COVID-19 VACCINATION ANALYSIS

Phase 5 submission document

Project Title: **Covid-19 vaccination analysis**

Phase 5: Project Documentation & Submission

Topic: *In this section we will document the complete project and prepare it for submission.*

Covid-19 Vaccination Analysis

Introduction:

- COVID-19, or Coronavirus Disease 2019, is a viral respiratory illness caused by the novel coronavirus, SARS-CoV-2. It was first identified in Wuhan, China, in late 2019 and rapidly spread around the world, leading to a global pandemic. This infectious disease has had a profound impact on public health, economies, and daily life, with symptoms ranging from mild respiratory issues to severe pneumonia. Various preventive measures, including vaccination, mask-wearing, and social distancing, have been employed to combat its spread and minimize its impact.

- Analyzing COVID-19 vaccination efforts involves examining various aspects of the global response to the pandemic. This includes the development, distribution, and administration of vaccines, as well as their impact on public health, economy, and society. Analysis can focus on vaccination rates, vaccine efficacy, equity in access, challenges in deployment, and the evolving strategies to combat the virus. It's a multidimensional topic with implications for healthcare, policy, and social well-being.

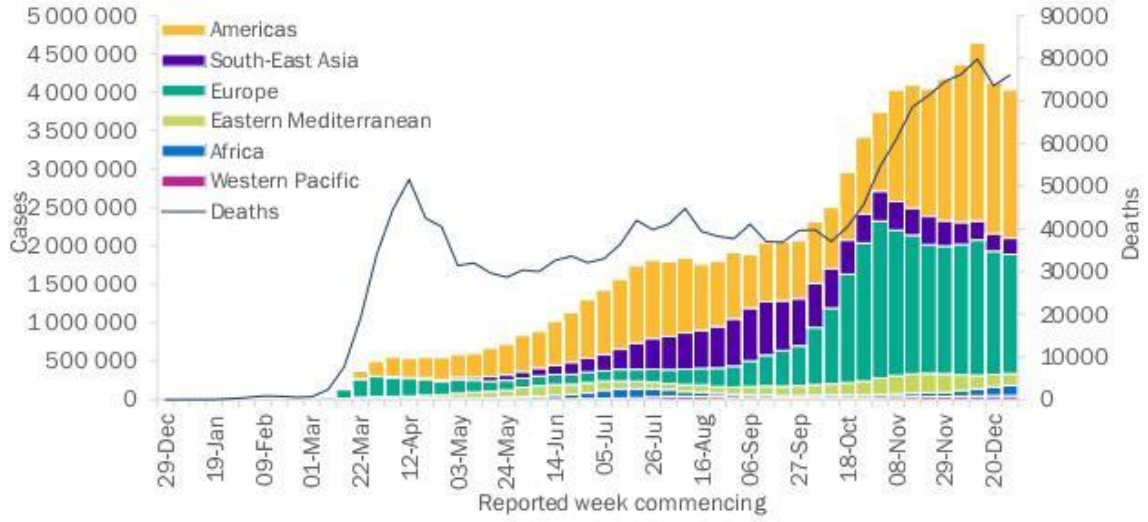
- The COVID-19 pandemic has brought about unprecedented challenges to the world. One of the most effective ways to control the spread of this disease is through the development and distribution of vaccines. Vaccines are essential in building immunity against the virus and preventing future outbreaks.

Given data set:

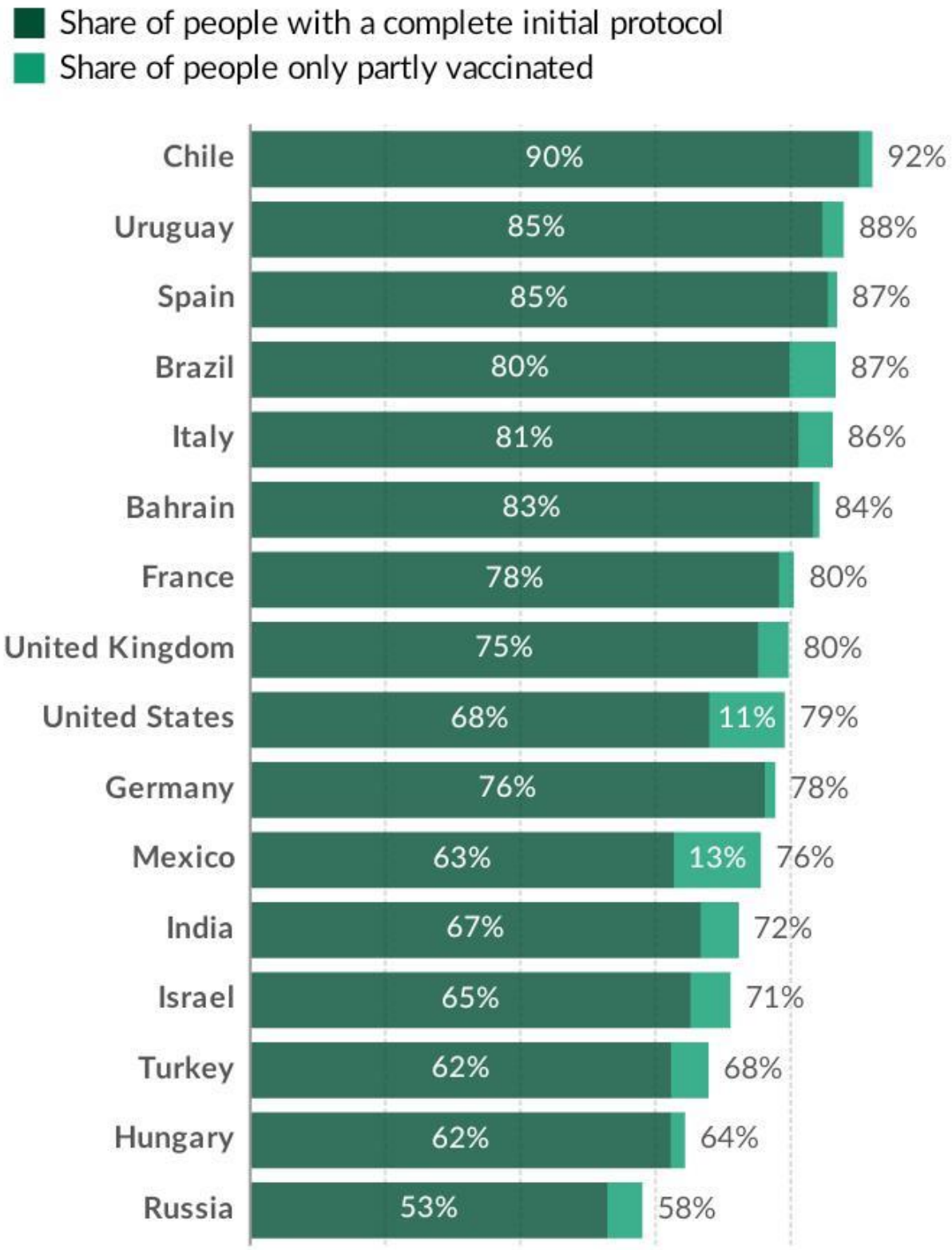
Name	Size	Format	Country	Link
AIforCOVID imaging archive	983 patients	DICOM	Italy	https://aiforcovid.radiomica.it/
MosMed COVID-19 Chest CT database	1110 patients	NIfTI	Russia	https://mosmed.ai/datasets/covid19_1110
SIRM database	68 patients	JPG	Italy	https://www.sirm.org/en/category/articles/covid-19-database/
Radiopaedia database	101 patients	JPG	Global	https://radiopaedia.org/articles/covid-19-3
BSTI COVID-19 database	59 patients	online only	UK	https://bit.ly/BSTICovid19_Teaching_Library
UCSD COVID-CT database	349 images from 216 patients	PNG	Global	https://github.com/UCSD-AI4H/COVID-CT
Coronacases.org	10 patients	online only	China	https://coronacases.org/
Eurorad database	50 patients	JPG	Global	https://www.eurorad.org/advanced-search?search=COVID
CT machine learning dataset	930 images from 461 patients, 20 volumes	JPG, NIfTI	Global	https://github.com/ieee8023/covid-chestxray-dataset
CT segmentation dataset 1	100 images from 40 patients	NIfTI	Italy	http://medicalsegmentation.com/covid19
CT segmentation dataset 2	800 slices from 9 patients	NIfTI	Global	http://medicalsegmentation.com/covid19
CT segmentation dataset 3	20 cases	NIfTI	Global	https://zenodo.org/record/3757476

The COVID-19 vaccination trend data was be analysed using Power BI

1. Below Bar chart shows that “The People Fully Vaccinated” by“Country”. In which “US” has highest People who are Fully Vaccinated.

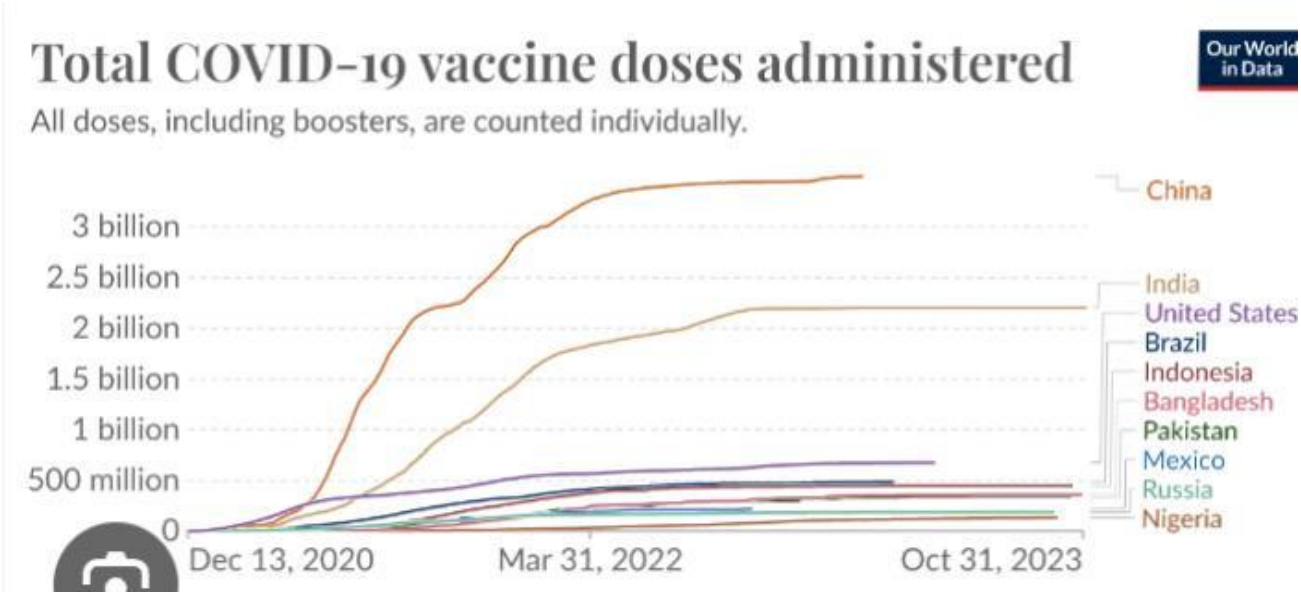


2. Below Tree map shows “The Total Vaccinations” by “Country” and details of the source.

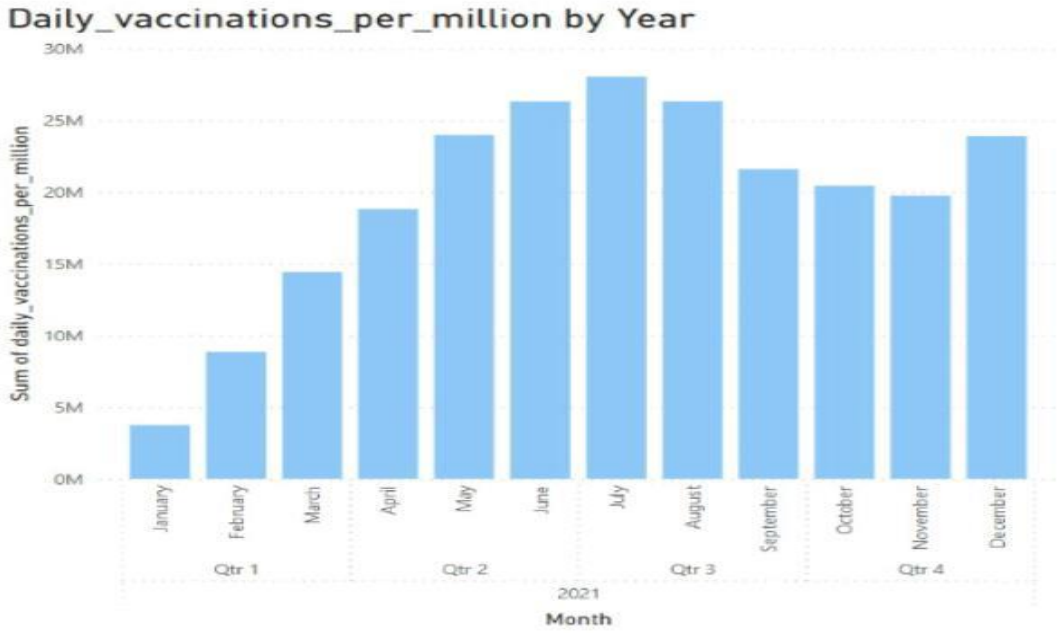


3. Below Area Chart shows the Daily vaccinations vs Date. We can see that July-2021, Aug-2021, Sept-2021, Jan-2022 highest no. of vaccinations.

Trend line shows as the time(months) passed the daily vaccination rate has increased.

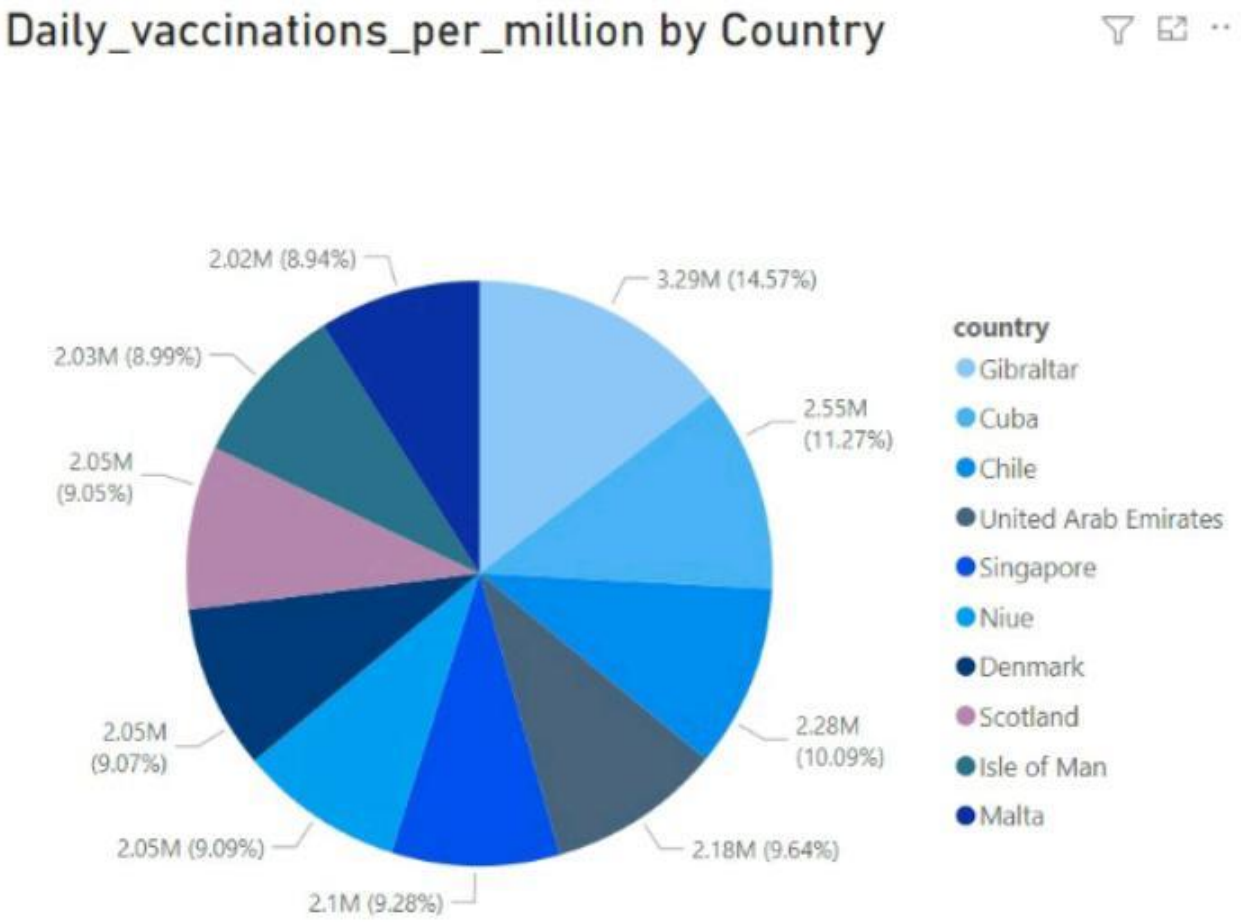


4. Below Bar Chart shows the Daily vaccinations per million vs Date. We can see that July-2021 highest no. of vaccinations per million.

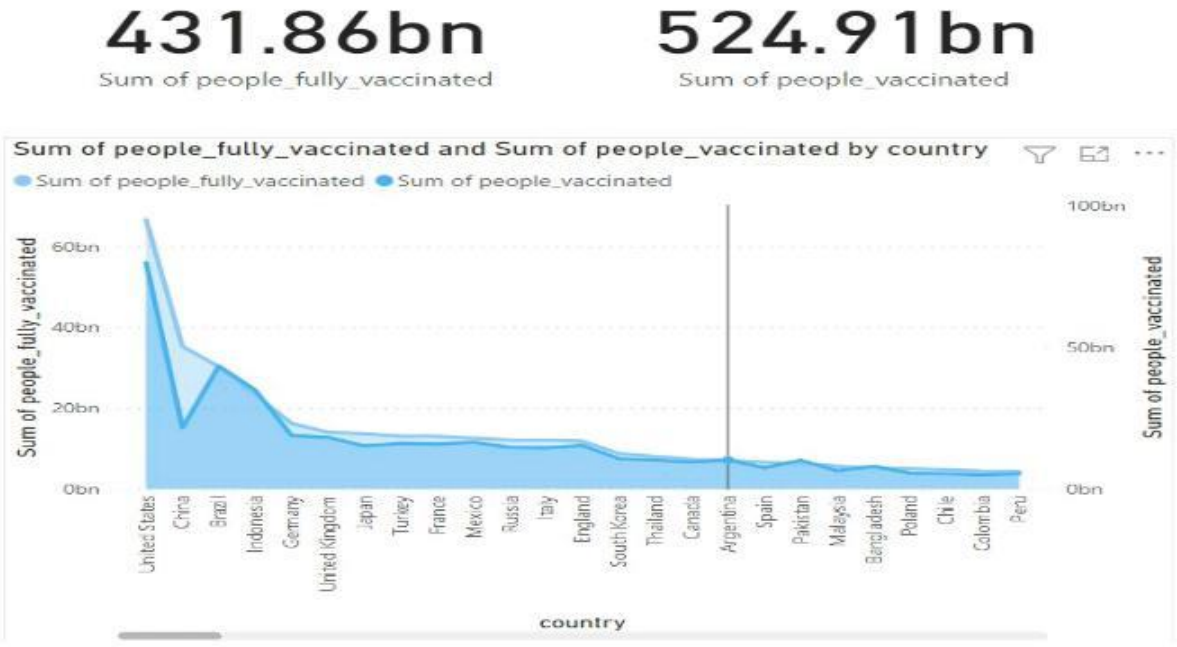


e

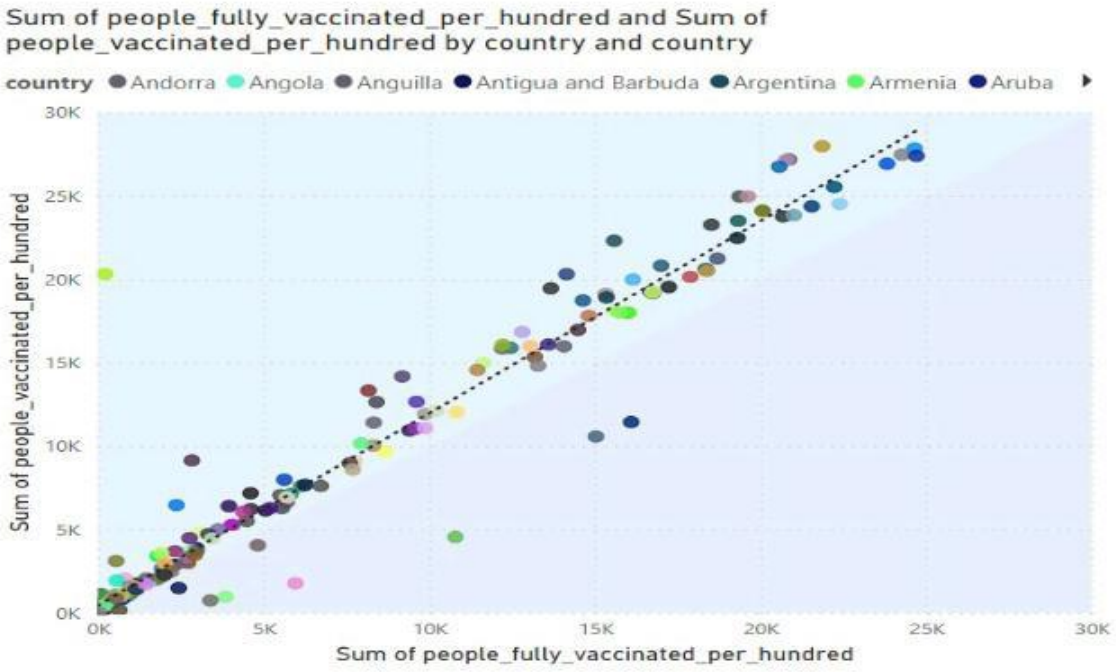
5. Below Pie Chart shows the Daily vaccinations per million vs Top 10 Country. We can see that Gibraltar has done highest no. of vaccinations per million.



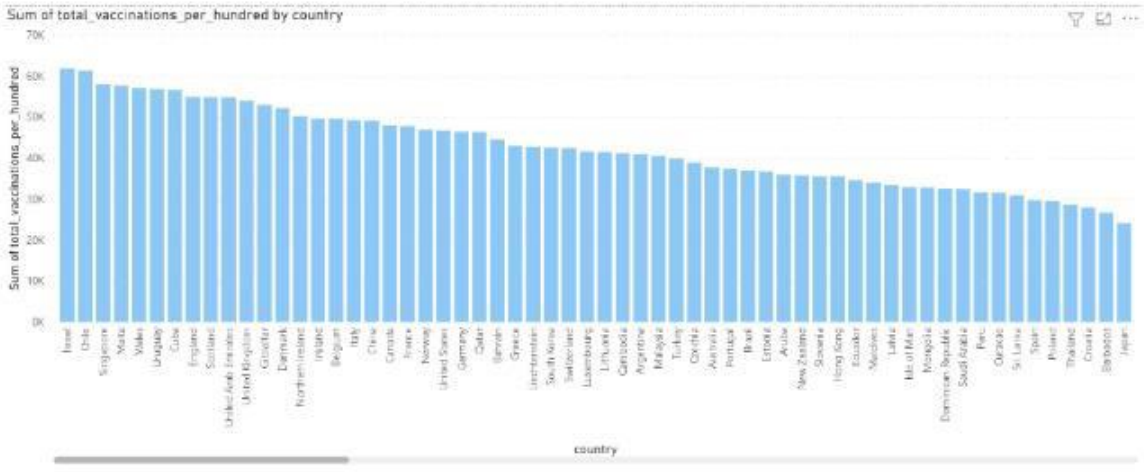
6. Below Area line Chart Shows Sum of people vaccinated and Sum of people fully vaccinated for individual country. United States had the highest Sum of people fully vaccinated and Pitcairn had the lowest Sum of people fully vaccinated.



7 . Below Scatter Chart shows the sum of people vaccinated per hundred and Sum of people fully vaccinated per hundred for individual country. Malta had the highest Sum of people fully vaccinated per hundred and Wales had the highest Sum of people vaccinated per hundred.



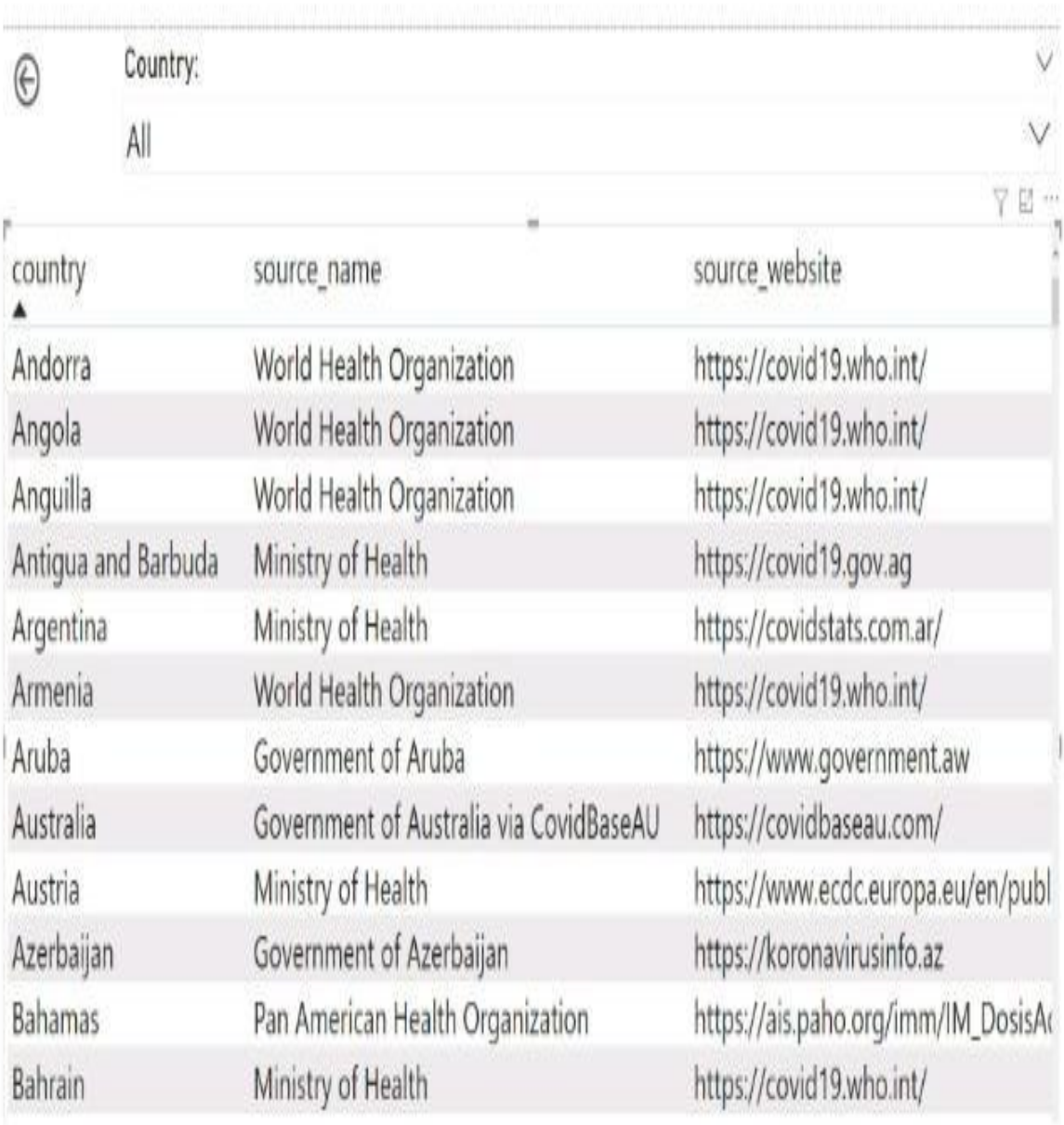
8. Below Bar chart shows “The Total Vaccinations per Hundred” by “Country”. In which “Israel” has highest Total Vaccinations per Hundred.



16,78,76,11,80,010.00
Sum of total_vaccinations

34,62,450.54
Sum of total_vaccinations_per_hundred

9. Below table shows the Country wise Source for vaccination.



country	source_name	source_website
Andorra	World Health Organization	https://covid19.who.int/
Angola	World Health Organization	https://covid19.who.int/
Anguilla	World Health Organization	https://covid19.who.int/
Antigua and Barbuda	Ministry of Health	https://covid19.gov.ag
Argentina	Ministry of Health	https://covidstats.com.ar/
Armenia	World Health Organization	https://covid19.who.int/
Aruba	Government of Aruba	https://www.government.aw
Australia	Government of Australia via CovidBaseAU	https://covidbaseau.com/
Austria	Ministry of Health	https://www.ecdc.europa.eu/en/publ
Azerbaijan	Government of Azerbaijan	https://koronavirusinfo.az
Bahamas	Pan American Health Organization	https://ais.paho.org/imm/IM_DosisA
Bahrain	Ministry of Health	https://covid19.who.int/

Program:

```
Import numpy as np

Import pandas as pd

Import matplotlib.pyplot as plt


Data = pd.read_csv('case_time_series.csv')


Y = data.iloc[61:,1].values

R = data.iloc[61:,3].values

D = data.iloc[61:,5].values

X = data.iloc[61:,0]


Plt.figure(figsize=(25,8))


Ax = plt.axes()
```

```
Ax.grid(linewidth=0.4, color='#8f8f8f')


Ax.set_facecolor("black")

Ax.set_xlabel('\nDate',size=25,color='#4bb4f2')

Ax.set_ylabel('Number of Confirmed Cases\n',

        Size=25,color='#4bb4f2')


Plt.xticks(rotation='vertical',size='20',color='white')

Plt.yticks(size=20,color='white')

Plt.tick_params(size=20,color='white')


For I,j in zip(X,Y):

   Ax.annotate(str(j),xy=(I,j+100),color='white',size='13')


Ax.annotate('Second Lockdown 15th April',
```

```
            Xy=(15.2, 860),

            Xytext=(19.9,500),

            Color='white',

            Size='25',

            Arrowprops=dict(color='white',

                        Linewidth=0.025))


Plt.title("COVID-19 IN : Daily Confirmed\n",

          Size=50,color='#28a9ff')

Ax.plot(X,Y,

        Color='#1F77B4',

        Marker='o',

        Linewidth=4,
```

Markersize=15,

Markeredgecolor='#035E9B')

Output:

DESIGN TO INNOVATION

1. Case Data:

Collecting data on the number of confirmed COVID-19 cases, including information on age, gender, and location of infected individuals.

2. Testing Data:

Recording data on the number of tests conducted, test results, and testing locations.

3. Hospitalization Data:

Tracking the number of COVID-19 patients admitted to hospitals, their conditions, and the availability of hospital resources.

4. Mortality Data:

Recording data on COVID-19-related deaths, including age, gender, and underlying conditions.

5. Contact Tracing Data:

Collecting information about individuals who may have been exposed to the virus, which is critical for containment.

6. Vaccination Data:

Monitoring the number of people vaccinated, vaccine types, and vaccination locations.

7. Genomic Sequencing:

Conducting genomic sequencing of the virus to monitor mutations and variations.

8. Public Health Measures:

Gathering data on the implementation and effectiveness of measures like lockdowns, mask mandates, and social distancing.

9. Data Sources:

Data was collected from various sources, including healthcare providers, testing centers, contact tracing apps, and government agencies.

Data was often reported through dashboards, reports, and visualizations to help inform the public and guide decision-making. Effective data collection and analysis were crucial in managing the COVID-19 pandemic and making informed decisions to mitigate its impact.

ABSTRACT

- The coronavirus is considered this century's most disruptive catastrophe and global concern.

- This disease has prompted extreme social, psychological and economic impacts affecting millions of people around the globe.

- COVID-19 is transmitted from one infected person's body to another through respiratory droplets.

- This virus proliferates when people breathe in air-contaminated space with droplets and microscopic airborne particles.

- This research aims to analyze automatic COVID-19 detection using machine learning techniques to build an intelligent webapplication.

- The dataset has been preprocessed by dropping null values, feature engineering, and synthetic oversampling (SMOTE) techniques.

- Next, we trained and evaluated different classifiers, i.e., logistic regression, random forest, decision tree, k-nearest neighbor, support vector machine (SVM), ensemble models (adaptive boosting and extreme gradient boosting) and deep learning (artificial neural network, convolutional neural network and long short-term memory) techniques.

- Explainable AI with the LIME framework has been applied to interpret the prediction results.

VACCINATION

Analyzing COVID-19 vaccination efforts is crucial in understanding and managing the pandemic. Here are key aspects of COVID-19 vaccination analysis:

1. Vaccine Efficacy: Ongoing analysis assesses the effectiveness of different COVID-19 vaccines in preventing infection, severe disease, and transmission.

2. Vaccine Coverage: Monitoring the percentage of the population that has received one or both vaccine doses and the impact of vaccination on community immunity.

3. Vaccine Distribution: Analyzing the distribution of vaccines to ensure equitable access and identifying underserved or at risk population.

4. Booster Shots: Assessing the need for booster doses and Analyzing their effectiveness in maintaining immunity.

5. Vaccine Hesitancy: Identifying factors contributing to Vaccine hesitancy and strategies to increase vaccine Acceptance.

6. Adverse Events: Monitoring and analyzing adverse events Associated with vaccination to ensure safety.

7. Variants: Analyzing the impact of emerging variants on Vaccine effectiveness and the need for updated vaccines.

8. Global Impact: Evaluating the distribution of vaccines Worldwide and their role in reducing the global spread of the Virus.

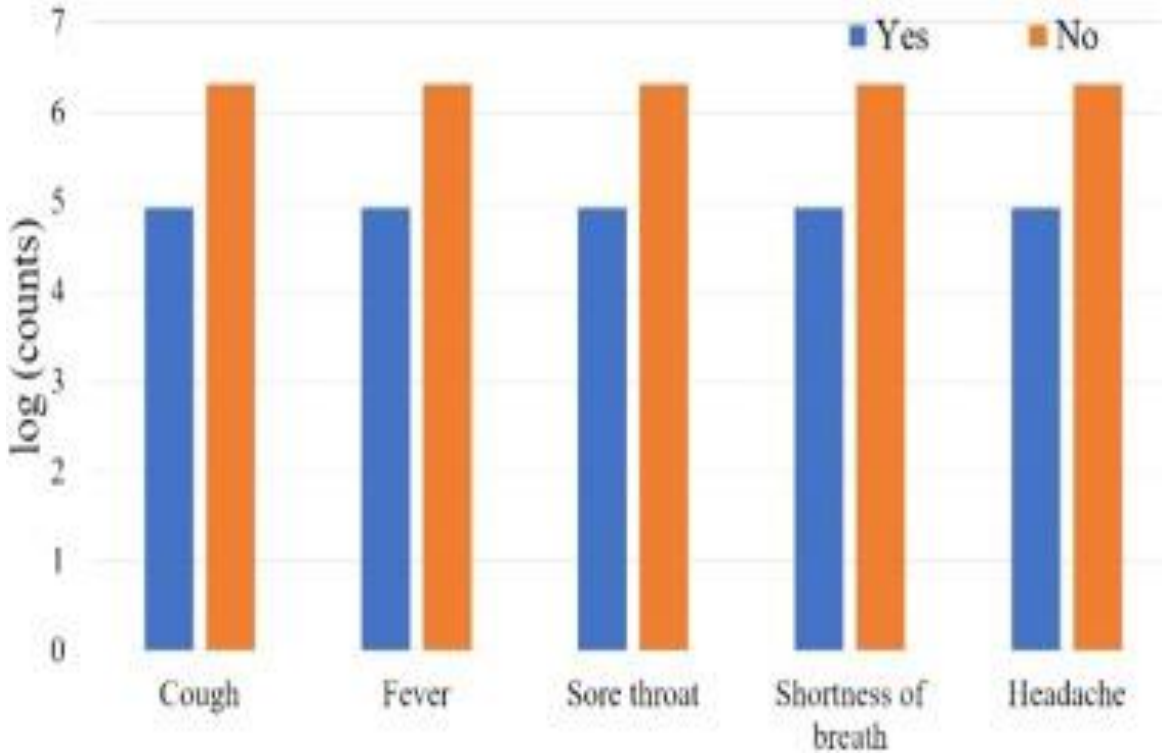
9. Public Health Measures: Analyzing how vaccination impacts The need for other public health measures, such as mask Wearingand social distanding .

10. Vaccine Equity: Ensuring fair distribution of vaccines to Low- and middle-income countries and addressing global Vaccine inequality.

11. Data Transparency: Maintaining transparency in vaccine Data reporting to build trust and confidence in vaccination Programs.

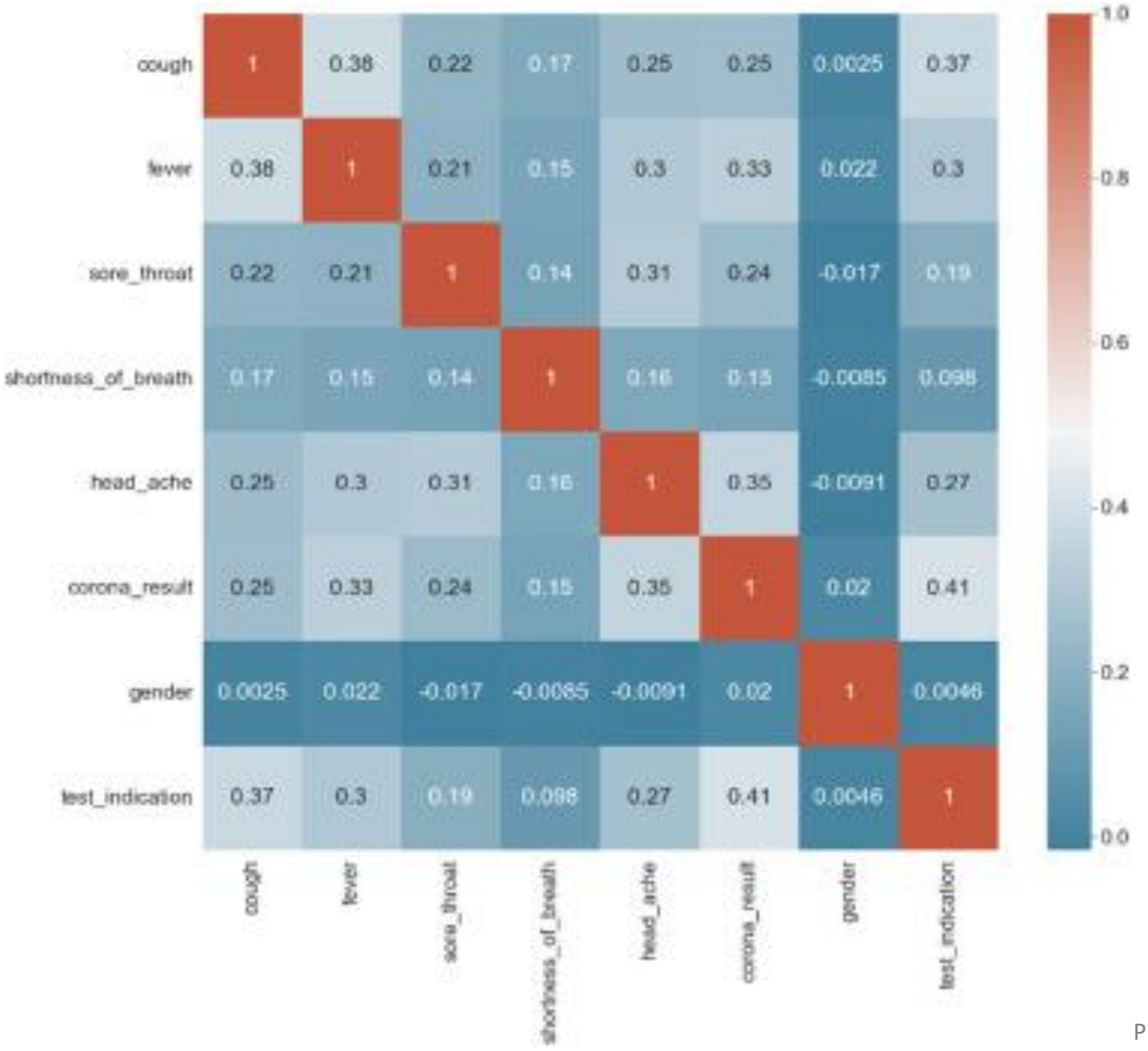
COVID-19 vaccination analysis is essential for adapting Vaccinatio strategies, improving vaccine distribution, and Ensuring the world's path to recovery from the pandemic.

BARCHART



COVID-19 TEST RESULTS IN TERMS OF THE FEATURES IN THE DATASET

The dataset is highly imbalanced as there is a higher number of negative cases and a significantly smaller number of positive cases with a ratio of 9.3:1.0. The linear dependence of various features of the used dataset has been measured with the Pearson correlation index



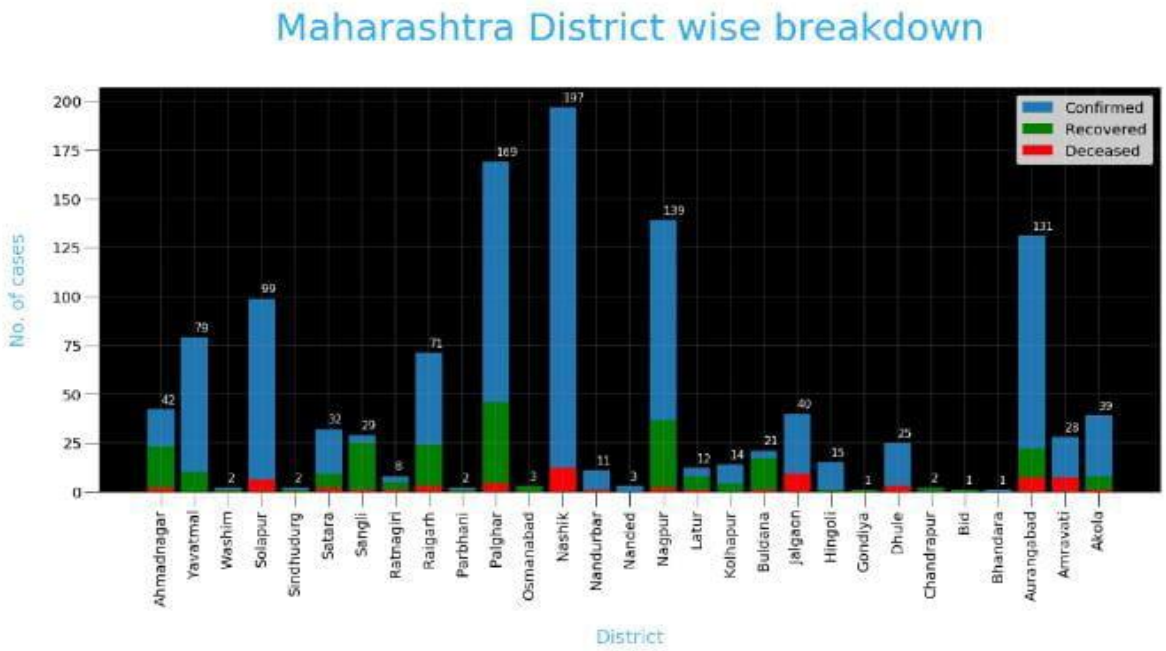
PYTHON PROGRAM

```
data = pd.read_csv('district.csv')
data.head()


re=data.iloc[:30,5].values

de=data.iloc[:30,4].values

co=data.iloc[:30,3].values

x=list(data.iloc[:30,0])


plt.figure(figsize=(25,10))

ax=plt.axes()


ax.set_facecolor('black')

ax.grid(linewidth=0.4, color='#8f8f8f')



plt.xticks(rotation='vertical',

           size='20',
```

```
        color='white')#ticks of X

plt.yticks(size='20',color='white')


ax.set_xlabel('\nDistrict',size=25,
            color='#4bb4f2')
ax.set_ylabel('No. of cases\n',size=25,
            color='#4bb4f2')


plt.tick_params(size=20,color='white')


ax.set_title('Maharashtra District wise breakdown\n',
            size=50,color='#28a9ff')

plt.bar(x,co,label='re')
plt.bar(x,re,label='re',color='green')
```

```
plt.bar(x,de,label='re',color='red')


for i,j in zip(x,co):
    ax.annotate(str(int(j)),
                xy=(i,j+3),
                color='white',
                size='15')

plt.legend(['Confirmed','Recovered','Deceased'],
           fontsize 20)
```

Output


Maharashtra District wise breakdown
Confirmed
Recovered
Deceased
No. of cases
District
Ahmadnagar
Yavatmal
Washim
Solapur
Sindhudurg
Satara
Sangli
Ratnagiri
Raigarh
Parbhani
Palghar
Osmanabad
Nashik
Nandurbar
Nanded
Nagpur
Latur
Kolhapur
Buldana
Jalgaon
Hingoli
Gondiya
Dhule
Chandrapur
Bid
Bhandara
Aurangabad
Amravati
Akola

CAUSES

The analysis of the causes and contributing factors to the Spread of COVID-19 involves various factors. While the Primary cause of the pandemic is the introduction of the novel Coronavirus (SARS-CoV-2) into the human population, several Factors have contributed to its rapid transmission and impact:

1. **Human-to-Human Transmission:**

   COVID-19 primarily spreads Through respiratory droplets when an infected person coughs, Sneezes, or talks. Close contact with infected individuals is a major Factor in transmission.

2. **Asymptomatic and Presymptomatic Spread:**

   People who are Infected with the virus but show no symptoms (asymptomatic) or Who have not yet developed symptoms (presymptomatic) can Unknowingly spread the virus.

3. **Variants:**

   The emergence of new variants of the virus has Impacted its transmission and potential resistance to immunity Developed through vaccination or prior infection.

4. **Global Travel:**

   International travel allowed the virus to spread Rapidly from its original epicenter in Wuhan, China, to other parts of the world.

5. **Public Health Measures:**

   The effectiveness of public health measures, such as mask-wearing,

social distancing, and lockdowns, in curbing the virus's spread varies depending on adherence and enforcement.

6. **Healthcare Infrastructure:**

The readiness and capacity of healthcare systems have played a role in managing and treating COVID-19 cases.

7. **Vaccination Rates:**

The availability and distribution of COVID-19 vaccines have a direct impact on the spread and severity of the virus.

8. **Public Behavior:**

Individual behavior, compliance with guidelines, and vaccine acceptance influence the spread of the virus.

9. **Variability in Government Responses:**

The response to the pandemic, including testing, contact tracing, and quarantine measures, has varied between countries and regions.

It's important to note that the situation and causes of COVID-19 continue to evolve as more is learned about the virus and its variants. Analysis of these causes helps inform public health strategies to control the spread of the disease.

IMPACTS OF COVID-19

The impact of COVID-19 analysis has been profound and has influenced various aspects of society and healthcare. Here are some key impacts:

1. Public Health Response:

Analysis of COVID-19 data has informed public health measures, such as social distancing, mask mandates, and lockdowns, to curb the spread of the virus and protect public health.

2. Vaccination Strategy:

Ongoing analysis of vaccine efficacy and distribution data has shaped vaccination strategies, ensuring that vaccines are distributed to those at highest risk first.

3. Resource Allocation:

Data analysis has guided the allocation of healthcare resources, including ventilators, ICU beds, and medical supplies, to regions with high case numbers.

4. Economic Impact:

The analysis of COVID-19's economic impact has influenced government stimulus packages and financial relief measures to support businesses and individuals affected by the pandemic.

5. Travel and Border Policies:

Data analysis has been crucial in determining travel restrictions, quarantine requirements, and border policies to prevent the spread of the virus across regions and countries.

6. Remote Work and Education:

The analysis of case data has led to the widespread adoption of

remote work and online education to reduce the risk of transmission.

7. Mental Health and Well-being:

Data analysis has highlighted the mental health impacts of the pandemic, leading to increased focus on mental health support and services.

8. Research and Vaccine Development:

Analysis of the virus's genetic sequence has informed vaccine development efforts, leading to the rapid development and distribution of COVID-19 vaccines.

9. Scientific Collaboration:

The pandemic prompted extensive global scientific collaboration, accelerating research and data sharing.

10. Future Pandemic Preparedness:

COVID-19 analysis has underscored the importance of pandemic preparedness and the need for improved surveillance, early warning systems, and research.

The impact of COVID-19 analysis is ongoing and has been central to the global response to the pandemic, helping to adapt strategies and mitigate its effect

PROBLEM DEFINITION AND DESIGN THINKING

Problem Definition:

The problem is to conduct an in-depth analysis of Covid-19 vaccine data, focusing on vaccine efficacy, distribution, and adverse effects. The goal is to provide insights that aid policymakers and health organizations in optimizing vaccine deployment strategies. This project involves data collection, data preprocessing, exploratory data analysis, statistical analysis, and visualization.

DESIGN THINKING:

1. **Data Collection:** Collect Covid-19 vaccine data from reputable sources like health organizations, government databases, and research publications.
2. **Data Preprocessing:** Clean and preprocess the data, handle missing values, and convert categorical features into numerical representations.
3. **Exploratory Data Analysis:** Explore the data to understand its characteristics, identify trends, and outliers.
4. **Statistical Analysis:**Perform statistical tests to analyze vaccine efficacy, adverse effects, and distribution across different populations.
5. **Visualization:** Create visualizations (e.g., bar plots, line charts, heatmaps) to present key findings and insights.
6. **Insights and Recommendations:** Provide actionable insights and recommendations based on the analysis to assist policymakers and health organizations.

Insights:

The analysis of the COVID-19 vaccination trend data provides the following insights:

• The vaccination efforts vary widely across different countries and regions.

• Some countries have made significant progress in vaccination, while others are lagging behind.

• We can see the Analyse of the sum of daily vaccinating details, fully vaccinating and vaccinating people details based on different countries and Date from the visualization.

Recommendations:

Based on the insights gained from the analysis, the following recommendations can be made:

• Improve the efficiency of the vaccination program in countries with low vaccination rates.

• Increase public awareness of the importance of vaccination and address any concerns or

misconceptions about vaccines.

• Collect day to day reports for better analysis.

• Like this dataset we can perform operations with various categories, city-wise or

regionwise.

• Collect city-wise and region-wise data to analyse, to get more in-depth insights

Methodology:

The methodology for this project involves the following steps:

1. Data Importing:

The given dataset of COVID-19 Country Vaccinations data was

collected and loaded to Power BI Desktop.

2. Data cleaning:

The data was cleaned and pre-processed by removing the missing values, duplicates. The columns contained null values have been replaced by 0 with the use of replace functions and started working on the data.

3. Data visualization:

The was visualized using Power BI Desktop software to identify trends and patterns using different kinds of charts, graphs, cards and table.

4. Data analysis:

The data was analysed to identify insights and recommendations.

Context

Data is collected daily from Our World in Data GitHub repository for covid-19, merged and uploaded. Country level vaccination data is gathered and assembled in one single file. Then, this data file is merged with locations data file to include vaccination sources information. A second file, with manufacturers information, is included.

Content

The data (country vaccinations) contains the following information:

Country- this is the country for which the vaccination information is provided;

Country ISO Code - ISO code for the country;

Date - date for the data entry; for some of the dates we have only the daily vaccinations, for others, only the (cumulative) total;

Total number of vaccinations - this is the absolute number of total immunizations in the country;

Total number of people vaccinated - a person, depending on the immunization scheme, will receive one or more (typically 2) vaccines; at a certain moment, the number of vaccination might be larger than the number of people;

Total number of people fully vaccinated - this is the number of people that received the entire set of immunization according to the immunization scheme (typically 2); at a certain moment in time, there might be a certain number of people that received one vaccine and another number (smaller) of people that received all vaccines in the scheme;

Daily vaccinations (raw) - for a certain data entry, the number of vaccination for that date/country;

Total vaccinations per hundred - ratio (in percent) between vaccination number and total population up to the date in the country;

Total number of people vaccinated per hundred - ratio (in percent) between population immunized and total population up to the date in the country;

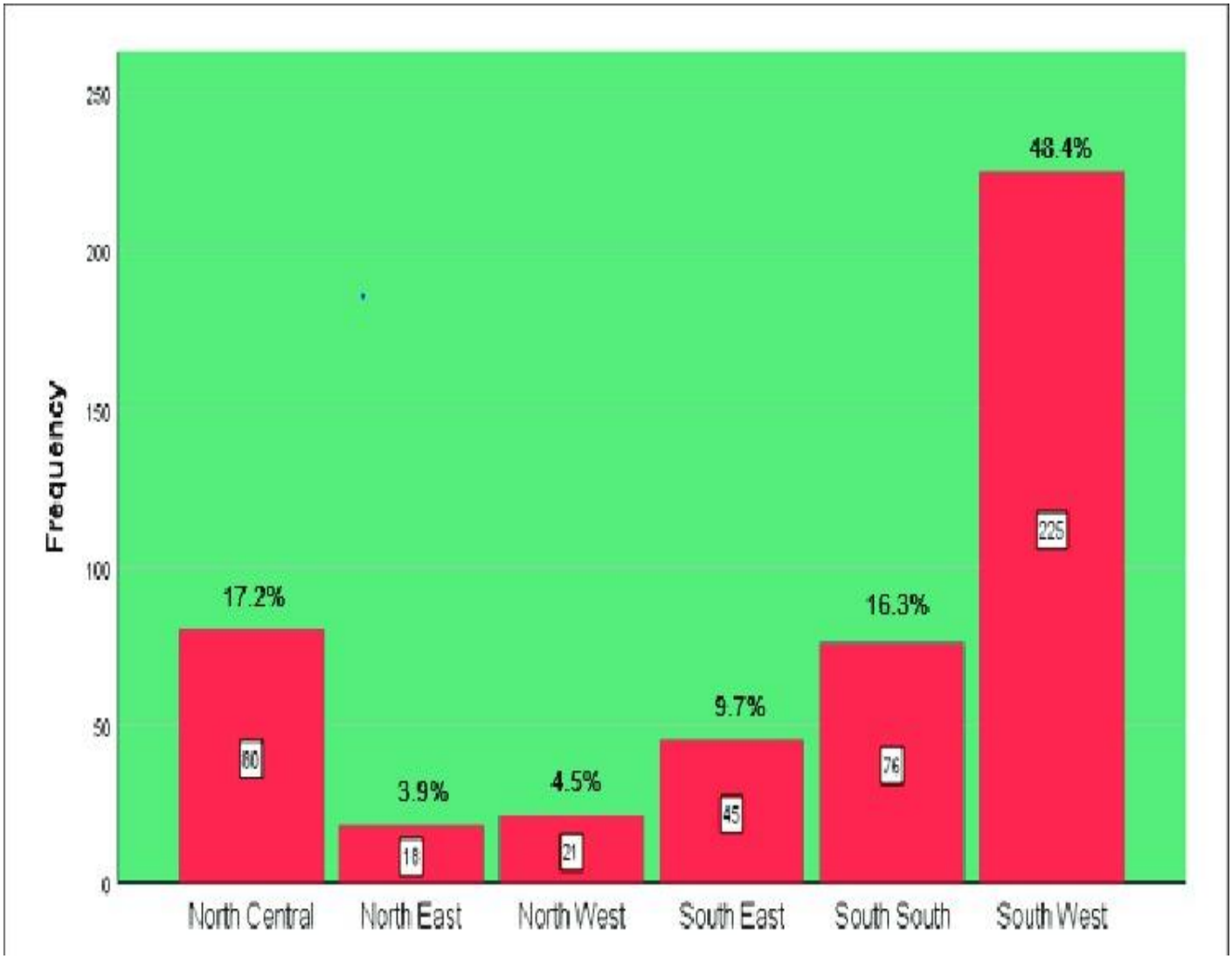
Total number of people fully vaccinated per hundred - ratio (in percent) between population fully immunized and total population up to the date in the country;

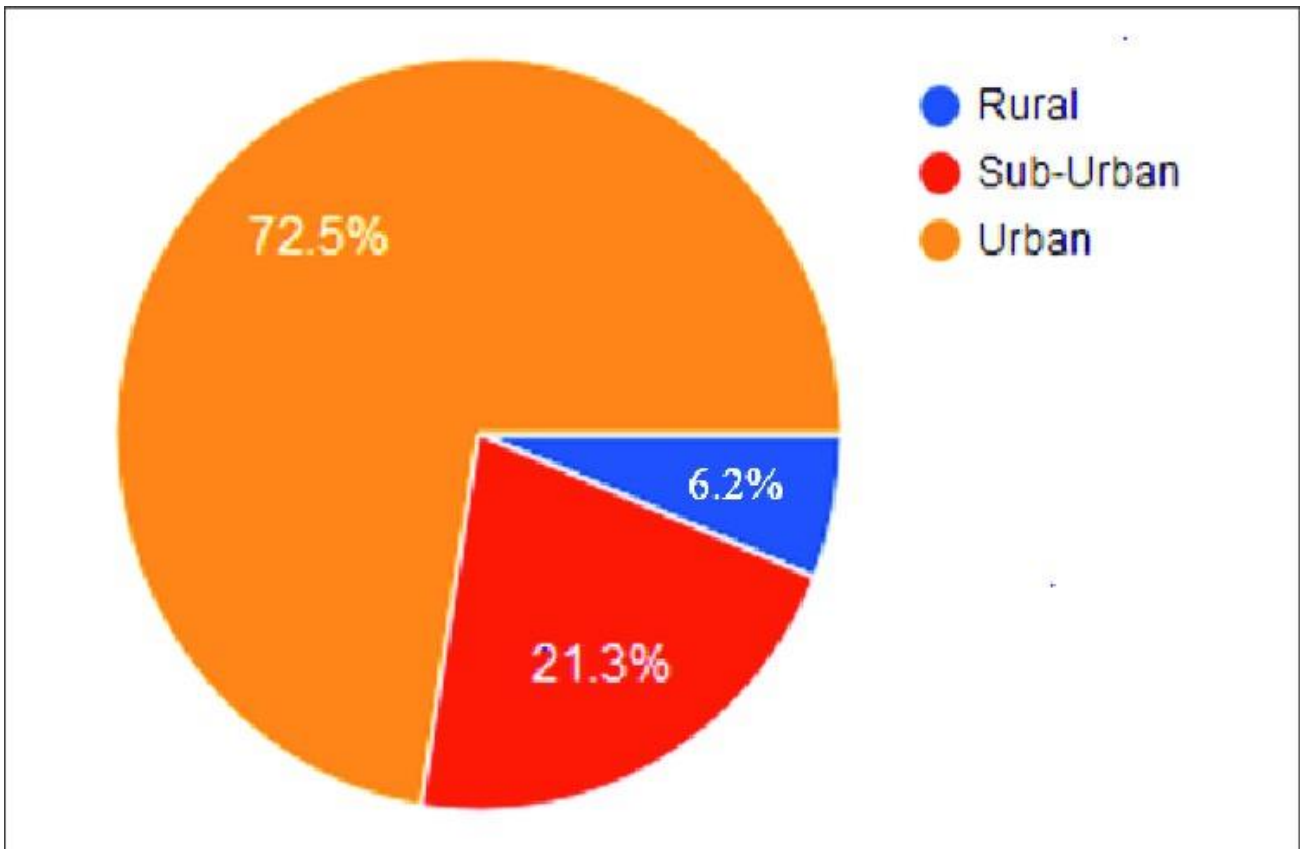
Number of vaccinations per day - number of daily vaccination for that day and country;

Daily vaccinations per million - ratio (in ppm) between vaccination number and total population for the current date in the country;

Vaccines used in the country - total number of vaccines used in the country (up to date);

Source name - source of the information (national authority, international organization, local organization etc.);

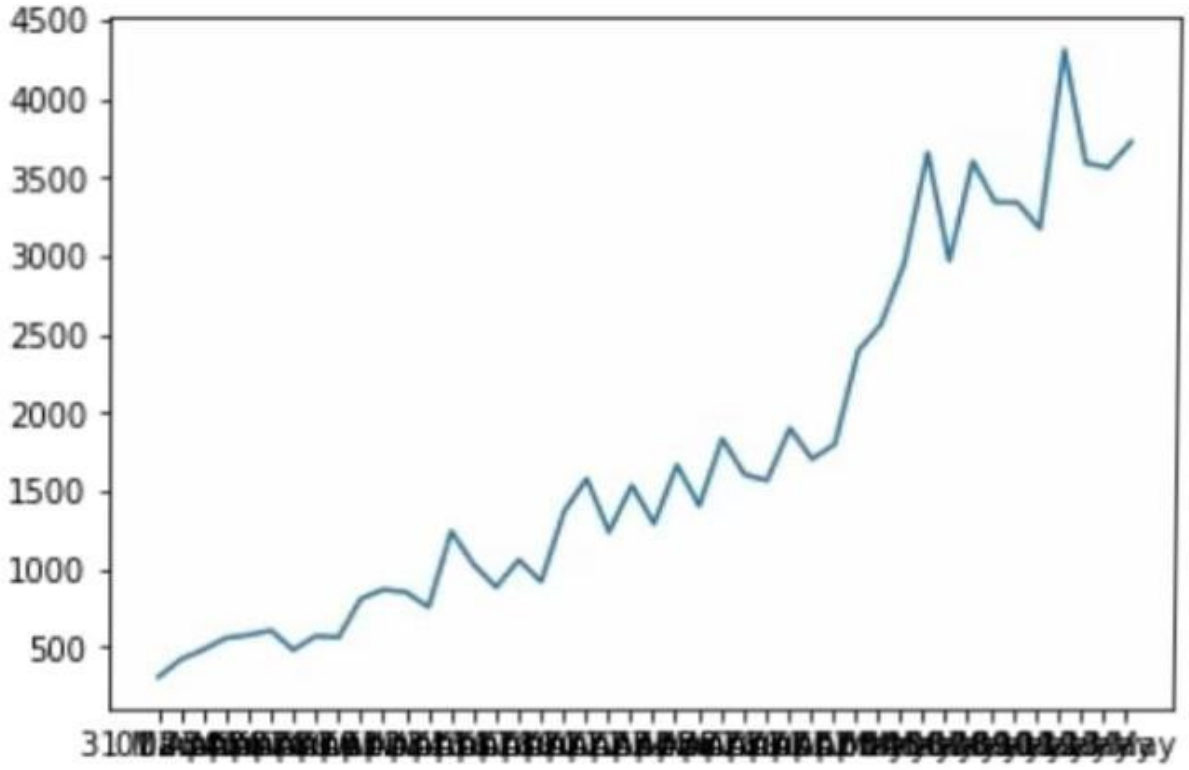

Frequency
250
200
150
100
50
0
17.2%
60
3.9%
18
4.5%
21
9.7%
45
16.3%
76
48.4%
225
North Central
North East
North West
South East
South South
South West


Rural
Sub-Urban
Urban
72.5%
6.2%
21.3%

Program

```
Import numpy as np

Import pandas as pd

Import matplotlib.pyplot as plt



Data = pd.read_csv('case_time_series.csv')


Y = data.iloc[61:,1].values

R = data.iloc[61:,3].values

D = data.iloc[61:,5].values

X = data.iloc[61:,0]

Plt.plot(X,Y)
```

Output



CONCLUSION

The conclusion of COVID-19 vaccination efforts is that vaccines have played a crucial role in reducing the spread of the virus, lessening the severity of illness, and saving lives. Vaccination campaigns have been instrumental in managing the pandemic, but ongoing monitoring, research, and boosters may still be needed to maintain protection and address emerging variants. It's important to follow public health guidelines and stay informed about the latest developments.