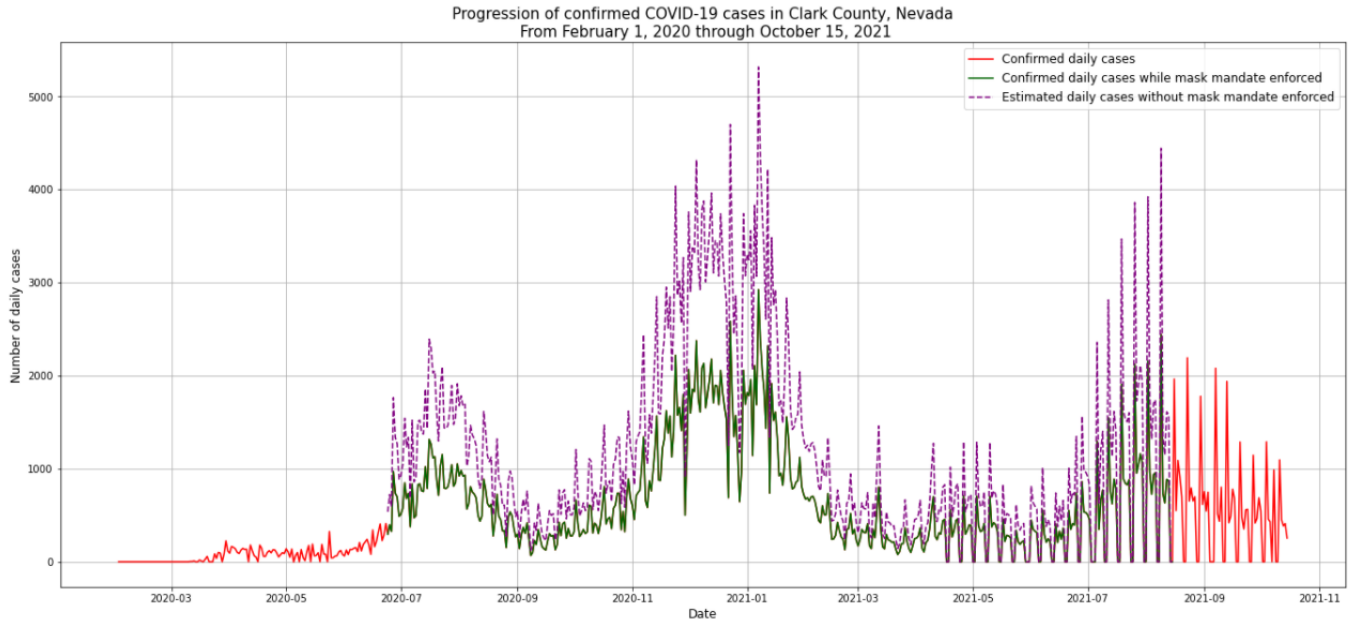
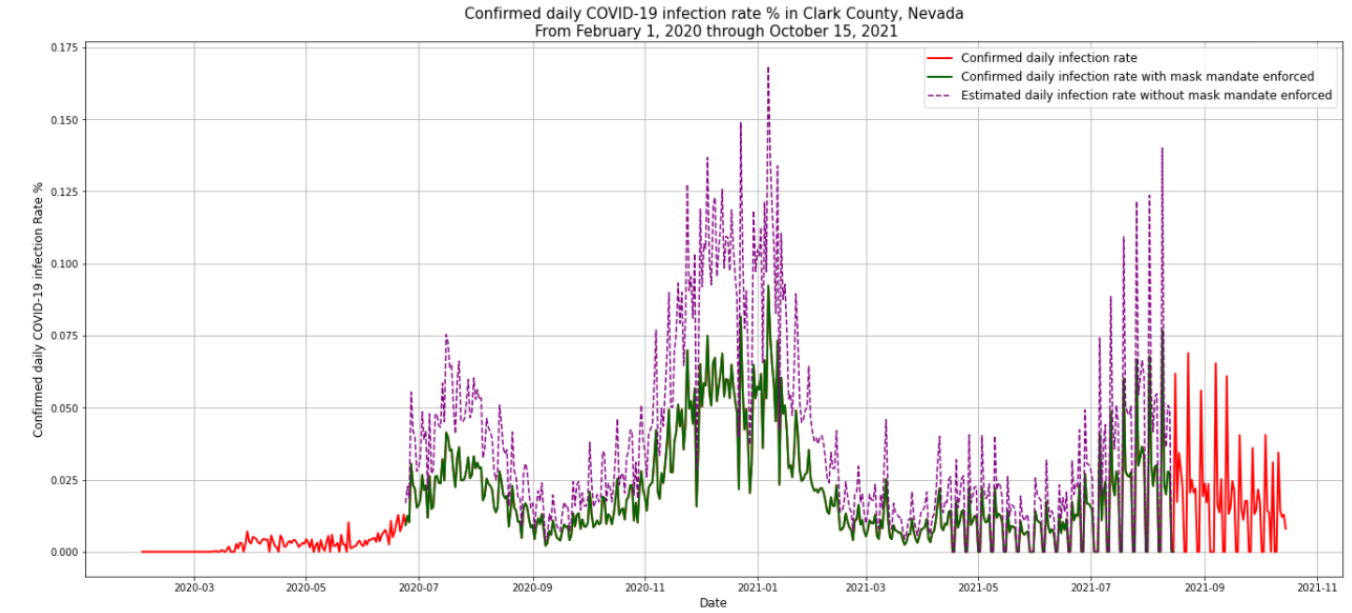
Dhruv Relwani
Data 512, UW MSDS
A4 Assignment

Images:





Explanation:
The above images show the progression of confirmed Covid-19 cases in Clark County, Nevada from Feb 1, 2020 to Oct 15, 2021. In the first image we can see the raw daily case count for confirmed Covid-19 infection, in a timeseries whereas in the second image we see changes in the derivative function of the rate of infection. On the Y-axes we have information about the daily

Covid cases in this county and on the X-axes we have the months represented for the pandemic.
The underlying data for these visualizations is a combination of 3 datasets:

1. The RAW_us_confirmed_cases.csv file from the Kaggle repository of John Hopkins University COVID-19 data. This is where I got information about confirmed Covid-19 case counts.
2. The CDC dataset of masking mandates by county. Here I got information about the dates during which a mask mandate was enforced in Clark county.
3. The New York Times mask compliance survey data. This dataset provided information about mask mandate compliance probabilities.

I combined these datasets into pandas dataframes and then selected just the relevant data for Clark county, NV for the timeframe of the assignment. I was only given cumulative case counts in the first dataset and hence had to calculate the daily counts myself by incrementally subtracting from the total count. Then using this info I was able to get a rate of change of cases every day for the second visualization.
One additional derived field in this data was the estimated case counts if the mask mandates were not enforced. I want to acknowledge that this modelling was inspired by Anushna Prakash's visualization pasted in the Slack channel. For this estimated case count I assumed a 55% efficacy of masks in preventing Covid-19 infection, a 10-day infection period from only the confirmed and alive cases (non-asymptomatic) and anyone in Clark county who was alive and not confirmed to be infected as "at-risk" for infection. The population data for this county was assumed by the Census data.

Reflection:
While answering the research question and developing the model I noticed that mask mandates were statistically significant in preventing worse outbreaks and curbing the rate of infection (from its potential) in Clark county. Even for a person who might consider masking as a non-deterministic atrocity with only a near 50% probability to help with containing Covid, this model suggests otherwise.
I learned a lot by viewing everyone else's visualizations in the Slack channel. As mentioned above, I learned how to calculate the estimated case count without mask mandates from Anushna's visualization (specifically what assumptions could simplify my work). Additionally, I saw Apoorv's visualization and learned that grid lines may be useful for representing this information in a timeseries for easy comprehension. I really liked some of the models which had accounted for standard deviations of rate of infection spread and would like to inculcate that in my project.