

Project Stage IV Team Report

Contributors: Alejandro Penaloza, Amantii Samson,
Quinn Tjin-a-soe, Peter Yuan

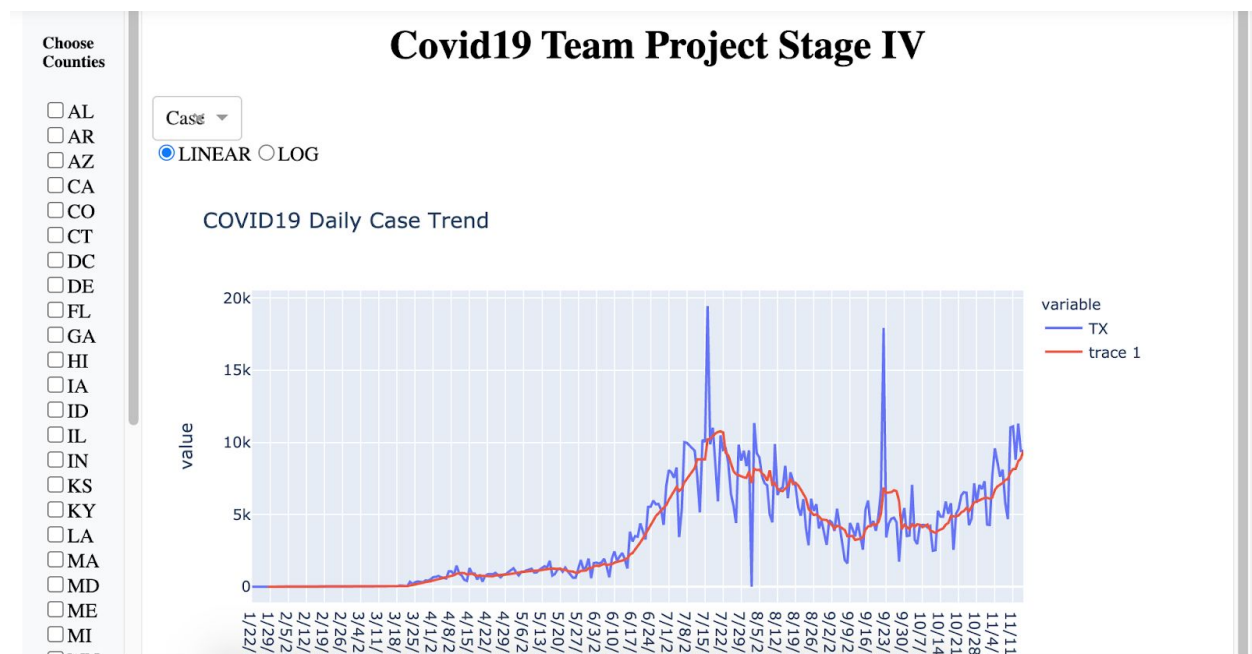
Goals:

“The final stage aims at developing a simple interactive dashboard based on the analysis you have done so far. In this we will be utilizing Plotly (<https://plotly.com/>) along with Dash (<https://plotly.com/dash/>) as our framework.”

Tasks:

- Task 1 (TEAM)
 - Allow for selection of date to show the trend of COVID-19 cases and deaths.
 - Allow for linear or log mode selection on the number of cases and deaths.
 - Incorporate your best model prediction trend line - Linear / Machine Learning (Decision Trees, Random Forest) and ARIMA you have identified in the previous stage.
- For this stage, we built datasets for the cases and deaths the United States broken down by states. Random Forest was the best model prediction line that was chosen to predict a day ahead in the data reflected by the dashboard screenshot below! The dashboard was then constructed in a way that would allow for users to select whether they wanted to see cases or deaths, what states they wanted to see it for, and whether they wanted it in a linear or log form!

Below is a screenshot of the dashboard showing the cases for Texas (Blue Line) along with the 7-day moving average (Red Line)

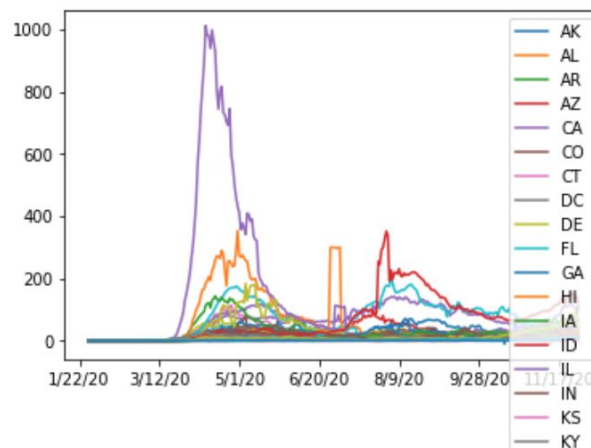


- Task 2 (TEAM)

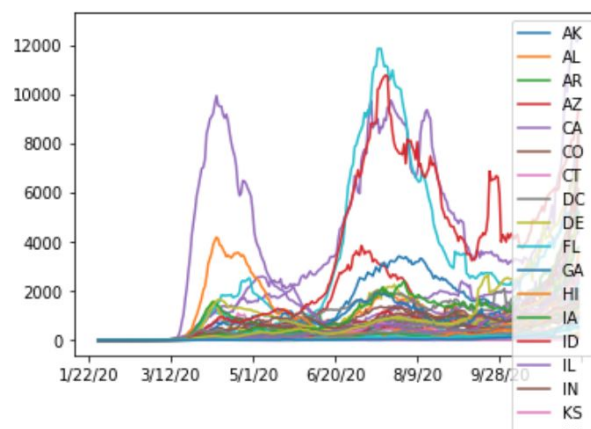
- Trend

- Plot the trend line using moving average
(https://en.wikipedia.org/wiki/Moving_average). Use a 7-day moving average.
- Done using a simple rolling() method call on the mean of the data frames of cases and deaths.

Deaths



Cases



- Map

- Show a US map of cases or deaths (normalized by population). Entire US with county level information. Use the mean of the county population to normalize for cases and deaths.
- Data from cases, population, and deaths were extracted and normalized by population. This data was then put into its own

data frame and used in a simple map of the US that shows the infections and deaths down to each county level!

