

A4: Common Analysis

I. The visualizations

How often do people wear masks in public in Macomb County, MI

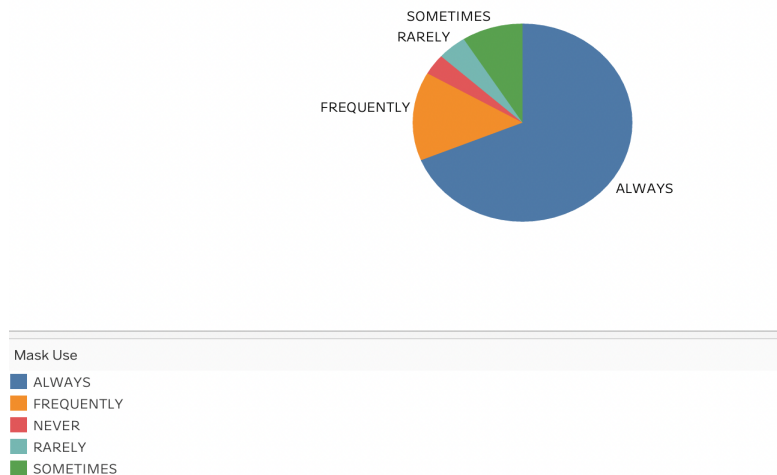


Figure 1: How often do people wear masks in public in Macomb County, MI

Confirmed COVID Cases in Macomb County, MI

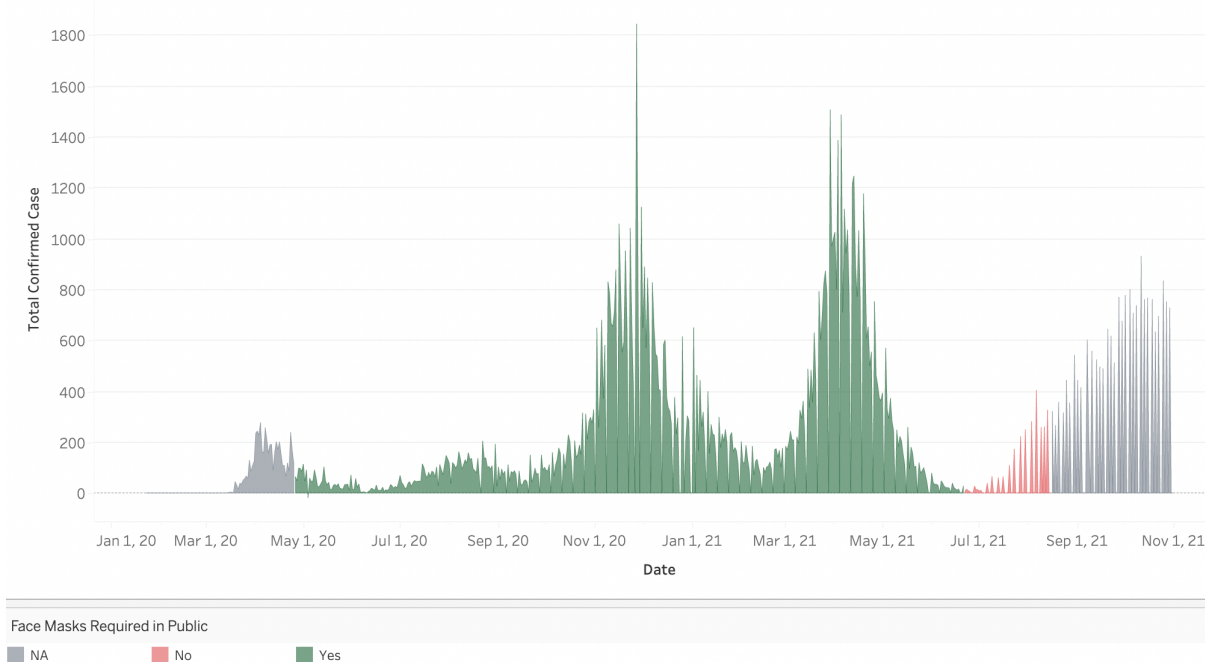


Figure 2: Confirmed COVID Cases in Macomb County, MI

II. About this visualizations

Figure 1 shows how people in Macomb County, Michigan would respond to the question “How often do you wear a mask in public when you expect to be within six feet of another person?”, where segments of the pie chart represent the estimated share of people in Macomb County. The data for Figure 1 is generated from a survey, and from this visualization, we can see that over 80% of the whole population in Macomb County wear masks in public frequently or always.

Figure 2 shows the total number of confirmed COVID-19 cases in Macomb County, Michigan from January 1, 2020 to present. The data was processed by extracting data of Macomb County, Michigan from the combined datasets of *the US confirmed COVID-19 cases from February 1, 2020 through October 15, 2021* from the Kaggle repository of John Hopkins University COVID-19 data, and *Public Mask Mandates From April 10, 2020 through August 15, 2021 by County by Day* from the CDC dataset. The number of confirmed cases for each day are plotted as a bar on the timeline. The bars are colored in three different colors, indicating whether wearing face masks in public is a requirement on that day. The red bars indicate that wearing a face mask in public is required, whereas the green bars indicate that wearing a face mask is not required. The gray bars in 2020 show that there is no such masking policy implemented by the official sources, likely because the pandemic just started. The gray bars in 2021 show that we do not have data showing if masks are required, but my assumption would be there is no masking policy once it gets lifted in July 2021.

Figure 2 is a time series showing the changes in the derivative function of the rate of infection. According to Wikipedia, the infection rate is the total number of infections, as known as the total confirmed cases, divided by the population at risk. Assuming the population at risk is the population of the whole county, and also assuming the population of the whole county stays roughly the same during this whole pandemic, the derivative function of the rate of infection would look the same as the derivative function of the total confirmed cases of COVID, which is shown by Figure 2.

The infection time series is a set of slopes, and the difference in the derivative function is significant, which indicates the effectiveness of masking policies. From Figure 2, we can see that although there are two peaks of the number of confirmed cases, the number of confirmed cases decreased twice, during the period that the masking policy was enforced. However, after the masking policy was lifted, the number of confirmed cases increased dramatically. Although there are many other factors and potential issues of what caused the number of confirmed cases to increase or decrease, we can conclude that the implementation of masking policies helped with the decreasing of the confirmed COVID-19 cases from February 1, 2020 through October 15, 2021.

III. About the collaborative research

It is inspiring to see other people's analyses and visualizations. Although I did not reuse any code or techniques from other people, it is fun to compare my results with other people's results. By comparing the results from different states, it raises the question of whether the geographic location, the climate, the culture, and the population plays an important role in 1) if the masking policy is hard to enforce, and 2) if either one of these factors plays a important role in the progression of confirmed COVID-19 cases.

One particular thing I learned (from the thread on Ed posted by Aaliyah Hanni) is how to define the rate of infection, and analyze the rate of infection from the data we have. The discussion regarding "population at risk" was very interesting, because who is "at risk" at a particular time point can depend on many factors -- too many factors for us to analyze without further research or investigation. Therefore, the assumptions I took, which is assuming the population at risk is the population of the whole county, and also assuming the population of the whole county stays roughly the same during this whole pandemic, should be sufficient for this analysis and can also simplify things a lot.