Data 512A Class Project Part 1 Reflection Statement

Find Rate of Infections

(*Credit to Arik's code snippet on course Slack channel for my data transformation part***)**

One of the most important things I learned from this assignment is finding the rate of infection from cumulative cases. My initial ideas to obtain the rate of infection is using the diff () function in Pandas to convert the cumulative cases to daily cases. However, after reading Arik's code snippets on course Slack channel, I realized that the rate of infection should be calculated by the first derivative of the cumulative data, and the gradient () function in NumPy allows for calculation of derivatives.

Change Point Detection

(*Credit to Detecting the Change Points in a Time Series for my change point detection*

Credit to TharunReddy for sharing the article on course Slack channel)

My previous experience for change point detection on time series was based on R changepoint package, and this is my first time try to identify significant change point using python. I found plenty of interesting strategies on course Slack channel shared by my classmates including: the Facebook Prophet, the Ruptures Package, and the Changefinder Package. Among all these methodologies, I found that the article *Detecting the Change Points in a Time Series* shared by TharunReddy has most clear explanations and detailed examples. With the help of the Ruptures package, I found the statistically significant change points for the derivative function of the rate of infections.