**COVID-19 ANALYSIS PROJECT REPORT**

Disha Shah

Shraddha Upadhyay

Nikhit Soares

Shalin Shanghavi

Mukul Pathak

**COVID-19: The Pandemic**

It is to no surprise that the Covid-19 virus has got the entire world on edge. As of 6th May,’20, 3,751,069 people are affected by it worldwide and the present situation says, there is more to come. In the US alone, 1,228,177 people are affected and 73,207 people are dead because of this virus. We, as a group of Business Analytics students, are interested in getting to know the situation better by analysing it through available data and have cumulated a small case study based on data collected via different data sources.

**DATA COLLECTION**

Data was obtained from various organisations like John Hopkins University repository, World Health Organisation, Centre of Disease Control and Prevention and Microsoft COVID-19 Tracker. Data obtained was from 22nd January and updated till 3rd May. Data consisted of the number of death cases, confirmed cases and recovered cases along with the mortality rate of the whole world. It also consisted of the cumulative tests that were conducted during this pandemic.

Data cleaning wasn’t a big problem during the analysis and so it gave us a big boost to start analysing and research on the pandemic.

**DATA ANALYSIS**

Due to extensive dataset, we have performed all visualizations using Excel and Python. Following are some of the analysis we performed for our project,

**Total Cases**

We have visualized the countries having the most number of confirmed, death and recovered cases from each continent

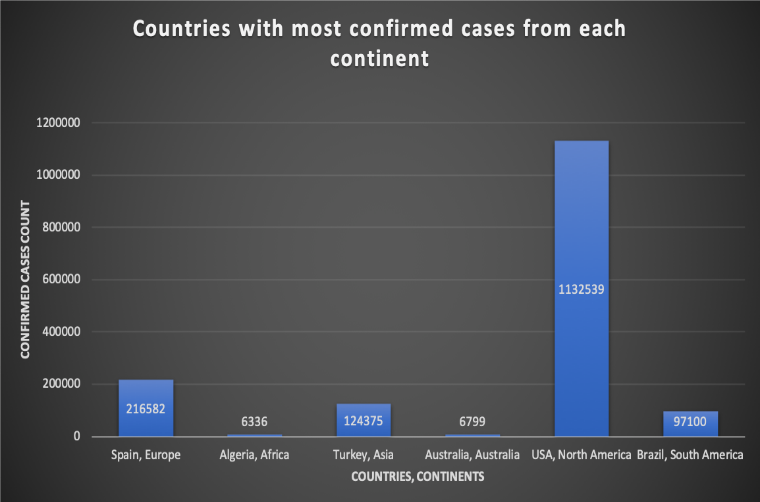
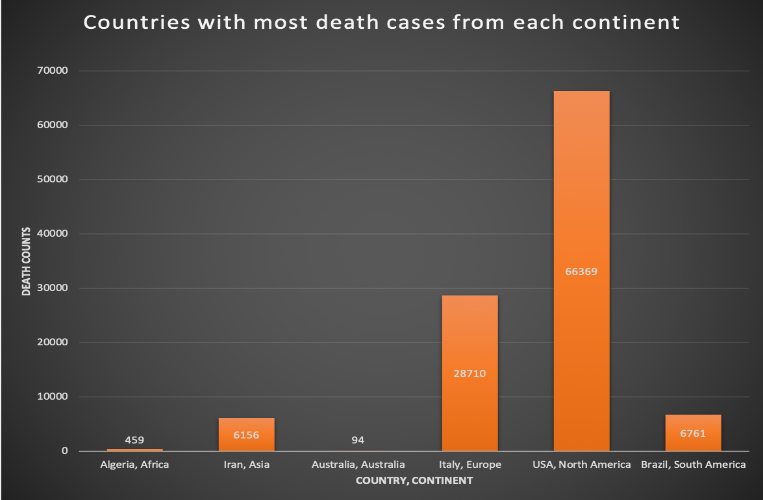
 

Fig 1. Total #Confirmed Cases Fig 2. Total #Death Cases

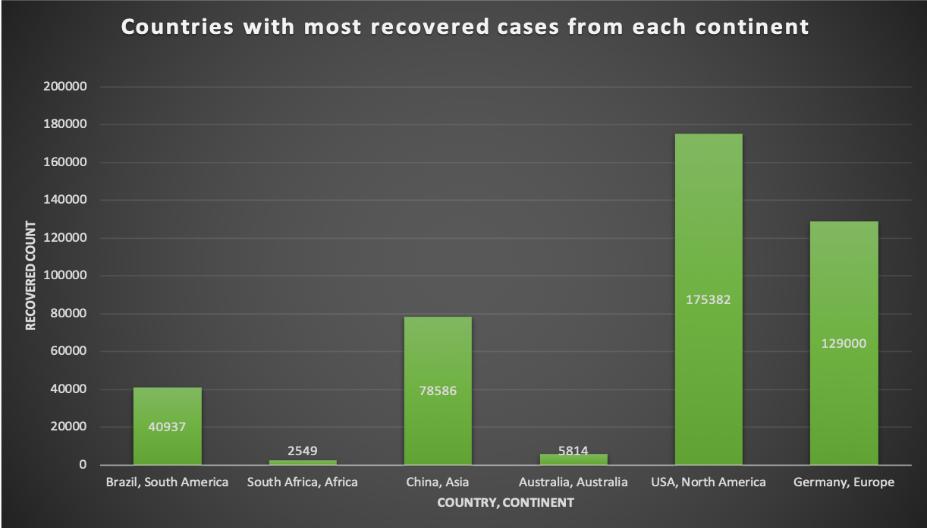


Fig 3. Total #Recovered Cases

Here we can see that North America has the most number of confirmed, death and recovered cases based on the current trend. Europe comes in line after North America for the greatest number of cases in every sector. Also, for the greatest number of these cases the countries in every continent are different for every sector.

**Mortality Rate**

Mortality rate measures the number of deaths due to the viral spread of Covid-19 over the whole world. The below graph depicts the rise of world mortality rate through time.

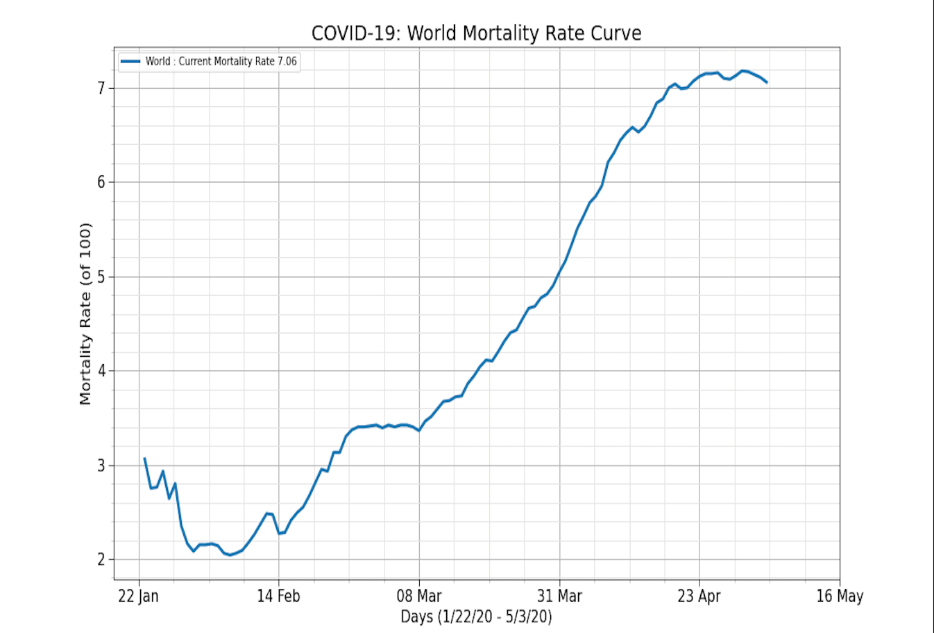


Fig 4. Mortality rate curve of the world

As we notice from the curve, the world mortality rate stays as low as 3% till early March. But it climbs uphill rapidly to reach 7% by early May. Recently, there are signs of flattening the curve. The below graph compares the mortality rate curves for different continents.

A close up of a map

Description automatically generated

Fig 5. Mortality rate by Continent

Surprisingly, despite being the inception continent of the disease, Asia managed to keep the mortality rate to a peak of 4.3%. As seen, Europe experiences the highest mortality rate of approximately 10%. The table below shows the summary statistics of different continents across the globe.

**A screenshot of a cell phone

Description automatically generated**

Table 1. Continent-wise summary

**Cases Tested across the world**

Here we analyse top countries from each continent having the most number of testing done for COVID-19 based on every sector.

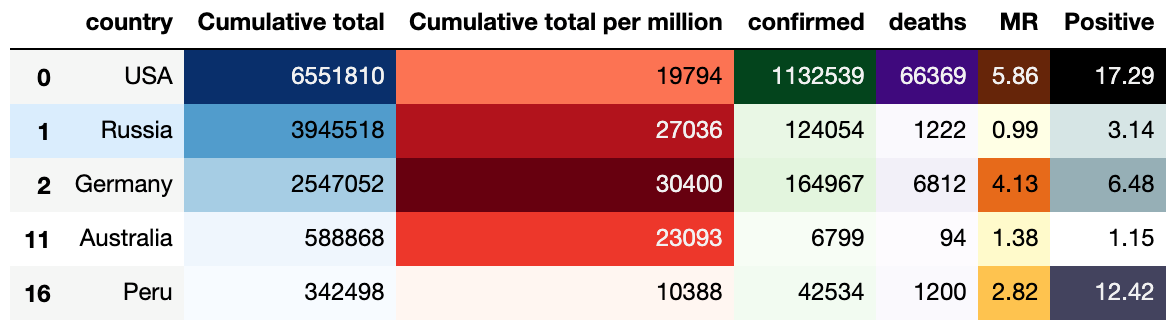


Table 2. Statistics of cases tested for top countries around the world

We can see that a cumulative total of 6551810 people in USA were tested for the virus out of which 1132539 were confirmed cases and a total of 66369 deaths were confirmed based on the current trend. We can also see that out of 100 people in USA at least 17.29% people are tested positive with a mortality rate of 5.86% which is the highest for any country.

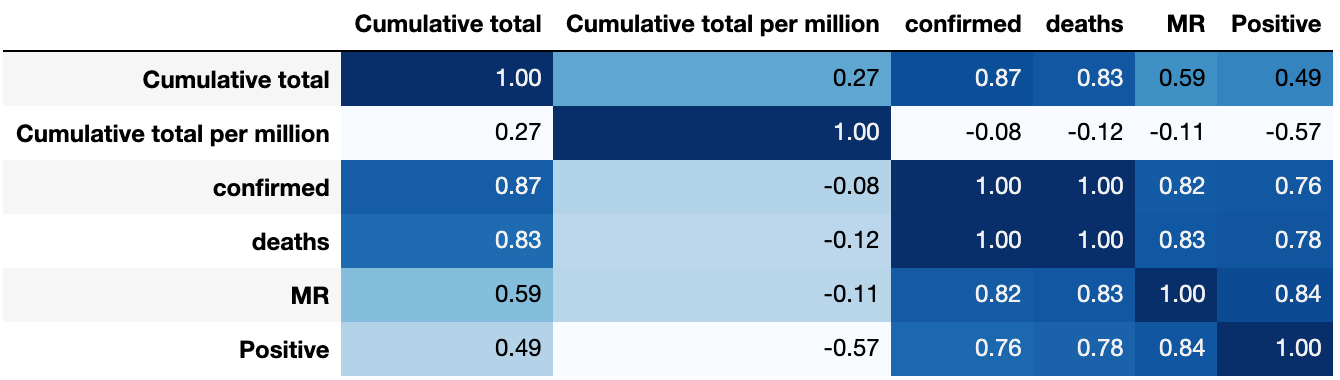


Table 3. Testing stats Correlation matrix

The above table is the correlation matrix which clearly shows that the tests conducted is highly correlated with deaths and confirmed case. It feels harsh but more tests should be conducted as more recoveries also take place due to more tests conducted. Strangely, cumulative total per million is negatively correlated with most of the variables.

**SUMMARY**

The below table shows the overall analysis for top Covid-19 affected countries of the world.

**A screenshot of a cell phone

Description automatically generated**

Table 4. Overall analysis for COVID-19

It is difficult to predict the accurate spread of the pandemic because of the large number of predictors playing their part as well as human intervention playing its role as well. We have tried to predict the number of deaths in the following 15 days according to the current trend.

A close up of a piece of paper

Description automatically generated

Fig 6. Future death cases prediction for next 15 days

**APPENDIX**

WHO: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>

CDC: <https://www.cdc.gov/coronavirus/2019-nCoV/index.html>

Microsoft COVID-19 Tracker: <https://www.bing.com/covid>

COVID-19 Tracker by Johns Hopkins University: <https://www.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6>