The random forest model for case counts and death counts are similar in that they both use race and the democratic political party as the highest ranked features for MSE and node purity. The case count model accounts for almost 95% of variation while the death count model accounts for about 56% of variation. However, neither of the models handle spatial variation. This is questionable, particularly for the case count model because almost all variation is explained by the random forest. Further investigation is needed here. It is possible that the remaining 6% of variation contains the spatial component, but it seems unlikely.

The suspected reason why the political party is a highly influential feature is because of New York City's influence on the data. New York City was the first location Covid was reported in the United States, there is a very high population density, and the city is primarily democratic. It would have made sense to run this analysis without New York City. It is interesting to see that race (white or black) has a strong influence on the model too. It is very difficult to speculate why because both fraction of black population and fraction of white population are within the top three on both models in MSE or node purity. In both models, however, fraction black population has the stronger effect.