iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.42438

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

A Study on COVID-19 Scenario Using Python Pandas

Mr. Yogesh Chaurasiya
MCA, Mumbai University Thane, India

Abstract: ***The World Health Organization (WHO) professed the coronavirus outburst a Public Health Emergency of International Concern (PHEIC) on 30th January 2020 and a pandemic on 11th March 2020. First part of the study incorporates presenting global pandemic scenarios and visual analysis of first five foremost countries (USA, Brazil, India, Russia, South Africa) of world affected with the pandemic based on data mining from John Hopkins university and Github Time series data using Python 3.7 and Pandas with other tools. Followed by that an analysis is done to comprehend situations between Singapore and countries with similar demography to understand the critical situations. The analysis started by uploading 3 different data types of all countries in terms of Confirmed, Death and Recovery cases starting from 22nd January, 2020 to 25th July 2020. This helps to summarise the trends of COVID-19 pandemic from initiation to further spread as population and GDP plays a vital role. Thus, we found that the countries with largest population and GDP is more prone to pandemic.***

Keyword: Trend analysis, COVID-19, data mining, global scenario, Indian scenario.

I. INTRODUCTION

The pandemic of Coronavirus (COVID-19) that incubated in province of Wuhan, China has by now been "disseminated" to more than 214 nations in all the continents,

budding in utmost of them by local spreading. Notably and significantly from unlike other epidemics, this one has exponential growth and replication across the globe (Kraemer, M. U., Yang, C. H., Gutierrez, B., Wu, C. H., Klein, B., Pigott, D. M., & Brownstein, J. S. 2020). The US was among the first few countries to report a COVID- 19 infection and at contemporary scenario is the most affected nation and ranking at first position (Atkeson, A. 2020). USA COVID 19 projections indicate that up to 150 million residents and at least 70 million people are likely to be infected. The forecasts were made by experts such as doctor Brian Monahan as reported by Axios. Also there had been increase in domestic violence (Boserup, B., McKenney, M., & Elkbuli, A. 2020). As of June 2020, Brazil has the second-highest number of confirmed COVID-19 cases in the world behind the United States (Dantas, G., Siciliano, B., França, B. B., da Silva, C. M., & Arbilla, G. (2020). Scientists predicted heavy traumatic situations in Brazil and in the worst-case scenario without measures to contain the virus (Bastos, S. B., & Cajueiro, D. O. 2020). They pointed out that a policy of social distancing was one of the most effective measures given the lack of a vaccine.

India is the third nation in row from June 2020 to become victim of pandemic, where the first case was found in the state of Kerala during midst of March 2020 (Tomar, A., & Gupta, N. 2020). Being a nation with large population it's a massive task to enforce proper preventive measures (Chatterjee, K., Chatterjee, K., Kumar, A., & Shankar, S. 2020). Even country is suffering from heavy economic loses. Next comes Russia which holds forth rank on most affected nations (Pramanik, M., Udmale, P., Bisht, P., Chowdhury, K., Szabo, S., & Pal, I. 2020). Apart from dealing with the consequences of a national-level slowdown, Russia will have to formulate its recovery plans amid a worldwide recession (Konarasinghe, K. M. U. B. 2020). While Africa was among the last regions the virus touched, with the first case of COVID-19 reported in Egypt on March 5, 2020 and currently has one of the highest case incidences on the continent (Gilbert, M., Pullano, G., Pinotti, F., Valdano, E., Poletto, C., Boëlle, P. Y., & Gutierrez, B. 2020). South African society is facing heavy challenges in the social, economic, health, environmental, and technological fronts (Wenham, C., Smith, J., & Morgan, R. 2020).

II. OBJECTIVES

1) To understand global scenario of COVID-19 as deadly disease via trend analysis.
2) To analysis the Lethality of Covid-19 and effectiveness of awareness campaigns launched by the government in five most affected states.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

III. RESEARCH METHODOLOGY

1) Mixed research (primary + secondary)
2) Explorative Approach
3) Testing theory through observation of past data (Trend analysis)
4) Literature review for analysis of secondary data

IV. RESEARCH FLOW

The research work flow consists of data extraction, data repository, data pre- processing, qualitative analysis of data, qualitative result of data, perform trend analysis and results & conclusions.

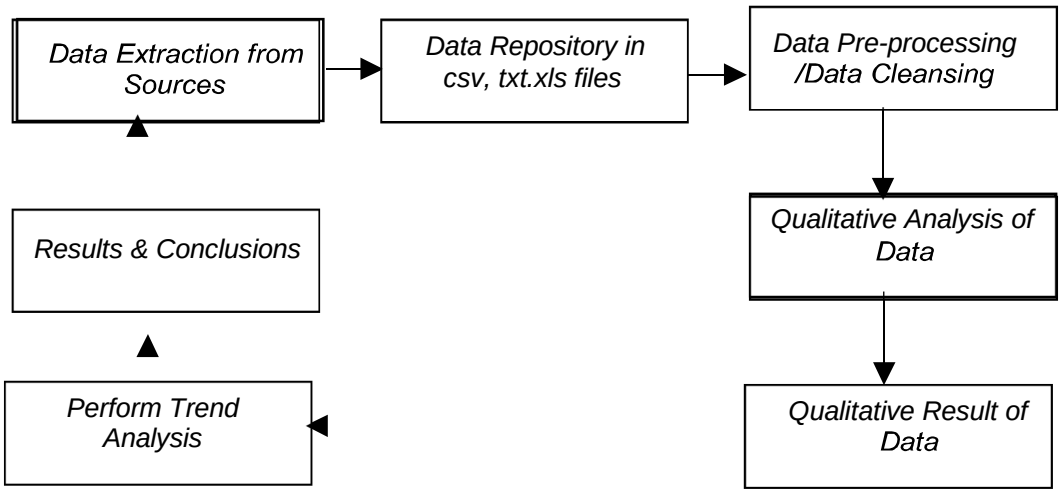


Figure 1 Research Flow Diagram

V. DATA ANALYSIS

The analysis started by uploading 3 different data types of all countries in terms of Confirmed, Death and Recovery cases starting from 22nd January, 2020 to 25th July 2020. This study involves trend analysis using analytical tools to identify critical scenario of pandemic crisis for first five nations of the world. We have used Data sources through data mining from John Hopkins university and Github Time series data. Technologies used are Python 3.7, Pandas, Folium, Numpy, Seaborn, Plotly.figure_factory, Plotly. subplots and Matplotlib. After loading the dataset, the first thing that needs to be checked is the shape of the databases, i.e identifying the number of rows and columns. The next thing that is to be done before doing the data analysis is to find the null values, Null values can degrade the accuracy and could give a sudden break to the flow of graph. This was done by doing a column reset.

```
[ ] confirmed  = df.groupby('Date').sum()['Confirmed'].reset_index()
    recovered  = df.groupby('Date').sum()['Recovered'].reset_index()
    deaths  = df.groupby('Date').sum()['Deaths'].reset_index()
```

```
df.isnull().sum()
```

```
Date            0
Province/State  0
Country         0
Lat             0
Long            0
Confirmed       0
Recovered       0
Deaths          0
Active          0
```

Exhibit 1 – Data analysis to check the null values

©IJRASET: All Rights are Reserved

After dealing with the null values, the overall sum was taken each country in terms of Confirmed cases, Deaths and Recovery. For better visualization, a treemap was plotted using the library offered by pandas which is known as **"plotly.treemap".**

No.	Country	TOTAL CASES	TOTAL DEATHS	TOTAL RECOVERED	TOTAL TEST	TEST/ MILLION POPULATION	TOTAL POPULATION
1	USA	42,49,779	1,48,521	20,28,361	5,25,22,208	1,58,615	33,11,29,317
2	BRAZIL	23,48,200	85,385	15,92,281	49,11,063	23,094	21,26,57,106
3	INDIA	13,51,164	31,542	8,57,976	1,58,49,068	11,478	1,38,08,63,473
4	RUSSIA	8,06,720	13,192	5,97,140	2,66,10,623	1,82,341	14,59,38,709
5	SOUTH AFRICA	4,21,996	6,343	2,45,771	26,84,488	45,228	5,93,55,008

Table 1 – Status of first five affected nations as of 25 July 2020 Source – Worldometer

Table 1 depicts the Recovered cases, Active cases and the Death rate of the first five affected countries. The active cases even show the amount of testing that has been done in developed countries and the recovery rate shows number of people that had mild symptoms of Coronavirus. The death rate was much controlled in few Asian and European countries after a sudden spike between March-April 2020 and followed by that there were devastating reports of increase in number of positive cases and deaths. Though the recovery rates are good but spread rate is much in comparison.

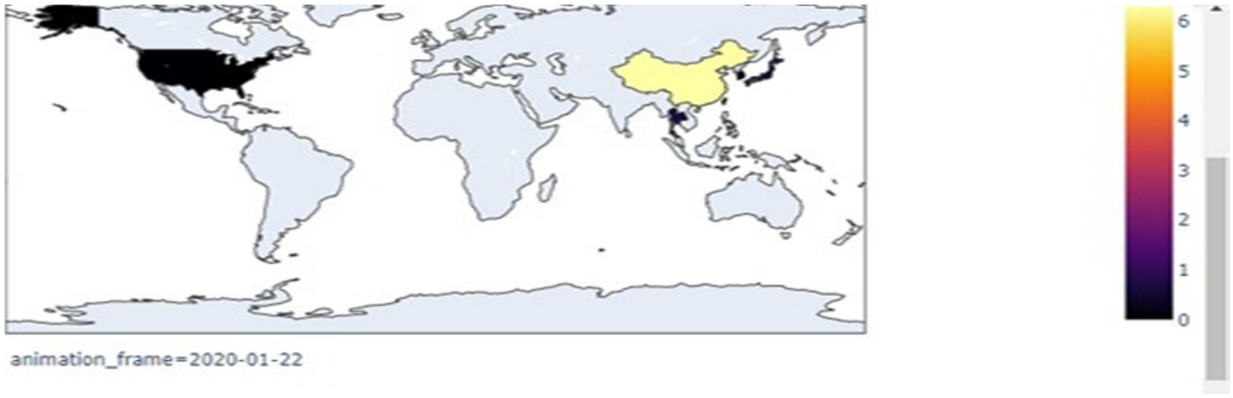


Exhibit 2 –Trend of coronavirus in the month of January 2020

The exhibit 2 depicts the origin of virus started from China, as at that time the confirmed cases and the mortality rate were increasing. Then had its further spread across globe.

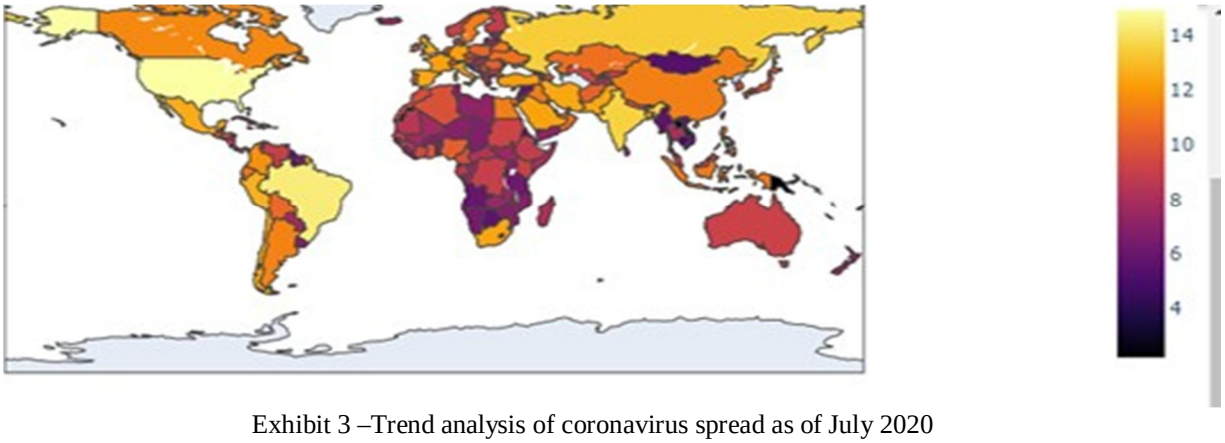


Exhibit 3 –Trend analysis of coronavirus spread as of July 2020

As the exhibit 3 depicts, the top 3 countries i.e US, Brazil and India are badly affected the coronavirus pandemic followed by Russia and South Africa. The USA leads so far in the number of deaths, recovery, total testing and total cases whereas India holds rank third but have the largest population so far demonstrating the better control and measurement so far. On the contrary there exist also a chance to sudden rise due to lack of proper concerns and awareness. Thus, its seen so far that most of these countries of the world are combating for having proper control and prohibition on further spread (Table 1).

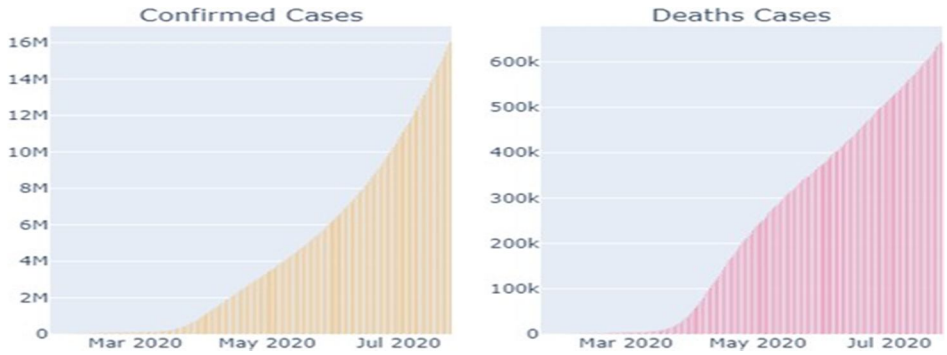


Exhibit 4 – Trend analysis of Confirmed Cases VS Death rate as of July 2020

The trend of Coronavirus started from its epicentre in the Chinese State of Wuhan, as country is densely populated the rate of infection increased by a huge margin. The virus managed to spread outside China as it is one of the most visited countries by the tourists, especially the people coming from EU. Confirmed cases rate is growing at a steady pace whereas as Death rate is growing at the steady pace as well.

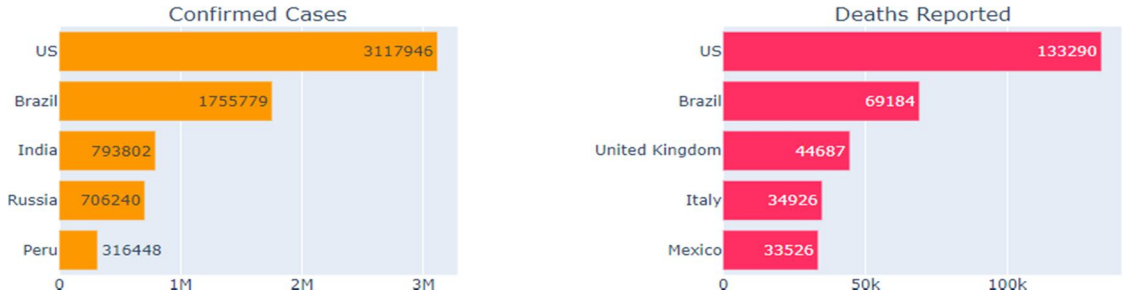


Exhibit 5 – Trend analysis of top five affected nations as of July 2020

The exhibit 5 depicts the top 5 country badly affected from COVID-19, these data are the sum of all cases of each country starting from Jan 22nd to July 25th. Using the tools of python 3.7 the countries were arranged in descending order. Following this data, we can even come up with a conclusion that countries with high population and higher Gross domestic Products (GDP) are more likely to have confirmed cases than other small countries. Thus, population and GDP tend to be the major factors of spread of the corona virus. As its understood that the nation with highest population and GDP would strive harder to move the sluggish economy amid COVID-19 pandemic. Thus, increasing the potential risk to further spread and chances of danger are more due to diversified climatic conditions, standard of living and literacy rate. Government of each nation tries the best to battle the pandemic but still the masses needs to understand the follow the awareness measurements of the government.

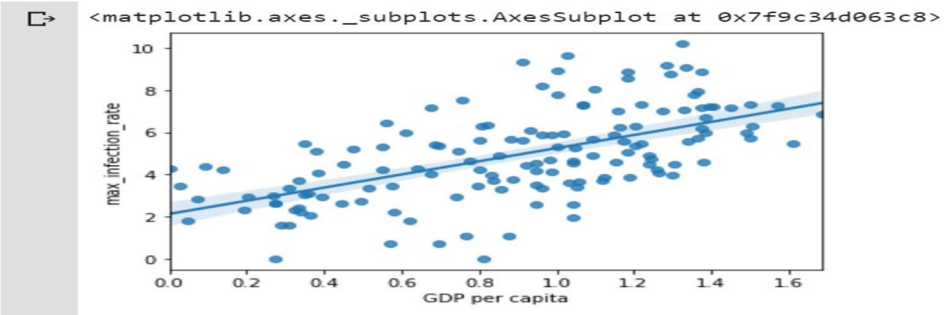


Exhibit 6 – Scatter plot showing relationship between maximum infection rate and GDP

The exhibit 6 depicts scatter plot shows the relation between max infection rate and GDP of a country. As observed the regression line on the above figure shows a positive trend which proves our hypothesis that rate of infection increases more in countries with high GDP.

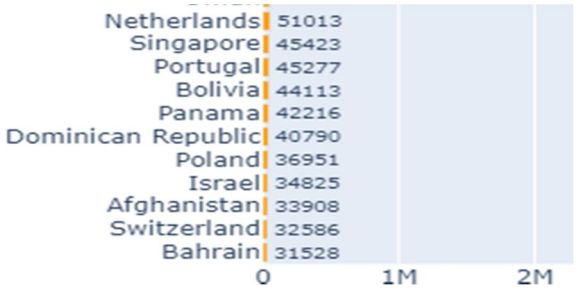


Exhibit 6 – Trend analysis of Singapore and other small countries

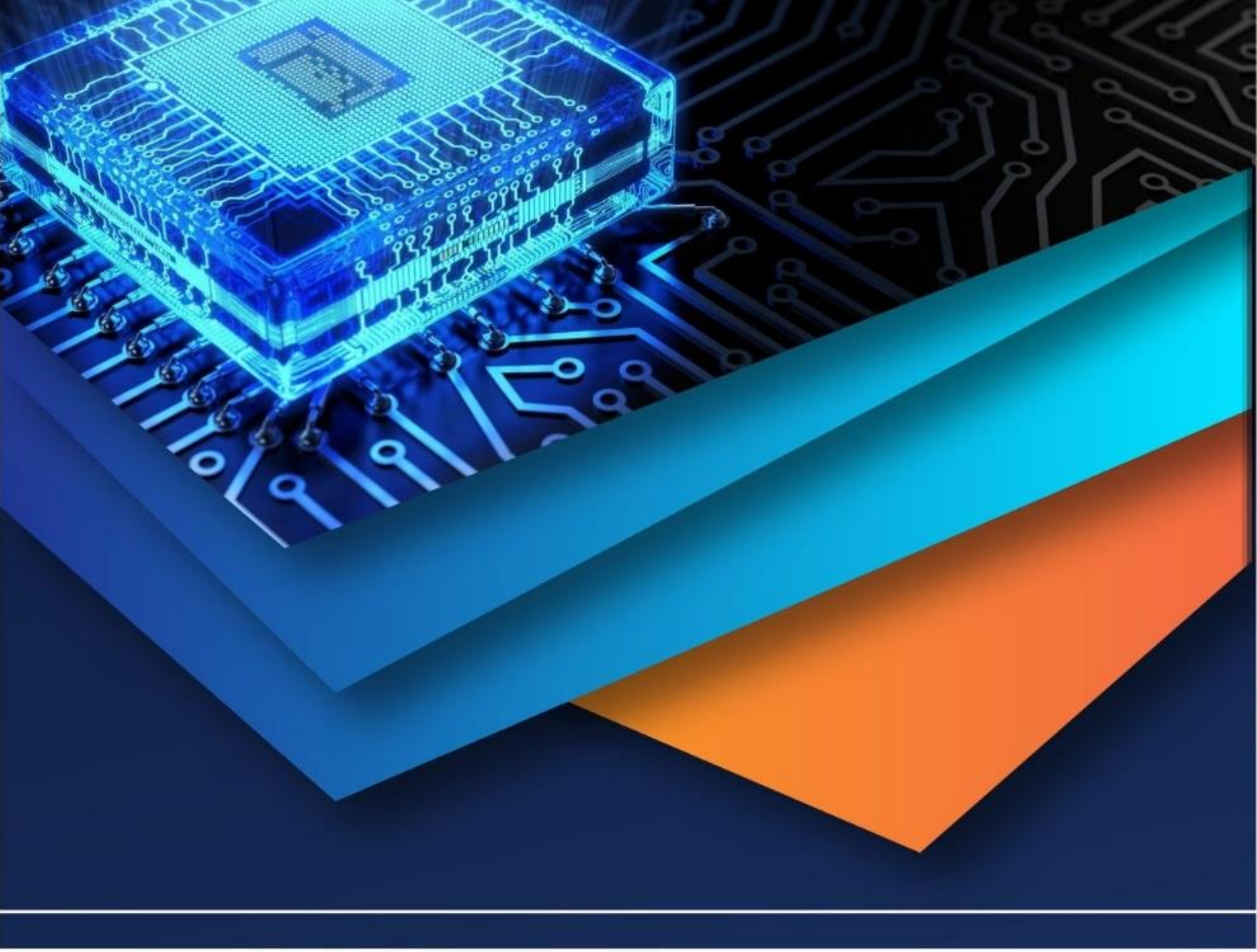
The geographic heterogeneity and associated intervention strategies highlight the need to monitor in real time the transmission potential of COVID-19. Singapore provides a unique case example for monitoring transmission, as there have been multiple disease clusters, yet transmission remains relatively continued. Singapore is a small country with a population of 56.4 lakhs, the government of Singapore have already increased the testing rate and encouraged people to follow social distancing. Singapore General Hospital (SGH) confirmed Singapore's first case of COVID-19 on 23 January 2020 as per Ministry of Health Singapore. Local transmission was subsequently reported and by 7 February 2020, Singapore raised the Disease Outbreak Response System Condition (DORSCON) alert level to orange, indicating that the disease is severe and spreads easily from person to person (but is not spread widely and is being contained)Singapore has always been one of the most visited places in Southeast Asia, that explains the rate of increase in confirmed cases of COVID 19.

VI. CONCLUSION

This study helps to understand that there exists a correlation between population and GDP while measuring the growth in the spread of corona virus in the five major affected nations taken in the study. Also, it is found that Singapore is worst affected nation when compared with counties with similar strengths. The coronavirus disease 2019 (COVID-19) outbreak has been designated a public health emergency of international concern. To prepare for a pandemic, hospitals need a strategy to manage their space, staff, and supplies so that optimum care is provided to patients. In addition, infection prevention measures need to be implemented to reduce in- hospital transmission. This trend analysis suggest that each nation must take valid steps to combat spread of disease and also presents a perilous scenario in front of mankind.

REFERENCES

[1] Atkeson, A. (2020). What will be the economic impact of covid-19 in the us? rough estimates of disease scenarios (No. w26867). National Bureau of Economic Research.
[2] Boserup, B., McKenney, M., & Elkbuli, A. (2020). Alarming trends in US domestic violence during the COVID-19 pandemic. The American Journal of Emergency Medicine.
[3] Bastos, S. B., & Cajueiro, D. O. (2020). Modeling and forecasting the Covid-19 pandemic in Brazil. arXiv preprintarXiv:2003.14288.
[4] Chatterjee, K., Chatterjee, K., Kumar, A., & Shankar, S. (2020). Healthcare impact of COVID-19 epidemic in India: A stochastic mathematical model. Medical Journal Armed Forces India.
[5] Dantas, G., Siciliano, B., França, B. B., da Silva, C. M., & Arbilla, G. (2020). The impact of COVID-19 partial lockdown on the air quality of the city of Rio de Janeiro, Brazil. Science of the Total Environment, 729, 139085.
[6] Gilbert, M., Pullano, G., Pinotti, F., Valdano, E., Poletto, C., Boëlle, P. Y., ... & Gutierrez, B. (2020). Preparedness and vulnerability of African countries against importations of COVID-19: a modelling study. The Lancet, 395(10227), 871-877.
[7] Kraemer, M. U., Yang, C. H., Gutierrez, B., Wu, C. H., Klein, B., Pigott, D. M., ... & Brownstein, J. S. (2020). The effect of human mobility and control measures on the COVID-19 epidemic in China. Science, 368(6490), 493-497.
[8] Konarasinghe, K. M. U. B. (2020). Modeling COVID-19 Epidemic of USA, UK and Russia. Journal of New Frontiers in Healthcare and Biological Sciences, 1(1), 1-14.
[9] Pramanik, M., Udmale, P., Bisht, P., Chowdhury, K., Szabo, S., & Pal, I. 2020. Climatic factors influence the spread of COVID-19 in Russia. International journal of environmental health research, 1-16.
[10] Tomar, A., & Gupta, N. (2020). Prediction for the spread of COVID-19 in India and effectiveness of preventive measures. Science of The Total Environment, 138762.
[11] Weber, A., Ianelli, F., & Goncalves, S. (2020). Trend analysis of the COVID-19 pandemic in China and the rest of the world. arXiv preprint arXiv:2003.09032.
[12] Wenham, C., Smith, J., & Morgan, R. (2020). COVID-19: the gendered impacts of the outbreak. The Lancet, 395(10227),846-848.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)