

Stage IV

The final stage of the project was to utilize the plotly package to create a dashboard to visualize all the trends and prediction for different states and counties across United States. All the functionalities of the dashboard are covered in depth in the presentation. Here I will present the high-level details about the dashboard.

To create the dashboard, I initially defined different functions to carry out multiple operations in our dataset. I have separate functions to extract county level, state level data from our superdataset. Along with that, following are all the functions with their respective purpose:

1. `extract_data` – Extract cases or deaths from superdataset (based on user's selection in the dashboard), convert wide format to long format data
2. `remove_before_first` – Remove all rows before first occurrence of cases or deaths in user's selected state or county
3. `rolling_avg` – Get the 7 days rolling average for new cases or deaths
4. `linear_normalize` – Normalize cases or deaths per county average population
5. `log_normalize` – Log normalize the number of cases or deaths
6. `remove_negatives` – Replace any negative values with 0
7. `linear_fit` – Fit the data into linear regression model and create prediction
8. `poly_fit` – Fit the data into polynomial regression model and create prediction

The final dataframe after all these steps are used differently based on different combination of options selected by the user on the dashboard. The figure (plot or US county level map) is returned to the calling function to present in the dashboard. I used plotly's `choropleth_mapbox` to present the data in map.

Following are the snapshots of the dashboard followed by a short description of what the figure represents:

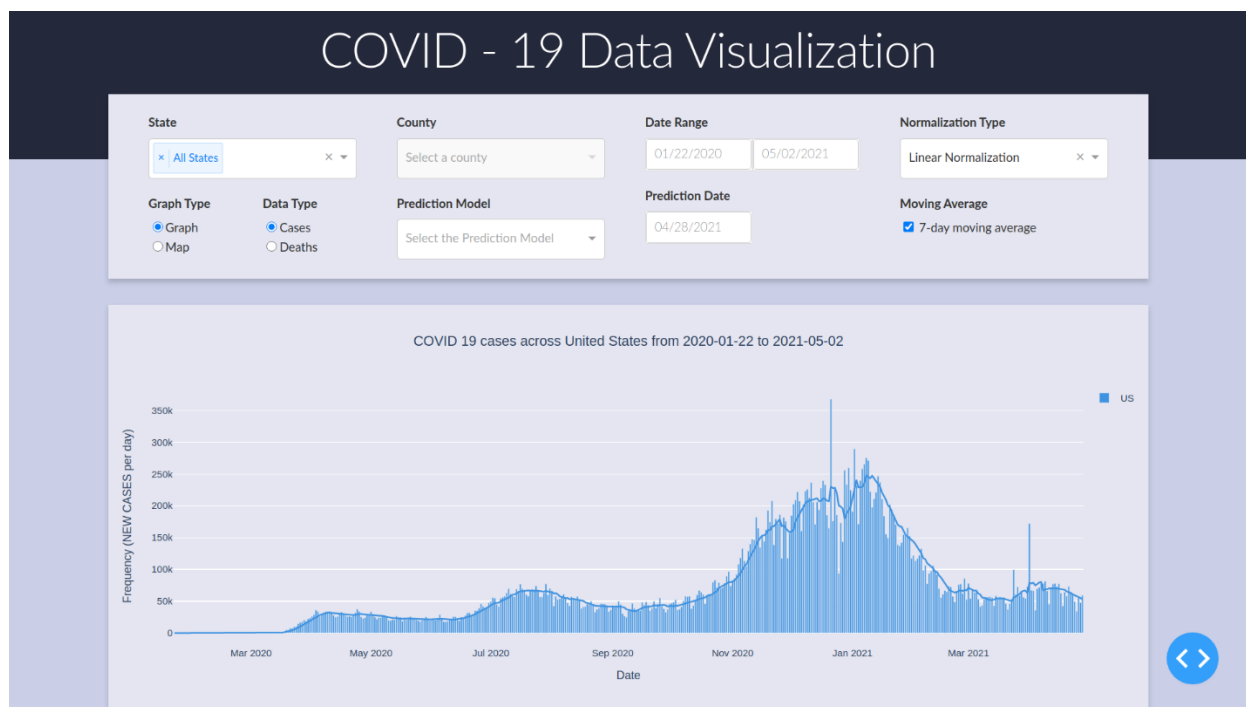


Figure 24: Landing page – Bar plot for daily new cases across US with rolling average

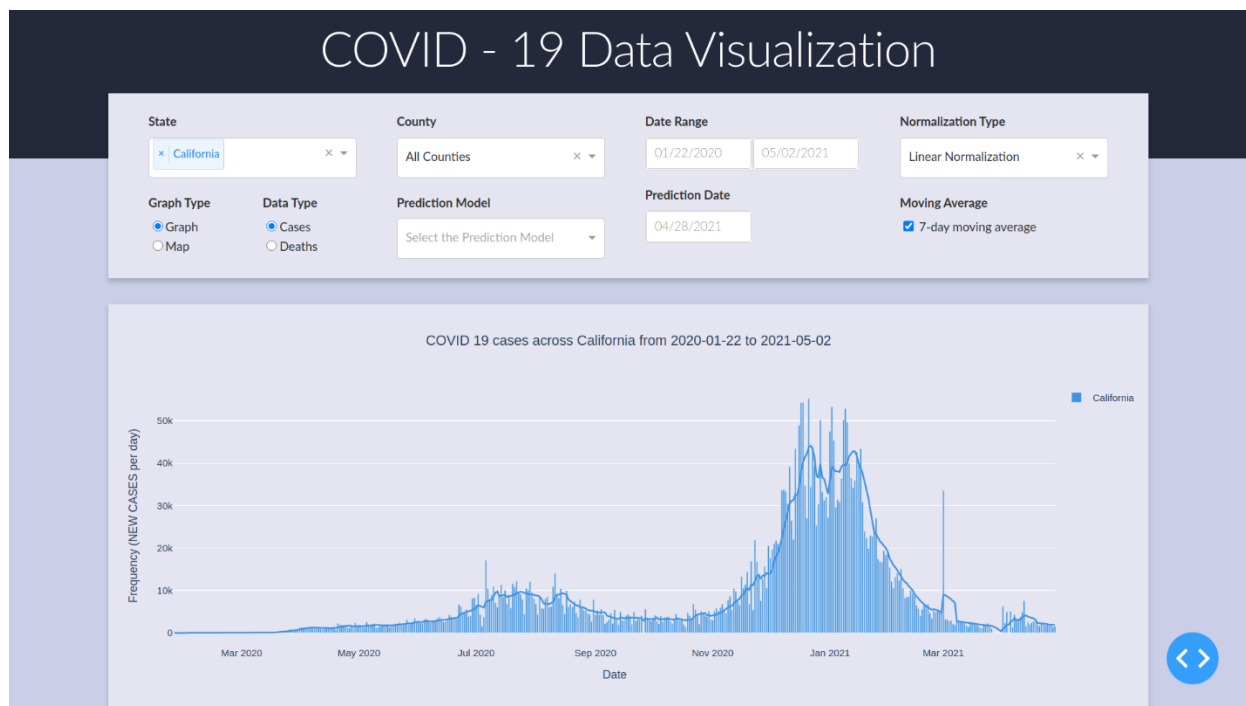


Figure 25: Bar plot with 7 day moving average for cases for California

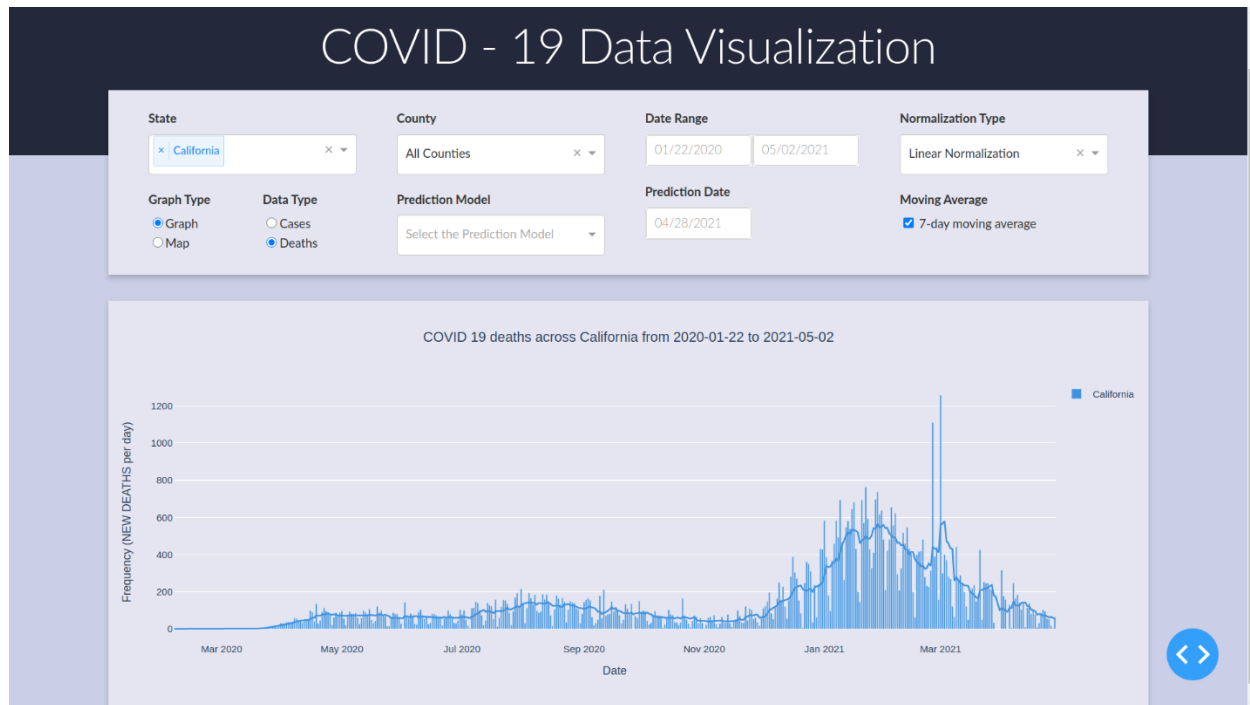


Figure 26: Bar plot with 7 day moving average for deaths for California

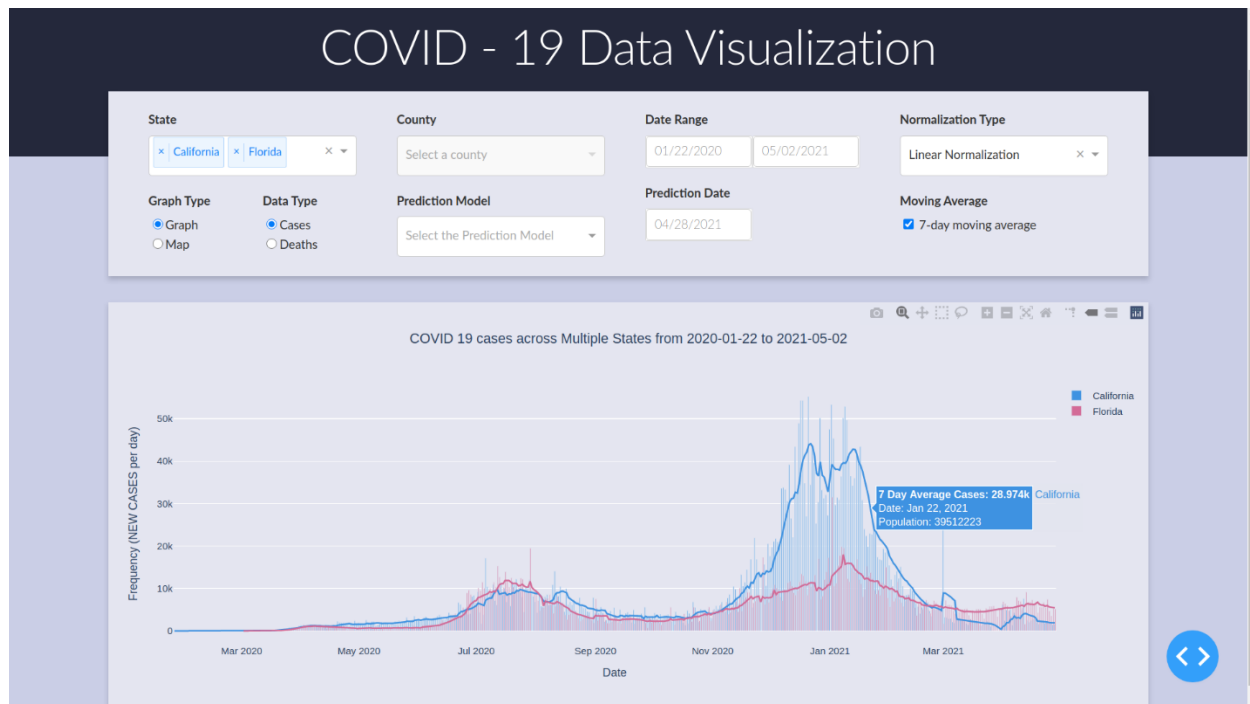


Figure 27: Bar plot with 7 day moving average for cases across multiple states

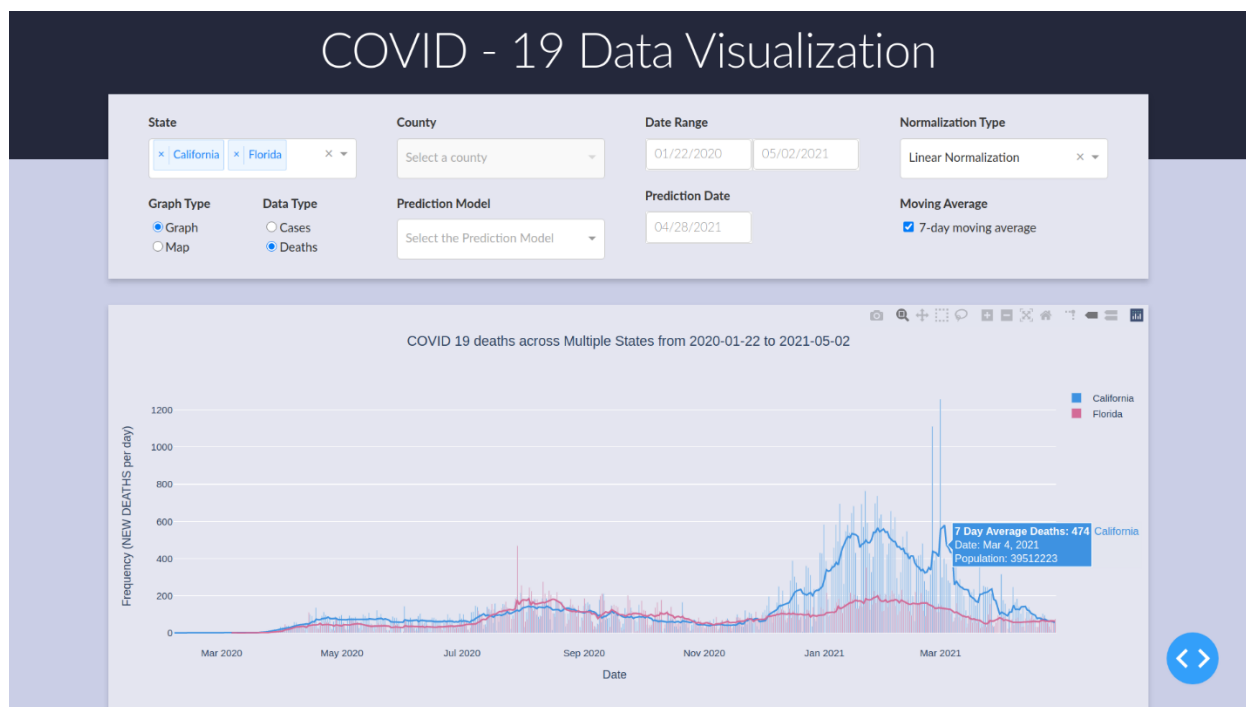


Figure 28: Bar plot with 7 day moving average for deaths across multiple states

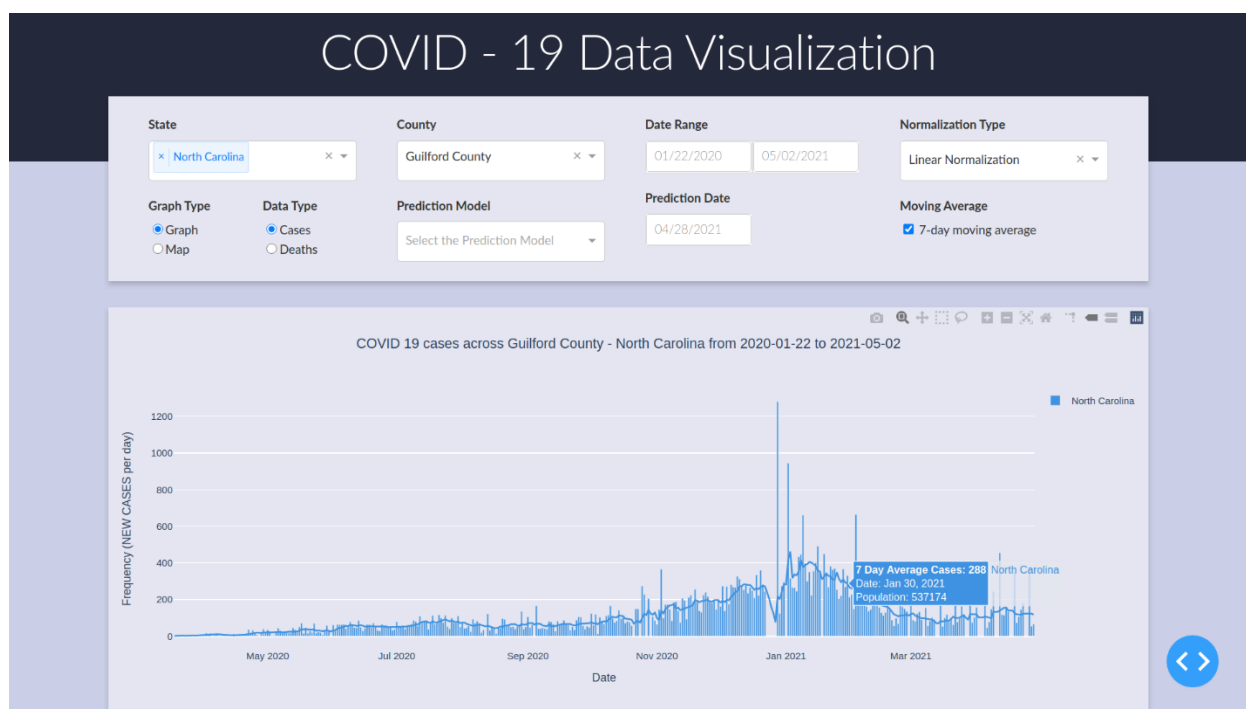


Figure 29: Bar plot with moving average for cases across Guilford county – NC

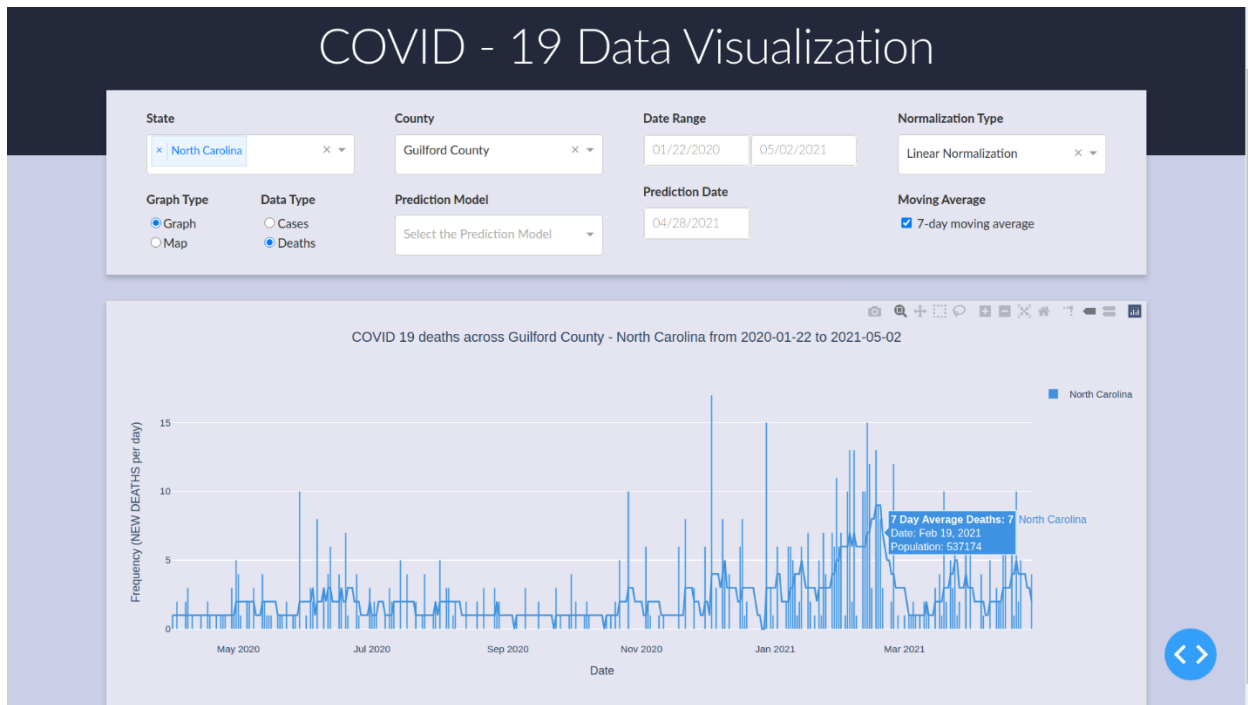


Figure 30: Bar plot with moving average for deaths across Guilford county – NC

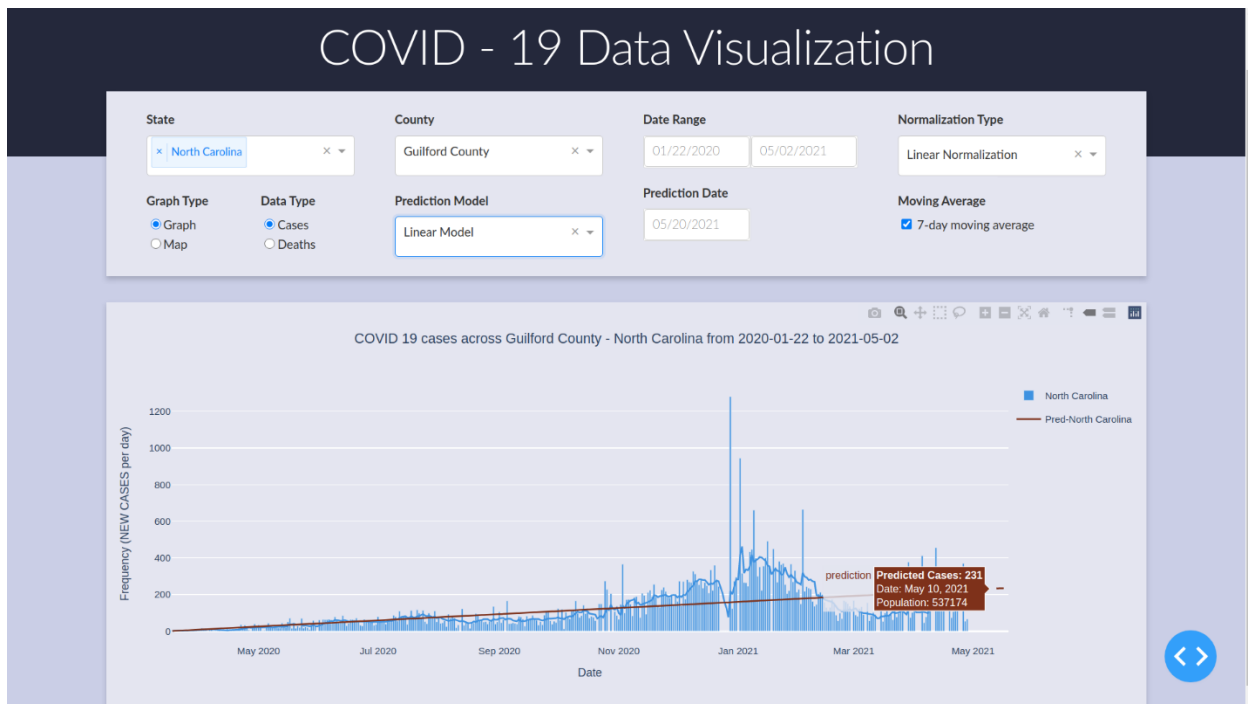


Figure 31: Prediction (till 05-20-2021) using linear model for Guilford County

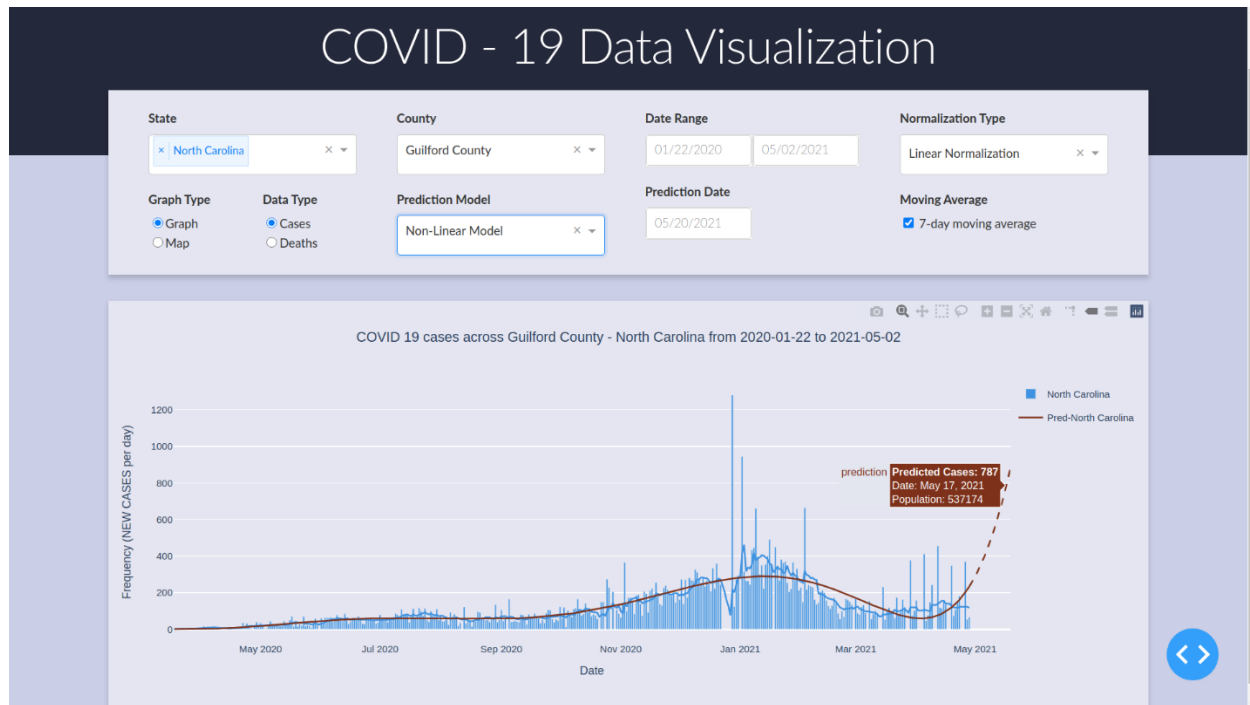


Figure 32: Prediction (till 05-20-2021) using non-linear model for Guilford County

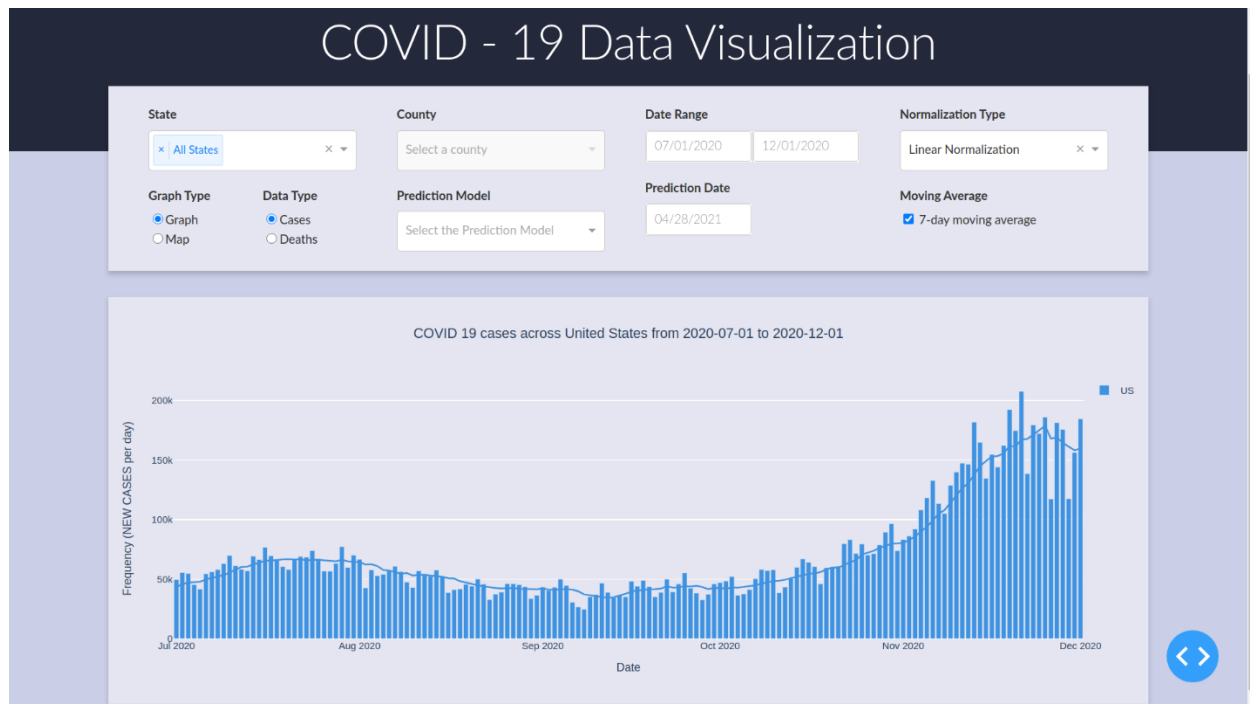


Figure 33: Data visualization between two user specified dates

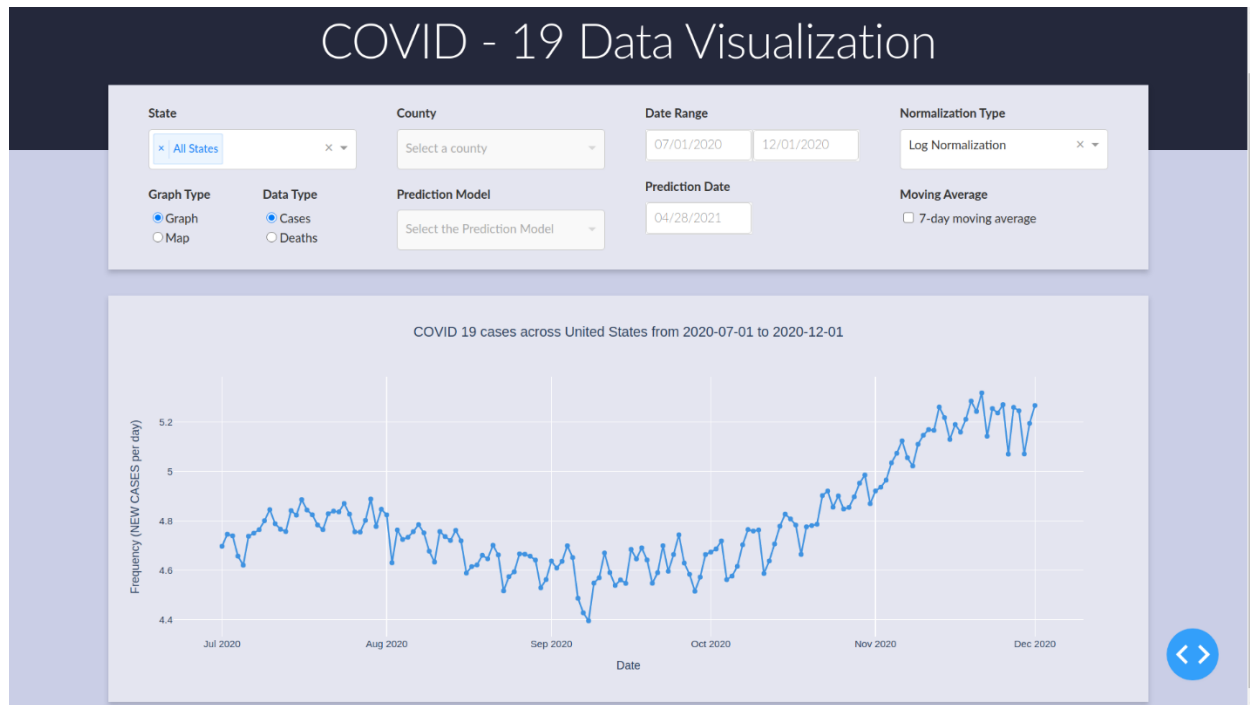


Figure 34: Log normalized data for cases across US between user specified dates

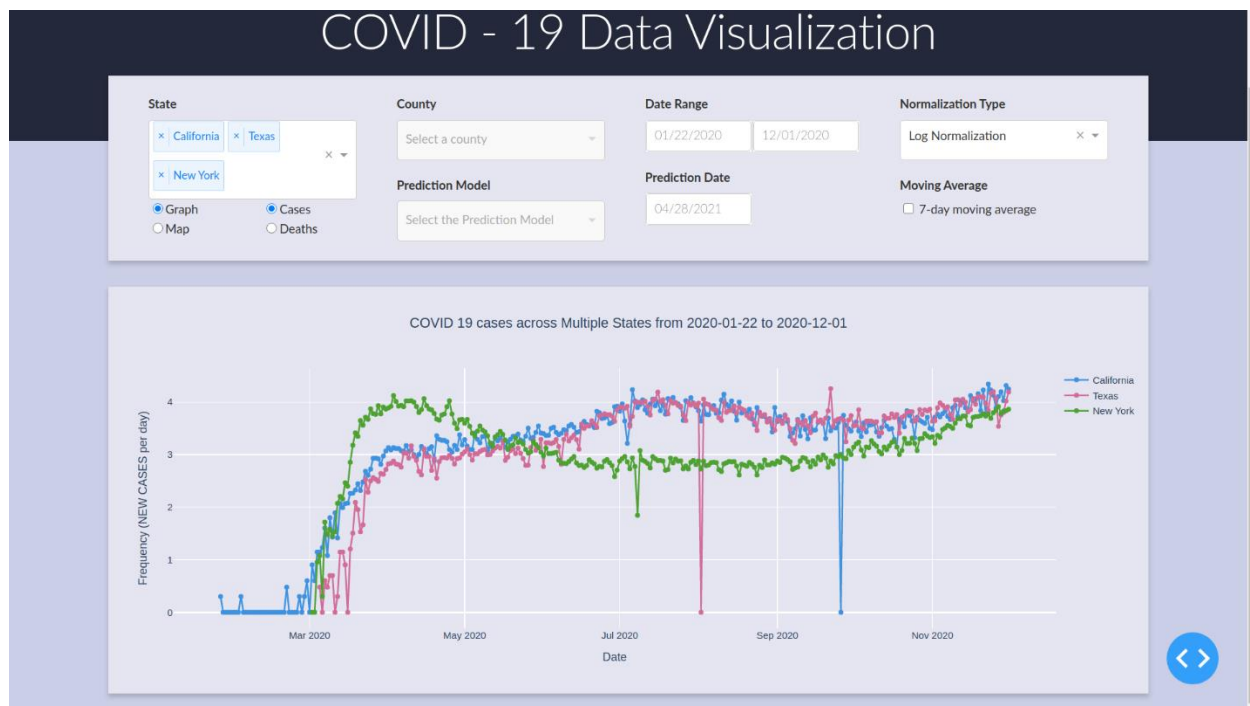


Figure 35: Log normalized data for multiple states

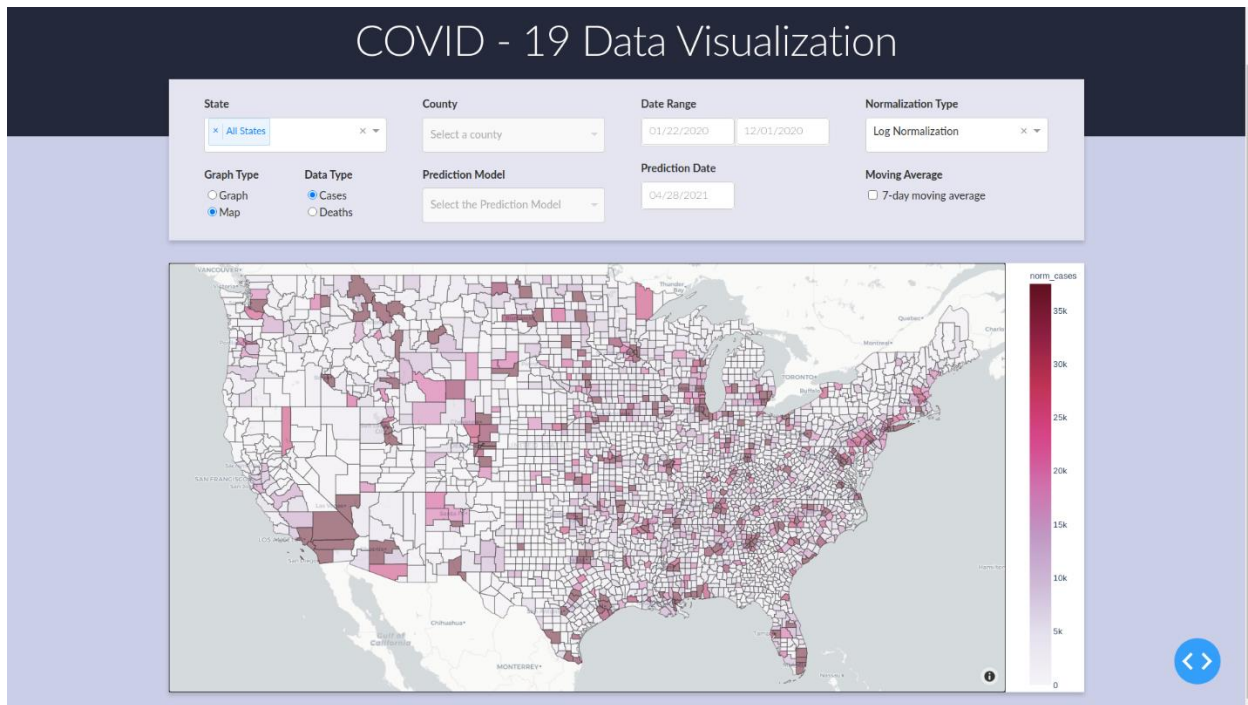


Figure 36: County Level map for united states

Here, the lighter states are the least suffered counties in terms of covid-19 cases and the darker states are the severely affected counties. The intensity from lighter to darker color represent the severity of infection in states from lower to higher severity.

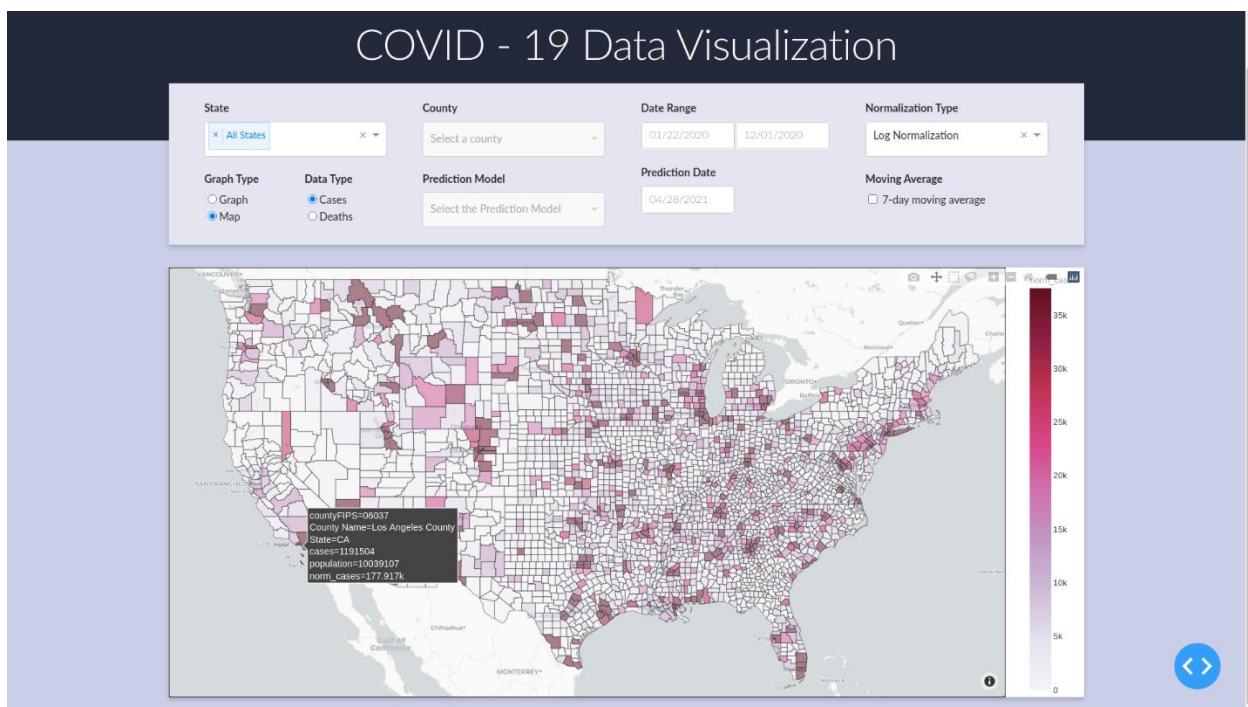


Figure 37: Hovering over a county present all information about that county

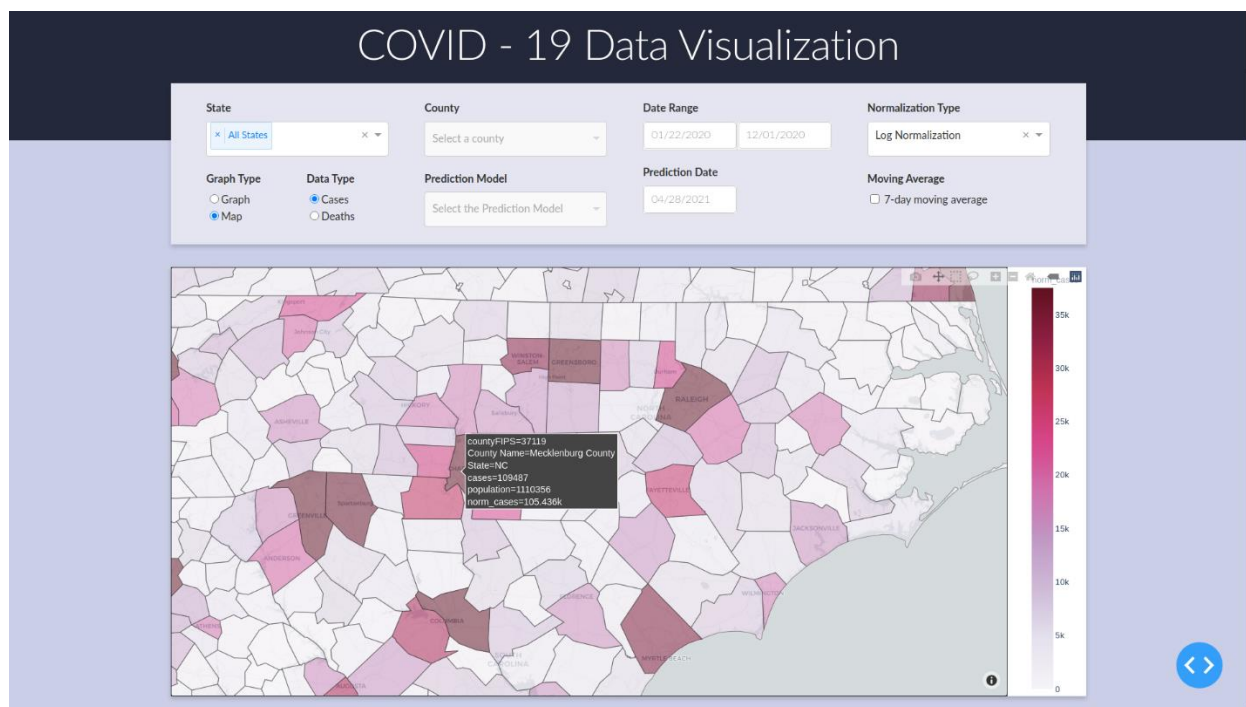


Figure 38: Severity in terms of cases in North Carolina

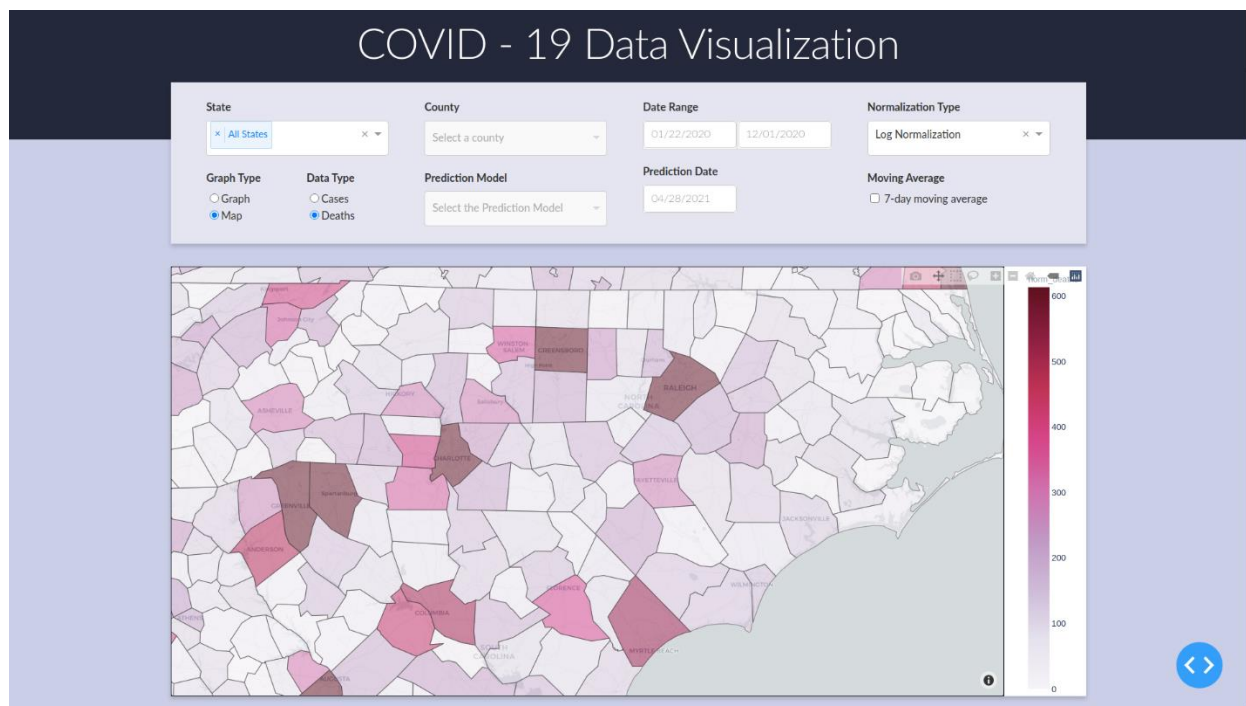


Figure 39: Severity in terms of deaths in North Carolina