**Abstract:**

This study aims to evaluate the impact of virus testing on new positive cases, deaths and case fatality rate (CFR) in the European countries. There are variations in how the testing has been done across the countries. Conclusion Massive population testing could have a significant effect on mortality in different ways. This analysis may help decision-makers to administer healthcare measures to limit the spread of the disease.

**## Introduction**

The COVID-19 pandemic is considered as the most crucial global health calamity of the century and the greatest challenge that the humankind faced since the 2nd World War. In December 2019, a new infectious respiratory disease emerged in Wuhan, Hubei province, China. It was named by the World Health Organization as COVID-19 (coronavirus disease 2019). A new class of virus, known as SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) has been found to be responsible for occurrence of this disease. In the history of mankind there are instances of severe outbreaks of diseases caused by several viruses. According to the report of the WHO, the current outbreak of COVID-19, has affected over 40 million people and killed more than 1.1 million people throughout the world. It has rapidly spread around the world, posing enormous health, economic, environmental and social challenges to the entire human population. The coronavirus severely disrupted the global economy causing lot of people to lose their jobs or loss in business. Most of the organizations have stopped the investments for a year because of COVID. On the other hand, digital platforms are getting more appreciation for the entertainment, education and advertising purpose. Almost all the nations are struggling to slow down the transmission of the disease by testing & treating patients, quarantine suspected people through contact tracing, restricting large gatherings, maintaining complete or partial lock down etc. Now there are clinically approved antiviral drugs or vaccines that are effective against COVID-19 which will out in market in a month.

With this study, we can relate the effect of testing on the new cases and deaths happening in the countries due to COVID-19. We expect that testing rate plays major role in examining the spread of corona virus and case fatality rate. The relation between different variables can be given by plotting various graphs for the confirmed cases, confirmed deaths and covid testing in the countries. Due importance is given to the geographical aspects as we know that there are few countries where COVID is spread badly and there are few countries where its almost over and they have started to open the markets for regular use. Also, the population of the country plays a vital role in determining the growth of covid due to the density of the population.

Here, the plan is to use the daily cases data along with the weekly testing data concentrated on Europe region. I have considered this data as we do not have daily testing data available and, we do not have the data for other regions in the csv files separated. Here, we assume that the external factors such that preventive measures taken by government for controlling the pandemic remains constant as we do not have the data for the same. Also, we cannot analyze the level of impact these measures have on the covid cases. It may happen that because of lockdown people roam very less on roads, typically for the daily needs. Hence, the spread of corona is avoided. Also, it may be possible that the covid patients are taking care of themselves at home and not going to hospitals for the testing. Also, we assume that the covid should with respect to time and after some time the trend should reverse. We are aiming to examine these assumptions in our analysis to show the linearity between the variables. As we are considering the case fatality rate weekly and daily, it is formulated by the weekly and daily deaths to the cases.

The report is divided into Data Pre-processing, Exploratory data analysis, hypothesis testing, regression modeling and residual plots. In the first section, the data is imported form the csv files provided, cleaned for the missing or false values and new columns are created based on the factual data in the same dataset. Here, we have transformed the daily data to weekly to get the final dataset for the analysis and modeling. The section exploratory data analysis gives us the behavior of each variable and their plots to show the relation between different variables. Also, it gives us the notion of how important predictor is for the prediction of response variable. Then we have tested for hypothesis we have assumed before modeling the data which shows us that whether our data is significant to process. If we can reject the null hypothesis then we can check which variable fits in the model using the various regression techniques. It is also important to check the variance inflation factor to statistically calculate the multicollinearity in the variables with the use of statmodels. We can plot the residuals to check whether the value we have fitted are significantly near to the actual values using difference residual plots.

## Conclusion

The above analysis outlines that mass testing increases the number of positive cases which gives the better understanding of how the virus is spreading across the country. Also, it will decrease the case fatality rate because the patients will get treated at the early stage of the virus. It will also reduce the spread of virus as the positive patients will be kept quarantined and under observation so that they cannot encounter others. It will also help the government to imply preventive measures such as lockdown or restrictive moments if the covid spread is too high. population of the country also plays major role in depicting the behaviour of new cases. It is hard to keep the testing rate high in dense countries than the one with low population. Increase in the number of cases creates the race for the hospitality, hence controlled testing is also necessary to not create chaos between the citizens. Case fatality rate and new cases varies a lot with the Countries as there are some countries with corona being highly spread and there are few where it is about to vanish. This makes the country factor most important in our model. We can illustrate that the countries which belong to group 4 that includes Spain and United Kingdom have the highest increase in the new cases followed by group 3 which includes Italy, Germany and France.

We have plotted various scenarios which tells us that the Case fatality rate decreases exponentially with the testing rate i.e. number of tests per lakh population. Number of tests also helps the researchers to study the virus and bring new methods of treatment which reduces the risk of like and hence reducing the case fatality rate. The main effect of testing was seen when mass testing has started. New cases have started to increase suddenly and same can be seen in the correlation plot. We have also learnt from the data that few countries in the Europe such as United Kingdom, Italy, Germany and Spain are highly affected by the corona virus. Hence, we can have the high spikes in the data along with some outliers which makes our model create large residuals though we have reduced them by grouping the countries into categories. Also, we can say that the number of deaths declines with the increase the number of cases.

There are numerous other factors such as demographics data of the country. It may be possible that one country has old people more than young. We know that old population have more probability to catch the corona virus. Corona virus spreads if social distancing is not maintained, so it would be tough in the highly dense countries to stop the spread of corona virus even if the restrictions are applied. Further, the healthcare facilities also play an important role in controlling the corona virus. As the cases increases, the number of beds required increases which may lead to lower the testing rate. Another study has shown that the environment conditions also has a key effect on the growth of coronavirus as they say higher temperatures lowers the effect of virus. In India, there is huge number of covid patients, but the CFR is very less, and this may be because of the temperature of the country. While, in countries like United Kingdom, Italy and France, though the population density is low, but temperature is also low causing more deaths due to the virus. Several studies predicted that by the end of the next year 80% of people will create antibodies for the corona virus themselves. Mass testing also helped researchers to get the vaccine quickly to stop the virus. The United Kingdom have approved the vaccines which will be available in the market in few weeks. Overall, predicting the effect of corona virus can be critical as there are several practical factors and government decisions included. But if these subjects are kept constant then the data along with the demographic information may prove helpful for prediction with the help of times series analysis.