Project Report on

"COVID - 19 ANALYSIS"

Submitted as Major Project as the award of the degree

Of

Master's of Computer Application
SILIGURI INSTITUTE OF TECHNOLOGY
DEPARTMENT OF COMPUTER APPLICATION



SILIGURI INSTITUTE OF TECHNOLOGY

Affiliated to Maulana Abul Kalam Azad University of Technology Salbari, District Darjeeling, Sukna, West Bengal - 734009

ACKNOWLEDGEMENT
<u>MCKNOWLEDGEMENT</u>
Success of any project depends largely on the encouragement and guidelines of many others. I take this sincere opportunity to express my gratitude to the people who have been instrumental in the successful completion of this project work.
I would like to show our greatest appreciation to Prof. Tumpa Banerjee. I always feel motivated and encouraged every time by her valuable advice and constant inspiration, without her encouragement and guidance this project would not have materialized.
2 Page

CERTIFICATE OF ORIGINALITY This is to certify that the project report entitled COVID-19 ANALYSIS submitted to Department of Computer Application, Siliguri Institute of Technology, affiliated to Maulana Abul Azad University of Technology (MAKAUT) for the Award of Master's of Computer Application (MCA), is an original work carried out by Mr./Ms. Shikha Kumari RollNo: 33671020017 and Mr./Ms. Subhajit Sarkar RollNo:: 33671020019 under the guidance of Prof. Tumpa Banerjee.
3 Page

SELF CERTIFICATE This is to certify that this project untitled "Covid 19 Analysis" submitted as Major Project for the award of the degree of MASTER OF COMPUTER APPLICATION is an authentic work which is done by me under the guidance of Mrs. Tumpa Banerjee. The			
matter embodied in this project work has not degree to the best of my knowledge and belief	t been submitted earlier for aw	•	
	Shikha Kumari 33671020017		
		4 Page	

SELF CERTIFICATE			
This is to certify that this project untitled "Control of the degree of MAST authentic work which is done by me under the matter embodied in this project work has not degree to the best of my knowledge and believed."	TER OF COMPUTER APPLICATION is an he guidance of Mrs. Tumpa Banerjee . The ot been submitted earlier for award of any		
	Subhajit Sarkar 33671020019		
	5 D 2 g 2		
	5 Page		

CONTENTS

- > Introduction
- Wave description
- Project category
- > Software requirement specification
- Project Modules
- > Algorithm
- > Dataset
- > World Plot
- > India Plot
- Covid-19 Vaccination
- Vaccination Plot of India
- Conclusion
- > Future scope and further enhancement of the project
- Bibliography

INTRODUCTION

The full name of COVID-2019 is the Coronavirus disease of 2019, which has created panic in the whole world. Novel COVID-2019 has been reported to be the most harmful and dangerous in the world since the 1918 H1N1 influenza epidemic. Based on the report of the World Health Organization, by April 10, 2020, a total of 15,225,252 case reports were filed and a total of 100,075 deaths occurred. Thus, it can be said that COVID-2019 has been spreading very fast since the first December 2020 to till date. Till now COVID-2019 has spread in 172 countries. At present, the highest number of cases has been found in the United States of America (USA). COVID2019 is a terrible contagious disease that results in very rapid movement from one person to another people. The COVID-2019 epidemic is a member of the family of Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2). Thus here it can be said that coronavirus is a contagious disease. The invention of the coronavirus was first discovered in 2002 and 2012 from China and Saudi Arabia respectively. Corona is a family of viruses that is responsible for diseases ranging from cold, cough, respiratory diseases and life-threatening diseases such as Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS). COVID-2019 was first invented by China in mid-December 2020. It was first found in the city of Wuhan, China [2]. According to some media reports, this COVID-2019 was found in China in mid-19 November. Therefore, we can say here that China did not reveal the correct information about this virus to the countries of the world. This was a serious matter. In a study with Jiang and his colleagues, it was found that the fatality rate for COVID-2019 is around 7.5%. These people have also found in their study that the fatality rate for persons in the age group of 70–79 is 8.0% whereas, for those above the age of 80 years, the fatality rate is around 14.8%. This study considers individuals above the age of 50 years with the highest risk of underlying illnesses such as diabetes, Parkinson's disease, and cardiovascular. A person suffering from COVID-2019 starts showing symptoms in 2-14 days. Due to this virus, the patient suffers from diseases like fever, cough, breathlessness, pneumonia, kidney failure, etc.. The coronavirus spreads very rapidly from one human to another by respiratory drops. This virus does not live long in the air. The virus does not spread through the air because it is not alive for long in the air.

WAVE DESCRIPTIONS

Covid-19 First wave

The first cases of COVID-19 in India were reported on 30 January 2020 in three towns of Kerala, among three Indian medical students who had returned from Wuhan, the epicenter of the pandemic. Lockdowns were announced in Kerala on 23 March, and in the rest of the country on 25 March. Infection rates started to drop in September. Daily cases peaked mid-September with over 90,000 cases reported per-day, dropping to below 15,000 in January 2021. A second wave beginning in March 2021 was much more devastating than the first, with shortages of vaccines, hospital beds, oxygen cylinders and other medical supplies in parts of the country. By late April, India led the world in new and active cases. On 30 April 2021, it became the first country to report over 400,000 new cases in a 24-hour period. Experts stated that the virus *may* reach an endemic stage in India rather than completely disappear; in late August 2021, Soumya Swaminathan said India may be in some stage of endemicity where the country learns to live with the virus.

Covid-19 Second wave

Like several parts of the world, especially Europe, India has been experiencing a massive surge of COVID-19 cases and deaths. As of April 10, 2021, India is the 3rd leading country based on the USA and Brazil's identified cases. Since the middle of March 2021, the second wave has started, and on April 09, the highest number of cases (144,829) has been identified in India. The major affected states are Maharashtra, Kerala, Karnataka, Andhra Pradesh, Tamil Nadu, Andhra Pradesh, Delhi, Uttar Pradesh, and West Bengal. Despite this high caseload, several national movements such as the farmers' movement, election in several states have been going on. Those can increase the risk of transmission of COVID-1.

PROJECT CATEGORY

The project category is Data Science based developed in Jupyter notebook and Google colab. Language used is Python.

This project is based on the Data Analytics Method. Here we first take some data in the form of csv file (Coma Separated values). Then by using data analytics method with the help of python language we analysis the data which contain various information about the country wise covid-19. From that csv file we provide a visualization of data in a bar graph and histogram. That any non-technical person can easily understand about the data.

This project provides visualization of the data in graphical manner .Here we use various types of python modules which contain some predefined built-in method. Like Numpy, Matplotlib , pandas, seaborn etc.

SYSTEM REQUIREMENTS:

The minimum requirements for such project are as follow:

HARDWARE:

- Computer or Laptop :: Intel
- Processor :: Quad Core or higher (Dual core is not best for this kind of work, but manageable)
- RAM :: 16GB RAM or higher (8GB RAM is okay but not for the performance we may want or expect)
- CACHE :: 512 MB
- Input Device :: Mouse and Keyboard

SOFTWARE:

The project requires the support of the following software for the database and other requirements:

- Windows 10 or higher operating system
- A high level programming language i.e. Python
- Python modules like NumPy, Matplotlib, Pandas, Seaborn etc.
- Online environment like Jupyter, Google colab

PROJECT MODULES:

Python is one of the most popular programming languages for task and it has replaced many languages in the industry. One of the reason is its vast collection of libraries. Python libraries that are used

- Pandas
- Matplotlib
- NumPy
- Seaborn
- Express
- Folium map
- > Choropleth map

PANDAS:

Pandas are a vast Python library used for the purpose of data analysis and manipulation and also for working with numerical tables or data frames and time series , thus being heavily used in for algorithmic trading using python . Pandas can be used for various functions importing .csv files , performing arithmetic operations in series , Boolean indexing , collecting information about a data frame etc.

Logic Process:

In order to be able to work with the data in Python , we'll need to read the csv file into a Pandas DataFrame . A DataFrame is a way to represent and work with tabular data . Tabular data has rows and columns , just like our csv file . To read in the data , we need to use the pandas.read_csv function .

Import pandas as pd Data = pd.read_csv (" master.csv ")

NUMPY:

NumPy or Numerical Python provides powerful implementations of large multi - dimensional arrays and matrices . The library consists of functions for complex array processing and high - level computations on these arrays . Some of the mathematical functions of this library include trigonometric functions (sin , cos , tan , radians) , hyperbolic functions (sinh , cosh , tanh) , logarithmic functions (log , log10 , log2) etc. Logic Process : It is common to need to reshape a one - dimensional array into a two - dimensional array with one column and multiple arrays . NumPy provides the reshape () function on the NumPy array object that can be used to reshape the data . The reshape () function takes a single argument that specifies the new shape of the array . Then we use Data = data . reshape (data . shape [0] , 1) Structures like lists and NumPy arrays can be sliced . This means that a subsequence of the structure can be indexed and retrieved . Then we use Data [from : to]

MATPLOTLIB:

It is a Python library used for plotting 2D structures like graphs , charts , histogram , scatter plots etc. Along with the other libraries which are used for computations , it becomes necessary to use matplotlib to represent that data in a graphical format usisng charts and graphs . Few of the functions of matplotlib include scatter (for scatter plots) , pie (for pie charts) , stackplot (for stacked area plot) , colorbar (to add a colorbar to the plot) etc. It has a module named pyplot which makes things easy for plotting by providing feature to control line styles , font properties , formatting axes etc.

Logic Process:

Matplotlib.pyplot is a collection of command style functions that make matplotlib work like MATLAB . Each pyplot function makes some changes to a figure : e.g. , creates a figure , creates a plotting area in a figure , plots some lines in a plotting area , decorates the plot with labels etc. In matplotlib.pyplot , various states are preserved across function calls , so that it keeps track of things like the current figure and plotting area and the plotting functions are directed to the current axes . If we provided a single list or array to the plot () command , matplotlib assumes it is a sequence of y values and automatically generates the x values for us.

SEABORN:

Seaborn is one of an amazing library for visualization of the graphical statistical plotting in Python. Seaborn provides many color palettes and defaults beautiful styles to make the creation of many statistical plots in Python more attractive.

Seaborn library aims to make a more attractive visualization of the central part of understanding and exploring data. It is built on the core of the matplotlib library and also provides dataset-oriented APIs.

Seaborn is also closely integrated with the Panda's data structures, and with this, we can easily jump between the different visual representations for a given variable to better understand the provided dataset.

Folium Maps:

Folium is a powerful Python library that helps you create several types of Leaflet maps. By default, Folium creates a map in a separate HTML file. Since Folium results are interactive, this library is very useful for dashboard building. You can also create inline Jupyter maps in Folium.

Folium builds on the data wrangling strengths of the Python ecosystem and the mapping strengths of the Leaflet.js library. Using Folium, you can manipulate your data in Python, then visualize it in a Leaflet map.

Choropleth map:

Plotly is a Python library that is very popular among data scientists to create interactive data visualizations. One of the visualizations available in Plotly is Choropleth Maps. Choropleth maps are used to plot maps with shaded or patterned areas which are proportional to a statistical variable. They are composed of colored polygons. They are used for representing spatial variations of a quantity.

<u>ALGORITHM</u>

- Data Collection
- Data Formatting
- Model Selection

Data Collection:

We have collected data sets related to population from online website. We have downloaded the csv files in which information

Data Formatting:

The collected data is formatted into suitable data sets. We check the co-linearity with mean suicide rate. The data sets which have co-linearity nearer to 1.0 has been selected.

Model Selection:

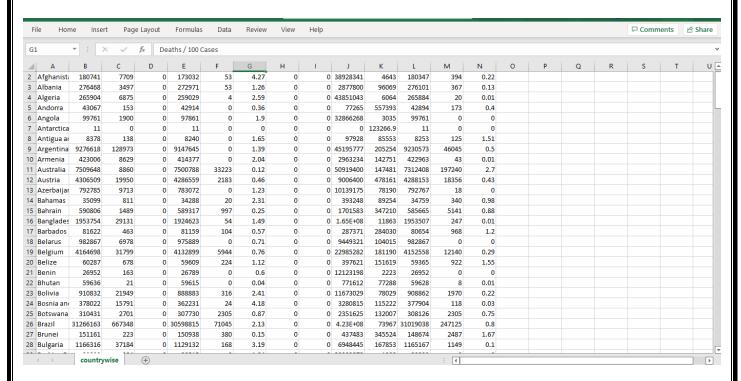
We have selected different models to minimize the error of the predicted value. The different models used are Pandas, NumPy ,Matplotlib , etc.

DATASET

A dataset is a collection of data. Mostly a data set refers to the contents of a single database table, or a single statistical data matrix, where every table column represents a particular variable, and each table row corresponds to a given members of the data set. The objective of project is to find the relations between the dataset to predict the future dataset.

The data set used in this project contains day-wise, country-wise Covid-19 details of all the countries of the world by various parameters from 2020 to 2022. This data is real and original.

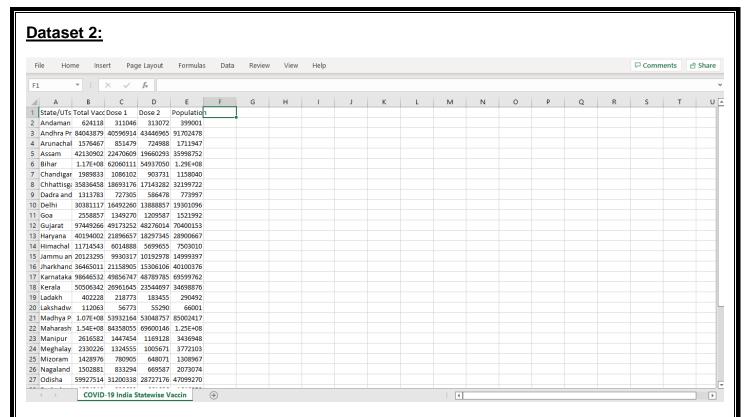
Dataset 1:



This DataSet is taken from Johns Hopkins University Center For Systems Science and Engineering website.

Dataset Features List:

- Country
- Confirmed
- Deaths
- Recovered
- New Cases
- Deaths/100k
- Recovered /100k
- Confirmed/100k



This dataset is taken from https://www.kaggle.com/datasets/anandhuh/covid19-india-statewise-vaccine-data.

DataSet Features List:

- State/UTs
- Total vaccination
- Dose 1
- Dose 2
- Total Population

WORLD PLOTS

Infected

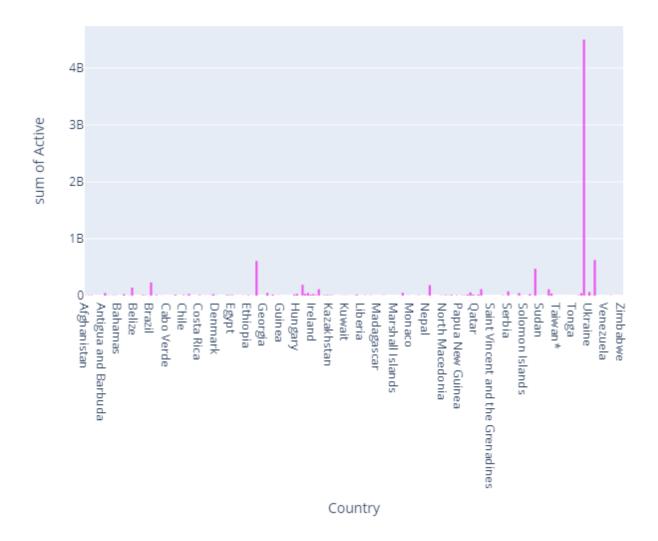


Fig. Histogram of covid-19 First wave Infected Cases (country-wise)

This bar graph shows the world first wave infected cases of covid-19. Here we see that Ukraine has the highest infected rate in first wave. The **COVID-19 pandemic in Ukraine** has resulted in 5,040,518 confirmed cases of COVID-19 and 112,459 deaths.

The virus was confirmed to have spread to Ukraine when the country's first case was confirmed to be hospitalized in Chernivtsi Oblast on 3 March 2020, a man who had travelled from Italy to Romania by plane and then arrived in Ukraine by car.

An emergency was declared on 20 March 2020 in Kyiv Oblast, Chernivtsi Oblast, Zhytomyr Oblast, Dnipropetrovsk Oblast, Ivano-Frankivsk Oblast, and the city of Kyiv.

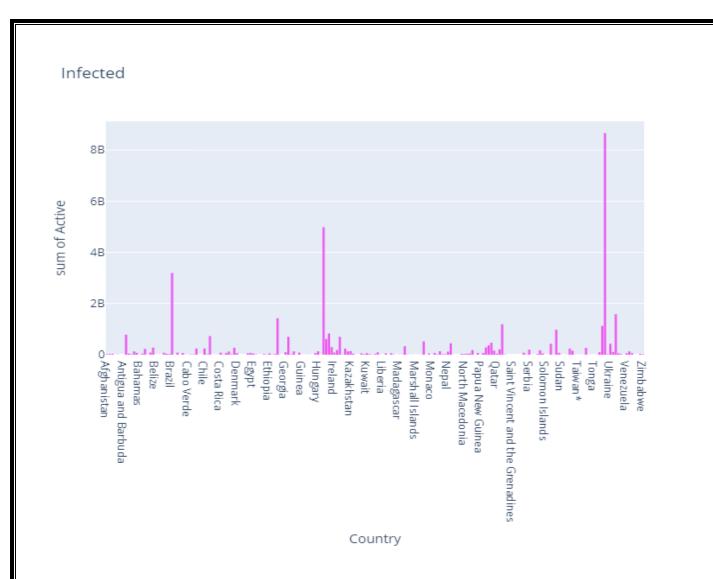


Fig. Histogram for covid-19 Second wave(country-wise)

This bar graph shows the World second wave of covid-19 infected cases. Again we have seen that Ukraine has the highest rate of covid cases followed by Ireland.

New infections and deaths started to break records by late October 2021. By then, a total of 2.8 million coronavirus cases and 64,936 COVID-19 related deaths had occurred in Ukraine.

Ukraine's ongoing vaccination program started on 24 February 2021 and from that day to 12 September 2021, 10,710,944 vaccinations were given in Ukraine (meaning 18% of the adult population of Ukraine had been vaccinated against COVID-19). About 44% of those vaccinated had been fully vaccinated In an August 2021 poll 56% of Ukrainians did not plan to be vaccinated. Demand for vaccinations multiplied sharply by late October 2021

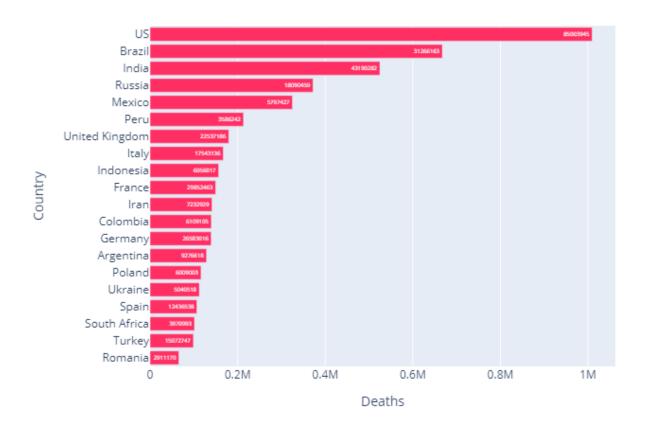


Fig. Bar graph of top 20 countries Death Rate of world total wave during Covid

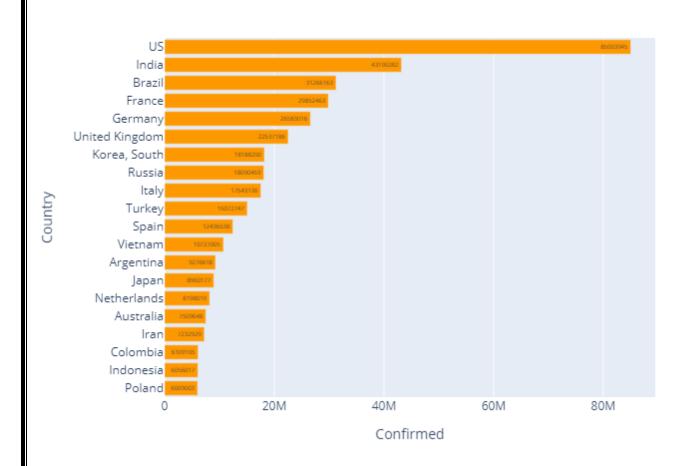


Fig. Bar graph of top 20 countries have confirmed cases of world total wave

According to this bar graph the US has the highest number of death cases and infected cases.

Since the first US case of the new coronavirus was reported in January 2020 in Washington state, more than 85 million people have been diagnosed. In United States of America, from 3 January 2020 to 14 June 2022, there have been 84,593,200 confirmed cases of COVID-19 with 1,001,598 deaths, reported to WHO. As of 2 June 2022, a total of 579,958,258 vaccine doses have been administered.

Mortality

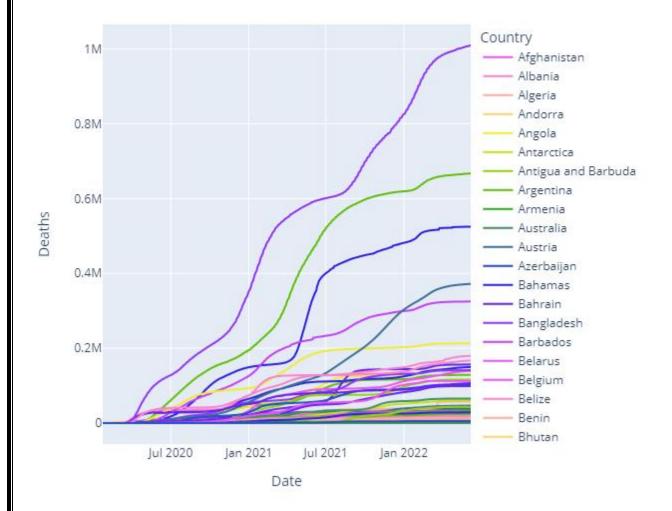


Fig. Line plot of all countries showing the Death rates

The COVID-19 pandemic in Bahrain is part of the worldwide pandemic of coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The virus was confirmed to have reached Bahrain on 21 February 2020.

Bahrain is reporting 1,302 new infections on average each day, 17% of the peak. There have been 597,184 infections and 1,489 coronavirus-related deaths reported in the country since the pandemic began.

Daywise Confirmed vs Recovery vs Deaths

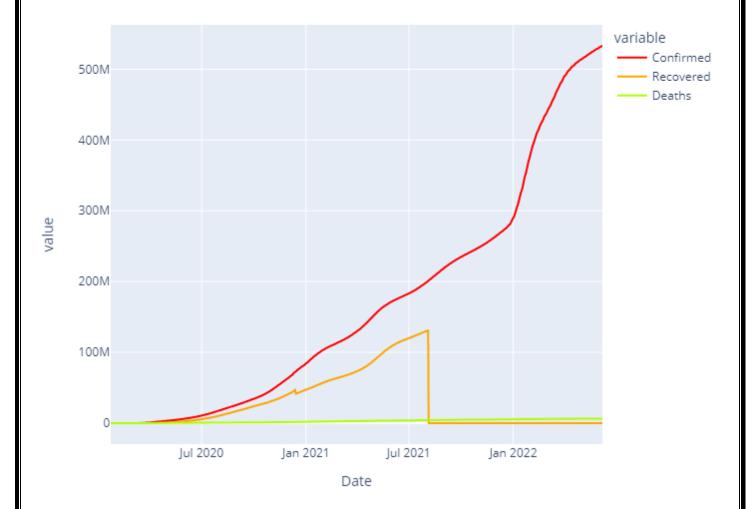


Fig. Histogram plot of World day-wise Confirmed vs Recovered vs Death cases



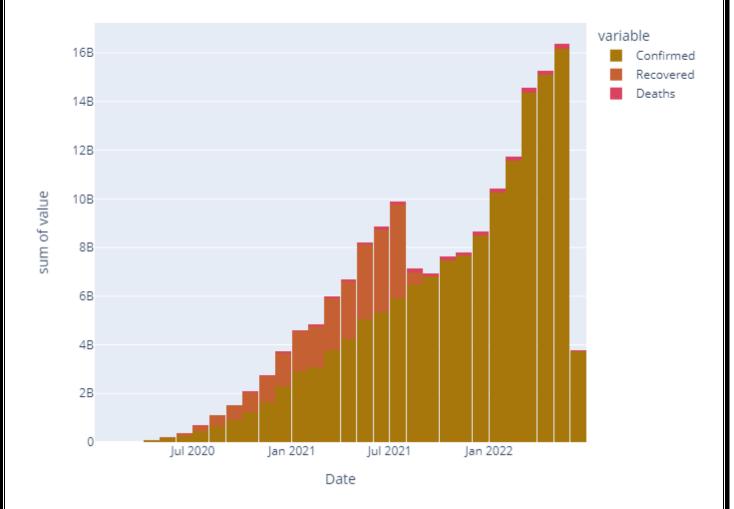


Fig. Histogram plot of World day-wise Confirmed vs Recovered vs Death cases

Cases over time

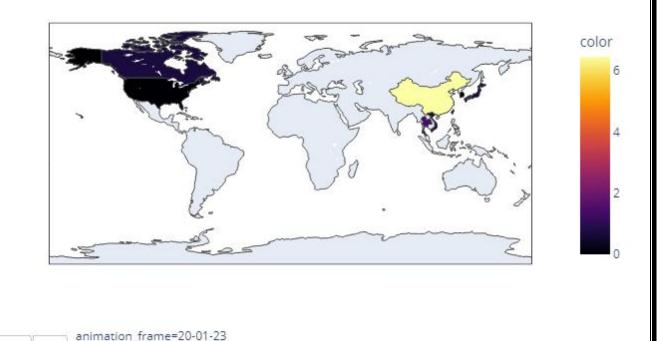


Fig. Choropleth map of Covid cases of World

20-01-23 20-05-09 20-08-24 20-12-09 21-03-26 21-07-11 21-10-26 22-02-10 22-05-28

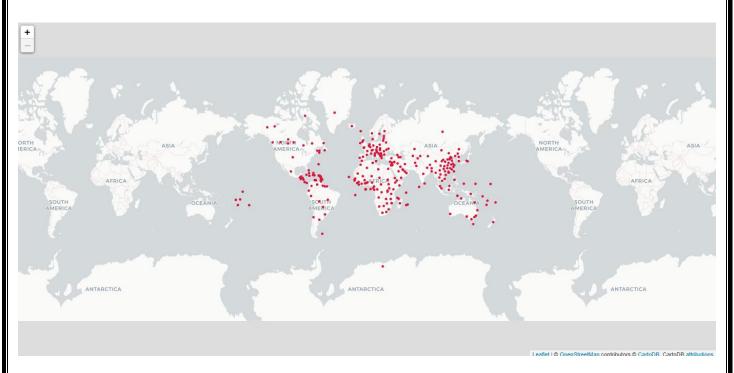


Fig. Folium map of covid 19 World cases

INDIA PLOT

The first case of COVID-19 in India was reported on January 30, 2020; the index patient was a student who had returned from Wuhan. Thereafter, only 2 more cases were reported in February. Subsequently, more cases came to the forefront in the month of March and there has been a surge in the number of cases since the latter half of April 2020. As of June 9, 2020, according to the Ministry of Health and Family Welfare (MoHFW), a total of 266 598 confirmed COVID-19 cases have been reported from 32 states/union territories.² Most of the cases have been reported from the states of Maharashtra, Tamil Nadu, Delhi, and Gujarat. Hitherto, the MoHFW have reported 7471 deaths due to COVID-19, translating into a case-fatality rate of 2.8%.

First Wave of Covid-19 in India

India First Wave Confirmed vs Recovered vs Deaths

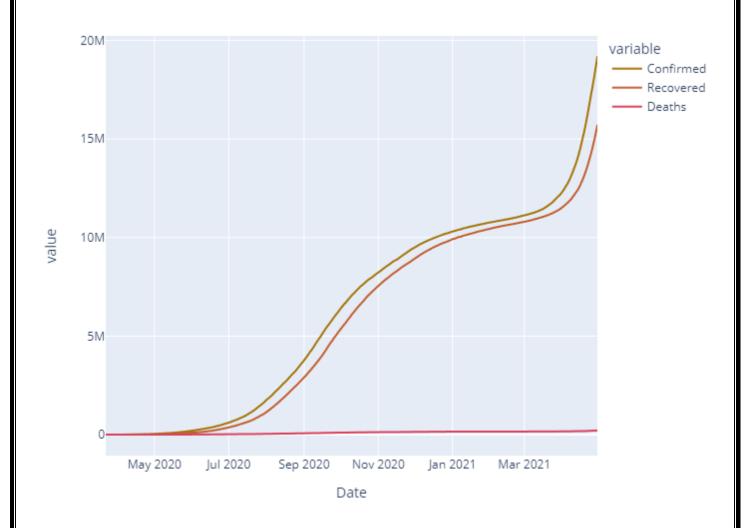


Fig. Line plot of India First Wave of Covid-19

The first cases of COVID-19 in India were reported on 30 January 2020 in three towns of Kerala, among three Indian medical students who had returned from Wuhan, the epicenter of the pandemic. Lockdowns were announced in Kerala on 23 March, and in the rest of the country on 25 March. Infection rates started to drop in September. Daily cases peaked mid-September with over 90,000 cases reported per-day, dropping to below 15,000 in January 2021. A second wave beginning in March 2021 was much more devastating than the first, with shortages of vaccines, hospital beds, oxygen cylinders and other medical supplies in parts of the country. By late April, India led the world in new and active cases. On 30 April 2021, it became the first country to report over 400,000 new cases in a 24-hour period.

Second Wave of Covid-19 in India

India Second Wave Confirmed vs Recovered vs Deaths

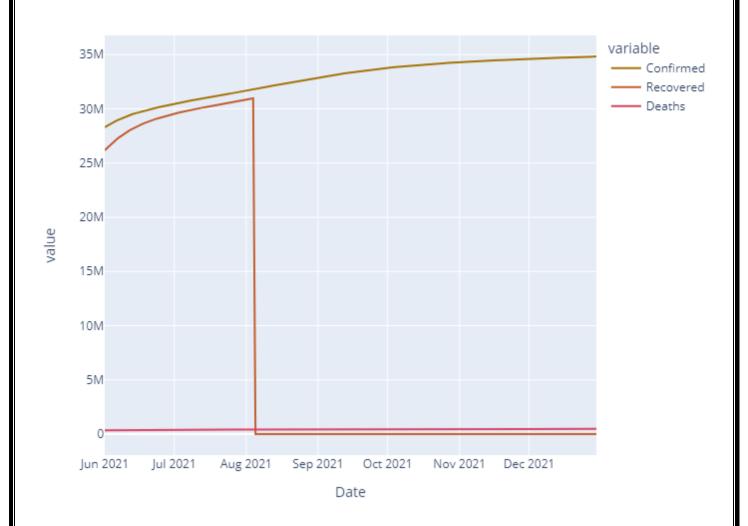


Fig. Line plot of India Second Wave of Covid-19

The wave of corona virus was spreading like a 'tsunami' in India. As on April 23, 2021, the COVID-19 cases have crossed the 15.9 million, with 1,85,000 deaths. The 2nd wave is evolving at a phenomenal speed as compared to the 1st wave. There could be several factors responsible for the increased number of cases in the second wave. It is observed that the mutant virus has more effective transmission capability and its incubation period is also lesser.

Total Cases of Covid-19 in India

India Total Confirmed vs Recovered vs Deaths

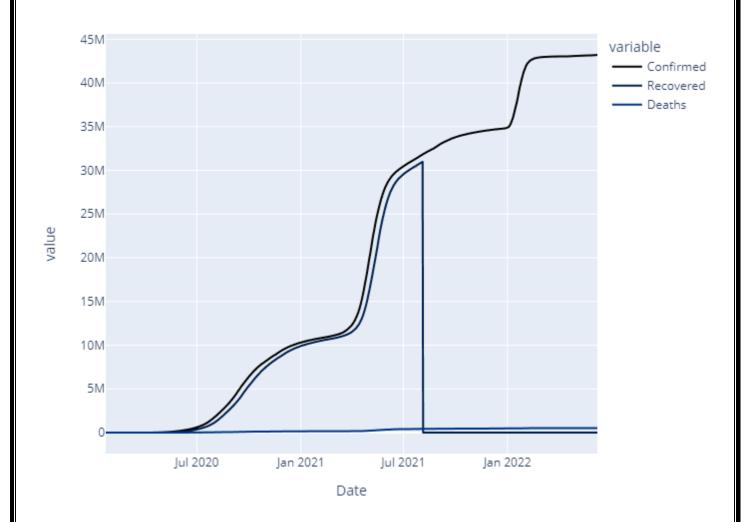


Fig. Line plot of India Total Wave of Covid-19

COVID-19 VACCINATION

Getting vaccinated could save your life. COVID-19 vaccines provide strong protection against serious illness, hospitalization and death. There is also some evidence that being vaccinated will make it less likely that you will pass the virus on to others, which means your decision to get the vaccine also protects those around you.

Even after getting vaccinated, keep taking precautions to protect yourself, family, friends and anyone else you may come into contact with. COVID-19 vaccines are highly effective, but some people will still get ill from COVID-19 after vaccination. There is also still a chance that you could also pass the virus on to others who are not vaccinated. Stay at least 1 metre away from other people, wear a properly fitted mask over your nose and mouth when you can't keep this distance, avoid poorly ventilated places and settings, clean your hands frequently, stay home if unwell and get tested, and stay informed about how much virus is circulating in the areas where you travel, live and work.

Vaccines

India has two approved COVID-19 vaccines: Covishield and Covaxin. Both of them were exported and used in foreign grants by the Government of India.

Covishield

On 1 January 2021, the Drug Controller General of India, approved the emergency or conditional use of Covishield. Covishield is developed by the University of Oxford and its spin-out company, Vaccitech.

Covaxin

On 2 January 2021, Covaxin India's first COVID-19 vaccine, developed by Bharat Biotech in association with the Indian Council of Medical Research and National Institute of Virology received approval from the Drug Controller General of India for its emergency or conditional usage.¹

VACCINATION PLOT OF INDIA

State-wise Population vs Dose 1

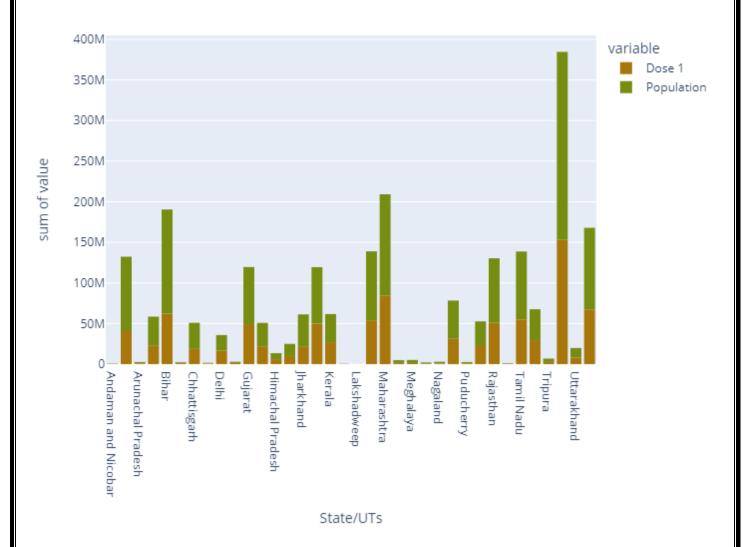


Fig. Histogram for India State-wise Vaccination Dose 1 vs Total Population

State-Wise Population vs Dose 2

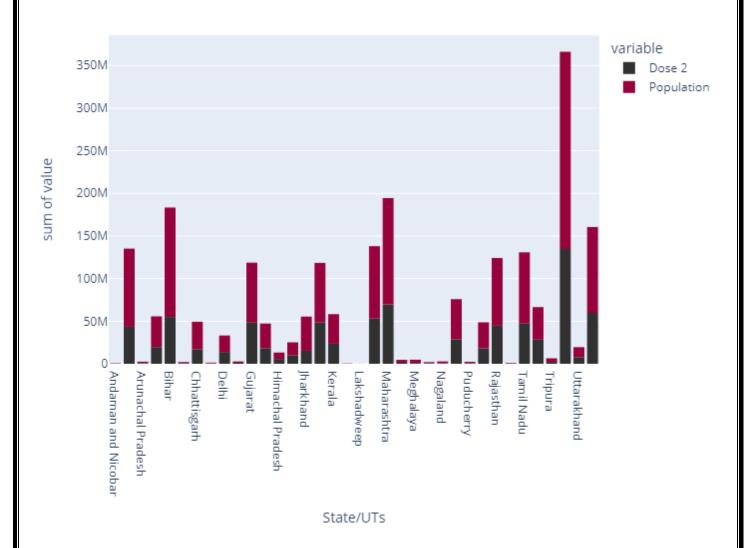


Fig. Histogram of India State-wise Covid Dose 2 vs Total Population

State Wise Total Population vs Total Vaccination

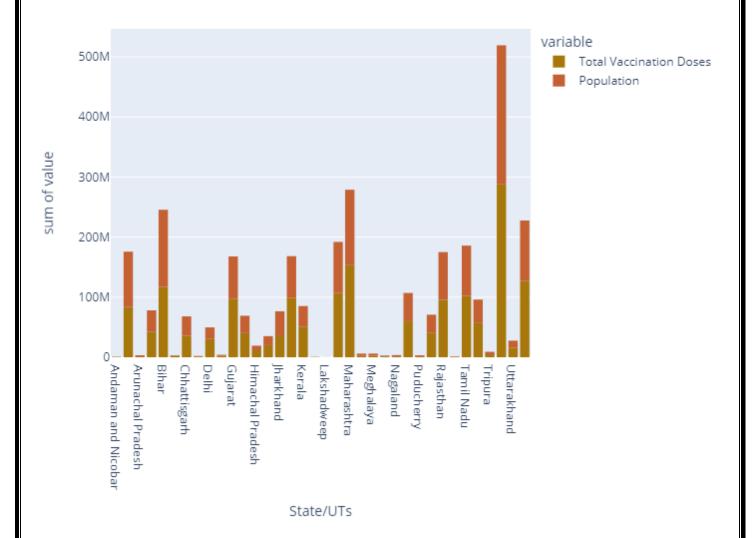


Fig. Histogram Of India State-wise Covid-19 Total Vaccination Dose vs Total Population

CONCLUSION

The results obtained gives us the clear vision about what type of population is highly affected by this problem. This model was designed in such a way that future modifications can be done easily. The following conclusions can be deduced from the development of the project:

- Automation of the entire system improves the efficiency.
- It provides a friendly graphical user interface which proves to be better when compared to the existing model.
- It gives appropriate access to the authorized users depending on their permissions.
- Updating of information becomes so easier.
- System security, data security and reliability are the striking features.

Future Scope:

The data which has been taken was limited. The project could be extended to more number of days. The Covid data was predicted by taking the last 2 years in data. The error can be minimized as well using other algorithms.

In future by this study for any group of population in given countries, the number of infected, deaths, recovery cases can be predicted. This can be used as a reference for evaluating the effectiveness of the preventive measures and policies that government took for reduction of covid cases.

Bibliography:

- Youtube Tutorials of Data Analysis
- Websites:

https://www.geeksforgeeks.org/python-programming-language/ https://www.tutorialspoint.com/python/python_modules.htm https://en.wikipedia.org/wiki/COVID-19_pandemic https://www.javatpoint.com/python-tutorial

