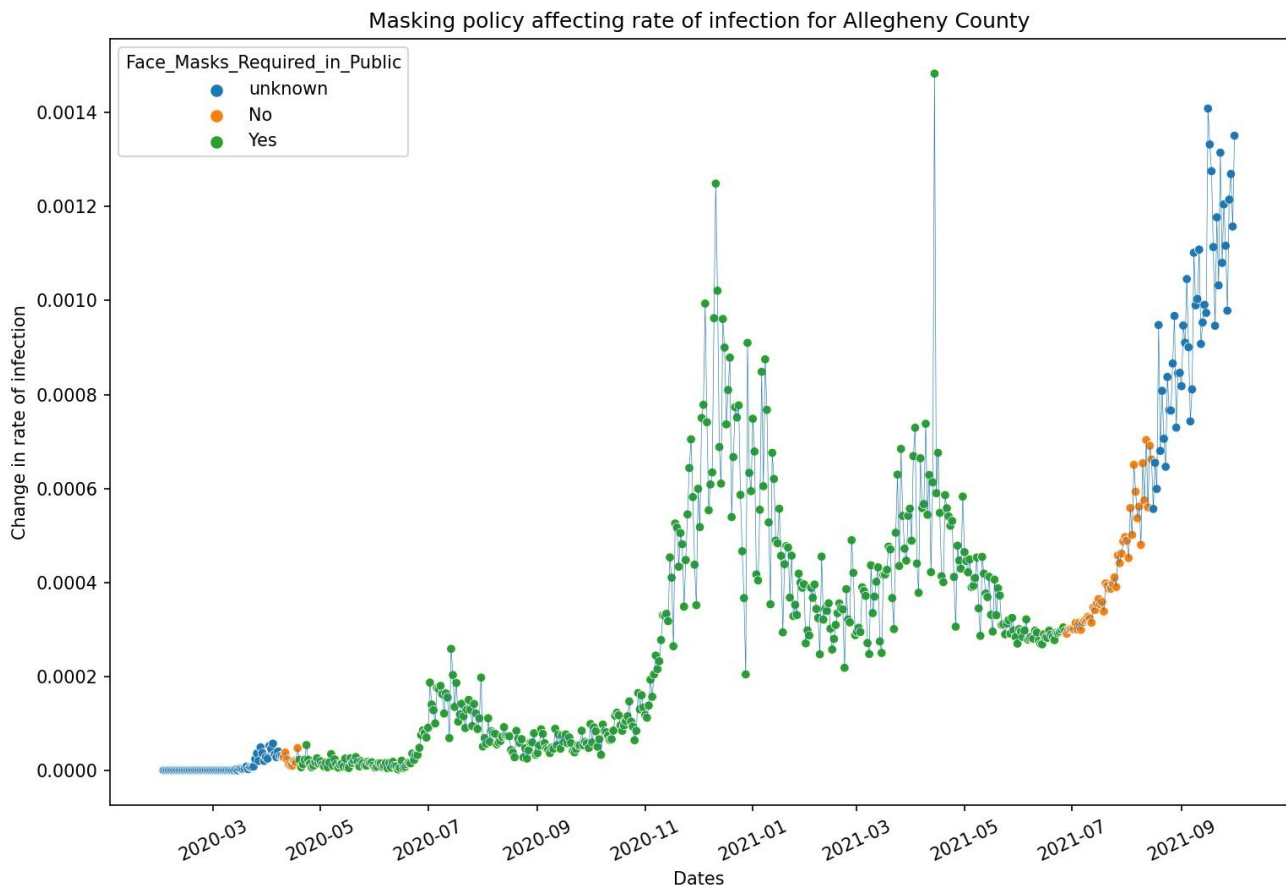


Explanation of Visualization

I performed the common analysis for Allegheny County and here is the final visualization displaying the masking policy over time and how that impacts the change in rate of infection.



Calculating rate of infection:

The rate in change of infection was calculated using the following equation:

$$\text{Rate of infection} = K \times \frac{\text{the number of infections}}{\text{the number of those at risk of infection}}$$

where k was taken to be 100 and the denominator was calculated by subtracting the current population with the number of deaths. Also the current population was updated at a daily level.

Getting change in rate of infection:

Once the rate of infection was obtained, the next step was to find the change by subtracting the rate of infection with the previous day. This helped me understand how the numbers were increasing or decreasing from the previous day.

Having this information can help us find out if there was an improvement based on the masking policies.

Creating Visualization:

After collecting all the required information, the above plot was generated by having date on the x-axis and change in rate of infection on the y-axis. Additionally, the hue was added to denote the status of masking policy.

Conclusion:

We can see that as the cases started to increase, the masking policy was introduced. There was a period of stable infection and then two high waves or peaks. As opposed to the expectations, some exceptions can be seen in this region with two days of extremely high cases.

The numbers decreased around 05-2021 the masking policy was changed to “no” and we can see another peak right after this time. This was expected as the infection spread would have increased with no masks.

Reflection on Collaborative Activities

Getting ideas and learning from others in the batch is a very important aspect in career growth. Having the opportunity to discuss the assignment with my colleagues on discord helped me learn a lot of new concepts which I probably would not have come across if I were working on my own. Here are some of the concepts we discussed over our chats on the discord channel-

1. Change points analysis

I had never worked with change points before. Charles and tharun shared their views, blog links and codes which started a healthy discussion about what are change points and how can that visualization be useful.

2. Rate of infection calculation

This was the idea which had the most discussion and views among the class. Everyone shared different points and a few blogs to understand what could be the different ways of looking at this. We finally discussed how the susceptibility of population is an important aspect for this calculator and determined the factors that would affect this.

Other than that we spent some quality time working on identifying the factors that might be the reason behind the change points we are seeing. Found out the number of days and recovery rates which could be used to better understand the results.

I did take into account some of the points I learned from this discussion. Here are the key takeaways I was able to implement on my analysis

- 1) After the discussion, I included the dataset about the number of deaths in my analysis as well. The idea of working with a susceptible population made logical sense and I used this population for the calculations.
- 2) Updating the rate of change of infection calculation after iterative discussions. We did agree that there can be multiple ways of reaching a formula and more than one point of view for this problem.

Overall we had some really good interactions within the group and all of us had some new ideas to share.