COVID-19 Vaccination and Infection in Fulton County, Georgia

Junhao Zhang

I. Introduction

Since the end of 2019, COVID-19 outbreak has lasted for almost two years, causing significant social and economic impacts. In addition to the direct deaths and spreading viruses, millions of people lost their jobs, students could not enjoy fully in-person education, and businesses are facing downtimes or even closure. With so many negative impacts, the only solution to solve this problem is to end the pandemic. Two most commonly used methods by the government are wearing masks and vaccinations. However, there are certain disadvantages to these two methods. For masks, wearing a mask is simply uncomfortable and has a small upfront cost. For vaccines, there are multiple brands of vaccines approved by the Centers for Disease Control and Prevention (CDC). These vaccinations are developed with solid medical and biological research processes. However, the vaccination rate has not increased much in recent months because some people do not accept these quickly developed vaccines and doubt its effect to prevent COVID-19 or any side effects it may cause.

With only a portion of the population willing to wear masks and/or to accept vaccination, it is worth studying **the effect of wearing masks and taking vaccination in our community and how they may affect the infection rate**. Answering this question can help government agencies and medical facilities to better predict the future spread of virus and make decisions on what to focus on regarding COVID-19.

II. Background/Related Work

The effectiveness of masks and COVID-19 vaccines has been well researched. For example, a blog post by Stanford Medicine has shown that "surgical masks reduce COVId-19 spread". Interim Estimates of Vaccine Effectiveness of Pfizer-BioNTech and Moderna COVID-19 Vaccines Among Health Care Personnel posted by CDC has done thorough research on popular vaccine brands and has shown these vaccines being very useful against COVID-19 infections. This gives us confidence that both wearing masks and taking vaccination provide reliable ways for us to end this pandemic. However, the impacts of vaccines on communities can be different from that on individuals due to social interactions, especially when there are many

people not willing to wear masks or be vaccinated. On one hand, vaccinated individuals may be those who trust CDC's approval, and thus may also be the ones following CDC guidelines before vaccination. These individuals are at low risk of being infected, so being vaccinated will not decrease the infection rate by too much. On the other hand, vaccinated individuals may also protect unvaccinated individuals by lowering the spread of virus, thus achieving greater effectiveness in reducing infection rates. Thus, it is important for us to focus on a larger scope and examine masks and vaccine's effectiveness on communities, not just individuals.

III. Methodology

To answer the research question, we gather COVID-19 data and use visualizations. We primarily use these datasets:

1. CDC COVID-19 Vaccinations Data by County. This data contains the percentage of population fully vaccinated vs. only one dose by age group and by county, as well as other information about vaccination status. It has Public Domain U.S. Government license and is available for public use.
2. The RAW_us_confirmed_cases.csv file from the Kaggle repository of John Hopkins University COVID-19 data. It contains daily confirmed cases by county. It has Attribution 4.0 International (CC BY 4.0) license that allows the public to share and adapt.
3. The CDC dataset of masking mandates by county. This data contains the local order status of mask mandates. It has Public Domain U.S. Government license and is available for public use.
4. The New York Times mask compliance survey data. This data includes voluntary responses to an online survey of frequency of wearing masks. This data is licensed under the same terms as our Coronavirus Data in the United States data and it is available for general public use.

None of the datasets contains any personal information for us to identify individuals, and due to the large population, the granularity of data is not too small for us to recover such information. To process the data, we load data and clean the data using Pandas library in Python. Infection data is delayed by 6 days for visualization purposes as the virus has an average incubation period of 5.6 days, as indicated by the CDC blog. We calculate the daily incremental cases from the cumulative counts, smooth the data by taking the 7 day average to lower periodicity (due to more confirmed cases on Monday when everyone takes COVID test), and divide by total population. Finally, visualizations are created using the Matplotlib library.

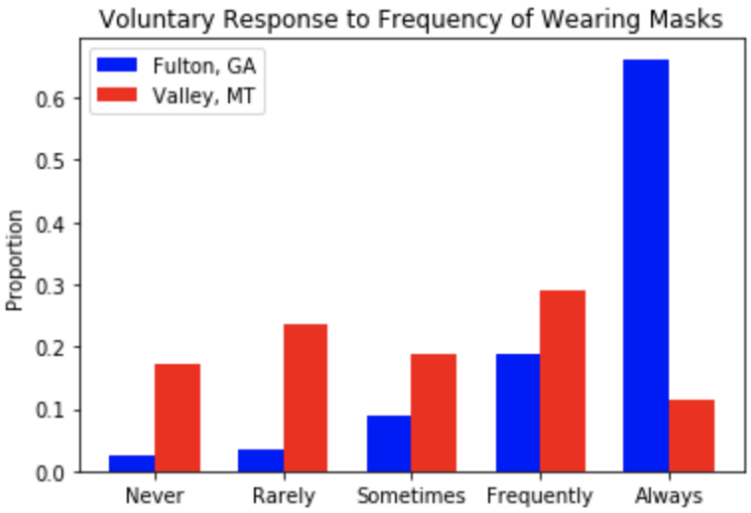
We decided to use data visualizations to answer the research question because it provides results most clearly and can engage non-technical audiences. It also will not violate any ethical considerations as the visualizations include aggregated data instead of individual data points of confirmed cases.

For the effects of wearing masks, we will first plot the cumulative infection rate and daily infection rate over time in Fulton County and color code the mandata status to compare the trend of infection with or without mask mandate. To reduce the effect of some random large events during the timeframe that affects the infection rate, we also compare Fulton County to Valley County because it has a much lower proportion of people wearing masks. We will compare their daily and cumulative infection rates using visualizations.

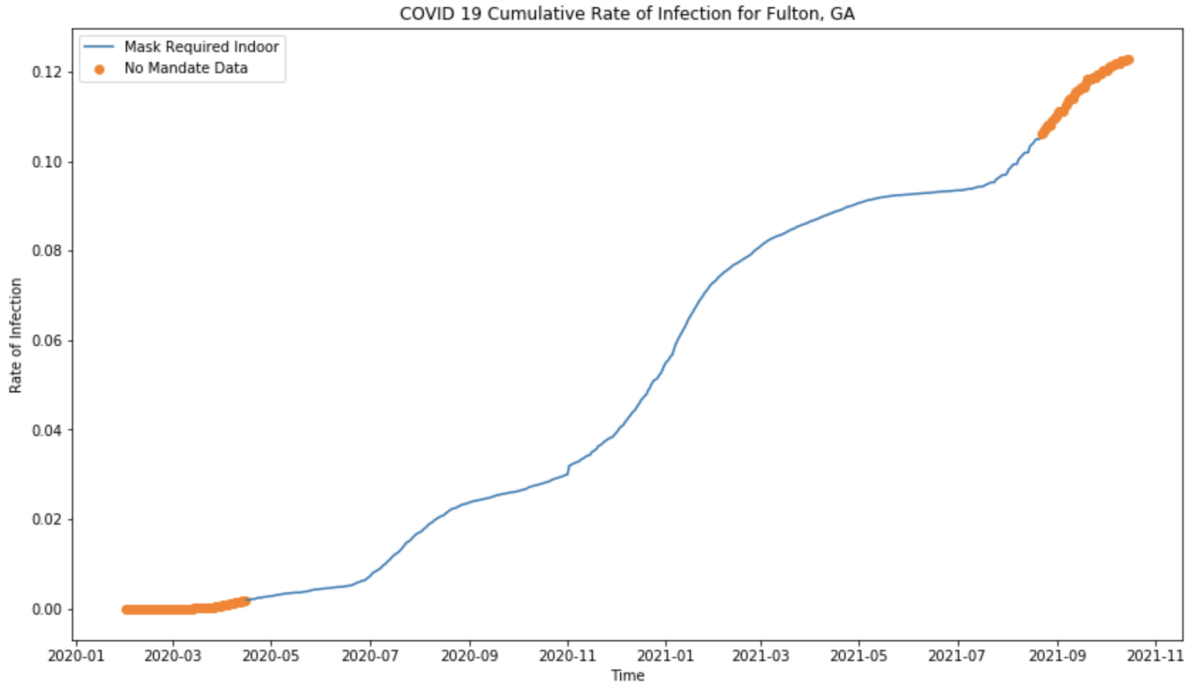
For the effects of vaccination, we will first plot the cumulative vaccination rate and daily infection rate over time in Fulton County to compare the trend of infection as more and more people become vaccinated. To reduce the effect of some random large events during the timeframe that decrease the infection rate, we also compare Fulton County to Valley County again because Valley County has a much higher vaccination rate. We will compare their vaccination rates and infection rates using visualizations.

IV. Findings

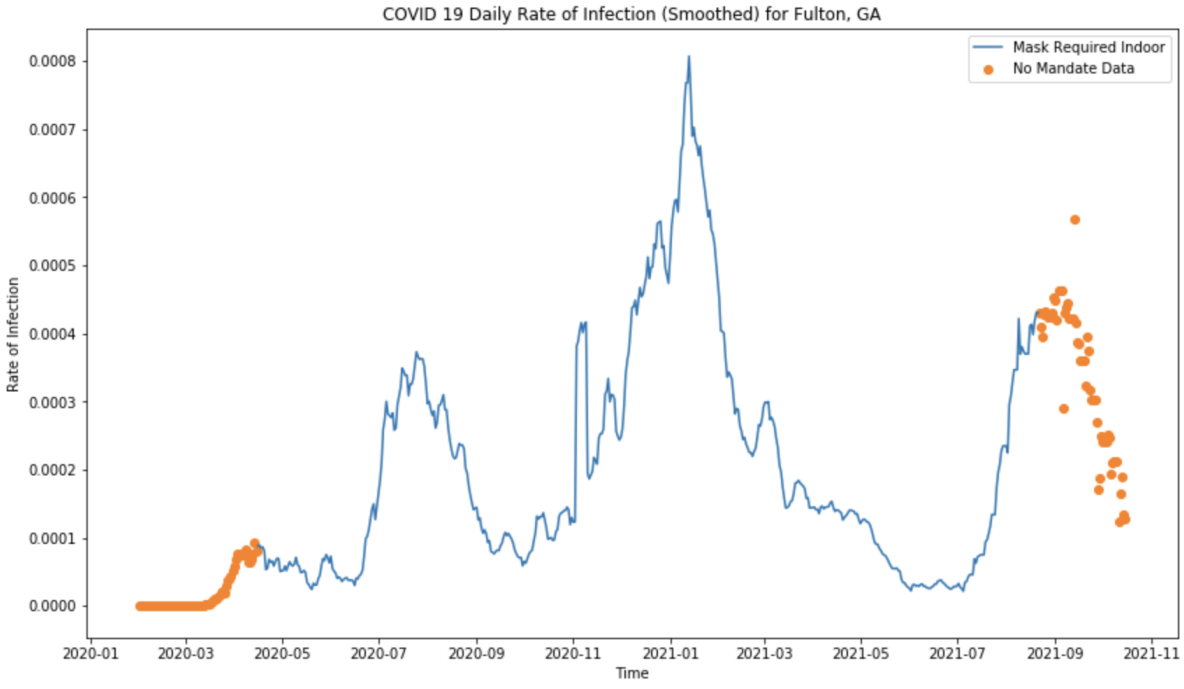
Mask Wearing Frequency Voluntary Response:



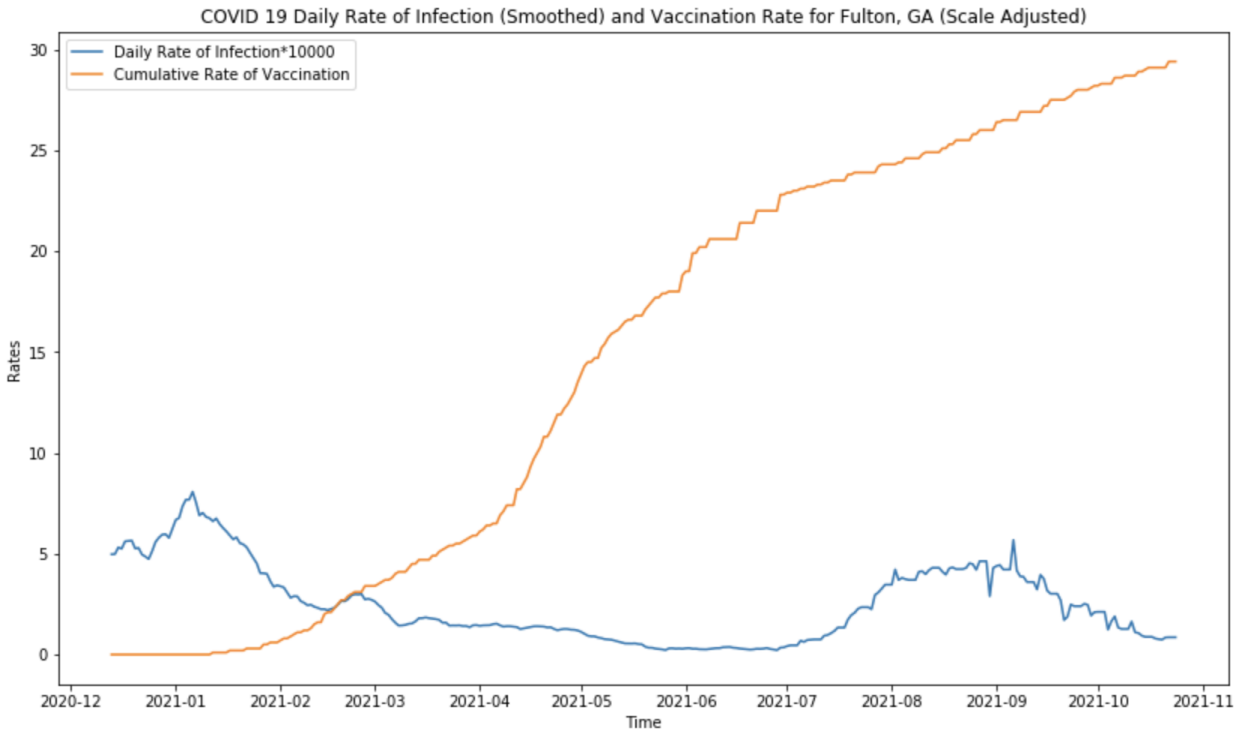
Fulton County COVID-19 Cumulative Infection Rate (with coded Mask Mandate):



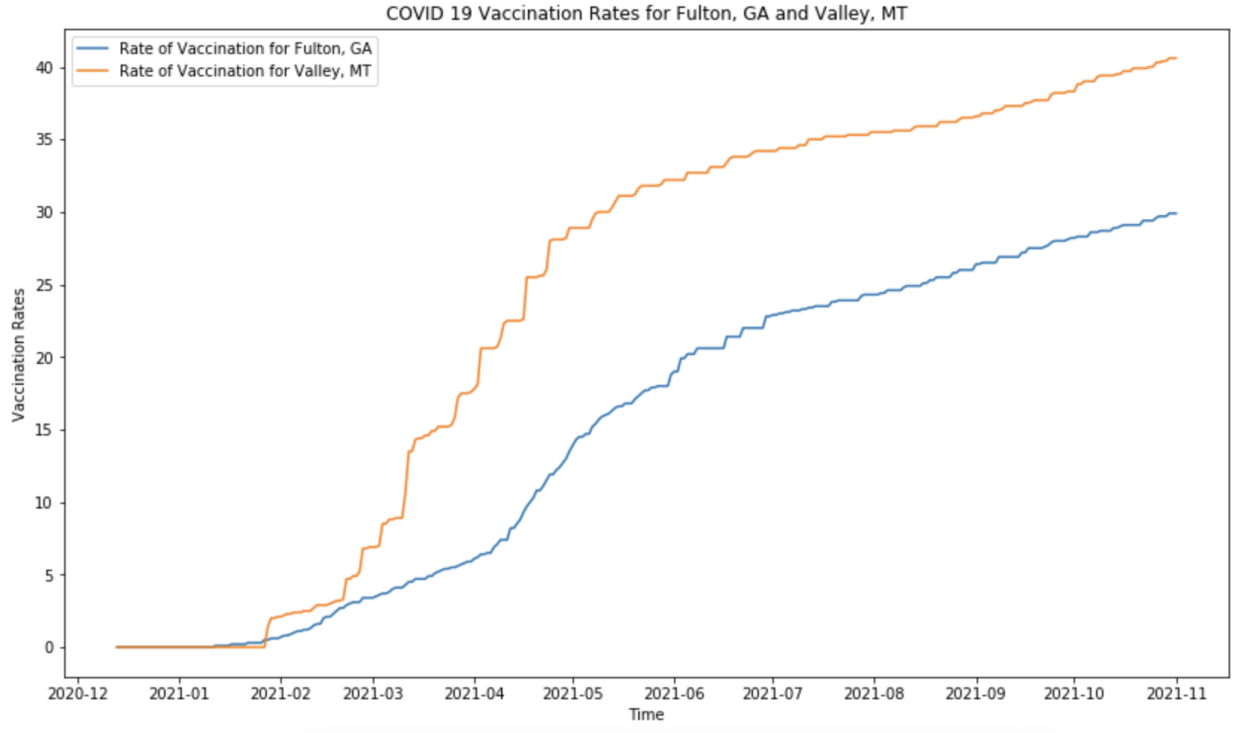
Fulton County COVID-19 Daily Infection Rate (with coded Mask Mandate):



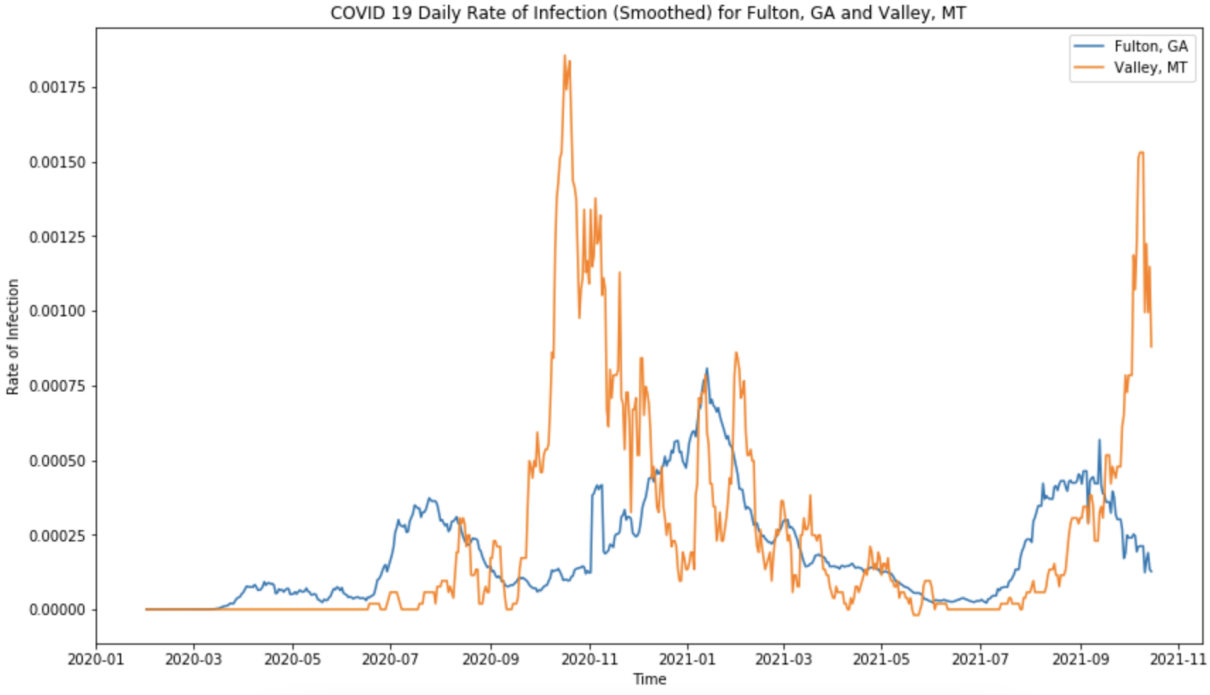
Fulton County COVID-19 Cumulative Vaccination Rate and Daily Infection Rate:



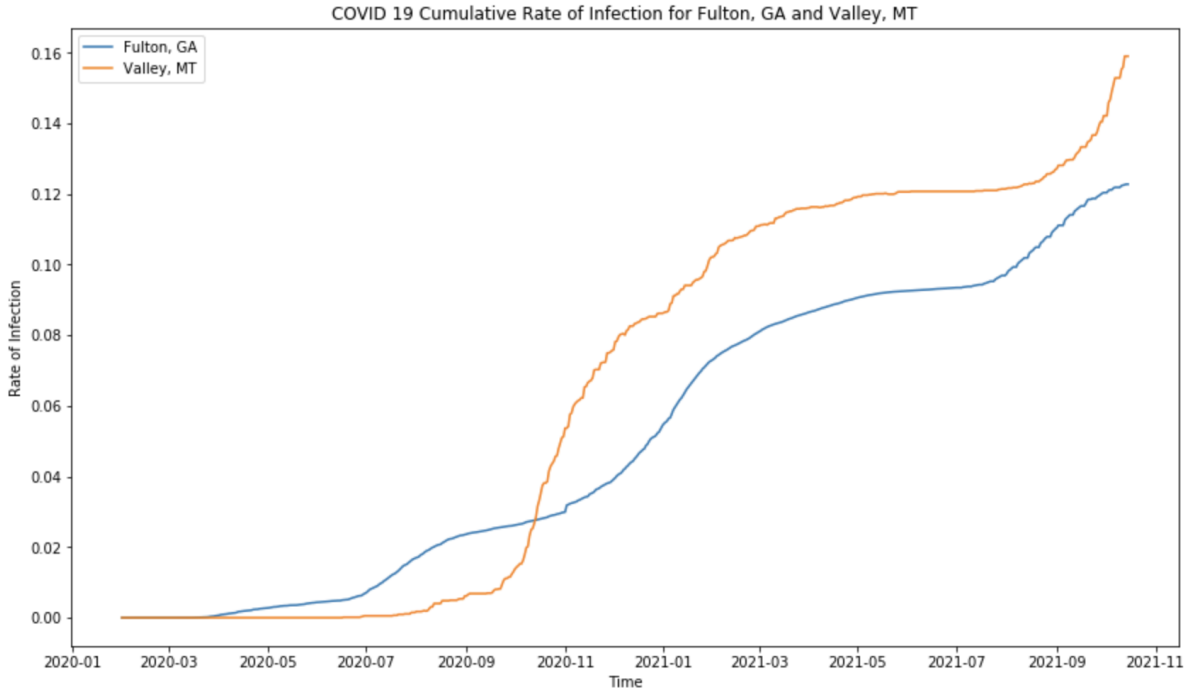
Fulton County and Valley County cumulative vaccination rate:



Fulton County and Valley County daily infection rate:



Fulton County and Valley County cumulative infection rate:



V. Discussion/Implications

From the visualization of Fulton County COVID-19 Cumulative Infection Rate (with coded Mask Mandate), we cannot draw clear conclusions due to lack of data on when public mandate is not required. However, by comparison with Valley County, it is clear that even though Valley County has a lower proportion of people wearing masks, there is not a clear sign of higher infection rate. Thus it seems that wearing masks does not reduce infection in a community. This might be due to the lack of consideration of people's frequencies of going outdoors and attending social events. People who do not wear masks may simply do so because they rarely go outdoor and thus do not necessarily need to wear masks.

From the visualization of Fulton County COVID-19 Cumulative Vaccination Rate and Daily Infection Rate, it does seem like there is a general downward trend of infection rate as cumulative vaccination rate increases. However, as mentioned in the method section, plots of time series are subject to the disturbance from random events that are not due to vaccination. Thus, a comparison of different counties is performed. Although Valley County has a vaccination rate of 40% at the end, much higher than Fulton County's 30%, there is no sign of lower infection rate for Valley County from either the cumulative or daily infection rates of the two counties. This result is not expected. One possible reason is that people who got vaccinated have been following CDC guidelines before being vaccinated, so their likelihood to get infected is low anyways. It is also possible that vaccination makes people more confident of immunity to the virus and thus they increase their social activities and social interactions.

Both results are surprising to me, but due to so many restrictions of data and assumptions made, as listed in the next section, these conclusions may not be valid. Future studies may also take into account geographical locations of large social events (such as elections) to improve the strength of this study. They may also study the reason behind this unexpected result, which can help the government agencies make correct decisions. For example, if it is determined that people socialize more after being vaccinated, then government officials may take actions to keep limiting dining in restaurants or closing some entertainment facilities from the public.

VI. Limitations

Below is a list of limitations in this study. Future research is encouraged to verify assumptions/break the limits to provide more reliable results.

- Plot over time series may include randomness from time events.
- Comparing only two counties may include much randomness in results.
- Visualization may not be the best statistical method to verify the relationship between vaccination and infection.

- Vaccinations impacts may be more significant for vaccination rates above 50% that are studied for Valley County.
- Data does not include information for any booster vaccination.
- Mask survey information includes only five categories and respondent's answer cannot be verified.
- Survey may include non-response bias into our study.
- County mandate data includes missing values that limits the date range of visualizations.
- The compounding effects of masks and vaccinations are not considered.

VII. Conclusion

In this study, we examined the effects of wearing masks and taking vaccination on infection rates in communities. We use data to create visualizations to display patterns. From the results, we may conclude that mask mandates and higher vaccination rates do not necessarily lead to lower infection rates in a community. This study, of course, comes with certain limitations and assumptions, so this conclusion can be further verified or denied in future studies. We hope that this data driven study can be useful for government agencies to make appropriate decisions to stop the spread of viruses and improve the lives of human beings, and can inform more researchers to take the advantage of this datafied era to provide more novel suggestions about ending COVID-19.

VIII. References

- Interim Estimates of Vaccine Effectiveness of Pfizer-BioNTech and Moderna COVID-19 Vaccines Among Health Care Personnel
- Surgical masks reduce COVID-19 spread, large-scale study shows
- Coronavirus Disease 2019 (COVID-19)
- Pandas
- Matplotlib

IX. Data Sources

- CDC COVID-19 Vaccinations Data by County
- RAW_us_confirmed_cases.csv from the Kaggle repository
- The CDC dataset of masking mandates by county
- The New York Times mask compliance survey data