Capstone Project - The Battle of Neighbourhoods against COVID19

Applied Data Science Capstone

OBJECTIVES

- The ongoing Covid19 Pandemic caused a serious distress across globe. The statistics and data related to COVID19 is humongous and is being evolved day by day.
- The recovery from COVID19 pandemic is largely dependent on fighting it as one neighbourhood in quickly vaccinating all populations across the globe.

Business Problem

- The objective of Capstone Project is to evaluate the impact of COVID19 on all countries and find how similar or dissimilar are the countries in terms of COVID19 impact.

Using Data Science, this project aims

1. To provide vaccination clusters based on clustering countries as per COVID19 impact for effective distribution
2. Based on Individual COVID19 response, find cluster of countries which require enhanced attention and resources

Data acquisition and cleaning

Data sources

- Coronavirus Country profiles are available in Coronavirus Pandemic (COVID-19) - Statistics and Research - Our World in Data.

- JSON data is further reduced to two separate Dataframes as below

1) **df_country_data** which consists each country's demographic and health profile
   - ['location', 'population', 'population_density','median_age', 'aged_65_older', 'aged_70_older','cardiovasc_death_rate', 'diabetes_prevalence','life_expectancy','hospital_beds_per_thousand','human_development_index','extreme_poverty']
1) **df_covid** which consists each country's covid data till data publication

Exploratory Data Analysis

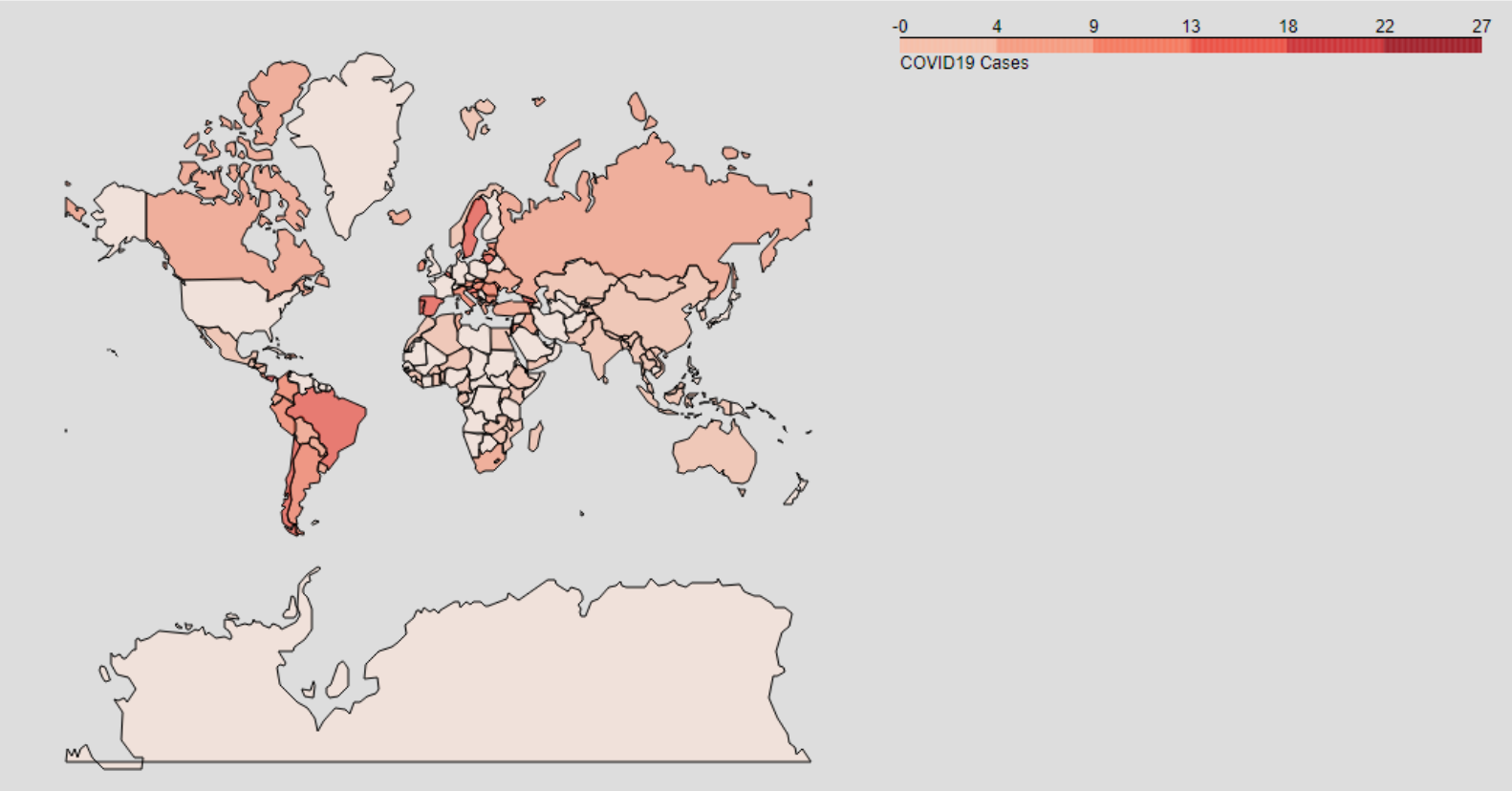

Figure 1 World map to show covid cases relative to population

Exploratory Data Analysis

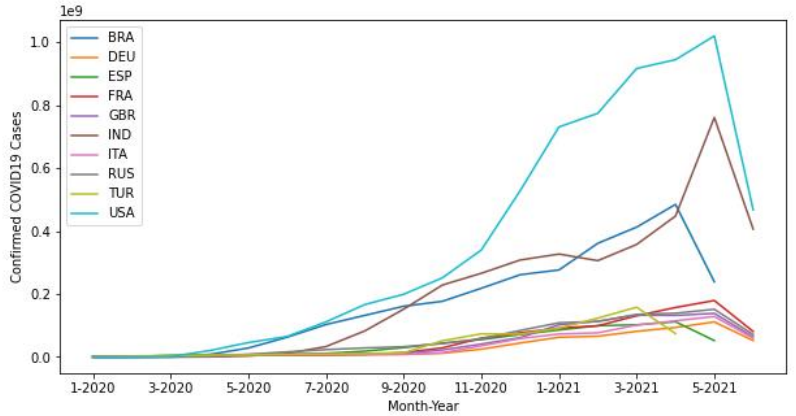

Figure 2 Plot Cumulative Confirmed Cases for Top 10 Countries

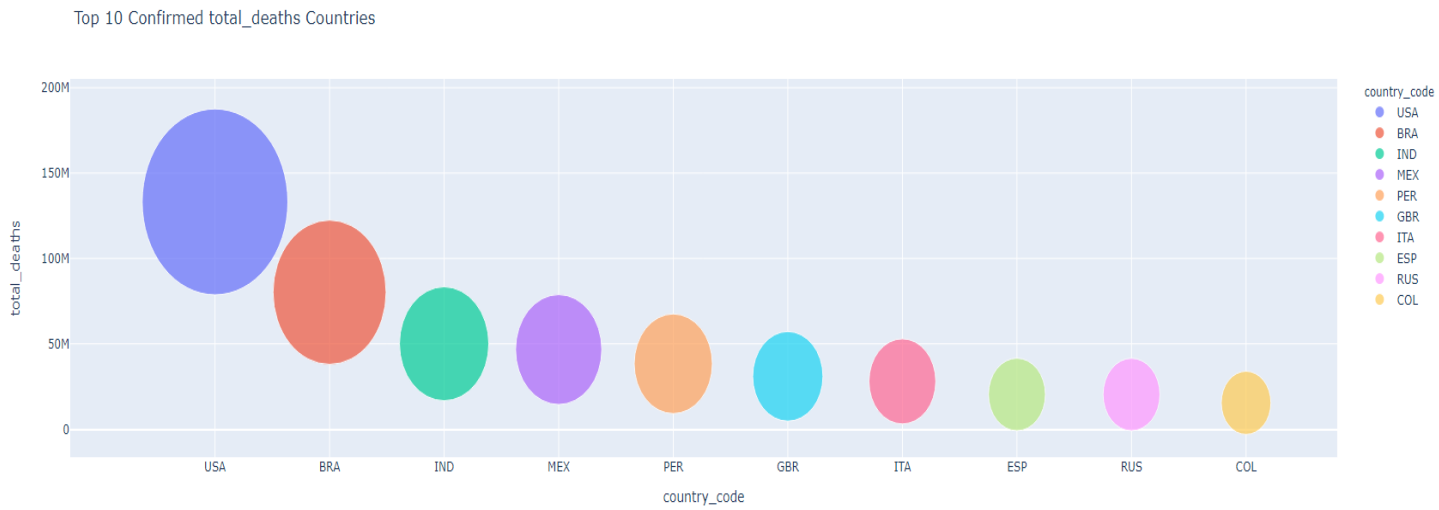

Figure 3 Top 10 Countries confirmed total_deaths countries

Feature selection

- For COVID19 impact on each country.
- **Data_final[['location','population','population_density','median_age','aged_65_older','aged_70_older','cardiovasc_death_rate','diabetes_prevalence','life_expectancy']]**
- COVID19 Response , these features as below are selected.
- **Data_final[['location','hospital_beds_per_thousand','human_development_index','extreme_poverty','total_cases','total_deaths','total_tests','people_fully_vaccinated','total_cases_per_million','total_deaths_per_million','people_fully_vaccinated_per_hundred']]**

Classification

Figure 4 World Map with Covid data for clustering countries

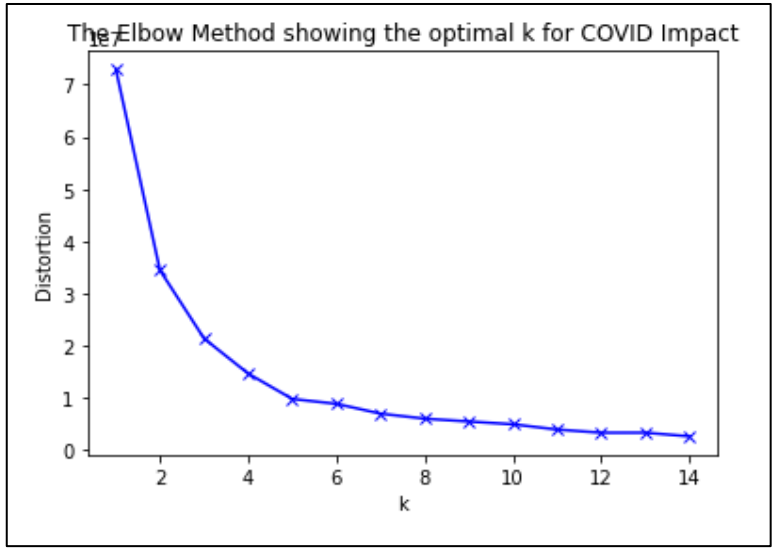

Figure 5 Kmeans Clustering of countries as per COVID19 Impact: Fit Kmean Clustering

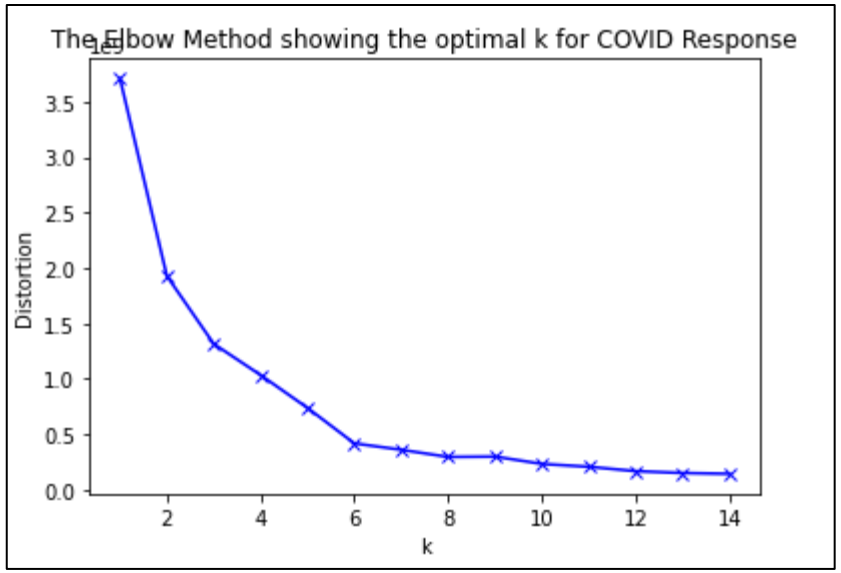

Figure 6 Kmeans Clustering of countries as per COVID19 Response: Determine optimal k

Conclusions

- 8 clusters were found based on COVID19 impact as shown in Figure 7. Similarly, 8 clusters were formed as per COVID19 response

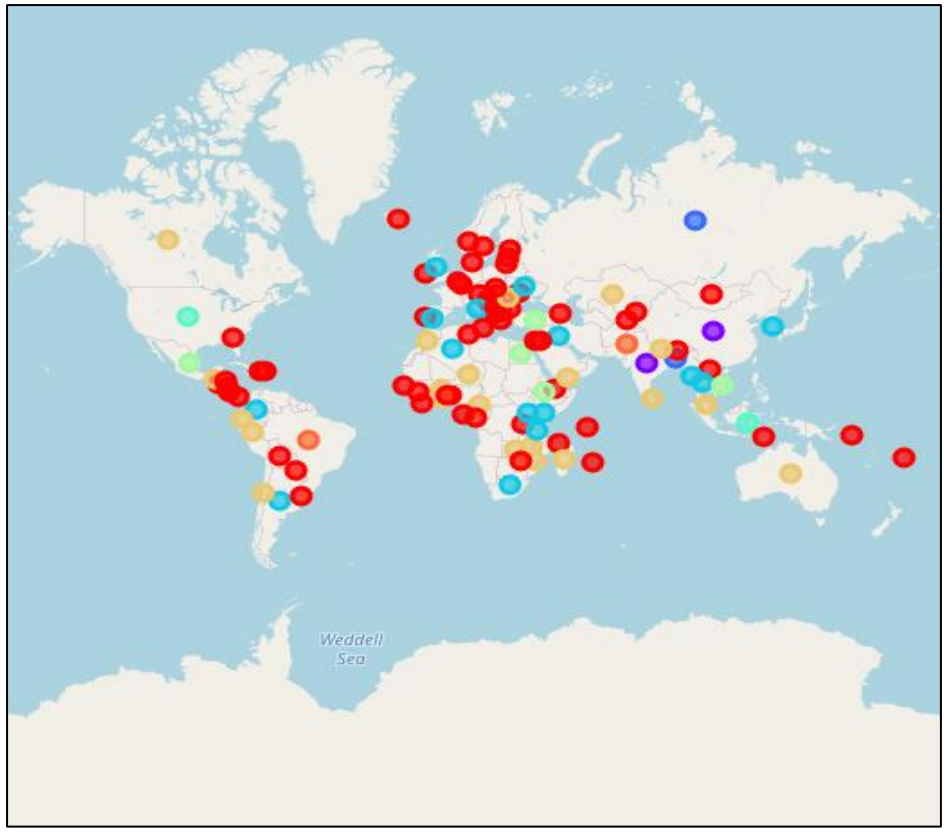

Figure 7 Clustering Results: Create a map of countries with similar COVID19 impact in terms of cases and deaths

Conclusions

- 8 clusters were found based on COVID19 impact as shown in Figure 7. Similarly, 8 clusters were formed as per COVID19 response

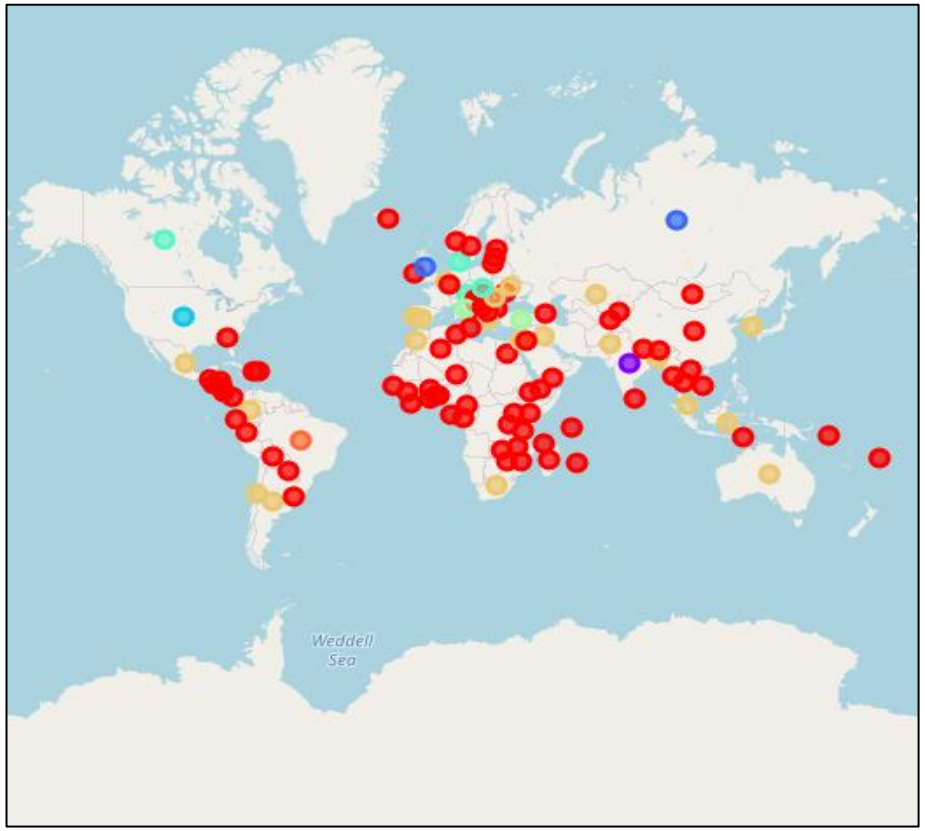

Figure 8 Clustering Results: Create a map of countries with similar COVID19 response in terms tests and Vaccinations