Data Analysis on Worldwide Covid19 Vaccinations

M.Nivas [1], K.UdaySurya Deveswar Reddy [2]
Vellore Institute of Technology, Chennai.

Email:
[1]maddukuri.nivas2019@vitstudent.ac.in
[2]udaysurya.deveswar2019@vitstudent.ac.in

Introduction

Coronavirus disease (COVID-19) is an infectious disease caused by a newly discovered coronavirus. A coronavirus identified in 2019, SARS-CoV-2, has caused a pandemic of respiratory illness, called COVID-19. It made every person work from home. Which causes more mental health. Many countries have been affected in many ways due to covid19.

India with a population of more than 1.34 billion will have difficulty in controlling the transmission of Covid-19 among its population. To overcome this pandemic situation the most preferable solution is vaccination.

With the huge data coming every day, this allows us to find how the vaccination is proceeding in every country, and how the count of new affecting cases and deaths varies based on the vaccination progress. And how the countries are varying deaths based on smokers, diabetes, and other medical issues. And finally, we will analyze our country India.

Table of Contents

2. **Data Preparation**
   a. Data Management
   b. Data Quality
   c. Data Preparation

3. **Exploratory Data Analytics**
   a. Analyzing country-wise total cases
   b. Analyzing the number of cases in all countries across the world
   c. Analyzing month-wise cases in top-4 affected countries
   d. Analyzing deaths per country
   e. Analyzing the number of deaths in all countries across the world
   f. Analyzing month-wise deaths in top-4 affecting countries
   g. Analyzing total vaccinations per country
   h. Analyzing the number of vaccinations in all countries across the world
   i. Analyzing the country-wise vaccinations per million
   j. Analyzing the number of cases and deaths by continent
   k. Pearson Correlation between Cases and Vaccinations
   l. Relation between cases and vaccinations in top 10 healthcare countries
   m. Analysis between covid factors, and country development indices
   n. Analysis between covid factors and different health death rates
   o. Correlation between Cases, Deaths, and Vaccinations
   p. Analyzing Covid-19 Cases in Top-5 Developed Countries

4. **Model Development**
   a. Arima Model
   b. Results
   c. Conclusion

1. Overview

The growth of the disease is moving from variations to variations which leads to the cause of very difficult to live in daily lifestyle. Most countries in the world are reeling under these difficult situations. So many people are affected by this disease and it causes many health problems. According to health organizations, India is one of the most affected countries in the world.
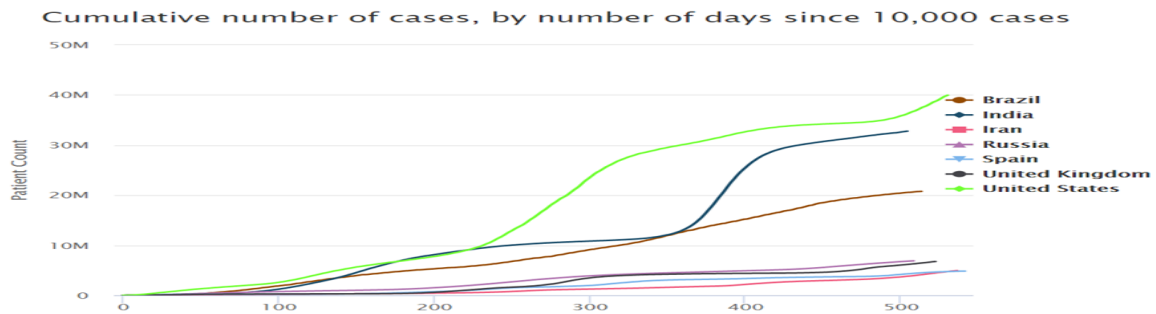

Cumulative number of cases, by number of days since 10,000 cases

Fig-1: Number of cases

Vaccines made a solution to control the seriousness of the pandemic all over the world. Among countries with more than 1 million people, South Korea is currently vaccinating faster than any other, administering a daily average of 1,536 doses per 100,000 people in the past seven days. Mongolia was previously the fastest, administering about 3,434 doses per 100,000 people per day. In India, 10.7% of people are fully vaccinated and 37% of people have got at least 1 dose of vaccine.

Problem Statement

To analyze any of the data, firstly we should be able to identify the path to solve the problem. Our main goal is to analyze covid19 vaccinations progress and to predict the time to complete the vaccination process for all the people. Also, we want to analyze how the count of new affecting cases and deaths varies based on the vaccination progress, how the countries are varying deaths based on smokers, diabetes, and other medical issues.To achieve the above mentioned targets, we should acquire knowledge of retrieving data from the dataset, data preparation, exploring the data, building the models, and ml-based algorithms to predict the values.This type of work will be helpful to the governments/ organizations. Though it will show the analysis on the problems it will help them to make more useful decisions.

Objective of the project

Our problem statement is to analyze the covid19 vaccination progress in a maximum number of countries and as well as state-wise in Indian countries. There are three main goals of this project.

1. Analysing the health problems causing death.

2. To analyze the vaccination progress between states in India.

3. To predict the time that will take to complete full vaccination in an Indian country.

Scope of the project

1. The scope of the study covers a major number of covid affected countries.

2. This study covers one-year data which was starting from 24th February 2021.

3. The study focuses on the factor's which data will be available that can be used for analysis.

Out of scope

 1. This study will not focus on developing the factors, it will focus on providing suggestions by predicting the information based on available data.

 2. This study covers one-year six months data which was starting from 24th February 2021.

Data source

The data for the project was obtained from COVID-19 Data Repository by the Centre for Systems Science and Engineering (CSSE) at Johns Hopkins University and the vaccination details are taken from National government reports. Currently, Johns Hopkins University is working on research regarding variants of covid19. Data is available for the following dimensions which we will use in analytics:

1. continent

2. new cases

3. new deaths

4. new cases per million

5. new deaths per million

6. people vaccinated

7. people fully vaccinated

8. new vaccinations

9. people vaccinated per hundred

10. people fully vaccinated per hundred

11. stringency index

12. population

13. median age

14. aged 65 older

15. aged 70 older

16. extreme poverty

17. cardiovascular death rate

18. diabetes prevalence

19. female smokers

20. male smokers

21. handwashing facilities

22. life expectancy

23. human development index

Tools and Technologies used

Tools Used are - R studio, Microsoft Excel.

Using graphs and box plots to visualize the data.

Techniques Used are handwashing - ARIMA algorithm, box plot, and graphs. Using R programming language and Microsoft Excel to analyze and predict the data.

Analytics approach

1. Data quality
2. Data cleaning and preparation
3. Explore the data
4. Study the variables
5. Split the variables (if needed)
6. Build the model
7. Divide into training and test set
8. Develop the model
9. Final model
10. Validation of data
11. Apply the algorithm
12. Predict the values using the main model

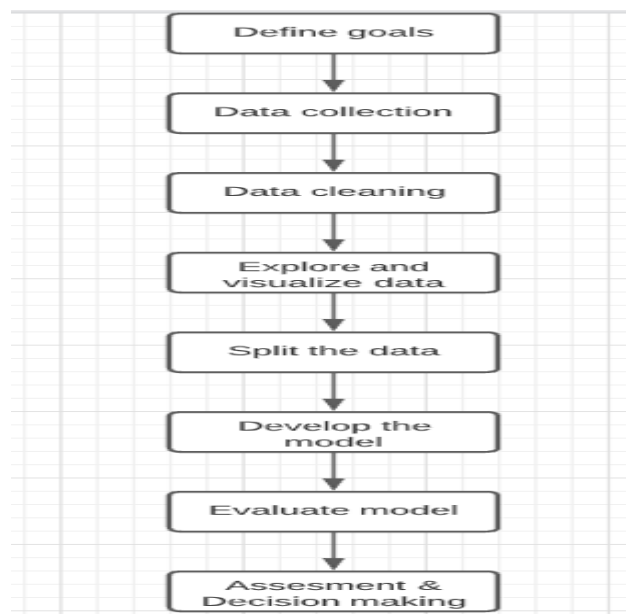**Stepwise analytical approach for our work:**

Fig-2: Process Flow

Limitations

There are a few restrictions that this study has w.r.t information and the strategy that can be utilized.

1. The dataset which we had taken for our analytics had some inconsistent data in a few attributes and some attributes will not be useful to achieve our goal. So, we are not using those particular attributes in our analysis.

2. This study covers one-year six months data which was starting from 24th February 2021.

3. This consideration will not center on creating the variables, it'll center on giving the proposals by predicting the data based on accessible information.

4. Our study mainly focuses on how far vaccinations impact the positivity rate of covid-19 rather than steep and gentle of the country's economical aspects.

2. Data Description and Preparation

Data Management

Data management is the practice of collecting, keeping, and utilizing information safely, effectively, and cost-effectively. We had extracted data for a one and half years across 26 variables. This was collected from COVID-19 Data Repository by the Centre for Systems Science and Engineering (CSSE) at Johns Hopkins University and the vaccination details are taken from National government reports. The data includes all the data of 190 countries. This data covers all the information about the number of cases per day, vaccinations per day, and the affected group ages.

By utilizing the above data, we analyzed the steep and gentle of vaccinations and cases. The visualizations which we used will give insights on country-wise vaccination progress, the impact of vaccinations in their respective countries across the world. We will perform a timeline prediction that conveys to us the flow of when the vaccination progress will be completed.

Attribute	Description	Data Type
continent	The landmass of the topographical area	Nominal
new cases	New confirmed cases of COVID-19	Continuous
new deaths	Unused passing's credited to COVID-19	Continuous
new cases per million	New affirmed cases of COVID-19 per 1,000,000 individuals	Continuous
new deaths per million	New deaths attributed to COVID-19 per 1,000,000 people	Continuous
people vaccinated	Total number of people who received at least one vaccine dose	Continuous
People fully vaccinated	Total number of people who received all doses prescribed by the vaccination protocol	Continuous
new vaccinations	New COVID-19 vaccination doses administered	Continuous
people vaccinated per hundred	Total number of people who received at least one vaccine dose per 100 people in the total population	Continuous
people fully vaccinated per hundred	Total number of COVID-19 vaccination doses administered per 100 people in the total population	Continuous

stringency index	Lockdown rate	Continuous
population	Population of country	Continuous
median age	The median age of the population	Continuous
aged 65 older	Share of the population that is 65 years and older	Continuous
aged 70 older	Share of the population that is 70 years and older	Continuous
extreme poverty	Share of the population living in extreme poverty	Continuous
cardiovascular death rate	The death rate from cardiovascular disease	Continuous
diabetes prevalence	Diabetes prevalence (% of population aged 20 to 79)	Continuous
female smokers	Share of women who smoke	Continuous
male smokers	Share of men who smoke	Continuous
handwashing facilities	Share of the population with basic hand washing facilities	Continuous
life expectancy	Life expectancy at birth	Continuous
human development index	A composite index measuring average achievement in three basic dimensions	Continuous

Table 1: List of attributes

Data Quality

Positive cases and several vaccinations data for certain days were missing. Some days had data for only a few of the variables. Data for the days where there was no data for key variables like new cases, new deaths, vaccinations, fully vaccinated. so, the missing values which can be calculated are imputed else the records have been removed.

For some attributes like total vaccinations and the number of vaccinations, we can't impute the values for those particular attributes. So, we didn't consider those particular attributes in our predictions and analysis.

This data contains 107647 rows and 26 columns. Each row corresponds to the day-by-day covid situation of each country. It includes all the data of 190 countries.

Data Preparation

Missing Values

1. Specific missing value treatment was used for particular records of attributes. For some of the attributes which data is missing at random has been replaced with 0 and for some attributes by using the attributes of the data values have been calculated, and for a few of the attributes, we used median as it was robust to outliers.

2. Data for which no data is available for key variables like new cases, new deaths, people vaccinated, people fully vaccinated, new Vaccinations have been replaced by zero as data is missing for only a few days.

3. Data for which no data is available for key variables like new cases per million columns, new deaths per million, people vaccinated per hundred, people fully vaccinated per hundred columns, stringency index column the values for records which missing values are present are calculated by using the attributes new cases, new deaths, and vaccination details. For variables like median age, life expectancy rate, cardiovascular rate the values have been imputed by using median as the median is robust to outliers.

Algorithm Used:

The data which we are using for analytics and prediction contains the information day-wise. It was time-series data. So, by using the ARIMA model we are predicting the time of day at which India will be fully vaccinated. An autoregressive integrated moving average, or ARIMA, a statistical analysis shows that uses time-series information to either get the information set or to predict future patterns. A measurable demonstration is autoregressive on the off chance that it predicts future values based on past values.

3. Exploratory Data Analysis

Exploratory data analysis insinuates the fundamental handle of performing starting examinations on data to discover designs, spot peculiarities, test hypotheses, and check assumptions with the help of rundown experiences and graphical representations.

Analyzing different countries' data and interpreting the situation of covid-19 cases and the vaccination progress. In this, we used various graphs to describe the insights and to compare the progress between different countries.

Analyzing country-wise total cases:

Analyzing country-wise total cases will help us to let us know the count of a total number of cases.

We used a simple bar graph to count the number of cases by each country. In the above bar plot, we are showing the top 15 countries with the highest number of cases. The count of cases was mapped on X-axis and country names were mapped on Y-axis.
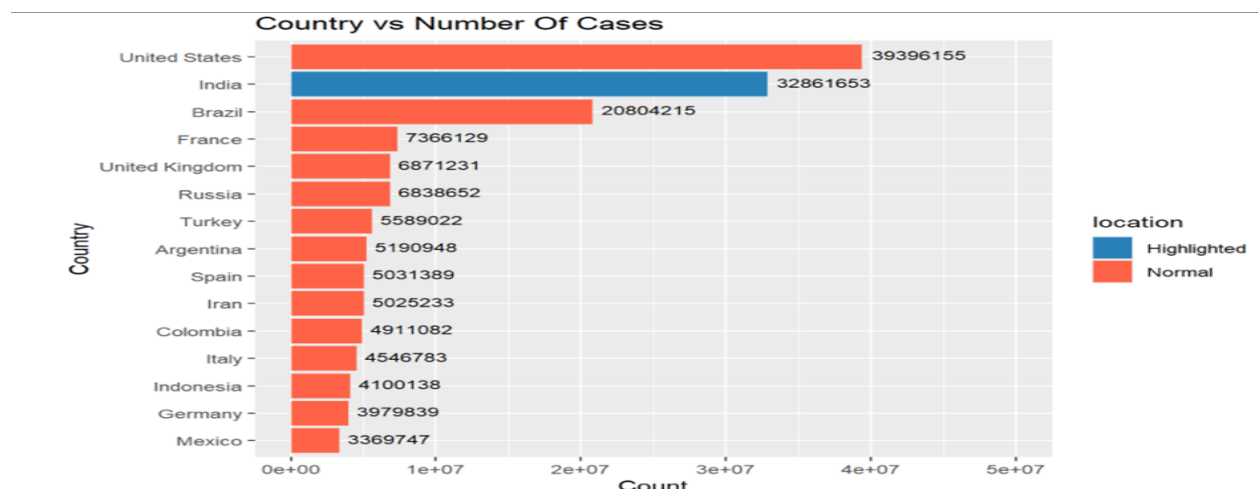

Fig-3: Plot showing top 15 countries with the highest number of positive cases.

Insights:

The bar highlighted shows the number of cases in India and the plot with normal shows the rest of all countries. Among all countries, the United States has the highest number of cases. United States, India, Brazil are the top 3 most affected countries in the world due to covid.

Analyzing the number of cases in all countries across the world:

In the previous plot, as we used a bar plot, we will be able to visualize only a limited number of countries. In this particular plot, we are using a world map so that we can visualize the number of cases in all countries.
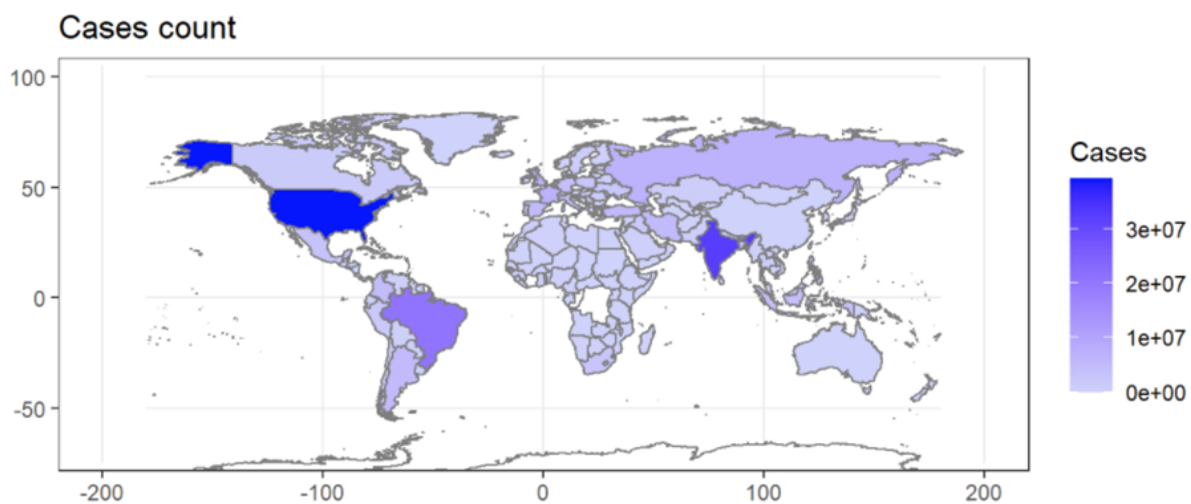
Fig-4: World map showing number of positive cases across the world.

Insights:

This world map shows the total number of positive cases in each country across the world. As the number of positive cases increases, the thickness of color will be more in the respective country region. This map shows us that India, the United States, Brazil, the United Kingdom, France, and Russia have been affected more than all other countries.

Analyzing month-wise cases in top-4 affected countries:

From previous visualizations, we know that the top-4 affected countries are the United States, India, Brazil, and France respectively. In this line plot, we are interpreting how cases are varying by different months. Months were mapped on the X-axis and the count of cases was mapped on Y-axis.
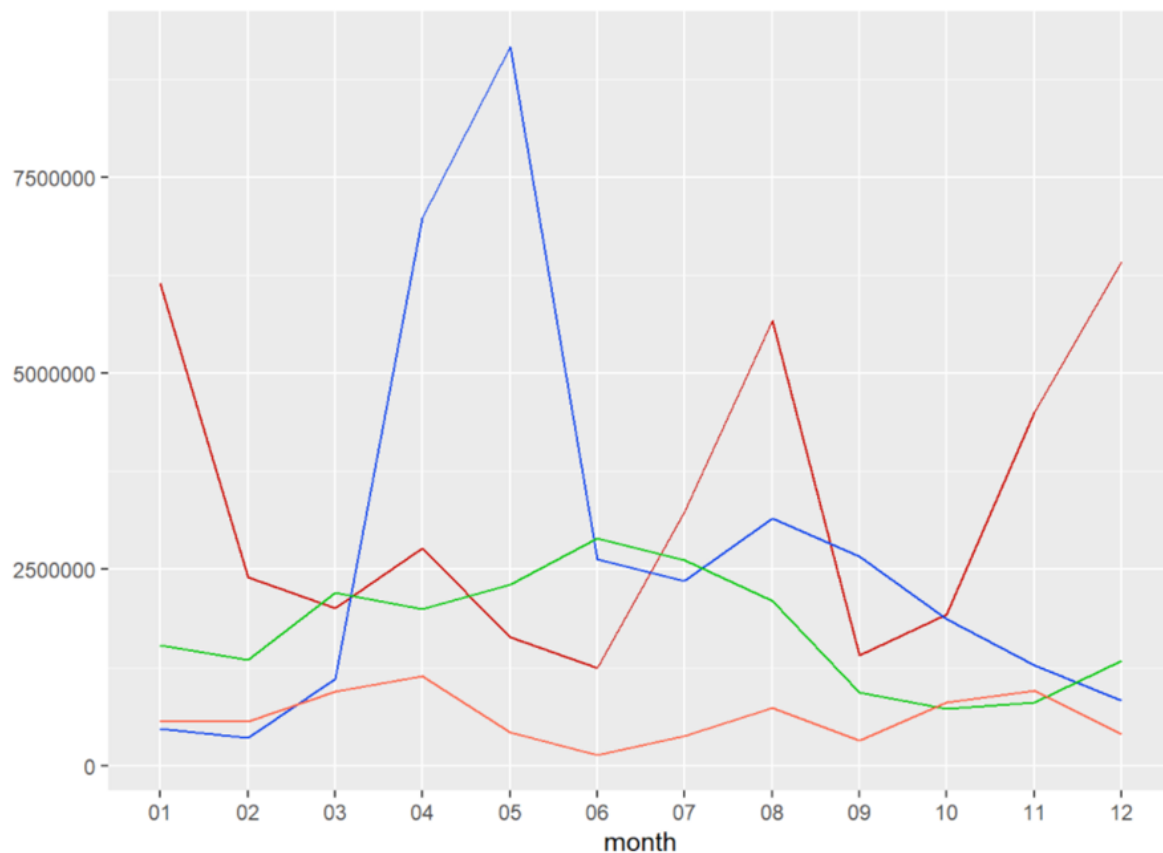

Fig-5: Month-wise cases analysis.

Insights:

In this particular line plot the color of lines represents their respective countries like blue represents India, green represents Brazil, red represents the United States. It shows that India has a gentle increase in positivity between February and May. America has a peak between June and September. Cases are increasing in the summer and rainy seasons. and decreasing in the winter season.

Analyzing deaths per country:

Analyzing country-wise total deaths will help us to let us know the count of the total number of deaths per country.

We used a simple bar graph to count the number of deaths in each country. In the above bar plot, we are showing the top 15 countries with the highest number of cases. The count of deaths was mapped on X-axis and country names were mapped on Y-axis.
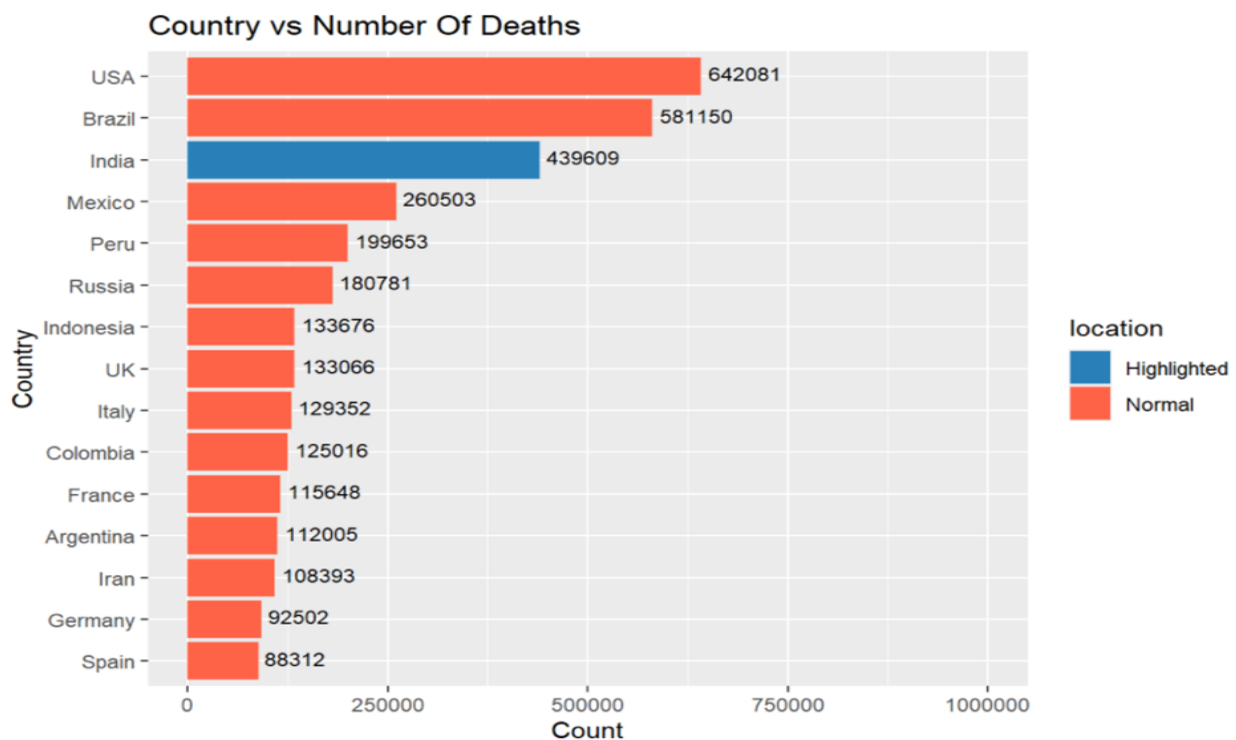

Fig-6: Deaths Per country.

Insights:

The bar highlighted shows the number of deaths in India and the plot with normal shows the rest of all countries. Among all countries, the United States has the highest number of deaths. The United States, Brazil, and India are the top 3 most affected countries(deaths) in the world due to covid. Compared between cases top-15 and deaths top-15, PERU was not present in the top-3 of cases but present at the top-5 in deaths. Turkey has the lowest death rate even though it presents in 7th place of cases, it was not present in the top-15 of deaths.

Analyzing the number of deaths in all countries across the world:

In the previous plot, as we used a bar plot, we will be able to visualize only a limited number of countries. In this particular plot,e are using a world map so that we can visualize the number of cases in all countries.
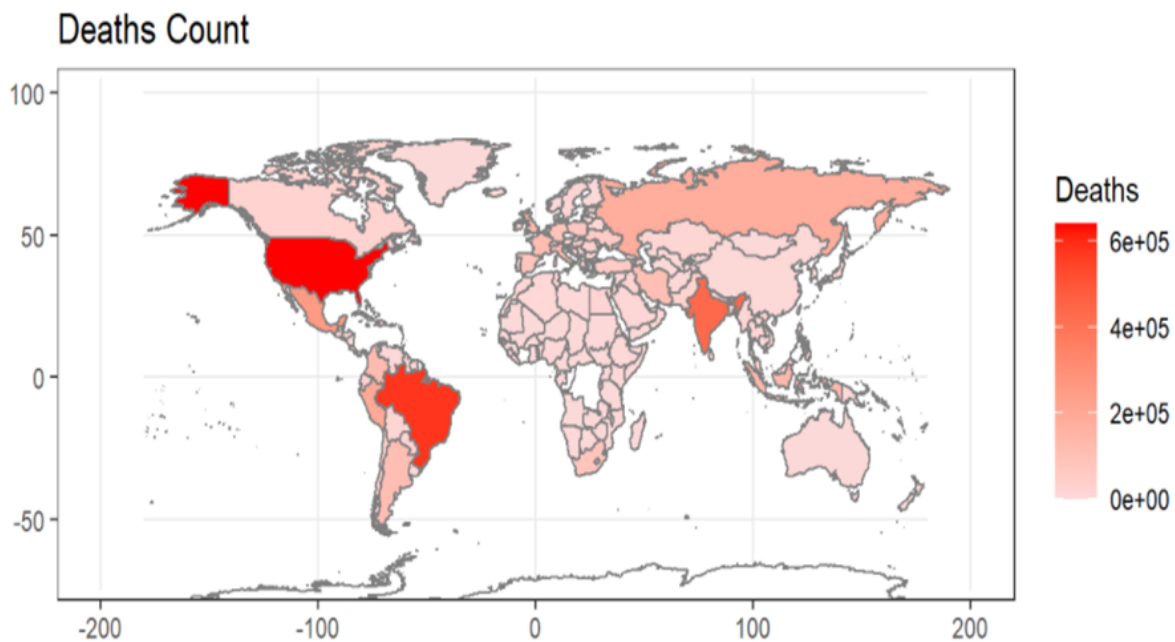

Fig-7: World map showing number of deaths across the world.

Insights:

This world map shows the total number of deaths in each country across the world. As the number of deaths increases, the thickness of color will be more in the respective country region. This map shows us that India, the United States, Brazil, Mexico, Peru, Russia have been affected more than all other countries in the count of deaths.

Analyzing month-wise deaths in top-4 affecting countries:

From previous visualizations, we know that the top-4 affected countries are the United States, India, Brazil, and France respectively. In this line plot, we are interpreting how cases are varying by different months. Months were mapped on the X-axis and the count of cases was mapped on Y-axis.
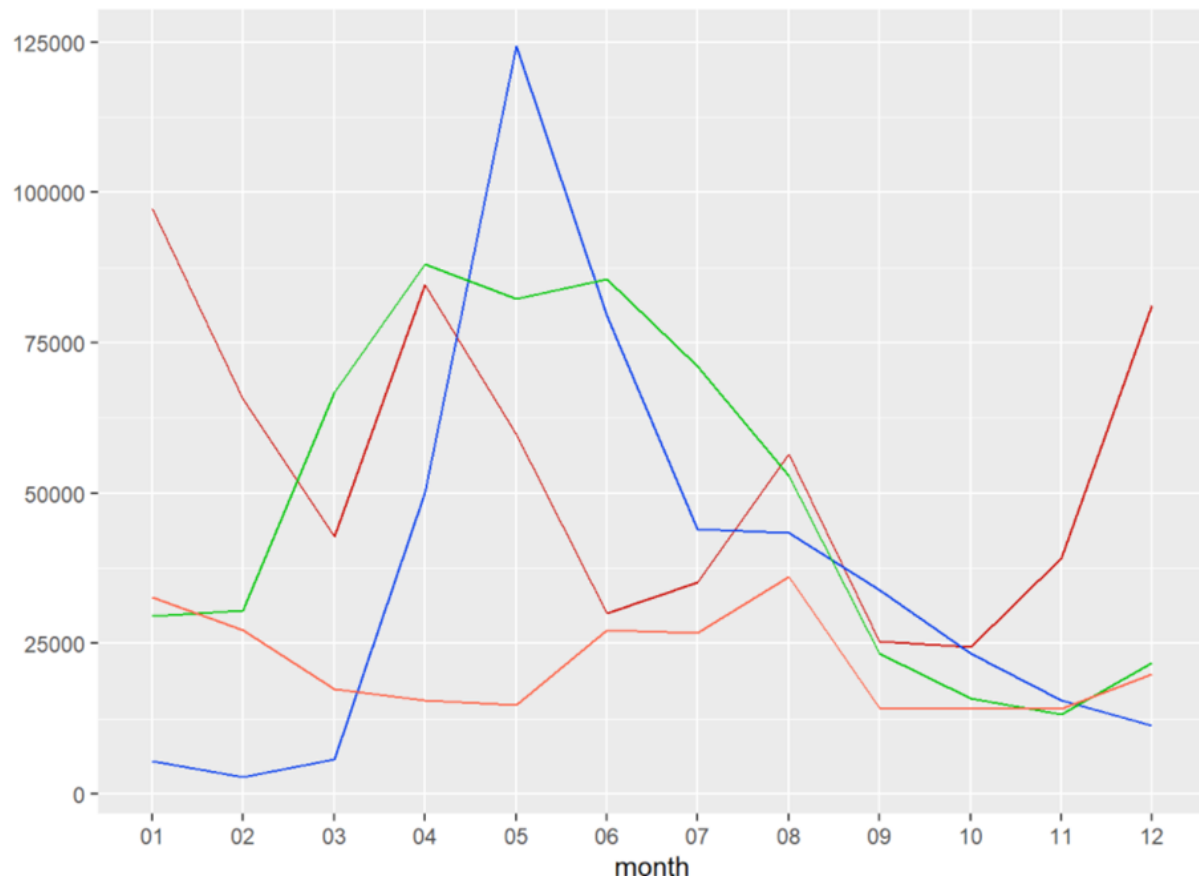

Fig-8: Month-wise deaths in top-4 affected countries.

Insights:

In this particular line plot the color of lines represents their respective countries like blue represents India, green represents Brazil, red represents the United States. It shows that India has a gentle increase in positivity between March and May. Deaths are increasing in the summer and rainy seasons. and decreasing in the winter season.

Analyzing total vaccinations per country:

Analyzing country-wise total vaccinations will help us to let us know the count of the total number of deaths per country.

We used a simple bar graph to count the total vaccinations by each country. In the above bar plot, we are showing the top 15 countries with the highest number of vaccinations. The count of vaccinations was mapped on X-axis and country names were mapped on Y-axis.
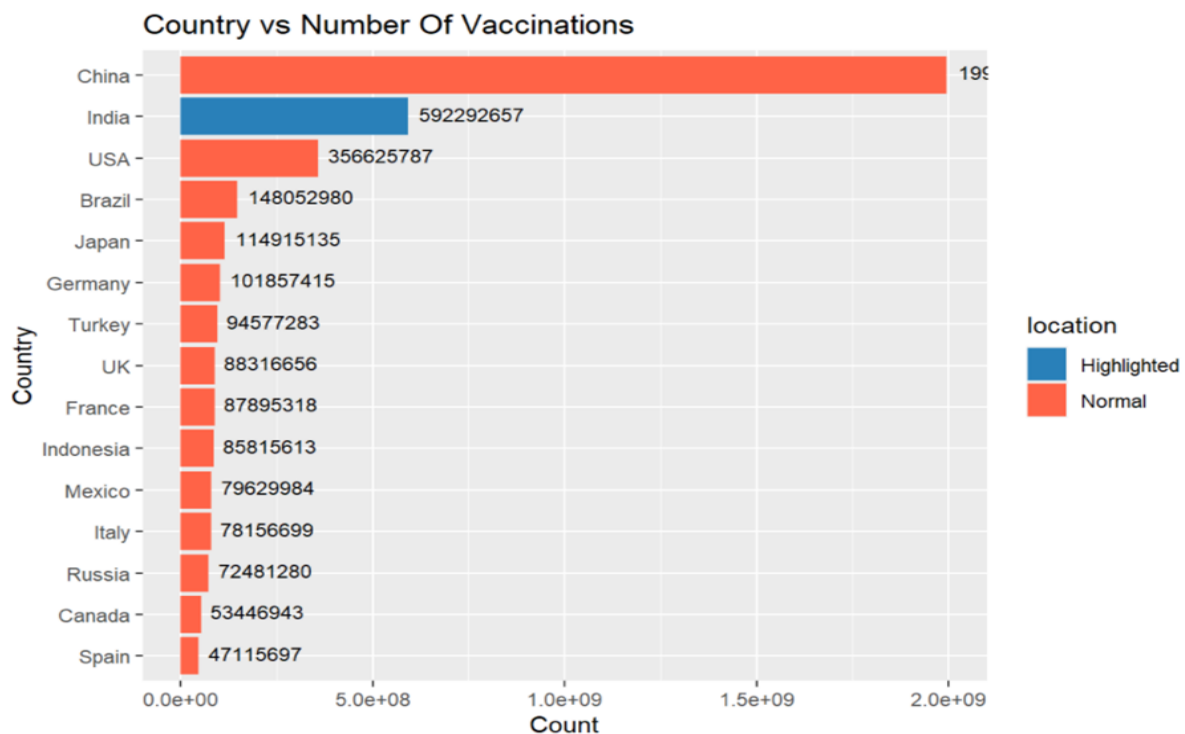

Fig-9: Country-wise vaccinations.

Insights:

The bar highlighted shows the number of deaths in India and the plot with normal shows the rest of all countries. Among all countries, China has the highest number of vaccinations that were given to their people. The top 5 countries in vaccinations are China, India, the USA, Brazil, and Japan. Even though the UK and Russia have their vaccinations they are placed below Brazil, Japan, etc. This is because of a lack of influence in the country.

Analyzing the number of vaccinations, handwashing in all countries across the world:

In the previous plot, as we used a bar plot, we will be able to visualize only a limited number of countries. In this particular plot, handwashing is using a world map so that we can visualize the number of cases in all countries.
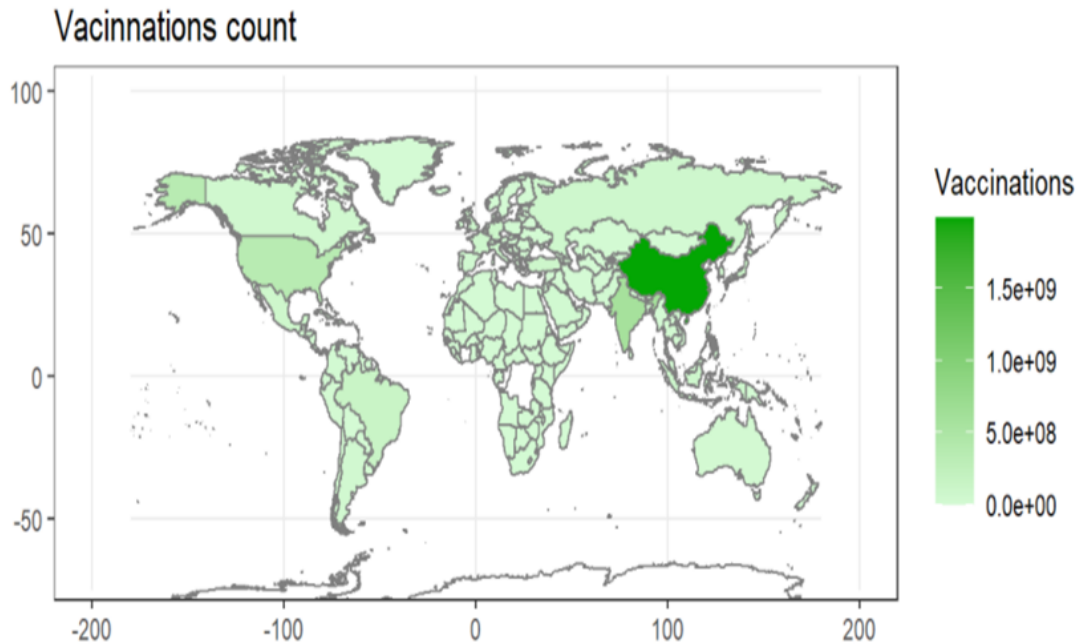

Fig-10: World map showing number of vaccinations across the world.

Insights:

This world map shows the total number of vaccinations given to the people in each country across the world. As the count increases, the thickness of color will be more in the respective country region. This map shows us that China, India, the United States, Brazil, and Japan have the highest number of vaccinations than all other countries in the world.

Analyzing the country-wise vaccinations per million:

Analyzing country-wise total vaccinations will help us to let know whether countries are doing vaccinations enough or it should be increased.

We used a simple bar graph to count the total vaccinations per million by each country. In the above bar plot, we are showing the top 15 countries with the highest number of vaccinations. The count of vaccinations was mapped on X-axis and country names were mapped on Y-axis.
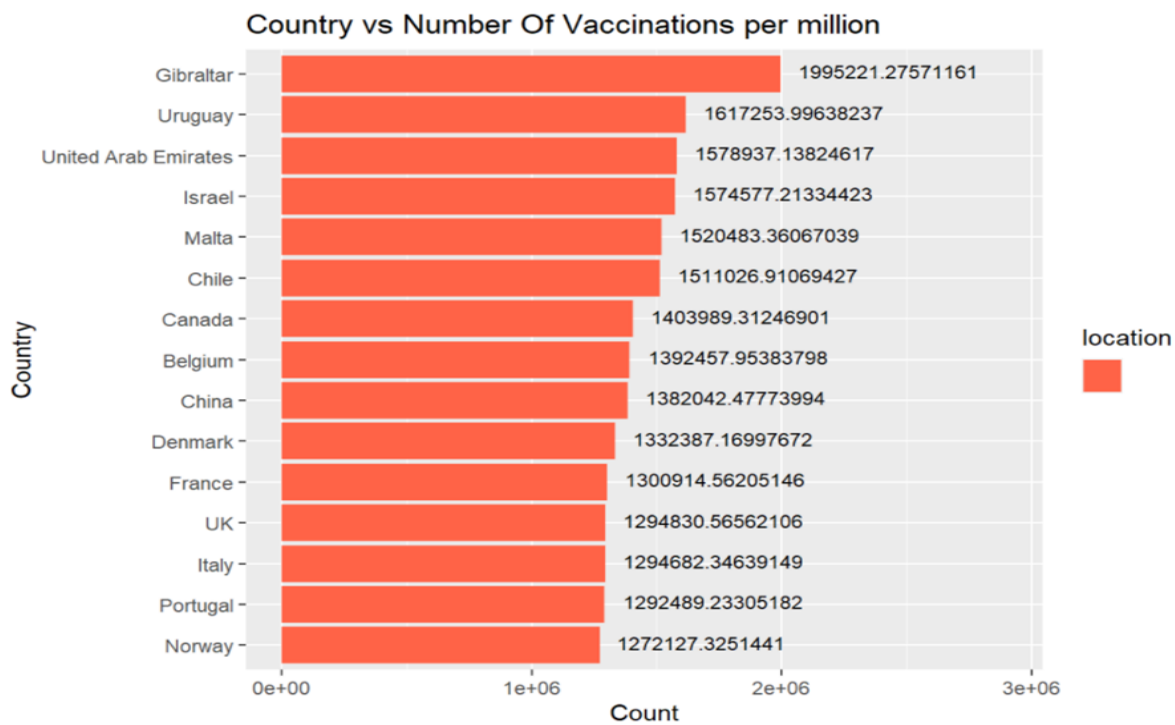

Country vs Number Of Vaccinations per million

Country	Count
Gibraltar	1995221.27571161
Uruguay	1617253.99638237
United Arab Emirates	1578937.13824617
Israel	1574577.21334423
Malta	1520483.36067039
Chile	1511026.91069427
Canada	1403989.31246901
Belgium	1392457.95383798
China	1382042.47773994
Denmark	1332387.16997672
France	1300914.56205146
UK	1294830.56562106
Italy	1294682.34639149
Portugal	1292489.23305182
Norway	1272127.3251441

Fig-11: Country-wise vaccinations per million.

Insights:

This graph shows the analysis on country-wise vaccinations per million. In this, the rank of India is 77. Chile, Canada, China, France, the UK, and Italy are the countries that present on top-15 of Vaccinations and Vaccinations according to their population. India is at 77th rank for the vaccinations compared to their population. This means India is not doing enough vaccinations according to its population.

Analyzing the number of cases and deaths by continent:

For analyzing those particular variables, we used a bubble plot that varies in size which insights the rate of increase and decrease. In this bubble plot, we mapped cases on X-axis and deaths on Y-axis. By visualizing this map we can draw the rate of death in positive cases and the impact of covid-19 in each continent.
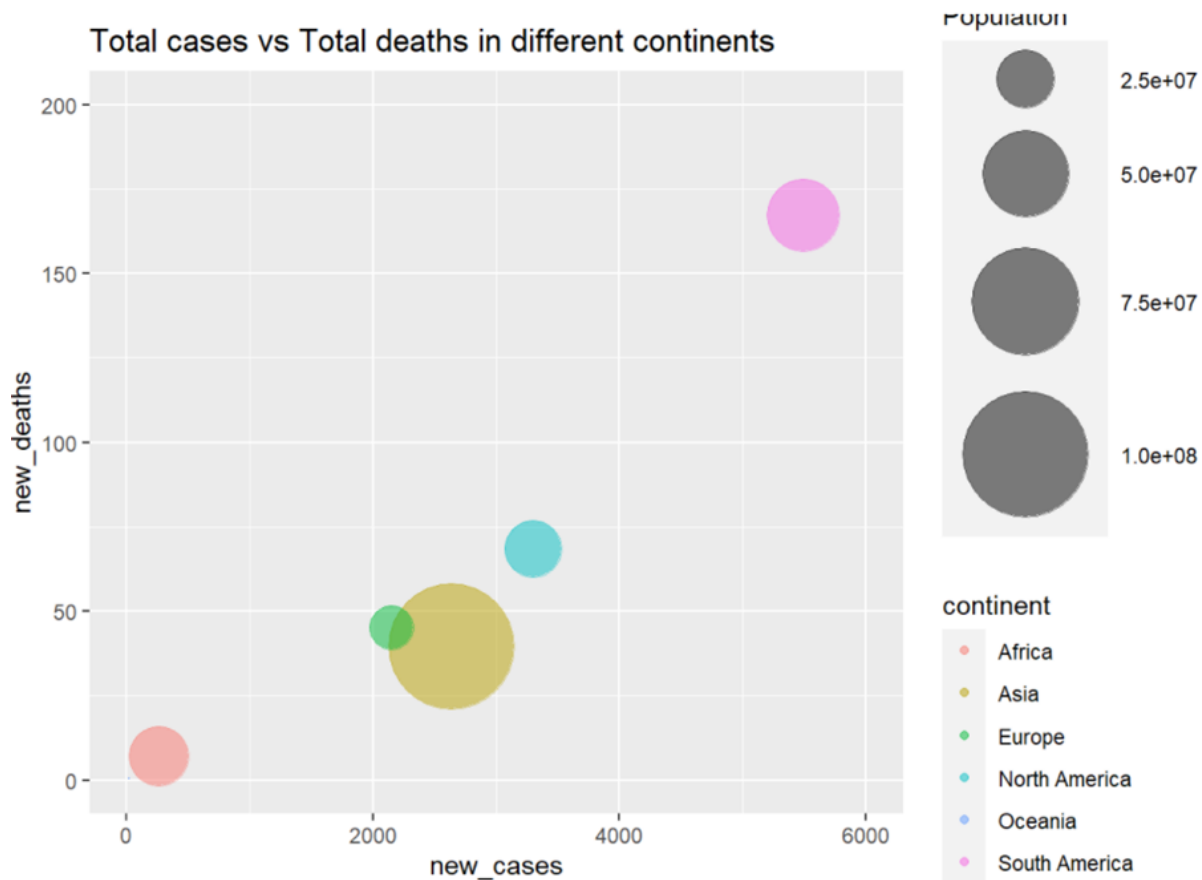

Fig-12: Total cases vs Total deaths (Continent wise).

Insights:

South America has a greater number of cases as well as death rates than all other continents. South America was the epi-center of covid-19 cases and deaths between 2020-2021. Even though the population is very high in Asian countries, the death rate is not very high when compared to other continents. Oceania manages to not affect much in those pandemic situations. Europe and North America have a moderate level of cases and deaths. At the starting stage of the pandemic Africa was not much affected by covid-19. But most of the cases and deaths are reported in recent months. This is because of HIV, which causes higher mortality rates.

Pearson Correlation between Cases and Vaccinations:

Pearson's relationship coefficient is the test measurements that measure the measurable relationship, or affiliation, between two continuous factors. It gives data around the size of the affiliation, or relationship, as well as the course of the relationship. In this, vaccinations are mapped on X-axis, and cases are mapped on Y-axis.
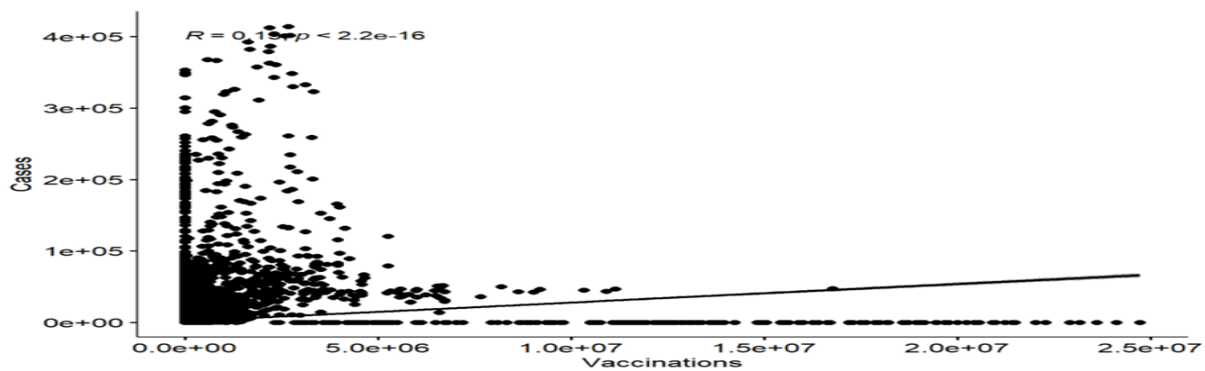

Fig-13: Pearson Correlation between cases and vaccinations.

Insights:

The value of the Pearson coefficient is 0.13 which suggests that it has a low positive correlation between those variables.
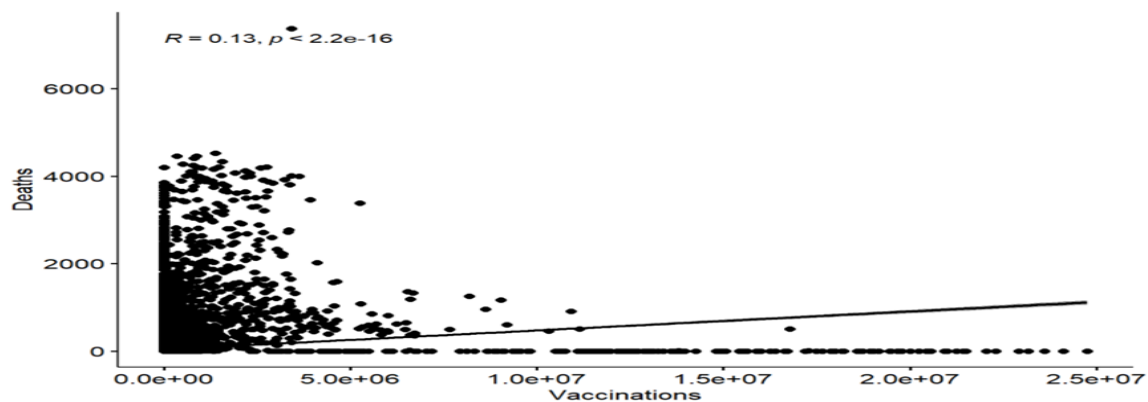
Pearson Correlation between deaths and Vaccinations:

Fig-14: Pearson Correlation between deaths and vaccinations.

Insights:

The value of the Pearson coefficient is 0.12 which suggests that it has a low positive correlation between those variables.

Relation between cases and vaccinations in top 10 healthcare countries:

To find the relation between cases and vaccinations we are using a scatter plot. Scatter plots are utilized to plot information focusing on a flat and a vertical hub within the endeavor to see how much one variable is influenced by another. The relationship between two factors is called their correlation. Here we used logarithmic because of skewness towards large values. In this plot case was mapped on X-axis and vaccinations were mapped on Y-axis.
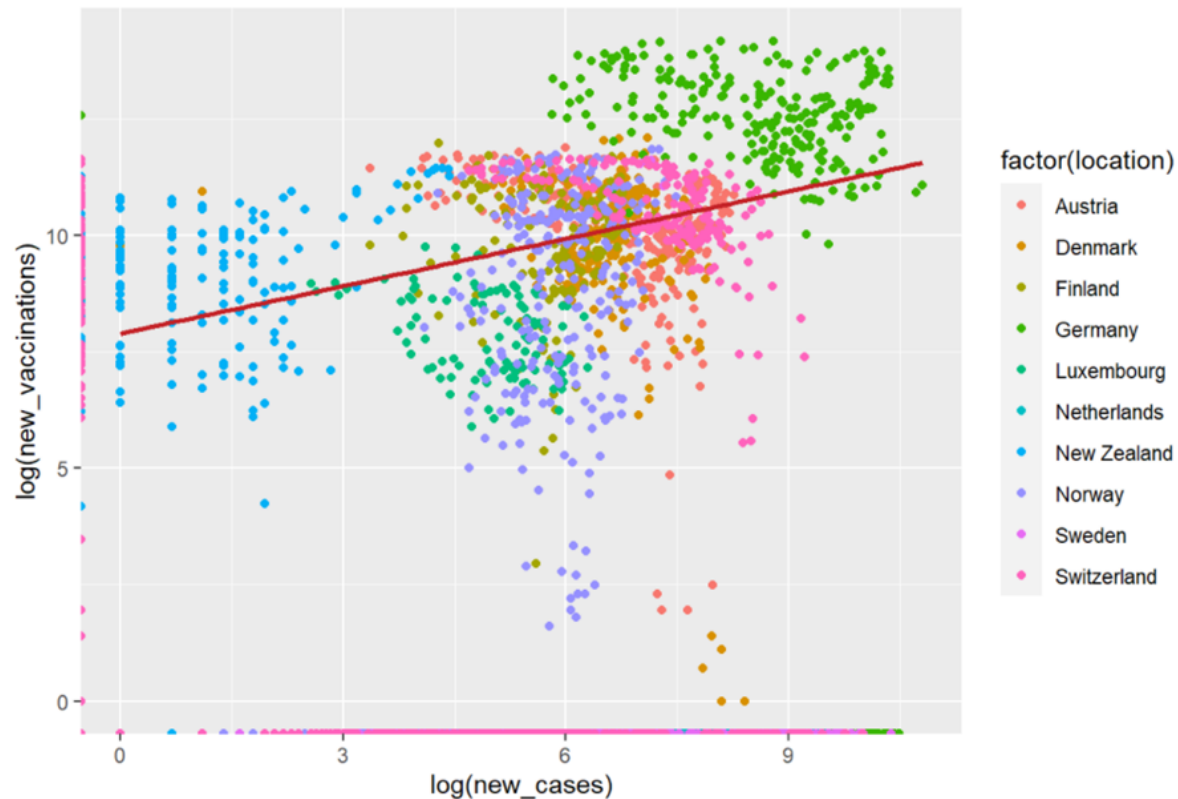

Fig-15: Scatter Plot between cases and vaccinations.

Insights:

This plot shows the relation between the vaccinations and cases. Here we used data from the top 10 health care countries. The vaccination process in Germany is going faster than in other countries. These two variables have a strong bond as the line of best fit suggests a positive relationship. As the count of cases increases day by day, the countries are managing to speed up the vaccination process.

Analysis between covid factors, and country development indices:

We used a scatter plot matrix to analyze the covid factors and the country development indices. A scatter plot matrix could be a grid (or framework) of scatter plots utilized to imagine bivariate connections between combinations of factors. Each diffuse plot within the network visualizes the relationship between a combination of factors, permitting numerous connections to be investigated in one chart.
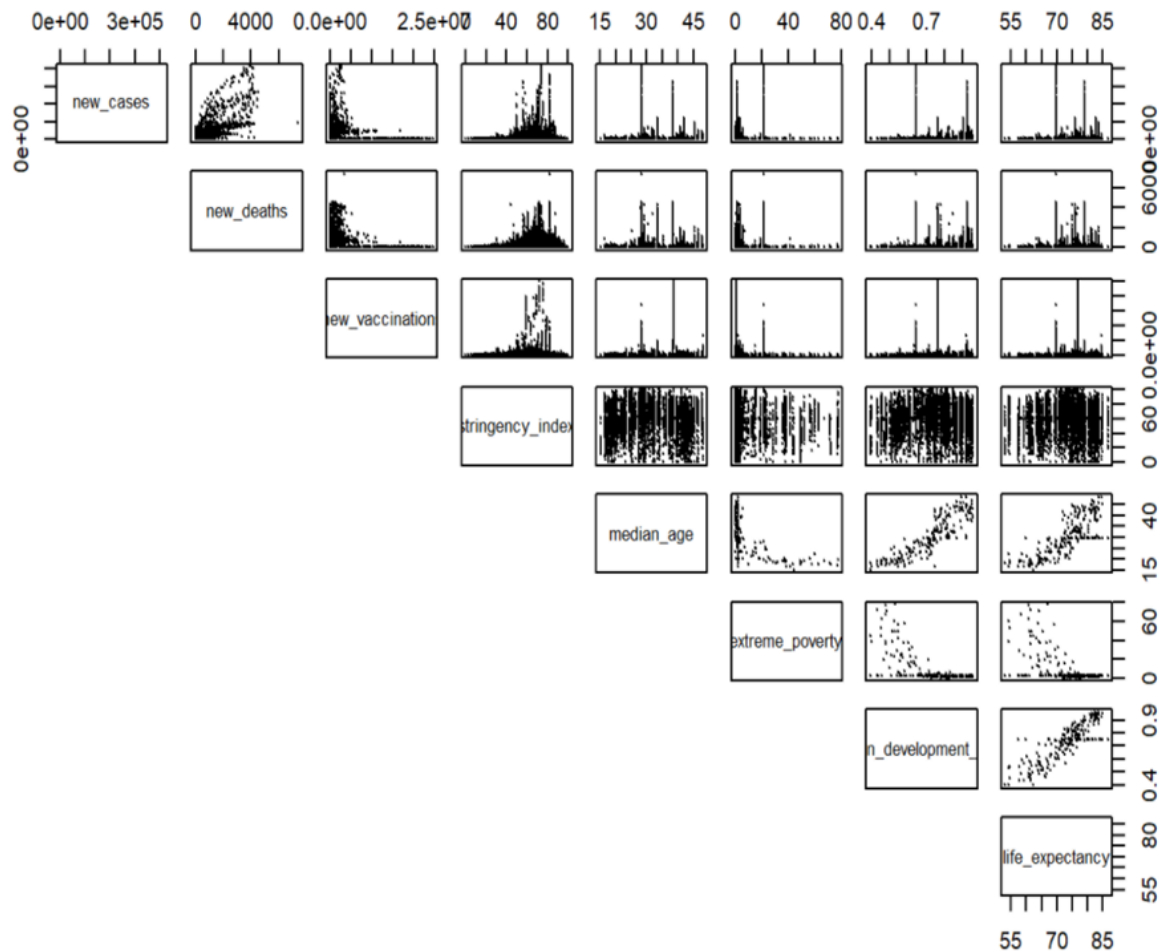

Fig-16: Analysis between covid factors and country development indices.

Insights:

The key inferences from the above scatter plot matrix are as the vaccination progress increases, deaths and cases are decreasing. As the stringency index increases, vaccinations are increasing. As extreme poverty increases, vaccinations are decreasing.

Analysis between covid factors and different health death rates:

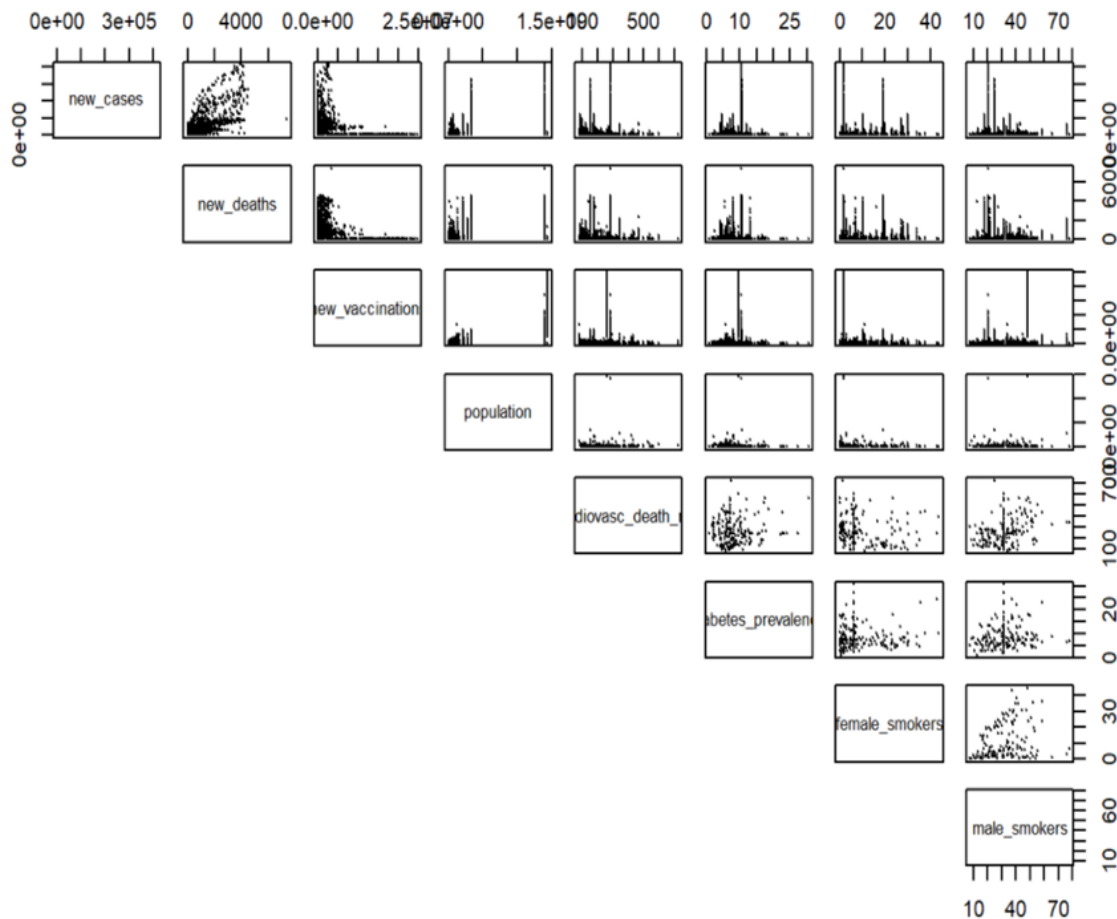We used a scatter plot matrix to analyze the covid factors and different health death rates.

Fig-17: Analysis between covid factors and different health death rates.

Insights:

As the population increases, cases, deaths, and vaccinations are increasing. As the Cardiovascular death rate and diabetes prevalence are increasing, vaccinations, deaths, and cases are increasing up to some point and decreasing after then. Since it is a rate, it depends on the population. So, we can say vaccinations, deaths, and cases were increasing.

Correlation between Cases, Deaths, and Vaccinations:

The correlation between cases and deaths is high. The correlation between cases and vaccinations is 0.13 which is almost equal to zero. This indicates they have a low positive correlation. The correlation between deaths and vaccinations is 0.12 which is almost equal to zero. This indicates they have a low positive correlation.

Analyzing Covid-19 Cases in Top-5 Developed Countries:

The top-5 developed countries are China, Ireland, Switzerland, Norway, Iceland. We used a line plot to insight how often the covid-19 cases are increasing and decreasing in those countries.
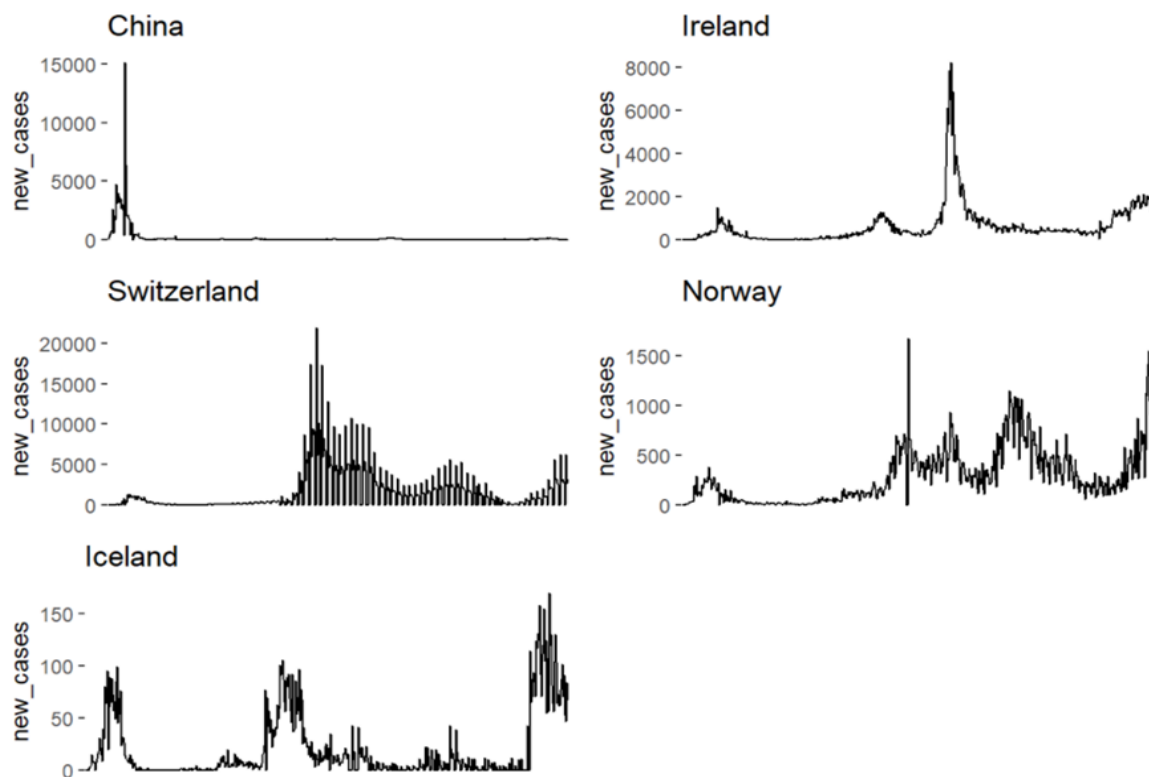

Fig-18: Analysis on Covid-19 Cases in Top-5 Developed Countries.

Insights:

From the above plot, we can infer that the rate of positive cases has been increasing day by day in Norway. At one moment Switzerland was affected more than the other top 4 developed countries. China managed to take certain measures to control the rate of cases and that gives a good result as they managed to not increase the rate. Iceland and Ireland are having peaks and valleys in their positive rate.

4. Model Development

● Arima model

The objective of the model development is to develop a model that will predict by what date the vaccination in India will be completed including all three doses (First dose, Second dose, Booster dose).

The arima model was trained on day-by-day data from 1st February 2020 to 1st September 2021. And it was used to predict the date by which all the doses including the booster dose will be completed.

To train the data, first, we filtered all the records in India. And we found the cumulative vaccinations which can be called total vaccinations up to that day.

```
## $pred
## Time Series:
## Start = c(2021, 248)
## End = c(2022, 247)
## Frequency = 365
##   [1]  601378960  610465263  619551567  628637870  637724173  646810476
##   [7]  655896780  664983083  674069386  683155689  692241993  701328296
##  [13]  710414599  719500902  728587206  737673509  746759812  755846115
##  [19]  764932419  774018722  783105025  792191328  801277632  810363935
##  [25]  819450238  828536541  837622845  846709148  855795451  864881754
##  [31]  873968058  883054361  892140664  901226967  910313271  919399574
##  [37]  928485877  937572180  946658484  955744787  964831090  973917393
##  [43]  983003697  992090000 1001176303 1010262606 1019348910 1028435213
##  [49] 1037521516 1046607819 1055694123 1064780426 1073866729 1082953032
##  [55] 1092039336 1101125639 1110211942 1119298245 1128384549 1137470852
##  [61] 1146557155 1155643458 1164729762 1173816065 1182902368 1191988671
##  [67] 1201074975 1210161278 1219247581 1228333884 1237420188 1246506491
##  [73] 1255592794 1264679097 1273765401 1282851704 1291938007 1301024310
##  [79] 1310110613 1319196917 1328283220 1337369523 1346455826 1355542130
##  [85] 1364628433 1373714736 1382801039 1391887343 1400973646 1410059949
##  [91] 1419146252 1428232556 1437318859 1446405162 1455491465 1464577769
##  [97] 1473664072 1482750375 1491836678 1500922982 1510009285 1519095588
## [103] 1528181891 1537268195 1546354498 1555440801 1564527104 1573613408
## [109] 1582699711 1591786014 1600872317 1609958621 1619044924 1628131227
## [115] 1637217530 1646303834 1655390137 1664476440 1673562743 1682649047
## [121] 1691735350 1700821653 1709907956 1718994260 1728080563 1737166866
## [127] 1746253169 1755339473 1764425776 1773512079 1782598382 1791684686
## [133] 1800770989 1809857292 1818943595 1828029899 1837116202 1846202505
```

```
## [139] 1855288808 1864375112 1873461415 1882547718 1891634021 1900720325
## [145] 1909806628 1918892931 1927979234 1937065538 1946151841 1955238144
## [151] 1964324447 1973410750 1982497054 1991583357 2000669660 2009755963
## [157] 2018842267 2027928570 2037014873 2046101176 2055187480 2064273783
## [163] 2073360086 2082446389 2091532693 2100618996 2109705299 2118791602
## [169] 2127877906 2136964209 2146050512 2155136815 2164223119 2173309422
## [175] 2182395725 2191482028 2200568332 2209654635 2218740938 2227827241
## [181] 2236913545 2245999848 2255086151 2264172454 2273258758 2282345061
## [187] 2291431364 2300517667 2309603971 2318690274 2327776577 2336862880
## [193] 2345949184 2355035487 2364121790 2373208093 2382294397 2391380700
## [199] 2400467003 2409553306 2418639610 2427725913 2436812216 2445898519
## [205] 2454984823 2464071126 2473157429 2482243732 2491330036 2500416339
## [211] 2509502642 2518588945 2527675249 2536761552 2545847855 2554934158
## [217] 2564020462 2573106765 2582193068 2591279371 2600365675 2609451978
## [223] 2618538281 2627624584 2636710888 2645797191 2654883494 2663969797
## [229] 2673056100 2682142404 2691228707 2700315010 2709401313 2718487617
## [235] 2727573920 2736660223 2745746526 2754832830 2763919133 2773005436
## [241] 2782091739 2791178043 2800264346 2809350649 2818436952 2827523256
## [247] 2836609559 2845695862 2854782165 2863868469 2872954772 2882041075
## [253] 2891127378 2900213682 2909299985 2918386288 2927472591 2936558895
## [259] 2945645198 2954731501 2963817804 2972904108 2981990411 2991076714
## [265] 3000163017 3009249321 3018335624 3027421927 3036508230 3045594534
## [271] 3054680837 3063767140 3072853443 3081939747 3091026050 3100112353
## [277] 3109198656 3118284960 3127371263 3136457566 3145543869 3154630173
## [283] 3163716476 3172802779 3181889082 3190975386 3200061689 3209147992
## [289] 3218234295 3227320599 3236406902 3245493205 3254579508 3263665812
## [295] 3272752115 3281838418 3290924721 3300011025 3309097328 3318183631
## [301] 3327269934 3336356238 3345442541 3354528844 3363615147 3372701450
## [307] 3381787754 3390874057 3399960360 3409046663 3418132967 3427219270
## [313] 3436305573 3445391876 3454478180 3463564483 3472650786 3481737089

## [319] 3490823393 3499909696 3508995999 3518082302 3527168606 3536254909
## [325] 3545341212 3554427515 3563513819 3572600122 3581686425 3590772728
## [331] 3599859032 3608945335 3618031638 3627117941 3636204245 3645290548
## [337] 3654376851 3663463154 3672549458 3681635761 3690722064 3699808367
## [343] 3708894671 3717980974 3727067277 3736153580 3745239884 3754326187
## [349] 3763412490 3772498793 3781585097 3790671400 3799757703 3808844006
## [355] 3817930310 3827016613 3836102916 3845189219 3854275523 3863361826
## [361] 3872448129 3881534432 3890620736 3899707039 3908793342
```

According to the data given, using the Arima model, after 221 days from 1st September 2021 which means 10th April 2022, the two doses of vaccination will be completed in India. And 365 days from 1st September 2021 which means 1st September 2022, all the doses including booster doses will be completed.

Forecasts from ARIMA(0,2,1)
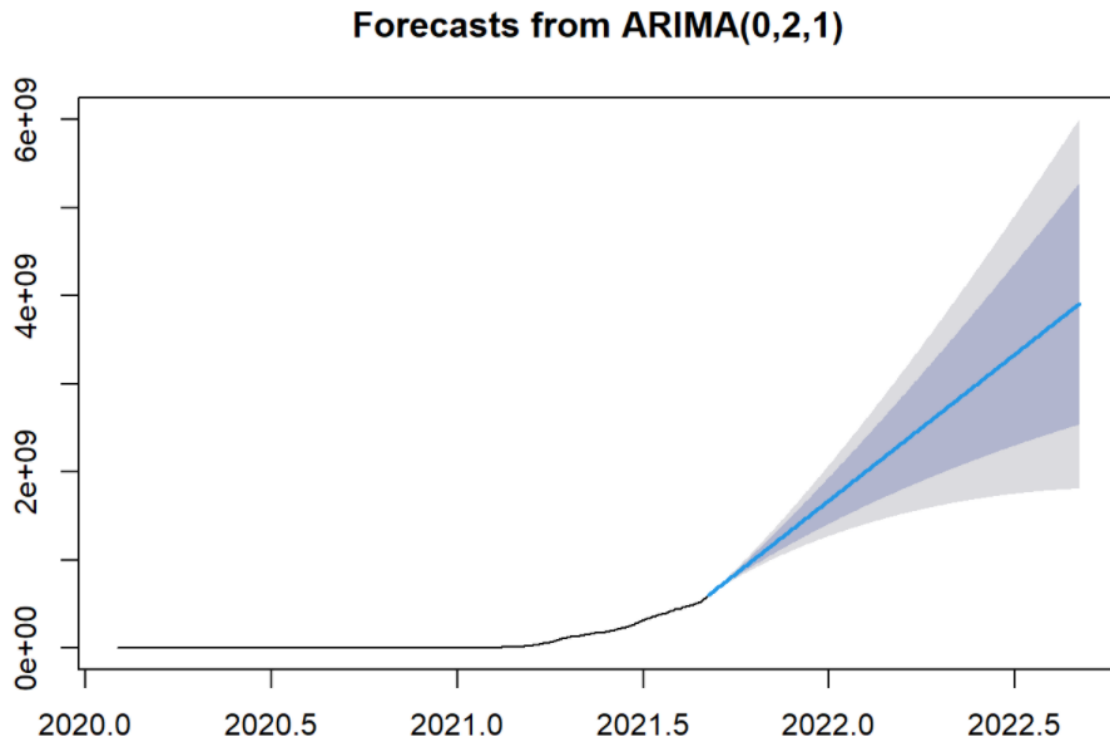
The above graph shows the timeline up to 390 crore vaccinations in India. From the graph, we can observe that the graph is increasing every day and approximately the vaccinations will be completed on 1st September 2022.

Conclusion

From the over-examination utilizing distinctive sorts of charts and expectations using a time series calculation called the Autoregressive Coordinates Moving Normal, it can be concluded that most of the nations within the world are taking great measures to urge out of widespread circumstances. From exploratory investigation ready to gather that as more individuals get immunized, the number of cases has been diminishing. In the nations which had a more prominent populace, the impact of widespread is more when compared to other nations. We anticipated the time at which India will get completely vaccinated. It appears that India will get completely immunized around April of 2022