Rivenbark

BAN 530

COVID Analysis: The Effect of Lockdown Stringency vs. Population Health Measures on Deaths Rates due to COVID-19

Executive Summary:  
 This analysis found some linkage between lockdown stringency and decreasing death rates due to COVID-19 in the United States and Mexico. However, when modeling for future predictions of deaths rates, lockdown stringency was not found to be a significant predictor of rising or falling death rates. Population health measures like cardiovascular death rate and diabetes prevalence were found to be stronger indicators of death rates, suggesting that population health has a greater influence on death rates due to COVID-19 than behavioral measures like lockdown.

1. Introduction:

This analysis looked the relationship between measures of lockdown stringency and death rates due to COVID-19 vs the relationship of population health measures (diabetes prevalence, cardiovascular death rate, age, smoking rates) and death rates due to COVID-19 in the United States and Mexico. This analysis found little relationship between lockdown measures and rising or falling death rates. There was statistical significance in the relationship between lockdown measures and death rates, but it was not found to be a significant predictor when building models to estimate future death rates.

Measures of the general health of the population were found to be better predictors of death rate than lockdown stringency. Cardiovascular death rate, in particular, was found to be a significant predictor when building models to estimate future death rates.

Lockdown stringency increasing was found to likely be a reaction to higher death rates than a cause of death rates decreasing, as lockdown stringency increases followed rising death rates due to COVID-19.

1. Data:

The data used is from a global database of COVID-19 metrics tracked since the start of the COVID-19 pandemic in February 2020 to March 2021. The COVID-19 disease, caused be the SARS-CoV-2 virus, presents very differently in different individuals. Some parts of the world have experienced much more deadly COVID-19 outbreaks than others. In order to determine whether this is because of quarantine and lockdown efforts or the general health and agedness of a population, this analysis looks at the effects of age and measures of health across populations versus the stringency index (a measure of the severity of lockdowns and quarantine regulations) on deaths per million people over time.

This analysis uses at data from two countries: the United States and Mexico. These were chosen since Mexico represents a country with a high fatality rate and the US represents a country with a low fatality rate, since the goal of is to lower the fatality rate as much as possible. Fatality rates are measured in new and total deaths per million (of a total population). The severity of lockdowns is measured using the Stringency Index variable. The age and health measures used are percent of population over 65 years old, percent of population over 70 years old, cardiovascular-related death rate, diabetes prevalence index, and the percentage of the population that smokes. Smoking rates were recorded as male smoking rates and female smoking rates. These were combined into smoking rates across populations by averaging the smoking rates of both. These variables were chosen, as it has been observed that the COVID-19 disease generally presents with more severe symptoms and greater mortality rates in aged and co-morbid individuals. The date variable in the data was changed from the general to date format in Excel for the analysis to compute properly.

1. Descriptive Analytics

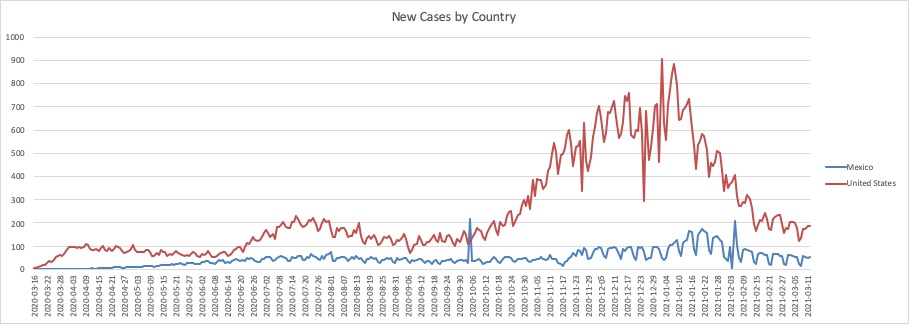


Figure 1: A time series graph showing new cases in the US and Mexico between 3/16/2020 and 3/11/2021.

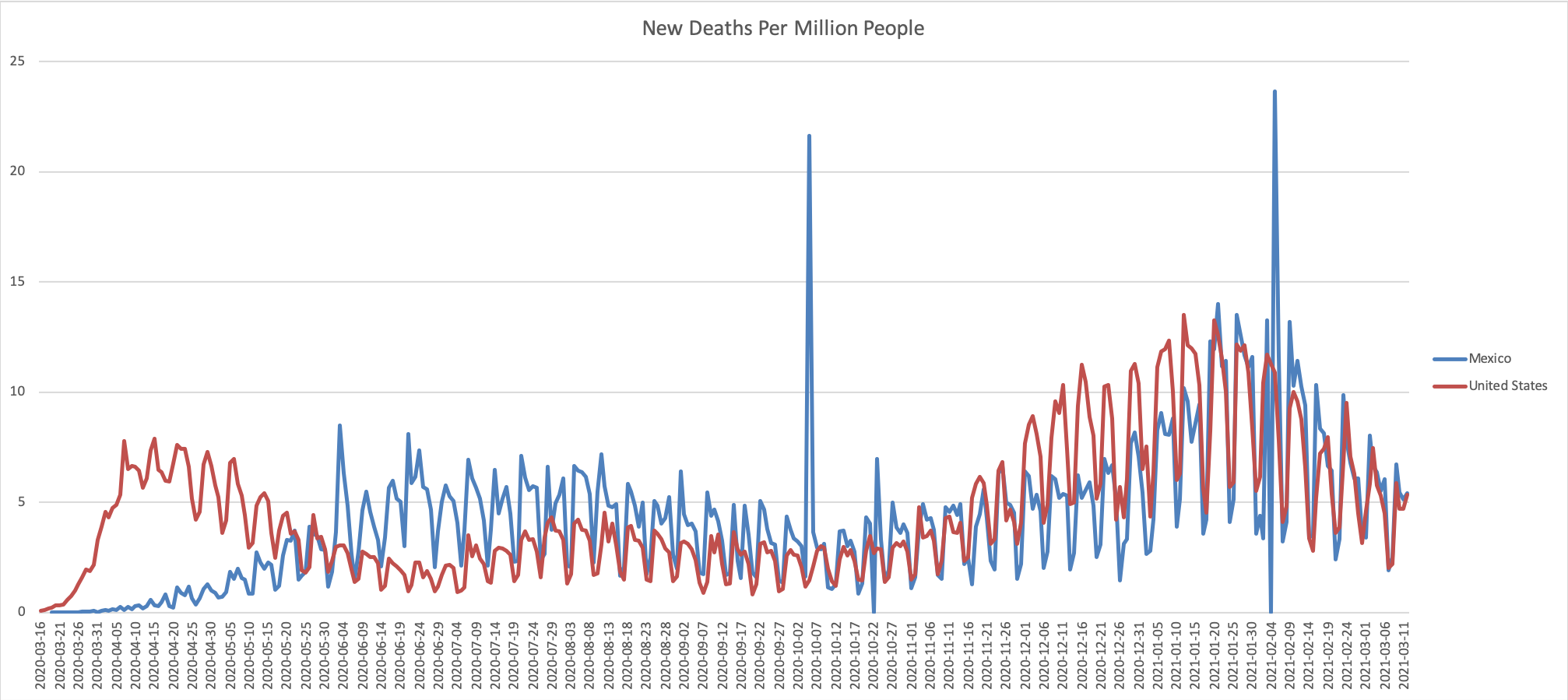


Figure 2: A time series graph showing new deaths per million in the US and Mexico between 3/16/2020 and 3/11/2021.

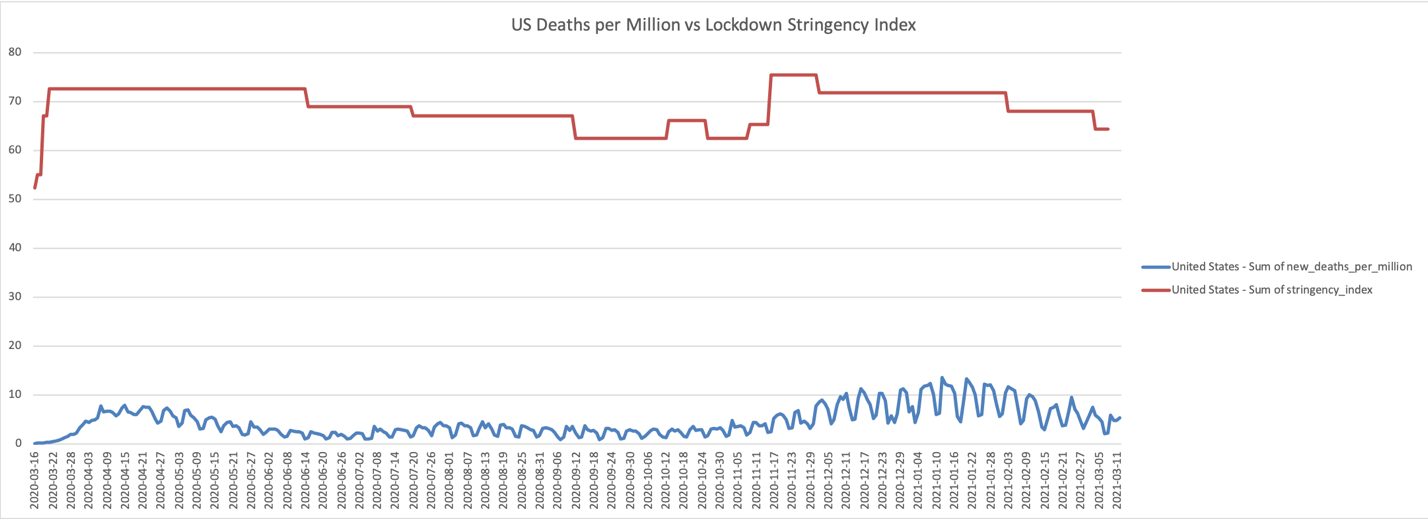


Figure 3: A time series graph showing new deaths per million vs lockdown stringency in the US between 3/16/2020 and 3/11/2021.

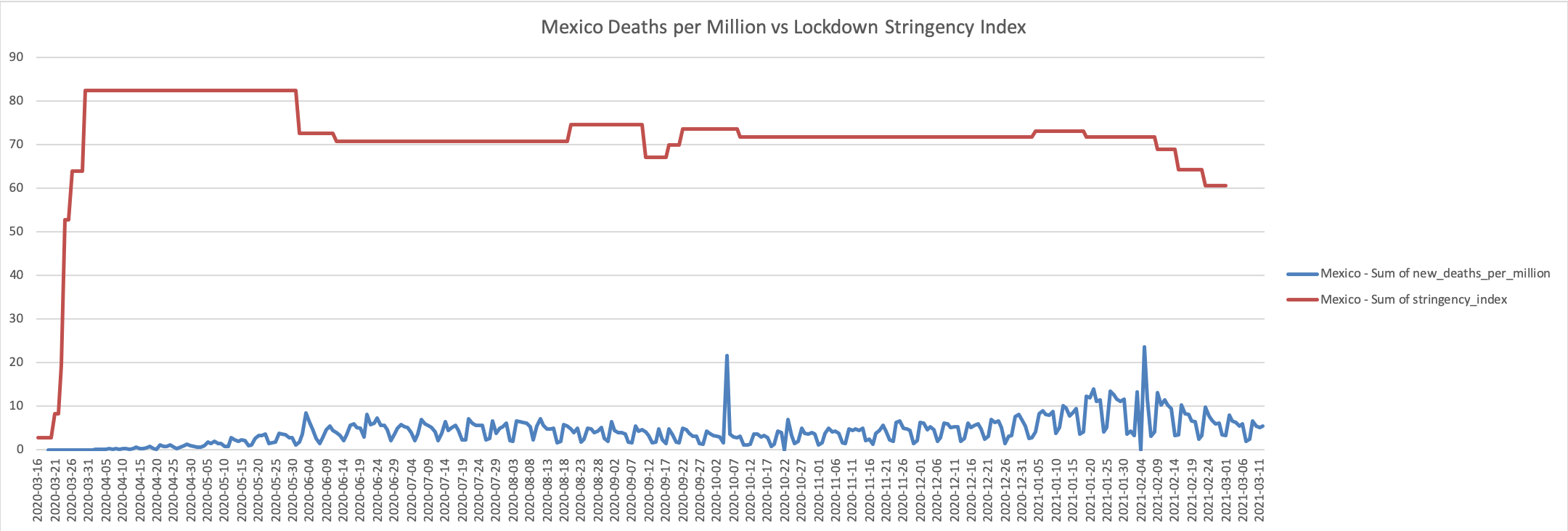


Figure 4: A time series graph showing new deaths per million vs lockdown stringency in Mexico between 3/16/2020 and 3/11/2021.

Looking at these two countries, the United States and Mexico we can see that Mexico was able to more effectively stop the spread of cases early in the Pandemic compared to the United States (Figure 1). Figure 2 shows that the United States also initially experienced a greater death rate due to COVID-19, but started to decrease steadily in the beginning of summer 2020, while Mexico’s death rate climbed until it levelled off around July and then stayed on a slight declined until about October. Around this time both case rates (Figure 1) and death rates (Figure 2) started to increase for both countries, as the regular flu virus season started.

Figure 3 shows deaths per million vs lockdown stringency in the United States. Around mid-summer, as lockdowns started to ease, cases also fell, but rose in the fall and winter, as lockdown stringency was raised back up. This is suggestive that lockdowns are a reaction to rising or falling cases, but not necessarily the cause of them. Figure 4 shows that Mexico had a very slow rise in death rate at the start of the Pandemic, as did the lockdown measures they put in place, shown by the slower rise in stringency index. This also suggests that lockdown stringency is a reaction to a rising or falling death rate, and not a cause of the rise or fall in the death rate.

1. Predictive Analytics

Growth factor

Growth factor is a measure of whether a variable is increasing or decreasing from the previous day and how quickly it is increasing or decreasing. The two main variables investigated were the growth factor of New Deaths per Million and New Cases per million in the United States and Mexico. The relationship between the Stringency Index and the Growth Factor of these variables over time is plotted below.

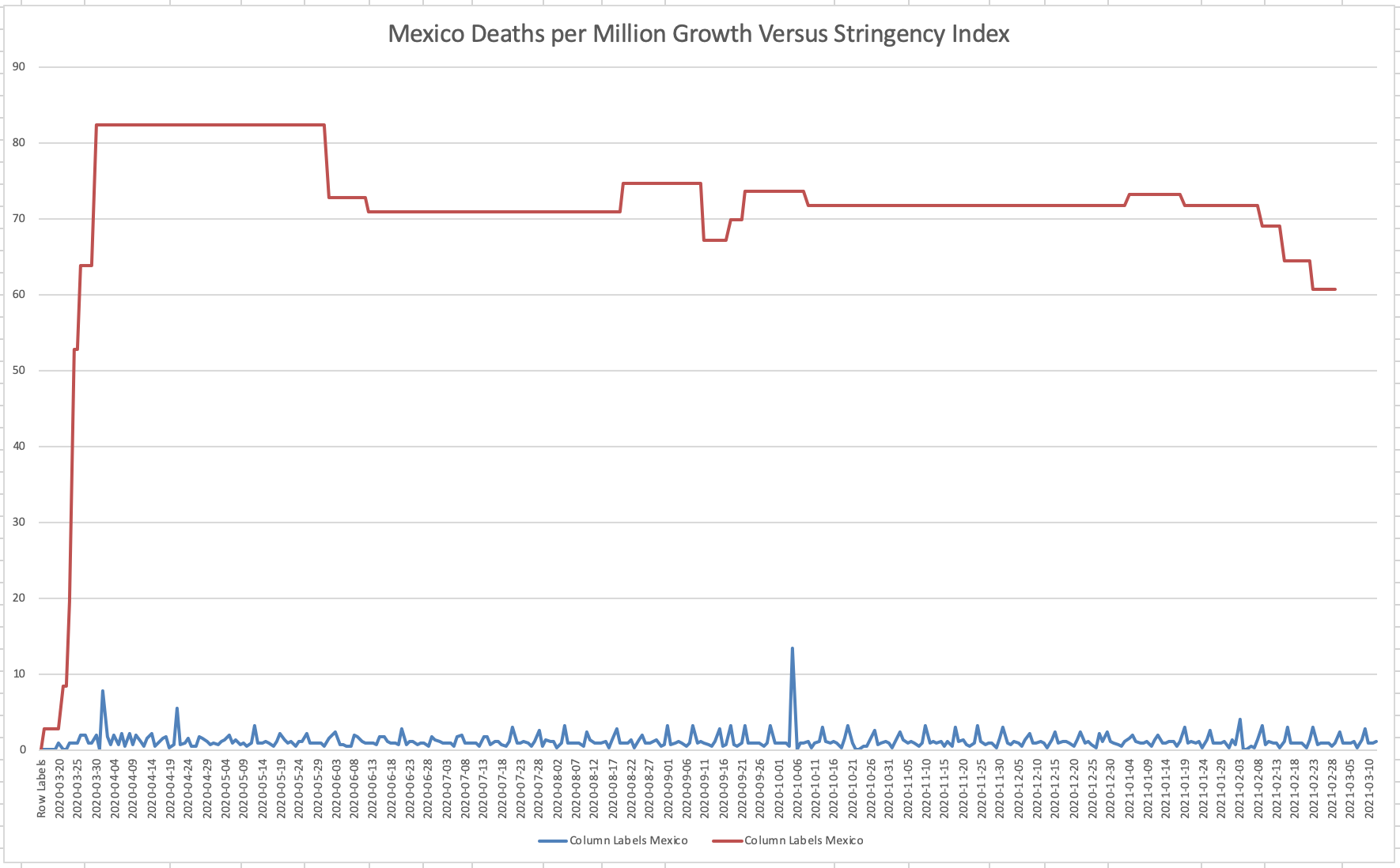


Figure 5: A time series graph showing new deaths per million growth rate vs lockdown stringency in Mexico between 3/16/2020 and 3/11/2021.

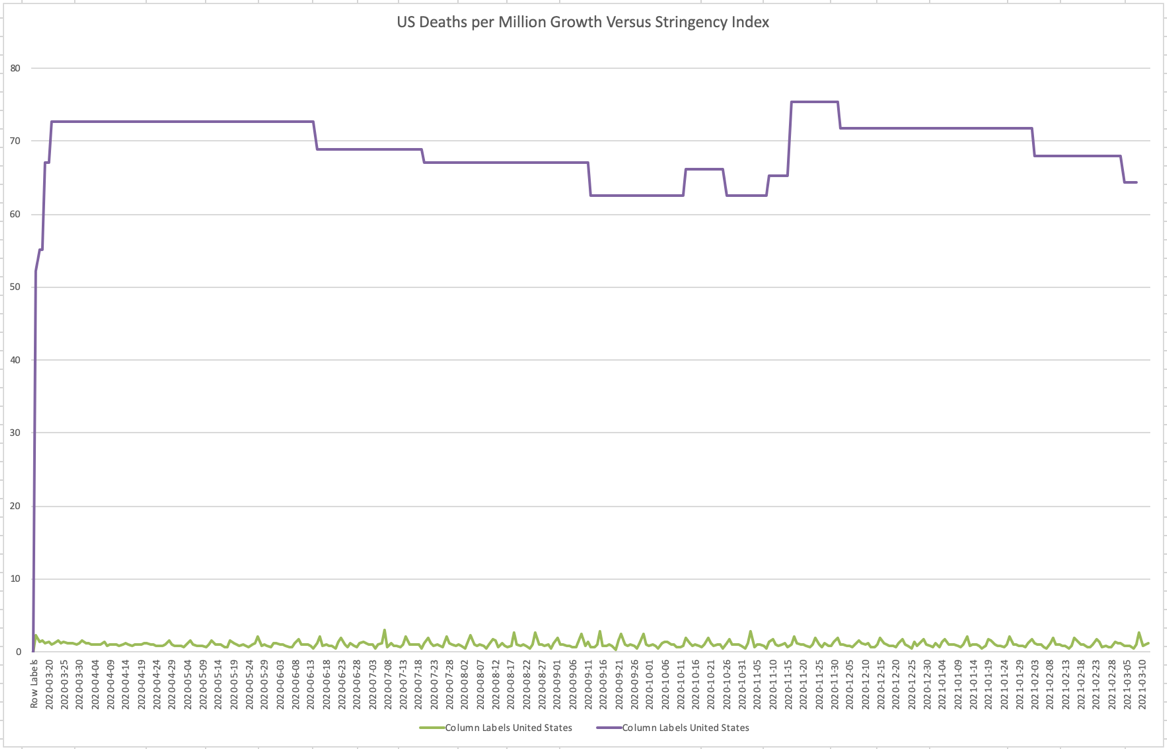


Figure 5: A time series graph showing new deaths per million growth rate vs lockdown stringency in the United States between 3/16/2020 and 3/11/2021.

These graphs imply that there is little relationship between the stringency of lockdowns and the growth rate of deaths per million due to COVID-19.

Clustering

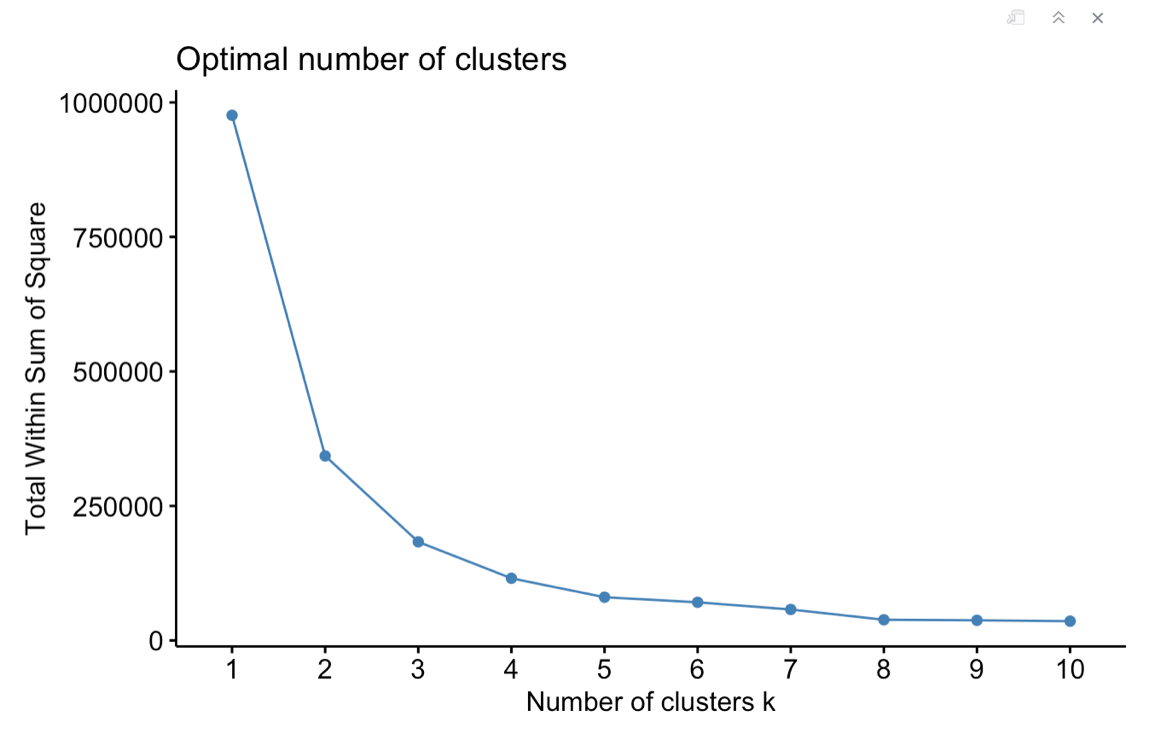


Figure 6: A graph identifying 3 clusters as the optimal number of clusters using the “elbow” method.

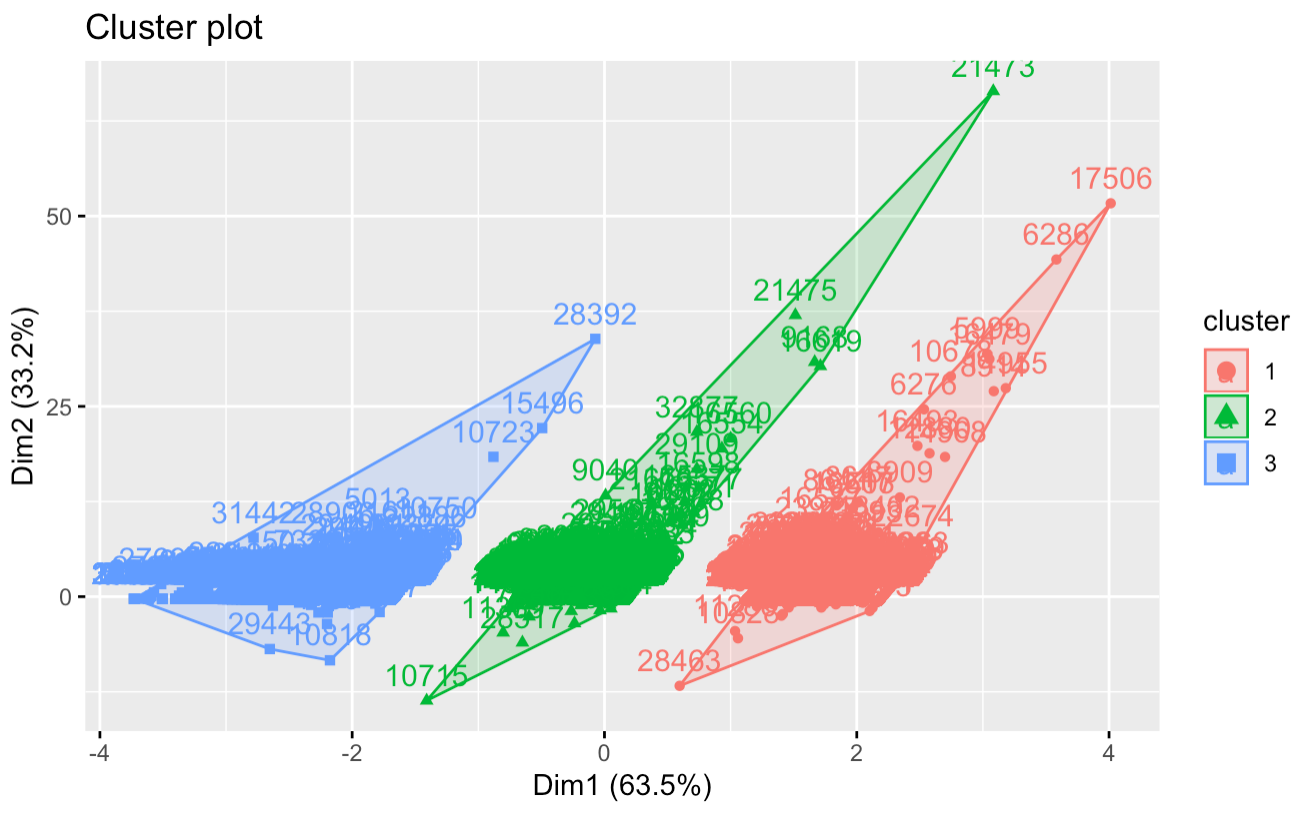


Figure 7: A cluster plot showing the relationship between deaths per million growth rate (x-axis) and lockdown stringency index (y-axis).

Figure 7 shows the relationship of observed deaths per million growth rate with lockdown stringency. This implies that lockdown stringency had little to do with increases of decreases in deaths per million.

The optimal number of clusters for this analysis was 3 (Figure 6). This implies that the different countries could be split into three groups based on low, medium and high death rates, but weren’t equally spread in severity.

Linear Regression

A simple linear regression model predicting New Deaths per million for the US and Mexico yielded a p-value of 9.66e-7, meaning that the relationship is very significant and linked. However, the intercept of 0.43 and slope of 0.05 suggest that there is a positive association between lockdown stringency and deaths per million. This is not an intuitive finding, suggesting that lockdown stringency is a reaction to increasing death (and case) rates, rather than death rates being influenced by lockdown stringency measures.

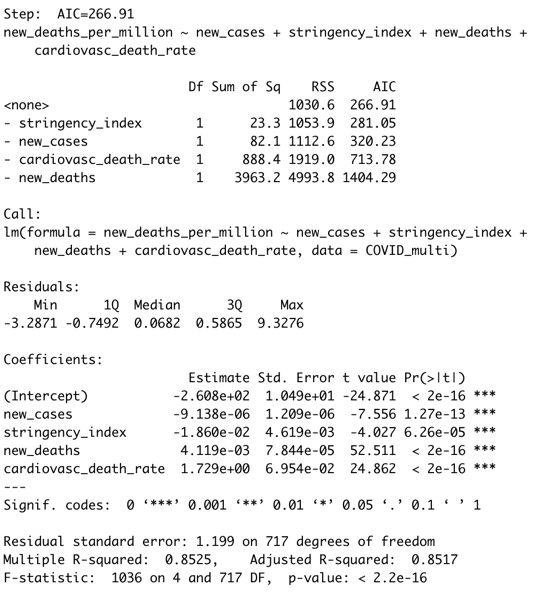


Figure 8: The results of the best multiple linear regression model predicting New Deaths per Million.

The best multiple linear regression model (Figure 8) yielded an AIC value of 266.91, a p-value of 2.2e-16 and an R-squared of 0.85 with relevant variables to predict New Deaths per Million being New Cases, Stringency Index, New Deaths and Cardiovascular Death Rate.

These variables all positively correlated with deaths per million, implying that as the general cardiovascular health of a population declines, death rates due to COVID will increase. It also implies that lockdown stringency is likely a reaction to rising or falling death rates, rather than death rates influencing lockdown stringency.

Machine Learning Algorithms

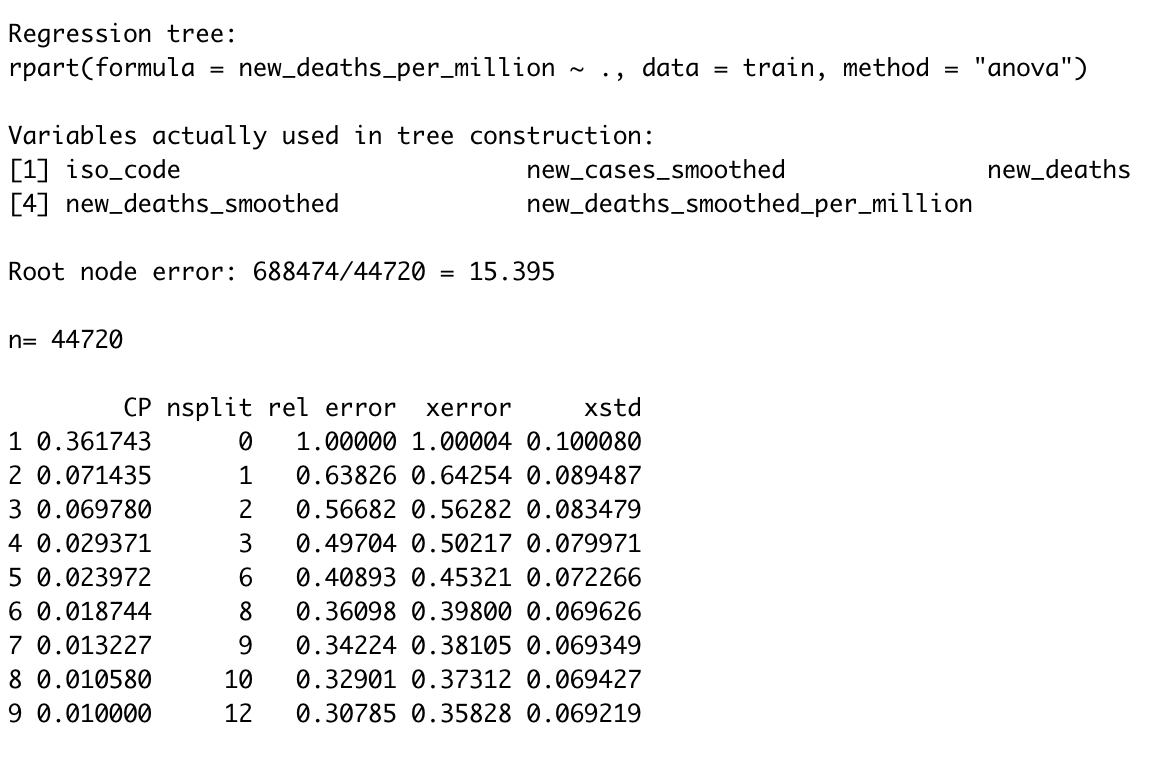


Figure 9: Results from a regression tree analysis predicting New Deaths per Million.

A classification tree algorithm (results in Figure 7) defined relevant variables to predict New Deaths per Million as Country Code, New Cases (smoothed), New Deaths, New Deaths (smoothed) and New Deaths per Million (smoothed). The test yielded an r-squared value of 0.69 on the training set and 0.64 on the test set, implying a fairly good explanation of the variation.

This implies that the most relevant variable to predict New Deaths is the Country, meaning that the environment plays a greater role in the death rate than human behavior or the general health of the population.

The multiple linear regression model is the chosen model, as it yielded the best r-squared value.

1. Prescriptive Analytics

Even if there are other contributing factors, lockdown stringency has played a role in lowering case rates of COVID-19 and part of lockdown measures are social distancing regulations. In order to help social distancing efforts during public events during COVID-19, a stadium has implemented social distancing policies to keep 2 empty seats between parties. Tickets can be purchased in sets of 1, 2 and 4. There is no discount for buying in bulk. Since the stadium can’t be filled, the manager wants to know how many sets of each ticket to sell in each seating zone to maximize profit and maximize capacity of the stadium.

In order to find the maximum capacity of the stadium under social distancing guidelines, this analysis set up a linear programming model that would take into account the social distancing requirements. Since there are no discounts for buying in bulk, this would also maximize profit.

Normally, there are 1500 seats available, but with social distancing of two chairs between each party, the maximum capacity of the stadium is 850. This would bring in $35,150 from ticket sales. Figure 10 shows the results of the capacity optimization and Figure 11 shows the results of the profit optimization. As shown, they both result in the same numbers and types of tickets being sold in each zone.

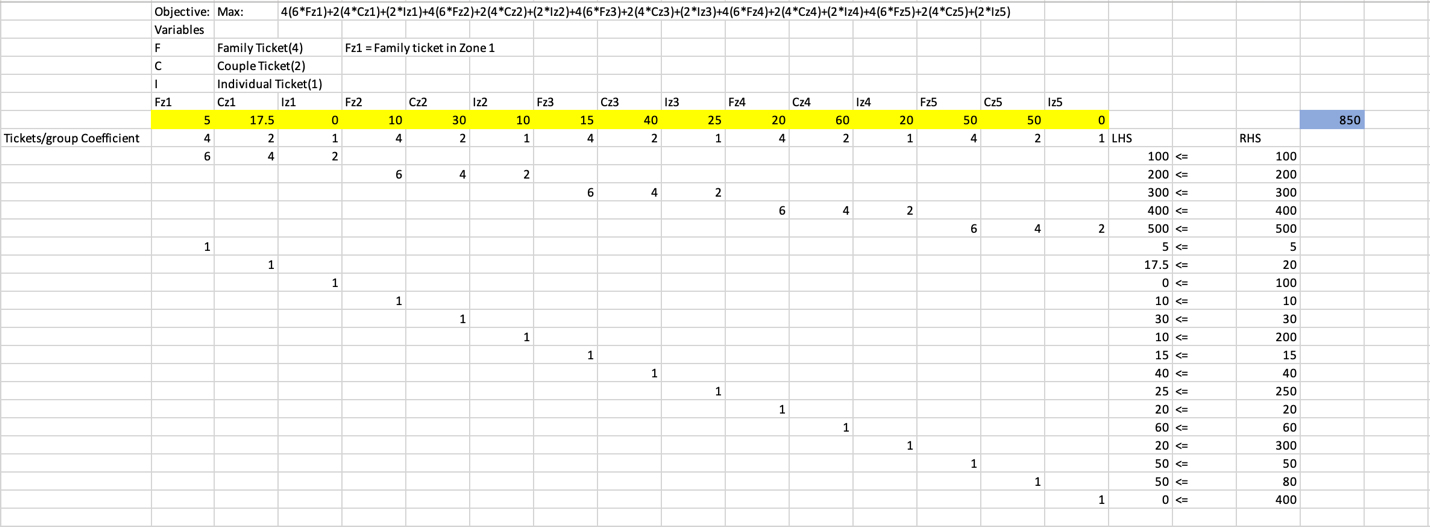


Figure 10: The linear programming model showing the maximum capacity of the stadium with social distancing regulations.

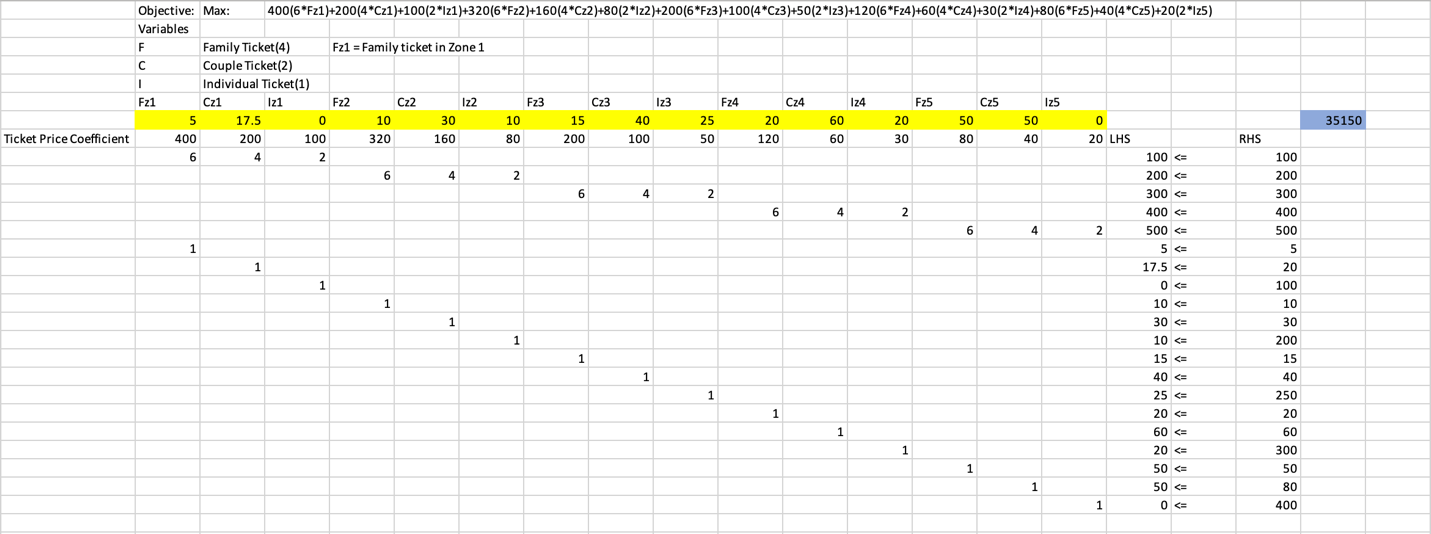


Figure 11: The linear programming model showing the maximum profit of the stadium with social distancing regulations.

1. Conclusion

The death rate due to COVID-19 is influenced by many factors, including weather, environment and the general health of a population. This analysis shows that more deaths due to COVID-19 can be attributed to population health measures than lockdown efforts in the United States and Mexico. Specifically, rates of diabetes and cardiovascular death were strong predictors of death rate. Interestingly, though COVID-19 attacks the respiratory system, smoking rates were not found to be strong predictors of death rate. The reasons for this could be explained in future analysis.

Another variable that was a strong predictor of death rates found in the predictive analytics section was Country. This implies that environment and weather plays a greater role in the spread of COVID-19 and the resulting death rates. It could also be the United States or Mexico had better handwashing, preventative measures than the other. This would be a good question for future analysis.

Lockdown stringency was not found to be a strong predictor of death rate. It is likely that a reverse causation relationship exists – that higher lockdown stringency is a reaction to increasing death rates, not a cause of increasing or decreasing death rates.

This analysis found that measures of the general health of the population (cardiovascular death rate, in particular) is a better predictor of death rate than lockdown stringency. The best course of action would be for healthy people to return to a more normal way of living and the more vulnerable to socially distance and avoid infection.