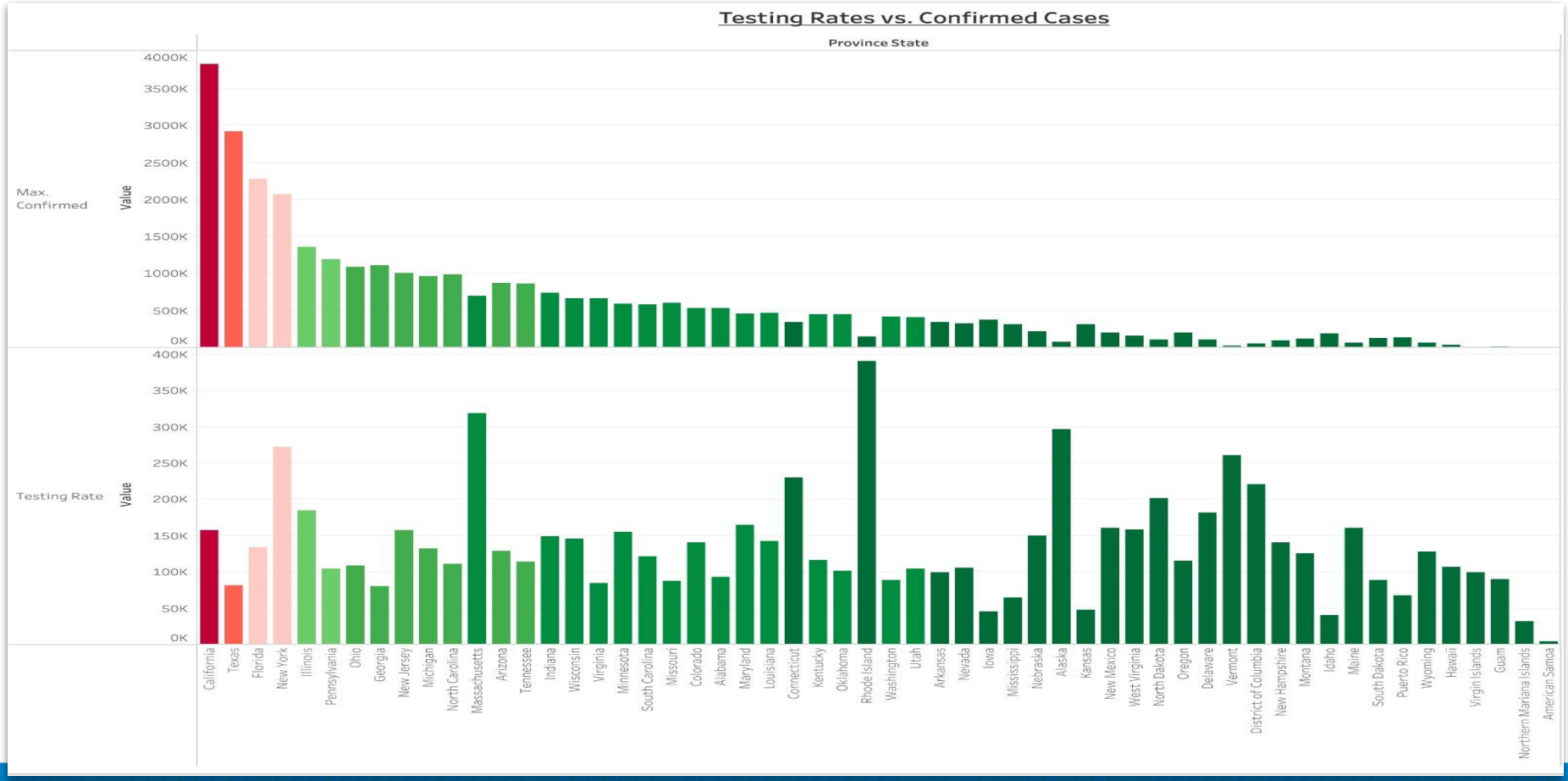
3. Do vaccines have actual effect on COVID cases?



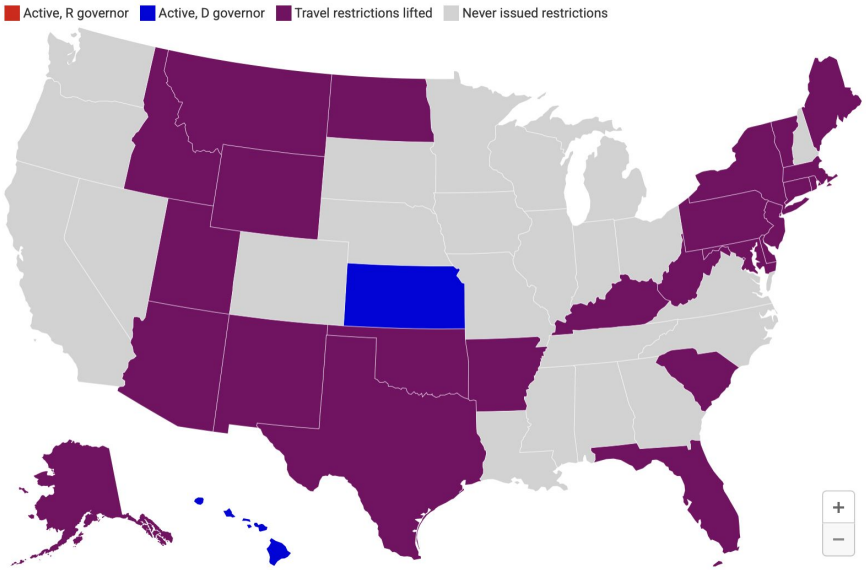


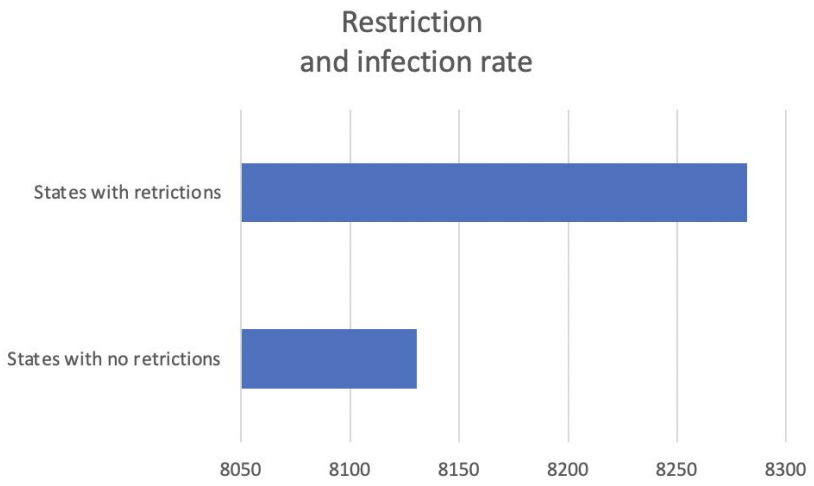


4. Do testing rates affect confirmed cases?



5. Do restrictions affect COVID cases?





(Source: Ballotpedia)

Conclusion

- Most of the states and territories with lower population have had higher percentage of confirmed cases
- The highest confirmed cases recorded were in states located on the coast. Logistic charges have been increasing constantly maybe due to the shipping docks running at a lower efficiency.
- Vaccines had a positive impact on case fatality ratio over the year for November 2020 vs November 2021 in New York and Texas
- We can conclude that the number of confirmed cases by state it is not directly related to the testing rate.
- States with restrictions have higher infection rates

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC

Attestation

All project team members listed below have contributed meaningfully to the project and have complete knowledge on any part of it.

Himanvesh Maddina

Elias Moreno

Chanmi Shin

Andrea Ho

Hyemin Park

Abid Shafiullah

Virat Pansuriya

THE GEORGE WASHINGTON UNIVERSITY

WASHINGTON, DC

THE GEORGE WASHINGTON UNIVERSITY

WASHINGTON, DC