

Data Analysis and Visualization of the COVID-19 Pandemic in the United States

Team 9: Kyle Lyon, Zeeshan Raza, Harshit
Aggarwal, Charlotte Ku, Lily Zeng, Tivon Johnson

Final Presentation
DNSC 6305_81

Outline

1

Methodology

Explain our teams approach to the project.

2

Data Sources

Identify sources of information and data collection.

3

Dimensional Model

Describe modeling: dimensional model and star schema.

4

Analytical Questions

Describe analytical questions and summarize the findings.

Methodology

What was team's our approach?

Methodology

Obtain Data

Retrieve data from the JHU Coronavirus Resource Center database.

Analysis

Perform dimensional analysis by structuring facts and measures to prepare data to answers business questions.

Star Schema

Create the star schema to provide a simple way to detail the design and optimize querying large data sets.

Insights

Write queries to answer analytical questions through use of SQL, Unix command line tools, Spark, and Python.

Visualization

Visualize data insights through visualizations tools such as Tableau and Excel.

Data Sources

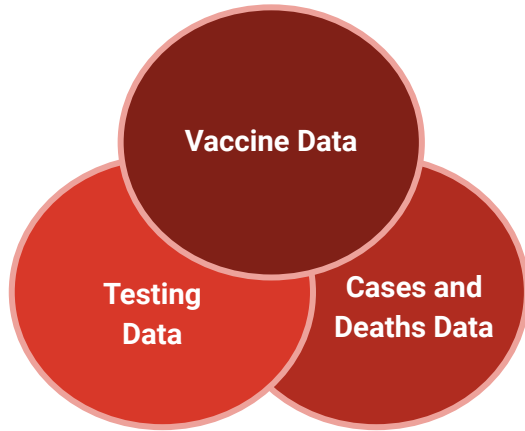
Where did we collect our data?

THE GEORGE
WASHINGTON
UNIVERSITY

WASHINGTON, DC

Data Sources

JHU Resource Center Datasets



U.S. Census Bureau Datasets

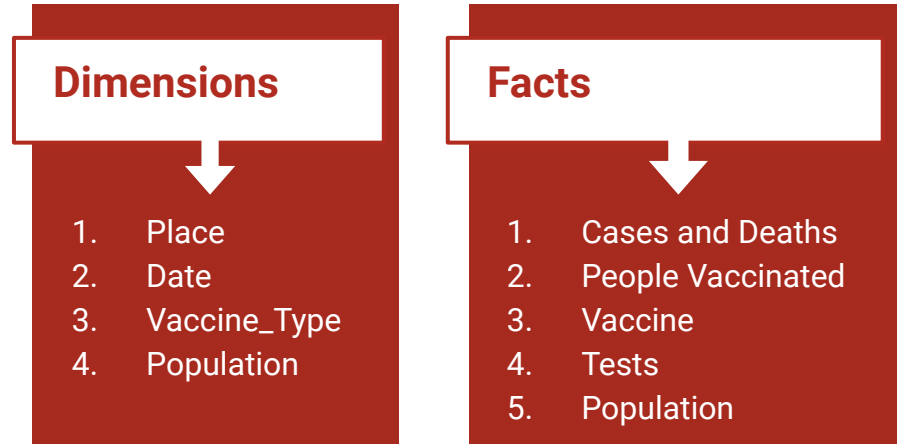


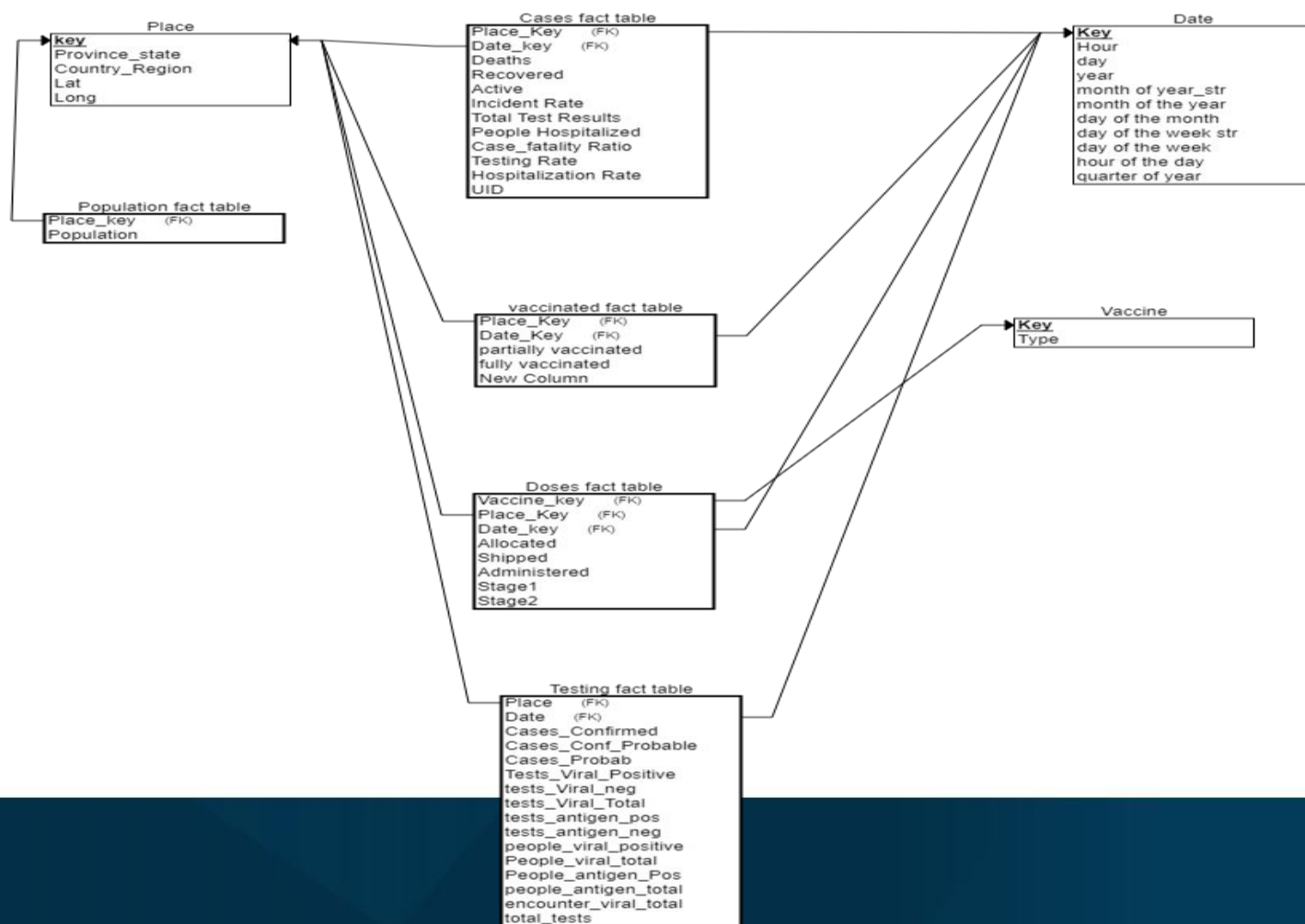
Dimensional Model

How did we structure our data?

Dimensional Model

- Dimensional modelling uses the concepts of facts (measures) and dimensions (context).
- This model includes three dimension tables and five fact tables.





Analytical Questions

What are our analytical questions and findings?

THE GEORGE
WASHINGTON
UNIVERSITY

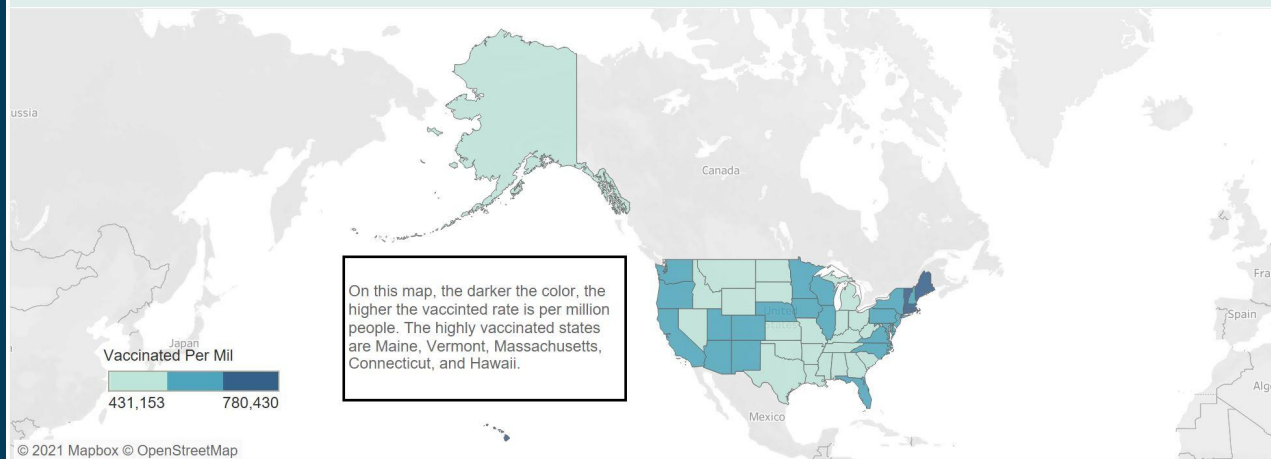
WASHINGTON, DC

Analytical Questions

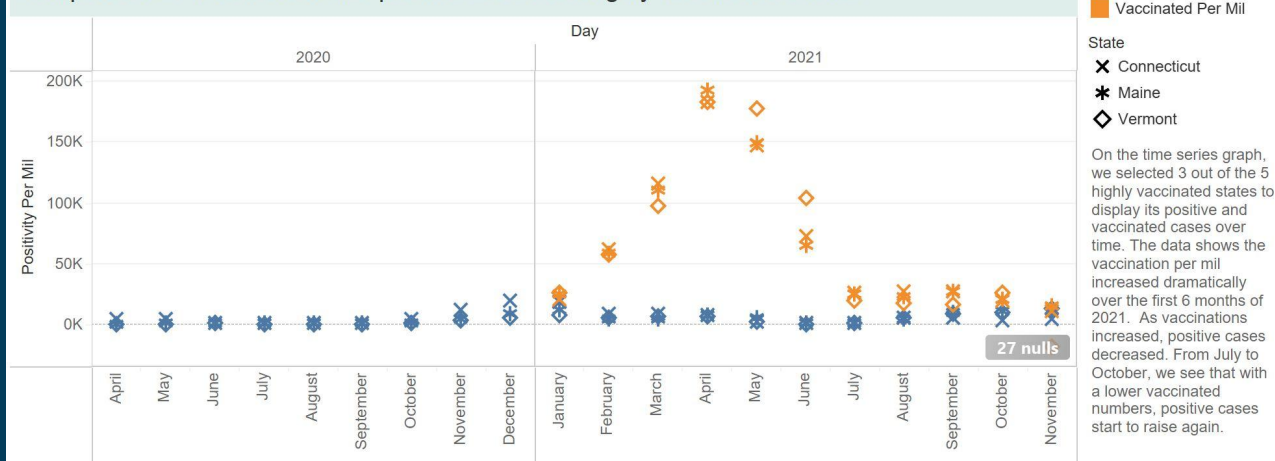
01	How do vaccinations affect the positivity and death rates by states?	<ul style="list-style-type: none">Analyzing how effective are vaccines at combating the pandemic
02	Does higher testing capacity lead to higher positive rates?	<ul style="list-style-type: none">Analyzing the relationship between testing and new cases
03	How does test type affect positivity rates?	<ul style="list-style-type: none">Comparing the effectiveness of Antigen testing versus PCR testing
04	What is the distribution of dosages allocated, administered, and shipped across states by quarters?	<ul style="list-style-type: none">Analyzing the adoption of vaccines by states

How does vaccination affect positivity and death rates?

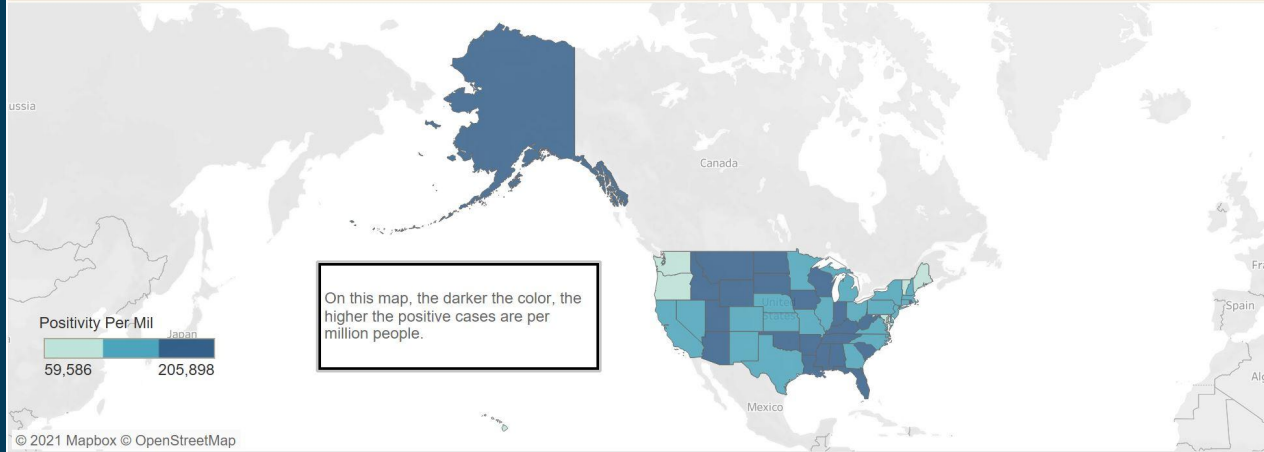
Vaccination per mil by state



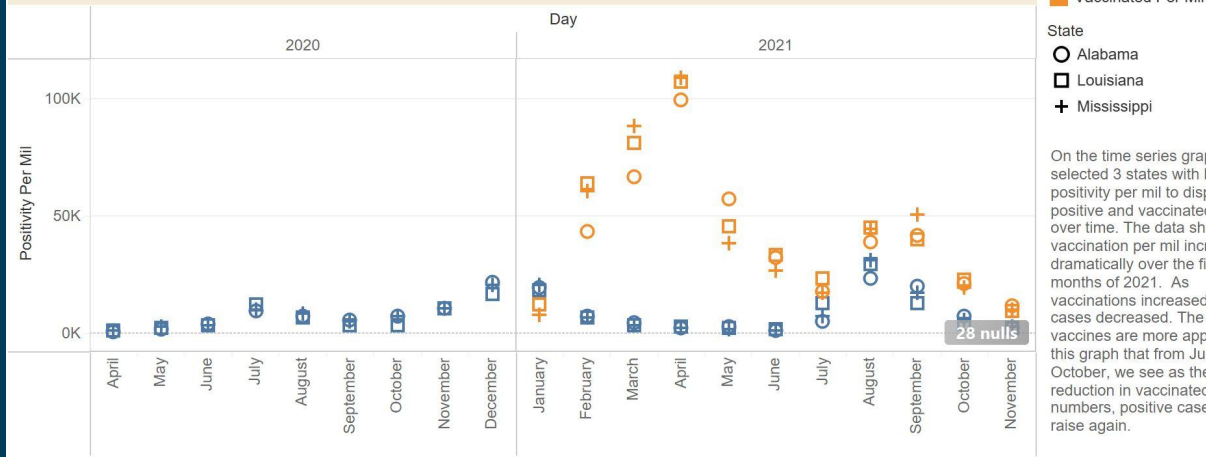
Comparison of vaccination and positive cases on highly vaccinated states



Positivity per mil by state

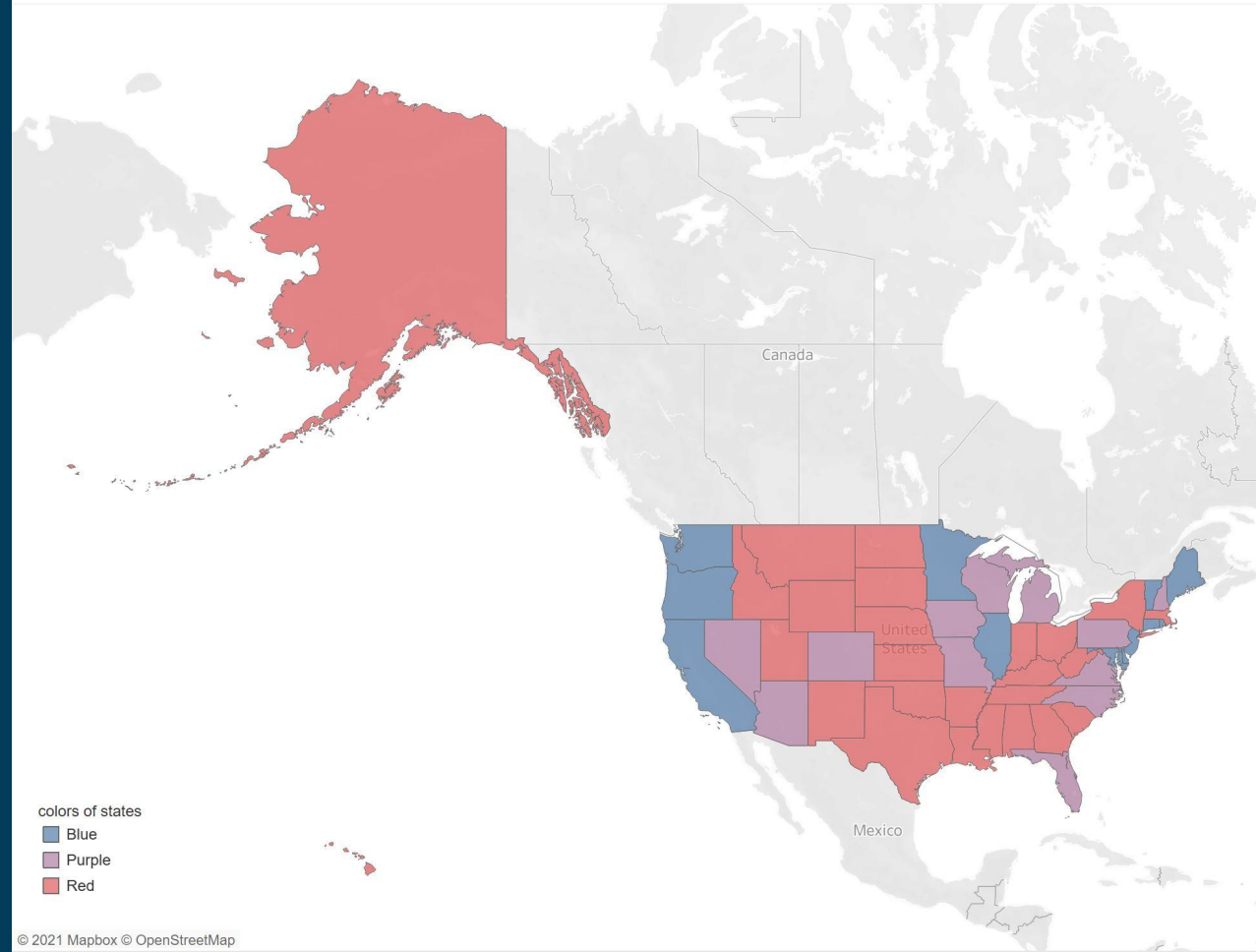


States with high positive cases per mil



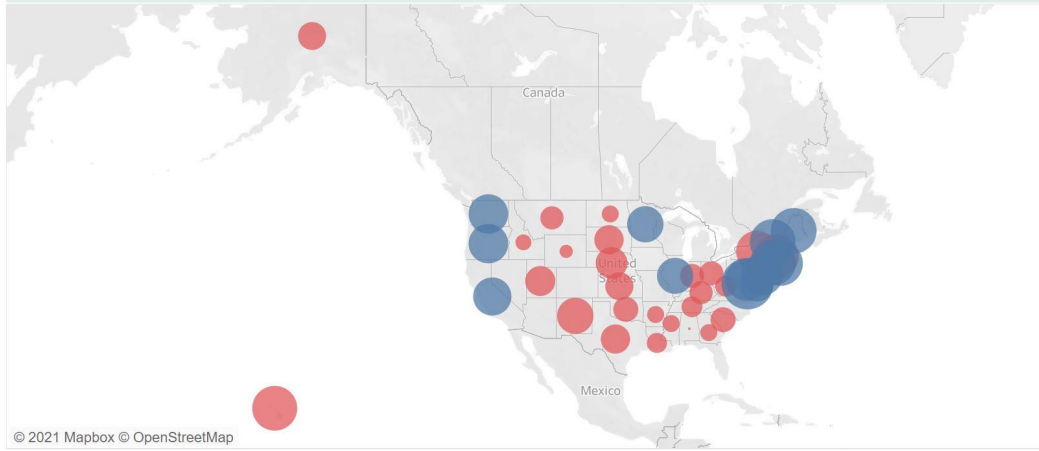
On the time series graph, we selected 3 states with high positivity per mil to display their positive and vaccinated cases over time. The data shows the vaccination per mil increased dramatically over the first 6 months of 2021. As vaccinations increased, positive cases decreased. The effect of vaccines are more apparent in this graph that from July to October, we see as the the reduction in vaccinated numbers, positive cases start to raise again.

State by Party Classification

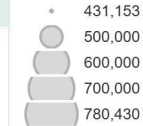


THE GEORGE
WASHINGTON
UNIVERSITY
WASHINGTON, DC

Vaccinations by color of state



Vaccinated Per Mil



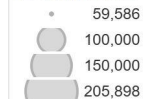
For the map on the top, the bigger the circle, the higher the vaccination rate per population per mil. We can see that the bigger circles, or the states with higher vaccination rate per million people are mostly the traditionally "Blue" states. Whereas the traditionally "Red" states are with smaller circles or lower vaccination rate per million people.

Party Division by State

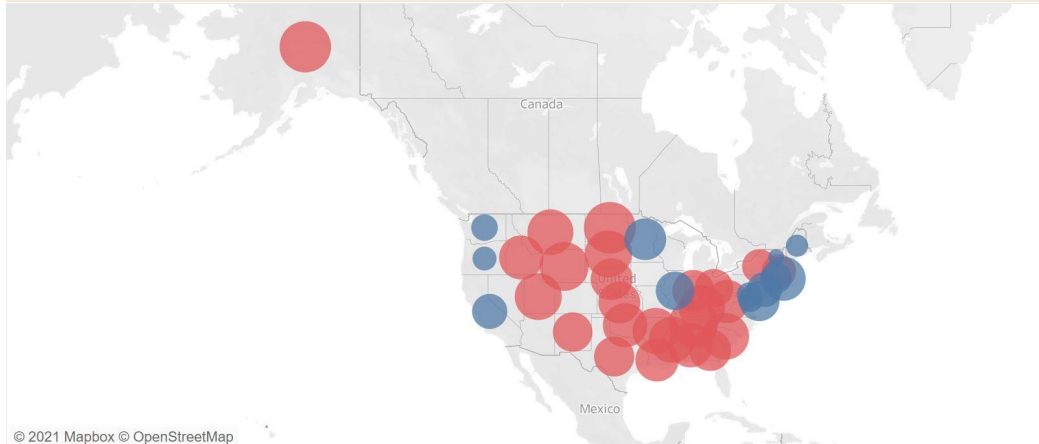


For the map on the bottom, the bigger the circle, the higher the positivity rate per million people. We can see that the bigger circles, or the states with higher positivity rate per million people, are mostly the traditionally "Red" states. Whereas the traditionally "Blue" states are with smaller circles or lower positivity rate per million people.

Positivity Per Mil

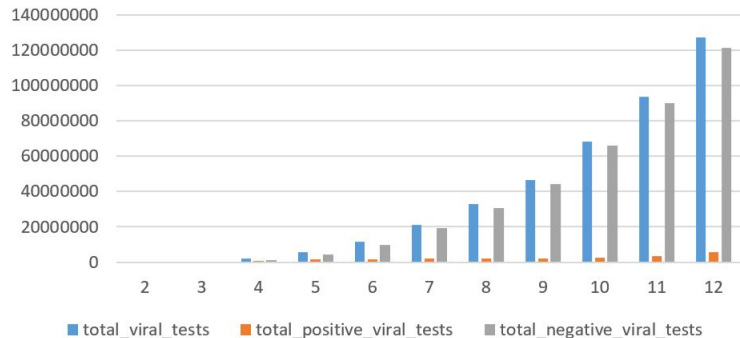


Positive Cases by color of state

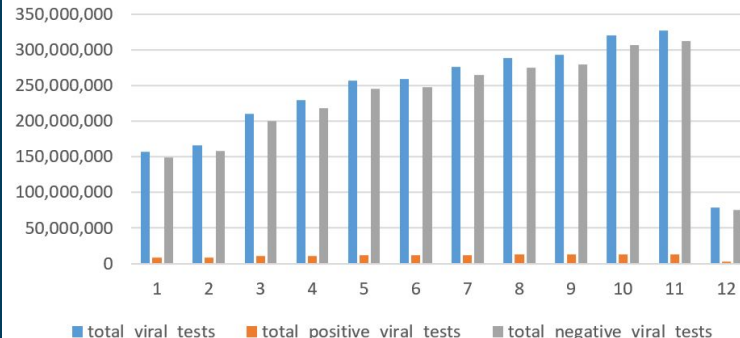


**Does higher testing
capacity lead to higher
positive rates?**

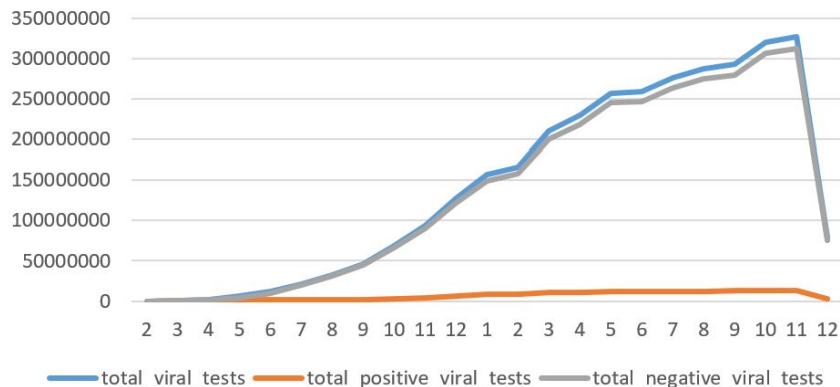
2020 Monthly Viral Testing Connecticut



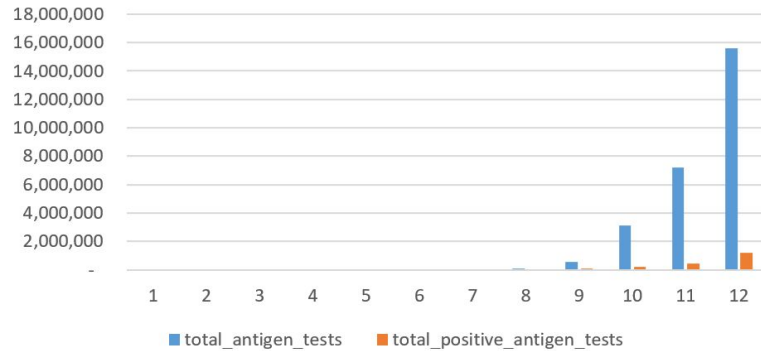
2021 Monthly Viral Testing Connecticut



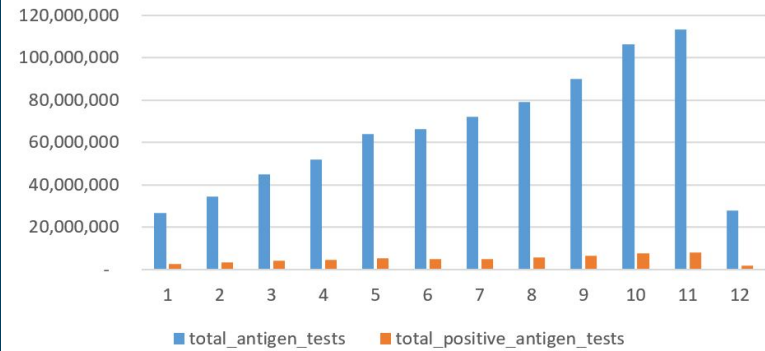
2020-2021 Monthly Viral Testing Connecticut



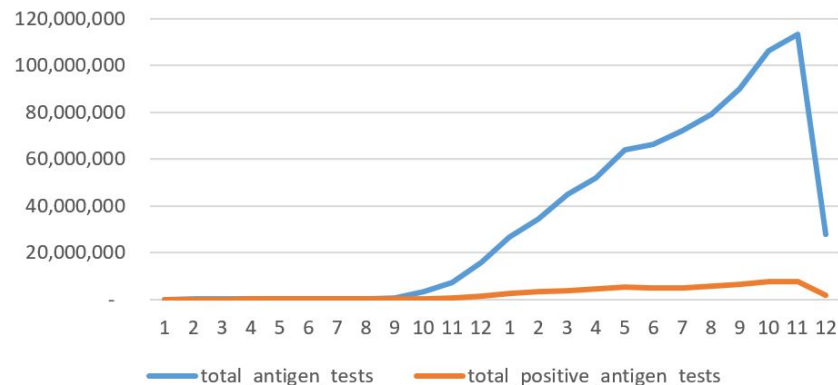
2020 Monthly Antigen Testing Virginia



2021 Monthly Antigen Testing Virginia



2020-2021 Monthly Antigen Testing Virginia



Conclusion

1. Findings require statistical analysis to validate observations.
2. Share findings with government officials to inform the public.

Team Attestation Statement

We have followed George Washington University's Code of Academic Integrity as well as the specific instructions for this final project. We affirm all work involved in completing this project represent our own work, without the use of any unpermitted aids or resources. We understand that there will no tolerance towards academic dishonesty and will lead to academic penalties.