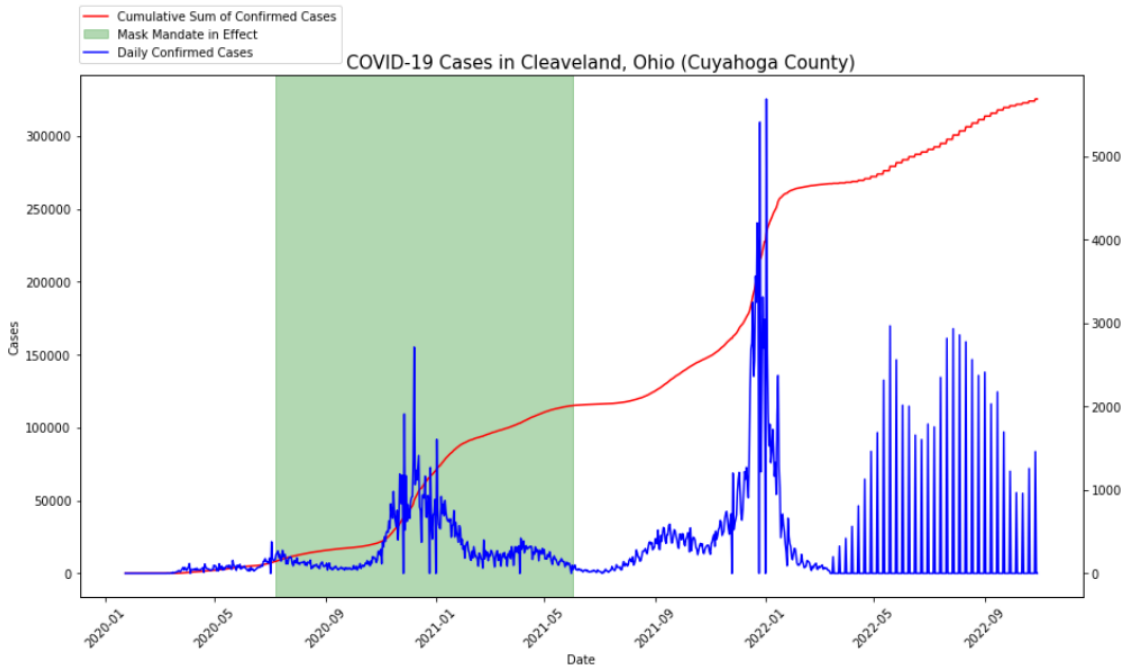
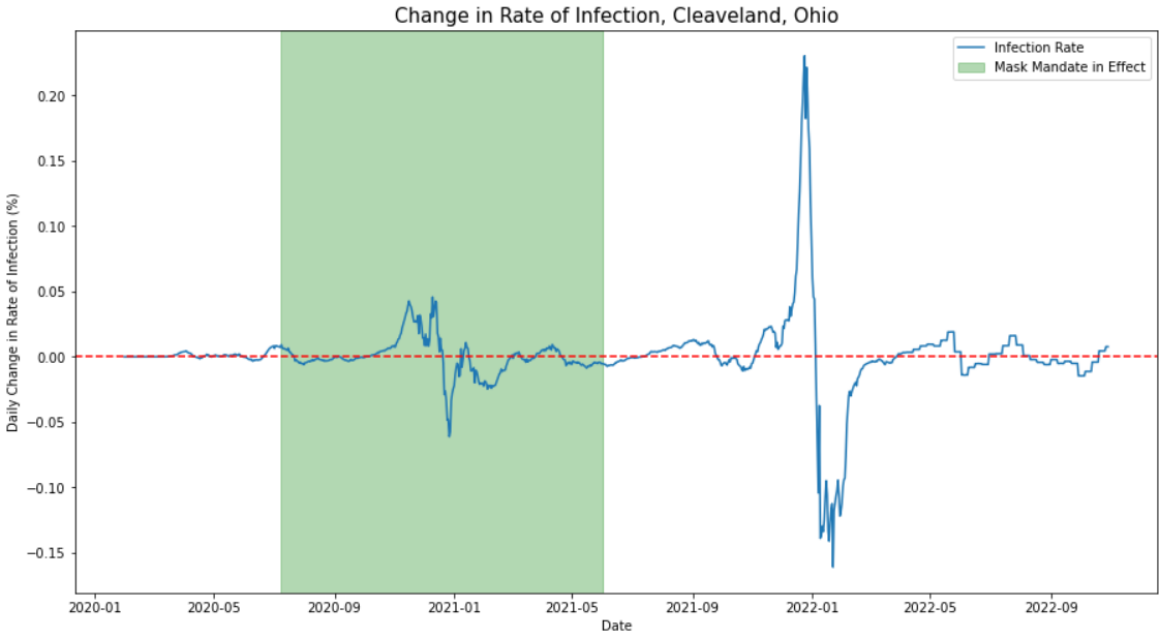
Visualization Explanation



The first figure I created shows the daily confirmed cases in blue with the cumulative sum of confirmed cases in red. The mask mandate period is highlighted in green with a range of July 2020 to June 2021. The x-axis represents the date and the y-axis represents the number of confirmed cases. We see that the daily confirmed cases have started going down shortly after the mask mandate has been placed, but then there is large resurgence of cases around November-January 2021 which is when the holidays were. People traveled and had family gatherings where they may not have worn their masks. The cases declined shortly after the holidays and remained relatively stable. After the mask mandate, the cases did not increase drastically until around November-January 2022. During this time, even more people traveled and hosted family gatherings because of the loosening restrictions.



This second graph represents the change of infection rate. The mask mandate period is highlighted in green with a range of July 2020 to June 2021. The x-axis represents the date and the y-axis represents the percentage of daily change in rate of infection. The daily infection rate is the proportion of the population at risk who are considered infected. Population at risk overall represents people who are not currently infected and this includes people who were infected previously. I applied a 7-day moving average for the daily confirmed cases and used the moving average to find the infection rate. After the mask mandate was put in place, we see that the change in infection rate has dropped below 0% and has stayed under 0% until about November 2020 and has stayed positive until it sharply dropped to -0.05% around January 2021. We see a similar trend in November 2021 to January 2022 but on a larger scale since the travel restrictions were looser and the mask mandate was no longer in place.

Collaboration Reflection

Collaborating with others really helped made me think more closely about how I want to do my analysis. A few people (Charles Reinertson, Tharun Reddy) in the cohort suggested plotting change points for the change of infection rates using Facebook Prophet and the Ruptures library. A few people (Arik Shurygin, Charles Reinertson) also showed how they preprocessed the data which greatly helped me get started with the project faster. I also got the idea to highlight the mask mandate period from other graphs that people shared (Eli Corpron). I was also initially unclear about what to do for the derivative function of the change in rate of

infection, but collaboration with Arik and Tharun cleared up the calculation for rate of infection and how it would be best to apply a 7-day rolling/moving average to detect the infection rate more accurately. While others have suggested more complex and scientific methods of creating their visualizations, I wanted to minimize the amount of data ink on my own visualizations to help people be able to read and interpret the graph more easily. I decided against including change points since this may have added too many lines to look at and may become overwhelming. Overall, collaboration with others helped me finish the project more quickly and I was able to generate insights that I would not have even considered without discussing the project with others. Discussing with others has allowed me to better understand various approaches and helped me keep in mind the different assumptions I would have to consider for my specific analysis. Some key assumptions I have made along the way are that we should factor in the delay in reporting confirmed cases and this is done with the rolling function which smooths out the change of rate of infection, we can test the derivative function by taking a second derivative, we will be ignoring vaccination statuses since vaccines mainly helped in reducing deaths, and we also assume that the infection rate is the same across the different variants.