Studying the Impact of the COVID Vaccination on the World Using Data Analytics

Arun Velu*
Pawan Whig**

Vivekananda Journal of Research
January - June 2021, Vol. 10, Issue 1, 147-160
ISSN 2319-8702(Print)
ISSN 2456-7574(Online)
Peer Reviewed Refereed Journal

http://www.vips.edu/vjr.php

Abstract

The new World Economic Forum-Ipsos vaccine trust study reveals that clear intent to have a vaccine for COVID-19 has improved in the United Kingdom and the United States, two countries where vaccines have begun to be delivered. In most other countries surveyed, however, vaccine trust has fallen. The primary reason people claim they won't get a vaccine for COVID-19 is that they fear side effects. Because of the short development time and the new technology, these vaccines will be used with several unresolved problems that can be clarified only by the passage of time. Technical issues related to the manufacture of billions of doses, as well as ethical issues related to the availability of these vaccines in even the poorest countries, are looming challenges. In order to ensure equal global access, safety of diverse subjects, and immunity against viral variants, we believe that more than one vaccine will be needed in the long run. In this research study, the impact of COVID vaccination on the world is studied using machine learning, and the trends found to be very interesting and presented in the Results section.

Keywords: *World Economic Forum ,COVID , vaccination*

INTRODUCTION

As the flare-up of the COVID-19 has become an overall pandemic, the ongoing investigations of epidemiological information are expected to set up the general public with better activity plans against the infection. Since the discovery of novel COVID-19 [1], the

* Director Equifax, Researcher , Atlanta, USA
Email:- arun.velu@equifax.com
** Dean Research, VIPS, New Delhi , India

world is fretfully battling with its spread. As of April 15, 2021, in light of the internationally shared live information by the Johns Hopkins dashboard, overall there are 138,976,244 affirmed cases, out of which **114,490,676** are recovered and 2,988,801 lost their lives [2]. Coronavirus has a place with the group of the SARS-CoV and MERS-CoV, where it starts with the underlying level indications of the common cold to serious degree of respiratory illnesses causing trouble in breathing, sleepiness, fever, and dry hack [3]. Prasad et al. [4] saw that the identification of the infection can be improved by imaging utilizing immunoelectron microscopy strategies [5]. Till date, itemized morphology and ultrastructure of this infection remain not completely comprehended, and there are no explicit antibodies or medicines for COVID-19. Nonetheless, numerous continuous clinical preliminaries are assessing likely medicines.

Man-made reasoning (AI) can help us in dealing with the issues that should be tended to be raised by the COVID-19 pandemic. It isn't just the advancement, in any case, that will influence yet rather the data and imaginativeness of the individuals who use it. In actuality, the COVID-19 crisis will most likely reveal a part of the critical deficiencies of AI. AI (ML), the current sort of AI, works by perceiving plans in chronicled preparing data. Individuals have a favored situation over AI. We can take in practices from one circumstance and apply them to novel conditions, drawing on our dynamic data to make the best hypotheses on what may work or what may happen. PC based knowledge systems, then again, need to pick up without any planning at whatever point the setting or task changes even insignificantly.

A wide family of distinct viruses comprise coronaviruses. Any of them trigger the common cold in populace. Others, including bats, camels, and horses, infect wildlife. But how did SARS-CoV-2, the new COVID-19-causing coronavirus, come into being? SARS-CoV-2 emerged in bats, researchers say. That's also how the Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS) coronaviruses began [6]. At one of Wuhan's open-air "wet markets," SARS-CoV-2 made the transition to humans. It's where consumers buy fresh meat and fish, even animals that are slaughtered on the spot. Any wet market sells wild or banned animals that may include cobras, wild boars, and raccoon dogs, [7]. Crowded environments can cause genes to be swapped by viruses from various species [8]. The virus often changes so much that it may begin to affect and propagate among individuals [9]. It affected people who had no close contact with wildlife, as SARS-CoV-2 spread both within and outside China. This meant that the virus was spread from one person to the next. In the U.S. and across the world, it is already spreading, suggesting people are inadvertently receiving and sharing the coronavirus. This rising dissemination worldwide is what is becoming a pandemic [10].

In 1965, scientists first observed a human coronavirus [11]. This produced a widespread cold. Researchers identified a group of related human and animal viruses later that decade and named them for their crown-like shape. The origins of how it can be transmitted is shown in Fig. 1

They will infect humans with seven coronaviruses [12]. The one that causes SARS established in 2002 in southern China and spread rapidly to 28 other countries. By July 2003, more than 8,000 people were sick, and 774 died. In 2004, a minor epidemic only affected four additional cases [13]. This coronavirus causes problems with fever, headache, and coughing, including cough and shortness of breath.

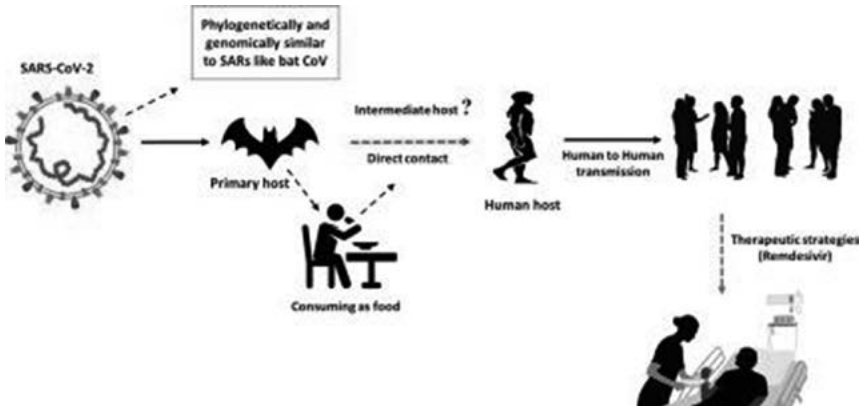


Figure 1 Proposed Origin of Corona Virus

As of April 2021 the total number of COVID cases were 138,976,244 with confirmed deaths of 2,988,801 with total 223 countries are infected as reported by WHO. The stats of COVID is shown in the Fig.2.

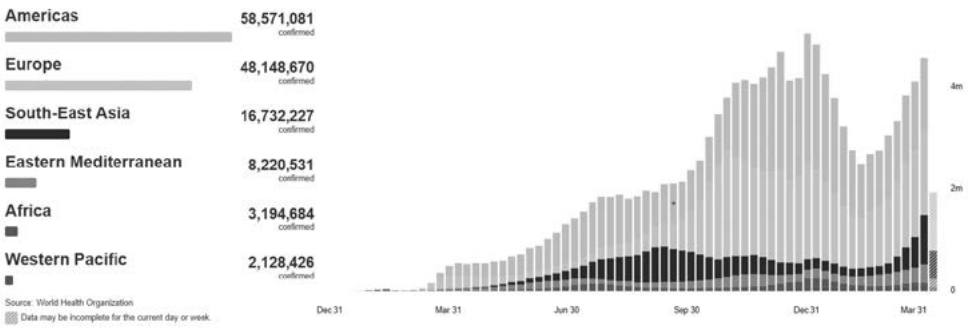


Figure 2 Stats of COVID till april 2021

DATA ANALYSIS USING MACHINE LEARNING

Machine learning is a data processing tool that automates the creation of analytical models. It belongs to the category of artificial intelligence focused on the concept that, with very little human interaction, computers can learn using data, recognize trends and make decisions. Machine learning today is not like machine learning in the past, thanks to emerging computing developments [14]. It was born from the identification of patterns and the idea that computers would learn to do complex tasks without being programmed; academics interested in artificial intelligence decided to see how computers could learn from results [15]. The iterative nature of machine learning is important because they are able to adapt independently when models are introduced to new data [16-20]. To generate accurate, repeatable decisions and outcomes, they learn from previous computations. It is a science that is not new, but has obtained new traction. In this research study we used various machine learning libraries like Pandas, NumPy, etc. for obtaining the impact of COVID vaccination on the world. The block diagram of the process involved is shown in Fig. 3.

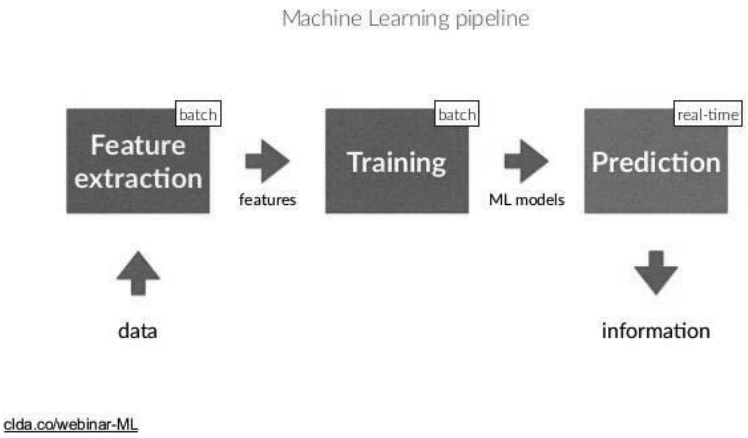


Figure 3 Block diagram of process involved

RESULT AND PROCESS-STEP WISE

For the data analysis Google Colab is used, “Colab” is a Google Analysis product, for short. Colab enables anybody through the browser to write and execute arbitrary python code, and is specifically well suited to computer learning, data processing and education. The detailed analysis is presented stepwise to understand the process in easy ways.

Step 1: Importing Packages

```
import pandas as pd
importnumpy as np
importplotly.express as px
importplotly.graph_objects as go
fromplotly.subplots import make_subplots
```

Step 2: Importing Dataset

```
data=pd.read_csv('country_vaccinations.csv')
data2=pd.read_csv('countries-aggregated.csv')
Step 3: Cleaning Dataset
data.dropna(subset=['daily_vaccinations'],inplace=True)
s=data['date'].str.split('-',expand=True)
data['Year']=s[0]
data['Month']=s[1]
data['Date']=s[2]
fig1=px.scatter_geo(data,color='vaccines',locationmode="ISO-3",locations="iso_
code",opacity=0.6,
hover_name="iso_code", size="daily_vaccinations",projection='conic equal
area',animation_group ="iso_code",color_continuous_scale='blackbody',
animation_frame="Date",scope='world',symbol='vaccines',template="plotly_
dark",title='Vaccination Count ' )
fig1.layout.updatemenus[0].buttons[0].args[1]["frame"]["duration"] = 400

fig1.update_geos(
landcolor="red",
oceancolor="#006994",
showocean=True,
lakecolor="Blue"
)

fig1.update_traces(
marker_coloraxis=None
)
fig1.show()
```

Number of vaccinations available in the world; the result obtained is shown in Fig 4

Figure 4 Various Vaccination count available round the world

It is easily seen that the influence of Pfizer/Biotech in the USA is immense and it tends to grow with time. We can also see CNBG/strong Sinovac's influence beginning in China. Sinovac takes a sudden roll on the 14th in Turkey and starts to develop afterwards.

Step 4: Analysis of most distributed vaccine

```
s=data.drop_duplicates(subset=['iso_code'])['vaccines'].apply(lambda x: x.split(','))
dic={}
fori in s:
for j in i :
if j[0]==' ':
        k= j[1:]
elifj[-1]==' ':
        k=j[:-1]
else:
        k=j
if k not in dic :
dic[k]=1
else:
dic[k]+=1
px.bar(x=list(dic.keys()),y=list(dic.values()),color=list(dic.keys()),template='plotly_
white',labels={'x':'Vaccine Name','y':'Total Count'})
```

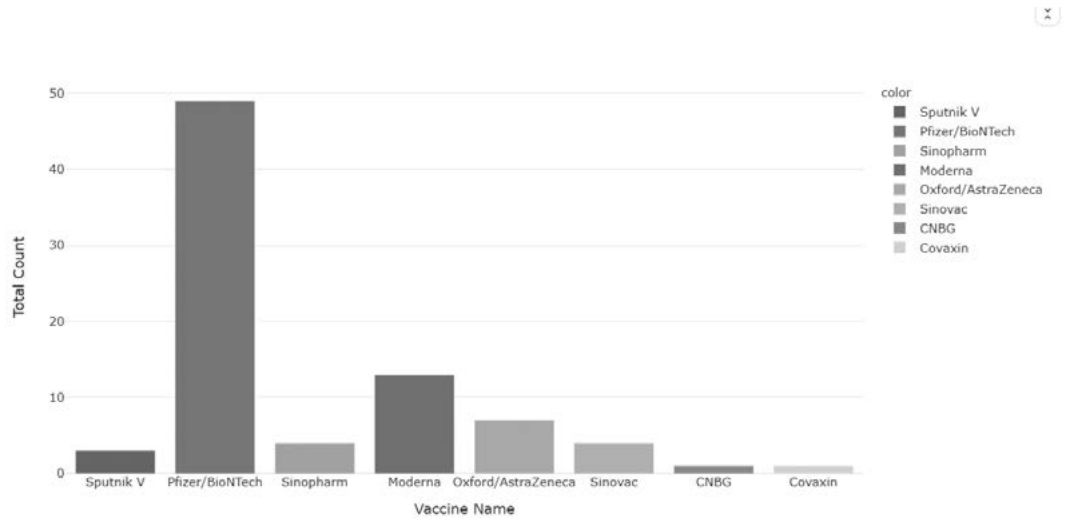


Figure 5 Result obtain in Step 4

It is clearly observed that Pfizer in January 2021, followed by Moderna and Sinovac, became the most common vaccine in the world.

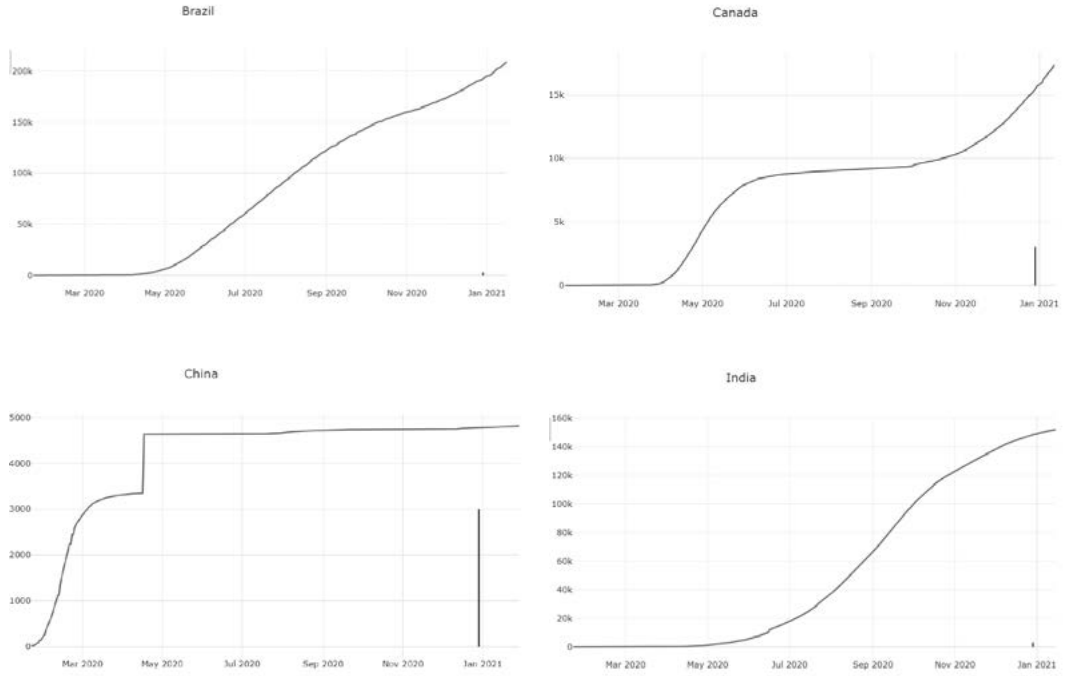
Step 5: Observe the deathrate after vaccine

```
arr=[]
index1=data.groupby([‘country’,’date’]).count().index
index2=data2.groupby([‘Country’,’Date’]).count().index
fori in index2:
ifi in index1:
arr.append(1)
else:
arr.append(0)
data2[‘Vaccine_is_there’]=arr
data2.head()
```

Out[14]:

	Confirmed	Recovered	Deaths	Vaccine_is_there
count	7.180800e+04	7.180800e+04	71808.000000	71808.000000
mean	1.440877e+05	8.692296e+04	3921.234305	0.020221
std	8.780563e+05	5.053113e+05	19220.500408	0.140755
min	0.000000e+00	0.000000e+00	0.000000	0.000000
25%	6.800000e+01	1.600000e+01	0.000000	0.000000
50%	2.645500e+03	1.290000e+03	53.000000	0.000000
75%	3.695150e+04	1.803075e+04	638.000000	0.000000
max	2.592928e+07	1.040915e+07	436678.000000	1.000000

Figure 6 Result obtain in step 5



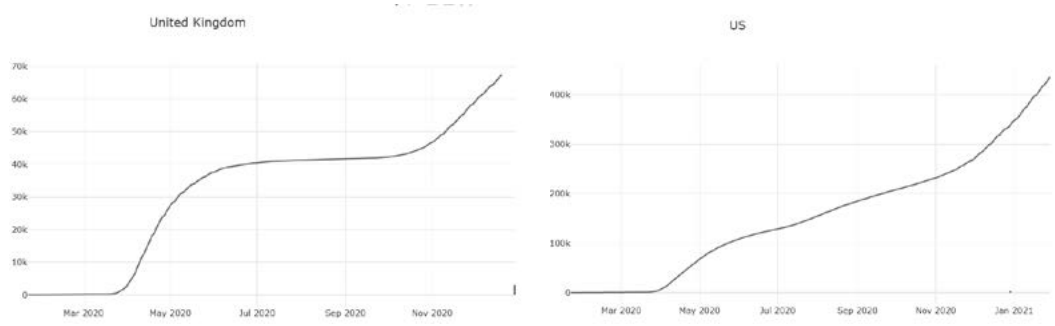


Figure 7 Results of different countries showing effect on deathrate after vaccine

A line is indicated in the figure to show when the vaccine was introduced in the country. Country names are indicated on the figure for better understanding of trends.

A variety of organizations are responsible for all of the vaccines delivered globally and many are currently being created while some are being phased out and some are waiting for clearance. Many nations, such as the UK, have chosen to use multiple vaccines to provide the people with the required doses.

However, where are the vaccines used?

It is found that

- The vaccines developed by Oxford/AstraZeneca and Pfizer/BioNTech are the most commonly distributed around the world.
- More vaccines tend to be nearby and certain vaccines appear to protect landless or nearby countries (i.e. Oxford/AstraZeneca, in itself, much of the continent of Africa).
- In the under-developed countries, the Oxford/AstraZeneca vaccine is more common. This is explainable because this vaccination is cost-free, so the countries with a lower GDP will get the vaccine.
- Covaxin vaccine is actually only being used in India, indicatinga potential scenario that it has not yet been licenced for use in other countries.
- China (Sinopharm/Beijing, Sinopharm/Wuhan, Sinovac) and Russia (EpiVacCorona, Sputnik V) are the only countries that do not use vaccines produced outside of their borders.

- According to the result, the United Arab Emirates and Hungary are currently rolling out the most vaccines, each coming from five different suppliers.

COVID-19 vaccines have been in clinical trials since mid-late 2020, and were formally rolled out and in use in December 2020. They have since been adopted in an increasing number of countries.

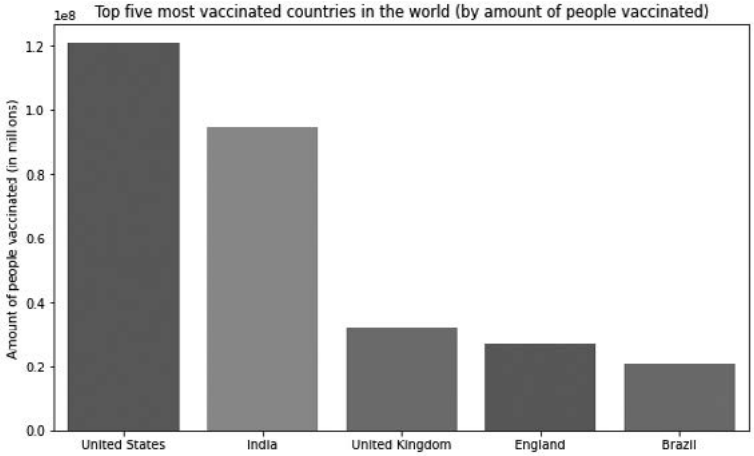


Figure 8 Top Five Most Vaccinated Countries in the world

At this time, several nations are in the middle of their vaccine campaigns. Although several countries are yet to begin vaccine campaigns, others are already on their way to vaccinating their people. What nations, on the other hand, have the highest and most successful vaccine programmes? The answer to the above question can be represented in Fig. 9-11.

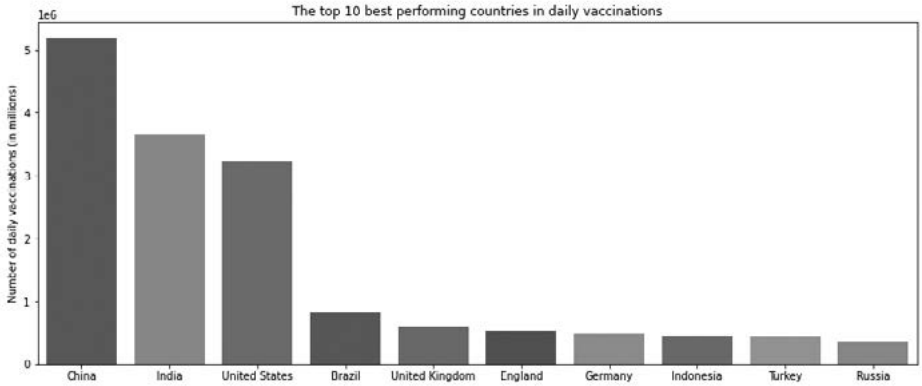


Figure 9 Top ten best performing countries in daily vaccinations

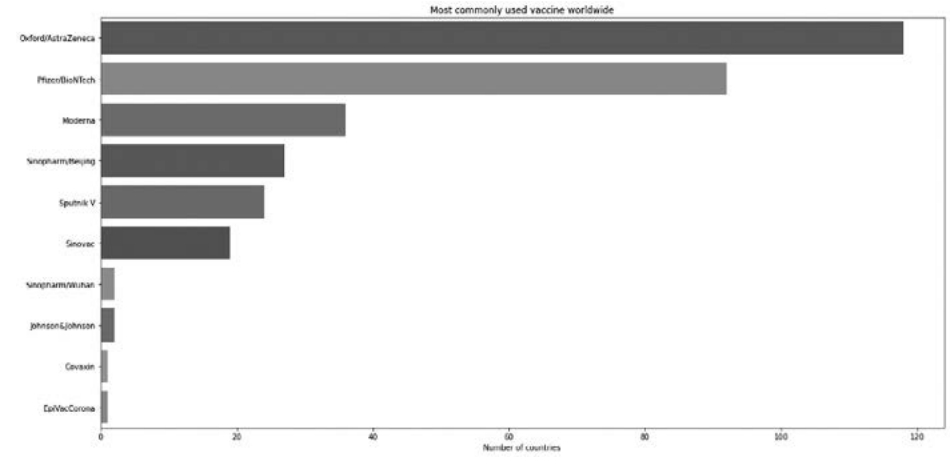


Figure 10 Most commonly used vaccine worldwide

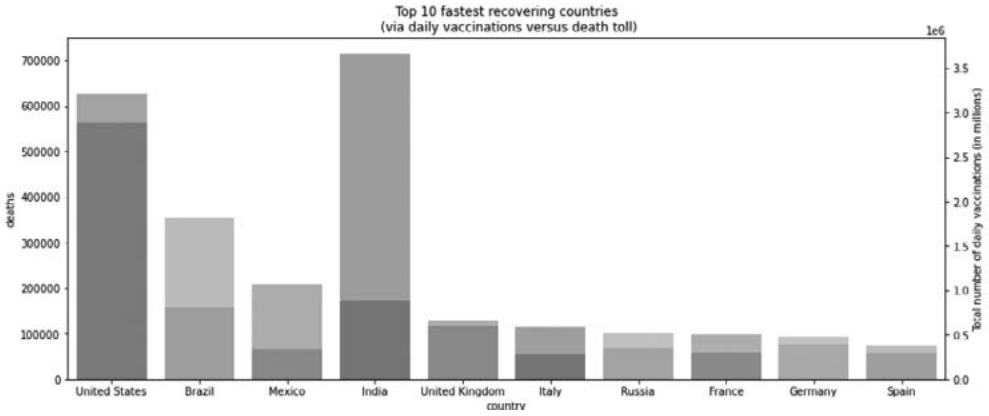


Figure 11 Top ten fastest recovering countries

The graph above depicts a very positive image of how vaccine programmes are improving around the world; not only is every country vaccinating more than their death tolls per day, but certain countries have reached significant margins for those that are vaccinated and their death tolls. India (with a massive gap margin of 173.4 percent *), the United States (138.8 percent *), Brazil (56 percent *), and Mexico (45.2 percent *) happen to be the highest performers in the top ten.

We've seen the vaccination numbers for each country, and they get increasingly greater as the population number grows. But what if we analyse from a different perspective, irrespective of population numbers? Are recovery rates related to the success of a vaccination

program, and is that dependent of what kind of country it is economically (i.e. developed, developing, under-developed)?

The answer to this question can be understood by Fig 12 given below

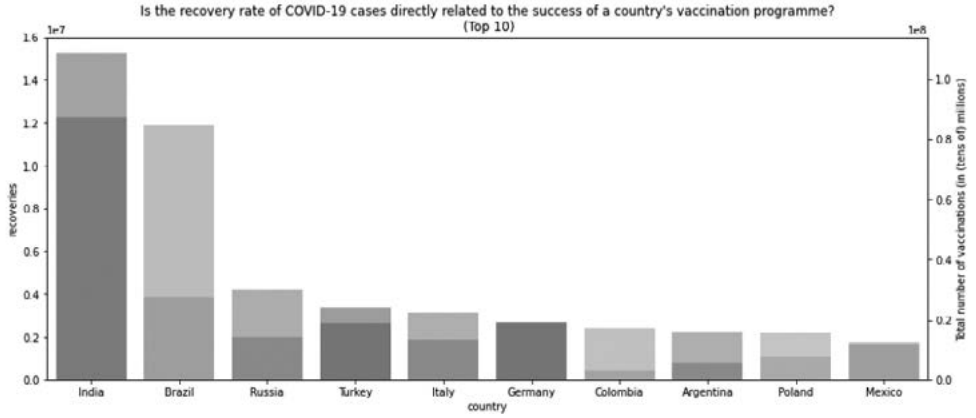


Figure 12 Relationship between recovery rate of COVID-19 with Vaccination Programme

CONCLUSION

In this top ten, we're seeing some newcomers. India, Brazil, Italy, and Mexico all feature on both maps, indicating that recovery rates are affecting the vaccine programme. The most intriguing aspect of this graph is that the wealthier countries (the United States, the United Kingdom, and so on) are not included in the top ten, implying that those countries were the worst hit and had poorer recovery rates (which makes sense, their death rates were higher). The results for Turkey and Germany are also interesting because the margins between recoveries and overall vaccinations are tiny. It could be their populations more compliant with lockout limits, resulting in less infections.

"Vaccination intent is highest in China, according to the survey, where 80 percent of respondents strongly or somewhat agreed with the statement "If a COVID-19 vaccine were available, I would get it. Countries with a reasonably high degree of intent include Brazil (78%), the United Kingdom (77%), Mexico (77%), Australia (75%) and South Korea (75 percent). Among the countries surveyed, South Africa (53 percent), Russia (43 percent) and France were those whose populations recorded the lowest intentions (40 percent). It is clearly observed from the Fig 7 that the death rate is either stagnant or becomes lower in almost all countries but it is still severe in US and UK where a new mutant of COVID 19 has been

found.

FUTURE SCOPE

This research study will be very helpful for the researchers working in the same field to analyse the impact of vaccination and the recovery rates. As more variants are showing up, it is critical to understand the impact of the vaccine efficacy on the variants.

REFERENCE

Singh, R., Singh, R., Bhatia, A. (2018). Sentiment analysis using Machine Learning technique to predict outbreaks and epidemics. Int.J.Adv.Sci.Res,3(2),19-24

Ting, Daniel Shu Wei, Lawrence Carin, Victor Dzau, and Tien Y. Wong. "Digital technology and COVID-19." Nature Medicine (2020): 1-3.

Benvenuto, Domenico, Marta Giovanetti, LazzaroVassallo, Silvia Angeletti, and Massimo Ciccozzi."Application the ARIMA model on the COVID-2019 epidemic dataset." Data in brief (2020):105340.

Deb, Soudeep, and ManidipaMajumdar. "A time series method to analyze incidence pattern and estimate reproduction number of COVID-19." arXiv preprint arXiv:2003.10655(2020).

Kucharski, Adam J., Timothy W. Russell, Charlie Diamond, Yang Liu, John Edmunds, Sebastian Funk, Rosalind M. Eggo et al. "Early dynamics of transmission and control of COVID-19: a mathematical modelling study." The lancet infectious diseases(2020).

Dey, Samrat Kumar, MdMahbubur Rahman, UmmeRaihan Siddiqi, and ArpitaHowlader. "Analyzing the Epidemiological Outbreak of COVID-19: A Visual Exploratory Data Analysis(EDA)Approach. Journal of Medical Virology(2020).

R. Meredith, Andrew S. Azman, Nicholas G. Reich, and Justin Lessler. " The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application." Annals of internal medicine (2020).

Singer, H. M. "Short-term predictions of country-specific COVID-19 infection rates based on power law scaling exponents."arXivpreprintarXiv:2003.11997(2020).

Awad, Mariette, and Rahul Khanna. "Support vector regression." In Efficient Learning Machines, pp. 67-80. Apress, Berkeley, CA,2015.

De Castro, Yohann, Fabrice Gamboa, Didier Henrion, Roxana Hess, and Jean-Bernard

Lasserre.”Approximate optimal designs for multivariate polynomial regression.”The Annals of Statistics 47, no. 1 (2019):127-155.

Greff,Klaus,RupeshK.Srivastava,JanKoutn´ık,BasR.Steunebrink,andJu¨rgen Schmidhuber. ”LSTM: A search space odyssey.” IEEE transactions on neural networks and learning systems 28, no. 10 (2016):2222-2232.

U. Ojha and S. Goel, “A study on prediction of breast cancer recurrence using data mining techniques,” *Proc. 7th Int. Conf. Conflu. 2017 Cloud Comput. Data Sci. Eng.*, pp. 527–530, 2017.

S. Ghosh, S. Mondal, and B. Ghosh, “A comparative study of breast cancer detection based on SVM and MLP BPN classifier,” *1st Int. Conf. Autom. Control. Energy Syst. - 2014, ACES 2014*, pp. 1–4, 2014.

A. Osareh and B. Shadgar, “Machine learning techniques to diagnose breast cancer,” *2010 5th Int. Symp. Heal. Informatics Bioinformatics, HIBIT 2010*, pp. 114–120, 2010.

D. Bazazeh and R. Shubair, “Comparative study of machine learning algorithms for breast cancer detection and diagnosis,” *Int. Conf. Electron. Devices, Syst. Appl.*, pp. 2–5, 2017.

M. S. B. M. Azmi and Z. C. Cob, “Breast cancer prediction based on backpropagation algorithm,” *Proceeding, 2010 IEEE Student Conf. Res. Dev. - Eng. Innov. Beyond, SCOReD 2010*, no. SCOReD, pp. 164–168, 2010.

B. M. Gayathri and C. P. Sumathi, “Comparative study of relevance vector machine with various machine learning techniques used for detecting breast cancer,” *2016 IEEE Int. Conf. Comput. Intell. Comput. Res. ICCIC 2016*, pp. 0–4, 2017.

Purva Agarwal and Pawan Whig, “Low Delay Based 4 Bit QSD Adder/Subtraction Number System by Reversible Logic Gate”, 2016 8th International Conference on Computational Intelligence and Communication Networks (CICN), IEEE Xplore: 26 October 2017.

Jacob B. Chacko ; Pawan Whig, “ Low Delay Based Full Adder/Subtractor by MIG and COG Reversible Logic Gate”, 2016 8th International Conference on Computational Intelligence and Communication Networks (CICN), IEEE Xplore: 26 October 2017

Ajay Rupani , Pawan Whig , GajendraSujediya and Piyush Vyas,”A robust technique for image processing based on interfacing of Raspberry-Pi and FPGA using IoT,” International Conference on Computer, Communications and Electronics (Comptelix), IEEE Xplore: 18 August 2017