

Eli Corpron

Data 512

11/3/22

Part 1: Reflection

I found the collaborative activities for this part of the project to be extremely helpful. At first I was feeling a little overwhelmed and wasn't sure where to start, but by talking with other people in the class I was able to develop a plan and decide on how to move forward. The two areas that I think this helped me the most were on how to approach cleaning the raw data and making it usable, along with how to approach a visualization that showcases "changes in the derivative function" and "whether the difference in the derivative function was significant".

For cleaning I found it very helpful to collaborate with other students in the class. That helped speed up the process in figuring out which fields were important and how they connected between data frames. What came back to bite me was that I hadn't completely understood all of the cleaning steps, so that when I started working on my visualization it turned out I hadn't actually cleaned my data correctly, and a lot of entries for the masking mandate were NA when they should have been filled out. This required me to then come through all of my cleaning steps, reorganize my code, and approach merging the CDC data and the John Hopkins data slightly differently so that it would work for me. In the end, this did challenge me to fully understand everything I was doing in the cleaning step, so I am still glad I worked with others on it. Arik was extremely helpful in deriving the `pasoDeathsByDate` dataframe and the `eIPasoMaskMandates`

dataframe, but I did come up with how to fully merge the mask mandate data in `pasoDeathsByDate` on my own.

The other area I found it extremely helpful to collaborate in was in designing the final figure. Initially I came up with plotting the deaths per day in different colors based on the current mask mandate. However, I was stuck on how to showcase the significant changes in the death rate. Charles brought up change point detection which started me thinking about it, but Tharun mentioned the `ruptures` package. I liked the way `ruptures` presents data, and thought that it would pair nicely with how I planned on displaying my data, so I looked into it more. From there I decided to use the Pelt method to display change points in my graph since it is quick and gives exact returns. I then picked a penalty parameter that I thought created some intuitive change points.

In the end, I found it extremely helpful to collaborate with other students. It helped me formulate a plan on creating a visualization, and also challenged me to fully understand what my code did and why. I appreciate the opportunity to work with others, and hope that I will get to do so in the future.