

# COVID 19 VACCINE ANALYSIS

## PHASE 5: SUBMISSION DOCUMENT

### PROJECT : COVID 19 VACCINE ANALYSIS

#### INTRODUCTION:

- *The COVID 19 pandemic caused due to the Corona virus devastated the world by causing several fatalities around the world.*
- *This virus originated in Wuhan, China in 2019 and was later spread throughout the world due to human contact in one way or the other.*
- *symptoms like breathing problems caused by damage to the lungs.*

#### PROBLEM STATEMENT:

- *A correlation study to assess the knowledge and self-expressed stigma regarding COVID-19 Outbreak among adults at selected society of Pune city.*



*Design Thinking:*

*Data Collection: Collect Covid-19 vaccine data from reputable sources like health organizations, government databases, and research publications.*

*Data Preprocessing: Clean and preprocess the data, handle missing values, and convert categorical features into numerical representations.*

*Exploratory Data Analysis: Explore the data to understand its characteristics, identify trends, and outliers.*

*Statistical Analysis: Perform statistical tests to analyze vaccine efficiency, adverse effects, and distribution across different populations.*

*Visualization: Create visualizations (e.g., bar plots, line charts, heat maps) to present key findings and insights.*

*Insights and Recommendations: Provide actionable insights and recommendations based on the analysis to assist policymakers and health organizations.*

### **STEPS OF DATA:**

**Data Importing :**

- *In power BI desktop with the help of the get data option import the CSV data which is named as country vaccinations and clicked load option.*

## Data Cleaning :

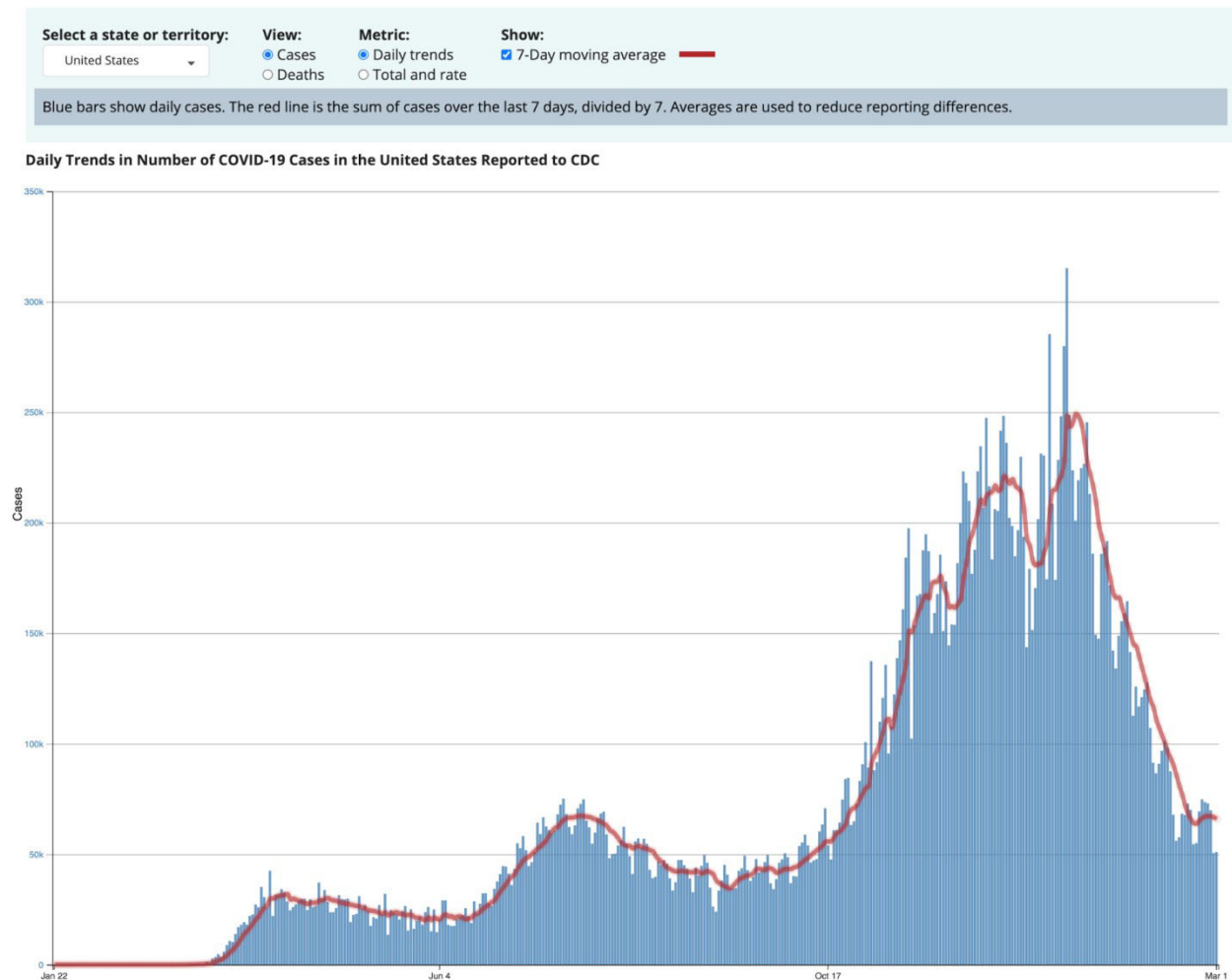
- *After loading the data and after analyzing the data | understood that there are 86512 rows and 15 columns.*
- *And in that some of the columns contained null values I have replaced the null values by 0 with the use of replace functions and started working on the data.*

## Visualizations :

In visualization part with the help of power BI desktop software I have used different kinds of charts, graphs, cards and table to display the data in the format which will be easy to understand.

## Analysis :

In the analysis part first | have analyzed the top 10 fully vaccinated countries by using area chart and have used the filter option to find the top countries and the result obtained as below,



From the below image we can able to come to know that India is the top country in terms full vaccination with 116 billion , followed by united states of America and china with 67 billion a35 billion respectively.

## IMPORTING LIBRARIES:

```
import numpy as np # linear algebra
```

```
import pandas as pd # data processing, CSV file
```

```
I/O (e.g. pd.read_csv)
```

```
import matplotlib.pyplot as plt
```

```
import seaborn as sns
```

```
import plotly.express as px
```

```
from plotly.offline import
```

```
download_plotlyjs,init_notebook_mode,plot,iplot
```

```
import plotly.graph_objects as go
```

```
import plotly.figure_factory as ff
```

```
from plotly.colors import n_colors
```

```
from wordcloud import
```

```
WordCloud,ImageColorGenerator
```

```
init_notebook_mode(connected=True)
```

```
from plotly.subplots import make_subplots
from pywaffle import Waffle
import warnings
warnings.filterwarnings('ignore')
```

## DATA ANALYSIS :

Here's a simplified step-by-step guide on how you might approach this:

### 1.Data collection:

- obtain a reliable dataset with covid19 vaccine information such as vaccination rates ,vaccine types and demographic data.

### 2. Data cleaning:

- remove duplicates and missing values Convert data types as needed Handle outliers and anomalies.

### 3.Descriptive statistics:

- Calculate basic statistics like means ,median ,standard deviation Create

summary tables and visualization to get a sense of the data.

#### 4. Visualization:

- Plot histogram for vaccine distributed  
Create bar charts to compare vaccine types or Manufactures.

#### 5. Correlation analysis:

- Use correlation coefficients and scatter plots to visualize these relationship.

#### 6. Time series analysis:

- If you have data over time analyze trends and seasonality in vaccination rates.

#### 7. Hypothesis testing:

- Test hypothesis related to vaccine efficacy.

#### 8. Geospatial analysis:



- If you have geographic data perform spatial analysis.

## 9. Machine learning:

- If you have enough data perform predictive models to forecast future vaccination rates.

## 10. Interpretation:

- Provide meaningful insights and conclusion from your Analysis.

country\_vaccinations - Microsoft Excel

country	iso_code	date	total_vacc	people_v	people_f	daily_vacc	daily_vacc	total_vacc	people_v	people_f	daily_vacc	vaccines	source_n	source_website
fghanistan	AFG	2/22/2021	0	0				1367	0	0		34	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	2/23/2021						1367				34	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	2/24/2021						1367				34	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	2/25/2021						1367				34	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	2/26/2021						1367				34	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	2/27/2021						1367				34	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	2/28/2021	8200	8200				1367	0.02	0.02		34	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	3/1/2021						1580				40	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	3/2/2021						1794				45	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	3/3/2021						2008				50	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	3/4/2021						2221				56	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	3/5/2021						2435				61	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	3/6/2021						2649				66	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	3/7/2021						2862				72	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	3/8/2021						2862				72	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	3/9/2021						2862				72	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	3/10/2021						2862				72	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	3/11/2021						2862				72	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	3/12/2021						2862				72	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	3/13/2021						2862				72	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	3/14/2021						2862				72	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	3/15/2021						2862				72	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	3/16/2021	54000	54000				2862	0.14	0.14		72	Johnson & World Health Organization	https://covid19.who.int/
fghanistan	AFG	3/17/2021						2882				72	Johnson & World Health Organization	https://covid19.who.int/

Dataset Link: <https://www.kaggle.com/datasets/gpreda/covid-world-vaccination-progress>

### Conclusion:

This is possible through a PHC approach that provides universal access to good-quality health services through empowered communities and multi-sectoral policy and action for health development.