Stage 4 Report

Alyiah:

For this stage I worked with the group generating the linear and polynomial models as well as finding the predictions for the US. I also generated the predictions for cases and deaths of pakistan. Overall while the US trends continued to decrease, the Pakistan trends increased.

Jacky:

For Stage IV, I worked together with the group when creating linear and non-linear polynomial regression models. When it came to country comparisons, I created models for the prediction of COVID trends in Bangladesh. When comparing Bangladesh to the United States, there is a huge difference in covid trends as Bangladesh sees roughly only double digit cases each day. The United States on the other hand is still coming out of the recent surge and still sees tens of thousands of cases each day. US trends seem to be on the decline while Bangladesh overall has seen low trends throughout the pandemic.

Annie:

For stage 4, I worked with the group when creating the prediction paths for polynomial regression. In the country comparisons, I created the polynomial regression and prediction path for Indonesia. The prediction path showed that Indonesia will reach about 51,000 cases and 1,700 deaths. Compared to the US, Indonesia had a more rapid increase in cases and deaths whereas the US was more gradual.

Sage:

I started off by creating the preliminary DataFrame for cases and deaths in the US. There was no need to crop any of the data in the US confirmed cases dataset since the first day of reported cases was the first date in the dataset. From the data, the first US COVID-19 death happened around 15 days after the start of the dataset, so I cropped those first few columns. I created a DataFrame for cases and deaths that included columns for days_since and new_cases. For the analysis of other countries with a similar population, I analyzed Brazil's curve showing that, with a prediction, there would be a decrease in cases. However, I found that the trend line was suffering from too high bias because it should have depicted a potential increase in cases. My individual findings showed that some counties in TX were at a higher risk of reaching peaks than other counties. By using a third order trend line on the top five highest-case-count counties in TX I found that Harris County's trend was going to rise rapidly.

Logan:

For stage 4 I assisted in creating the preliminary DataFrame for cases and deaths in the US. I also assisted in plotting linear and non-linear graphs of our datasets as well as contributed input to analysis of the data presented by the generated graphs. I assisted in creating our prediction graphs and also did analysis on the country of Nigeria. It appears that Nigeria closely mirrors the US in cases and to a lesser degree deaths. In both scenarios the graphs suffer from high bias and appear to trend down despite the data ending on an upward trend.