

Covid Vaccine Analysis

Project



Infectious disease pandemics such as influenza (flu) and COVID-19 continue to have a major social and economic impact on affected populations.

Abstract

Objective The COVID-19 vaccination program in India started after the first wave of infections had almost subsided. In this work, the objective is to perform a statewise analysis to assess the impact of vaccination during the second COVID-19 wave in India. A total of 21 states are chosen for the analysis encompassing 97% of the Indian population.

Methods We use the generalized Gompertz curve to study the COVID-19 outbreak. The generalized Gompertz model is then modified to study the impact of vaccination. The modified model considers the cumulative daily number of individuals having the first and second shots of the vaccine in each state as explanatory variables.

Table of contents

- o Aim
- o Introduction
- o Problem statement
- o Methodology
- o Analysis (Data sheets pertaining to it)
- o Insights
- o Recommendations.
- o Conclusions

Aim:-

This project aims to analyze data in detail on vaccinated people and fully vaccinated by year and country.

Introduction:-

The main objective of this project is to analyze the data on, COVID-19 Vaccinations. Through the Analysis of data, we can find out some important insights.

Problem Statement:-

- We can find out how many people are completely vaccinated.
- Total vaccinations by each year and then by country. In this analysis, we can see the number of people vaccinated per hundred.

Methodology:-

Step 1: Data cleaning

In data cleaning I cleaned the data with the transform data option present in Power BI because the dataset was not cleaned and included some missing values, so with the help of the duplicates function in Power BI, I cleaned the data and replace the null values with 0 and then started working on it.

How to protect ourselves & others

9 important COVID-19 prevention measures



- 01** Stay home and self-isolate if you feel unwell, even with mild symptoms



- 02** Clean hands frequently with soap & water for 40 seconds or with alcohol-based hand rub



- 03** Cover your nose and mouth with a disposable tissue or flexed elbow when you cough or sneeze



- 04** Avoid touching your eyes, nose and mouth



- 05** Maintain a minimum physical distance of at least 1 metre from others



- 06** Stay away from crowds and avoid poorly ventilated indoor spaces



- 07** Use a fabric mask where physical distancing of at least 1 metre is not possible



- 08** Use a medical / surgical mask if you may be at higher risk (age, medical conditions)



- 09** Regularly clean & disinfect frequently touched surfaces

Step 2: Data interpretation

In this data interpretation step, I found out some important information about the data set.

- The data set was almost clear and completed.
- I replaced the null values with 0.
- There is total of 86512 rows and 15 columns.

Step 3: Visualization

In visualization, I took the help of Power BI Desktop software to make graphs and charts here some relevant graphs and charts are attached.

Analysis (Data sheets pertaining to it):-

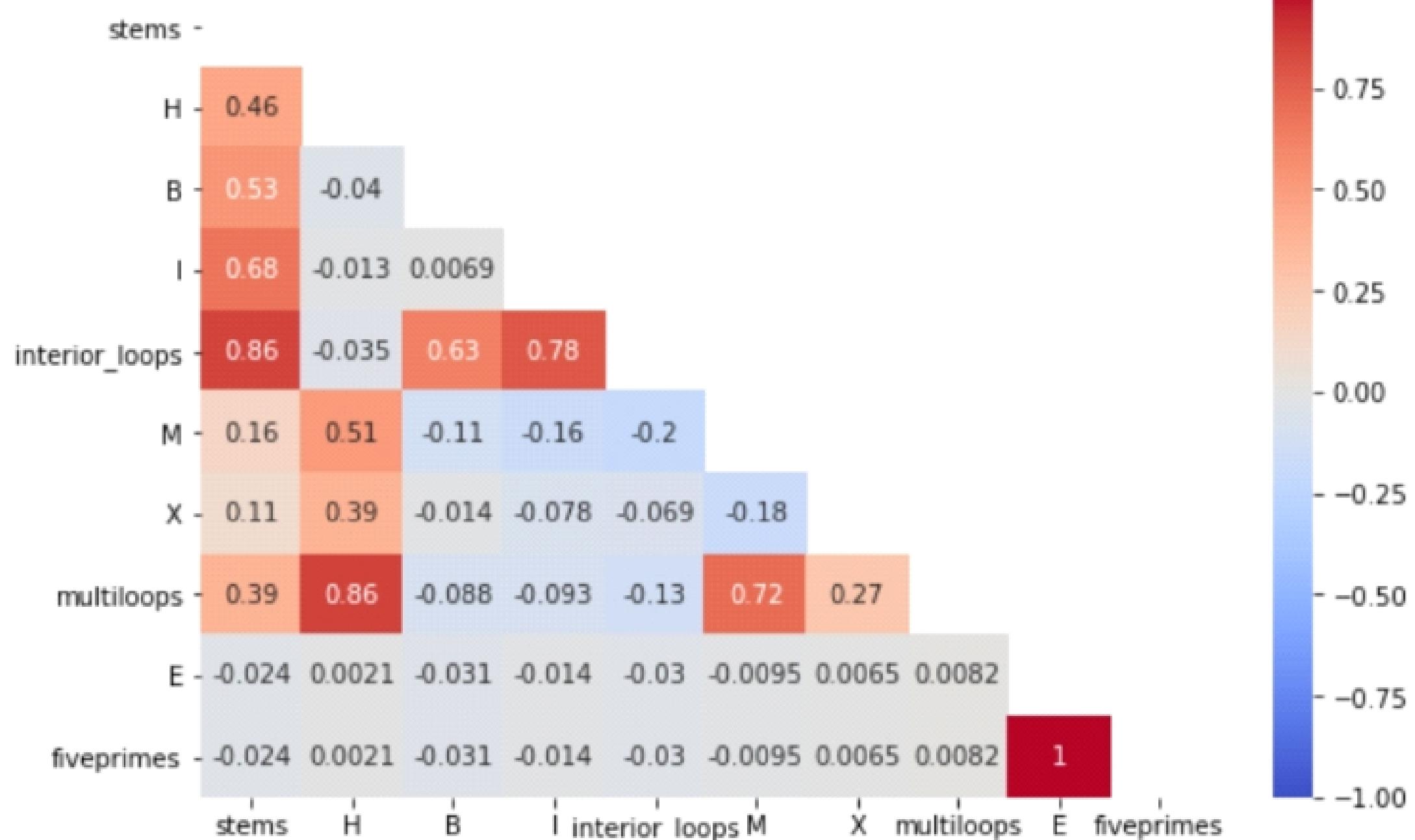
In observation found some interesting points that are mentioned below.

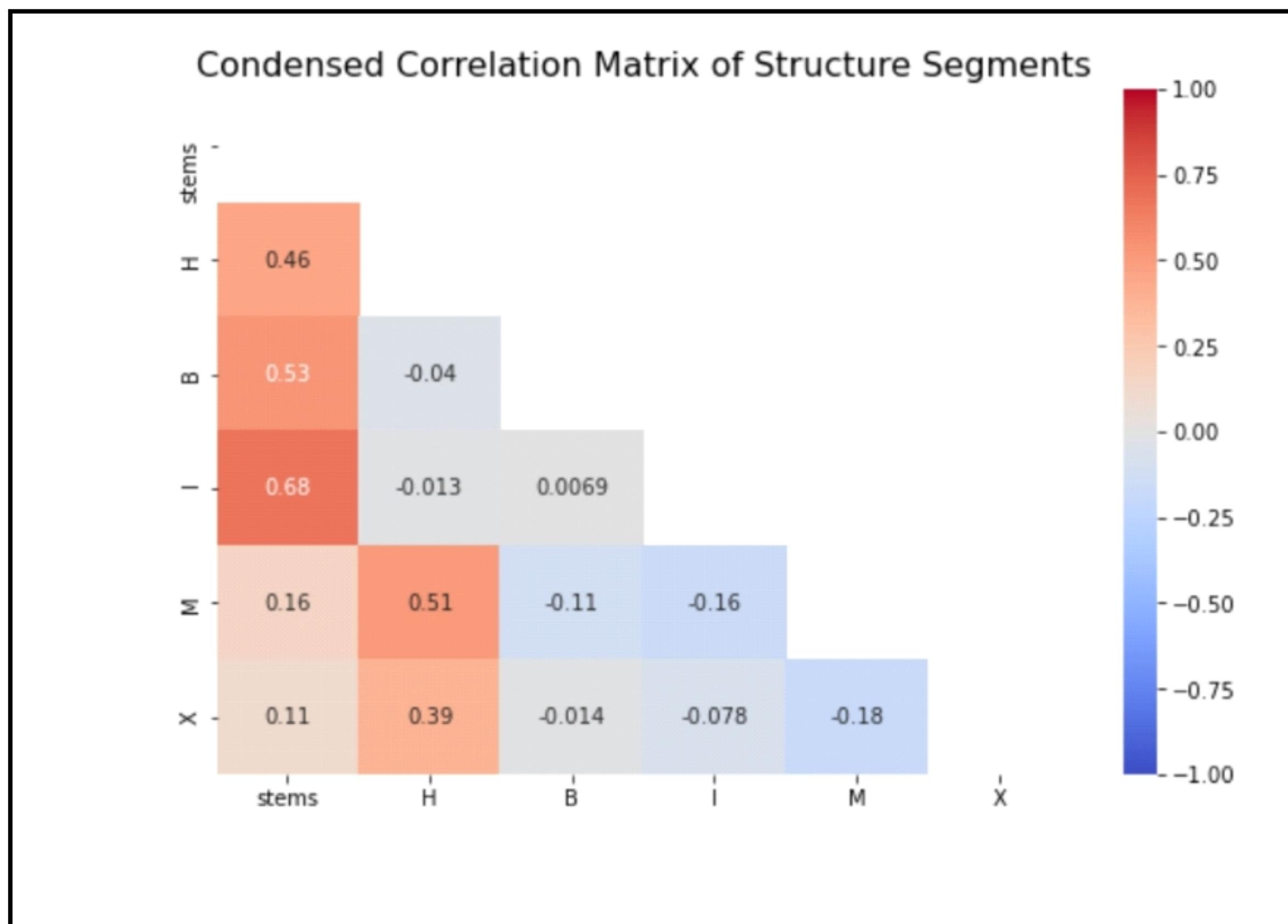
- According to data on people fully vaccinated in the year.
- The dataset available is from 2020 to 2021
- Total vaccination by the country in map visualization.
- Total Highest Daily vaccination is 9223133948 in 2021.
- Total Highest vaccination is 704952663700 in China.
- Highest number of vaccinations by the source name is the Ministry of Health is 7.41K.
- Top 10 countries with fully vaccinated:
- Top 10 Daily vaccination in the countries:

Top 10 countries with the top 10 vaccine

Top 10 Countries where most people

Correlation Matrix of Structure Segments

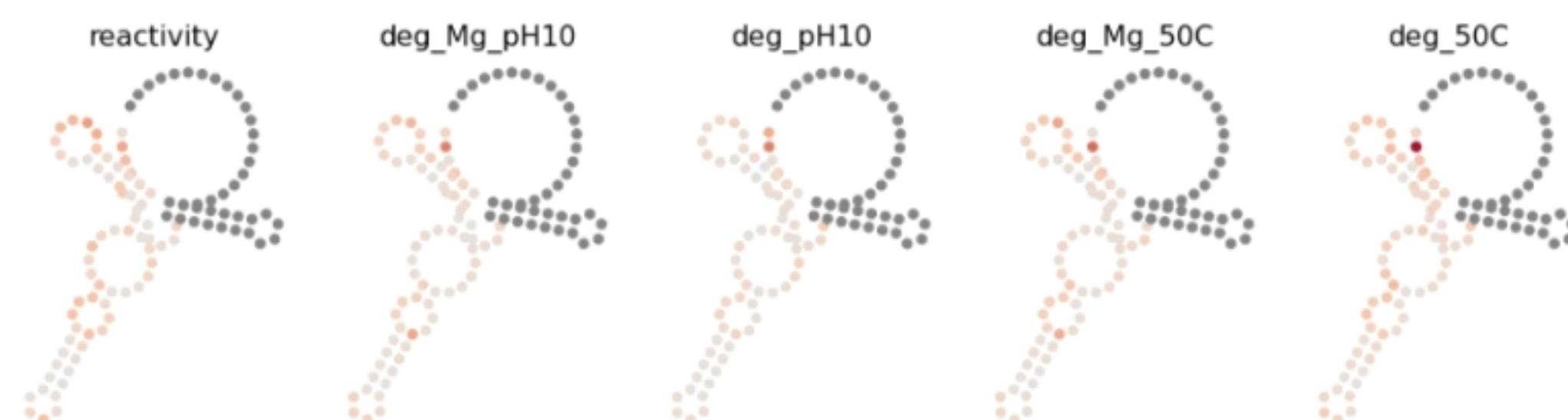




Insights:-

- In conclusion, we can take a look at the final Report for further analysis.
- We can see the top 10 sources.
- Countries by fully vaccination etc.

OpenVaccine: EDA & Feature Engineering with Forgi



reactivity deg_Mg_pH10 deg_pH10 deg_Mg_50C deg_50C



reactivity deg_Mg_pH10 deg_pH10 deg_Mg_50C deg_50C



reactivity deg_Mg_pH10 deg_pH10 deg_Mg_50C deg_50C

Current global situation

- 75 million cases
- 5 countries with highest cumulative number of cases
 1. United States of America
 2. India
 3. Brazil
 4. Russian Federation
 5. France

- 1,68 million deaths
- 5 countries with highest cumulative number of death
 1. United States of America
 2. Brazil
 3. India
 4. Mexico
 5. Italy



Global level

Availability: One of the most critical challenges for LMICs currently is to get sufficient supplies of vaccine and to sustain the level of supplies needed. In September last year, Oxfam reported that the richer nations - with just 13% of the world's population - secured more than half of the supplies of the five leading vaccine candidates at that time. The COVAX Vaccine Global Access Facility (COVAX) has been established to ensure rapid, fair and equitable distribution of the vaccine across the globe. The success of COVAX will be a key determinant for ensuring availability of vaccines in LMICs with 67 low income countries relying on COVAX for their vaccine supplies.

Affordability: The price of vaccines will certainly be a major factor in countries' ability to roll out mass vaccination. Prices of the approved vaccines vary widely, reported as ranging from USD 6 to USD 74 per dose. COVAX alone will not be able to meet the demand of LMICs for COVID-19 vaccines and countries' ability and affordability to secure additional vaccine, possibility at a higher cost, will be critical for vaccine roll-out, especially at the early stage.

National level

Political commitment: Strong political commitment is crucial for the success of a massive initiative such as the roll-out of Covid-19 vaccines. A recent panel discussion of members of parliaments from several countries, including LMICs, led by Lord Boateng of the UK Parliament identified cross-party, financial, legal, civil-society and grassroots commitments as the cornerstones for success of national immunization. These domestic political commitments are equally, if not more,

Individual level

Vaccine hesitancy: Covid-19 vaccination hesitancy is strong and widespread in many parts of the world. Thankfully, prior evidences suggests that in LMICs vaccine coverage is mainly correlated with supply side factors such as vaccine availability, whereas in high income countries it is more correlated with personal belief. However, increasing rumours linked to social and digital media are leading to growing mistrust of vaccines and vaccine trials in LMICs. Effective and targeted community engagement and communication campaigns can help in mitigating vaccine hesitancy in countries where the risk is substantial.

Documentation:

1. Introduction:

- Briefly introduce the project, its goals, and the context of COVID-19.

2. Data Collection:

- Describe the data sources you used, including their reliability and update frequency.

3. Data Preprocessing:

- Explain how you cleaned and prepared the data for analysis.

4. Analysis Methods:

- Detail the statistical and analytical methods you used, including any models or algorithms.

5. Results and Insights:

- Present your findings with clear visualizations and explanations.

6. Discussion:

- Interpret the results and discuss their implications. Consider the limitations of your analysis.

python code integration

```
In [2]: import pandas as pd

In [16]: import pandas as pd
df = pd.read_csv(r"C:\Users\Machines\Downloads\country_vaccinations.csv\country_vaccinations.csv")
print(df)

      country iso_code     date  total_vaccinations \
0    Afghanistan    AFG 2021-02-22              0.0
1    Afghanistan    AFG 2021-02-23            NaN
2    Afghanistan    AFG 2021-02-24            NaN
3    Afghanistan    AFG 2021-02-25            NaN
4    Afghanistan    AFG 2021-02-26            NaN
...
86507   Zimbabwe    ZWE 2022-03-25  8691642.0
86508   Zimbabwe    ZWE 2022-03-26  8791728.0
86509   Zimbabwe    ZWE 2022-03-27  8845039.0
86510   Zimbabwe    ZWE 2022-03-28  8934360.0
86511   Zimbabwe    ZWE 2022-03-29  9039729.0

  people_vaccinated  people_fully_vaccinated  daily_vaccinations_raw \
0                  0.0                      NaN                      NaN
1                  NaN                      NaN                      NaN
2                  NaN                      NaN                      NaN
3                  NaN                      NaN                      NaN
4                  NaN                      NaN                      NaN
...
86507    4814582.0            3473523.0            139213.0
86508    4886242.0            3487962.0            100086.0
86509    4918147.0            3493763.0            53311.0
86510    4975433.0            3501493.0            89321.0
86511    5053114.0            3510256.0            105369.0

  daily_vaccinations  total_vaccinations_per_hundred \
0                  NaN                      0.00
1                 1367.0                      NaN
2                 1367.0                      NaN
3                 1367.0                      NaN
4                 1367.0                      NaN
...
86507    69579.0                     57.59
86508    83429.0                     58.25
86509    90629.0                     58.61
86510   100614.0                     59.20
86511   103751.0                     59.90

  people_vaccinated_per_hundred  people_fully_vaccinated_per_hundred \
0                      0.00                      NaN
1                      NaN                      NaN
```

```
86507    Zimbabwe    ZWE 2022-03-25      8691642.0
86508    Zimbabwe    ZWE 2022-03-26      8791728.0
86509    Zimbabwe    ZWE 2022-03-27      8845039.0
86510    Zimbabwe    ZWE 2022-03-28      8934360.0
86511    Zimbabwe    ZWE 2022-03-29      9039729.0

    people_vaccinated  people_fully_vaccinated  daily_vaccinations_raw \
0                  0.0                      0.0          0.0
1                  0.0                      0.0          0.0
2                  0.0                      0.0          0.0
3                  0.0                      0.0          0.0
4                  0.0                      0.0          0.0
...
86507      4814582.0            3473523.0        139213.0
86508      4886242.0            3487962.0        100086.0
86509      4918147.0            3493763.0        53311.0
86510      4975433.0            3501493.0        89321.0
86511      5053114.0            3510256.0        105369.0

    daily_vaccinations  total_vaccinations_per_hundred \
0                  0.0                      0.00
1                 1367.0                     0.00
2                 1367.0                     0.00
3                 1367.0                     0.00
4                 1367.0                     0.00
...
86507      69579.0                      57.59
86508      83429.0                      58.25
86509      90629.0                      58.61
86510     100614.0                      59.20
86511     103751.0                      59.90

    people_vaccinated_per_hundred  people_fully_vaccinated_per_hundred \
0                          0.00                      0.00
1                          0.00                      0.00
2                          0.00                      0.00
3                          0.00                      0.00
4                          0.00                      0.00
...
86507      31.90                      23.02
86508      32.38                      23.11
86509      32.59                      23.15
86510      32.97                      23.20
86511      33.48                      23.26

    daily_vaccinations_per_million \
0                      0.0
1                      34.0
2                      34.0
3                      34.0
```

```
0 https://covid19.who.int/
1 https://covid19.who.int/
2 https://covid19.who.int/
3 https://covid19.who.int/
4 https://covid19.who.int/
...
86507 https://www.arcgis.com/home/webmap/viewer.html...
86508 https://www.arcgis.com/home/webmap/viewer.html...
86509 https://www.arcgis.com/home/webmap/viewer.html...
86510 https://www.arcgis.com/home/webmap/viewer.html...
86511 https://www.arcgis.com/home/webmap/viewer.html...
86512 https://www.arcgis.com/home/webmap/viewer.html...
```

[86512 rows x 15 columns]

In [17]: `df.head()`

Out[17]:

	country	iso_code	date	total_vaccinations	people_vaccinated	people_fully_vaccinated	daily_vaccinations_raw	daily_vaccinations	total_vaccinations_per_100k
0	Afghanistan	AFG	2021-02-22	0.0	0.0	NaN	NaN	NaN	NaN
1	Afghanistan	AFG	2021-02-23	NaN	NaN	NaN	NaN	1367.0	
2	Afghanistan	AFG	2021-02-24	NaN	NaN	NaN	NaN	1367.0	
3	Afghanistan	AFG	2021-02-25	NaN	NaN	NaN	NaN	1367.0	
4	Afghanistan	AFG	2021-02-26	NaN	NaN	NaN	NaN	1367.0	

In [23]: `#filling missing values
df.fillna(0, inplace=True)
print(df.head())`

```
country iso_code      date total_vaccinations people_vaccinated \
0 Afghanistan AFG 2021-02-22          0.0          0.0
1 Afghanistan AFG 2021-02-23          0.0          0.0
2 Afghanistan AFG 2021-02-24          0.0          0.0
3 Afghanistan AFG 2021-02-25          0.0          0.0
4 Afghanistan AFG 2021-02-26          0.0          0.0

people_fully_vaccinated daily_vaccinations_raw daily_vaccinations \
0                      0.0                      0.0          0.0
```

```
4 Afghanistan      AFG  2021-02-26      0.0      0.0
   people_fully_vaccinated  daily_vaccinations_raw  daily_vaccinations \
0           0.0                  0.0                  0.0
1           0.0                  0.0                1367.0
2           0.0                  0.0                1367.0
3           0.0                  0.0                1367.0
4           0.0                  0.0                1367.0

   total_vaccinations_per_hundred  people_vaccinated_per_hundred \
0                   0.0                  0.0
1                   0.0                  0.0
2                   0.0                  0.0
3                   0.0                  0.0
4                   0.0                  0.0

   people_fully_vaccinated_per_hundred  daily_vaccinations_per_million \
0                           0.0                  0.0
1                           0.0                 34.0
2                           0.0                 34.0
3                           0.0                 34.0
4                           0.0                 34.0

   vaccines \
0 Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
1 Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
2 Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
3 Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
4 Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...

   source_name          source_website
0 World Health Organization  https://covid19.who.int/
1 World Health Organization  https://covid19.who.int/
2 World Health Organization  https://covid19.who.int/
3 World Health Organization  https://covid19.who.int/
4 World Health Organization  https://covid19.who.int/
```

```
In [24]: #Processing Data
#eliminating missing value
print(df.dropna())
```

```
country iso_code      date  total_vaccinations \
0    Afghanistan      AFG  2021-02-22      0.0
1    Afghanistan      AFG  2021-02-23      0.0
2    Afghanistan      AFG  2021-02-24      0.0
3    Afghanistan      AFG  2021-02-25      0.0
4    Afghanistan      AFG  2021-02-26      0.0
...
86507    Zimbabwe      ZWE  2022-03-25  8691642.0
86508    Zimbabwe      ZWE  2022-03-26  8791728.0
86509    Zimbabwe      ZWE  2022-03-27  8845039.0
```

```
1                      34.0
2                      34.0
3                      34.0
4                      34.0
...
86507                  ...
86508                  4610.0
86509                  5528.0
86510                  6005.0
86511                  6667.0
86511                  6874.0

                           vaccines \
0      Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
1      Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
2      Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
3      Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
4      Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
...
86507  Oxford/AstraZeneca, Sinopharm/Beijing, Sinovac...
86508  Oxford/AstraZeneca, Sinopharm/Beijing, Sinovac...
86509  Oxford/AstraZeneca, Sinopharm/Beijing, Sinovac...
86510  Oxford/AstraZeneca, Sinopharm/Beijing, Sinovac...
86511  Oxford/AstraZeneca, Sinopharm/Beijing, Sinovac...

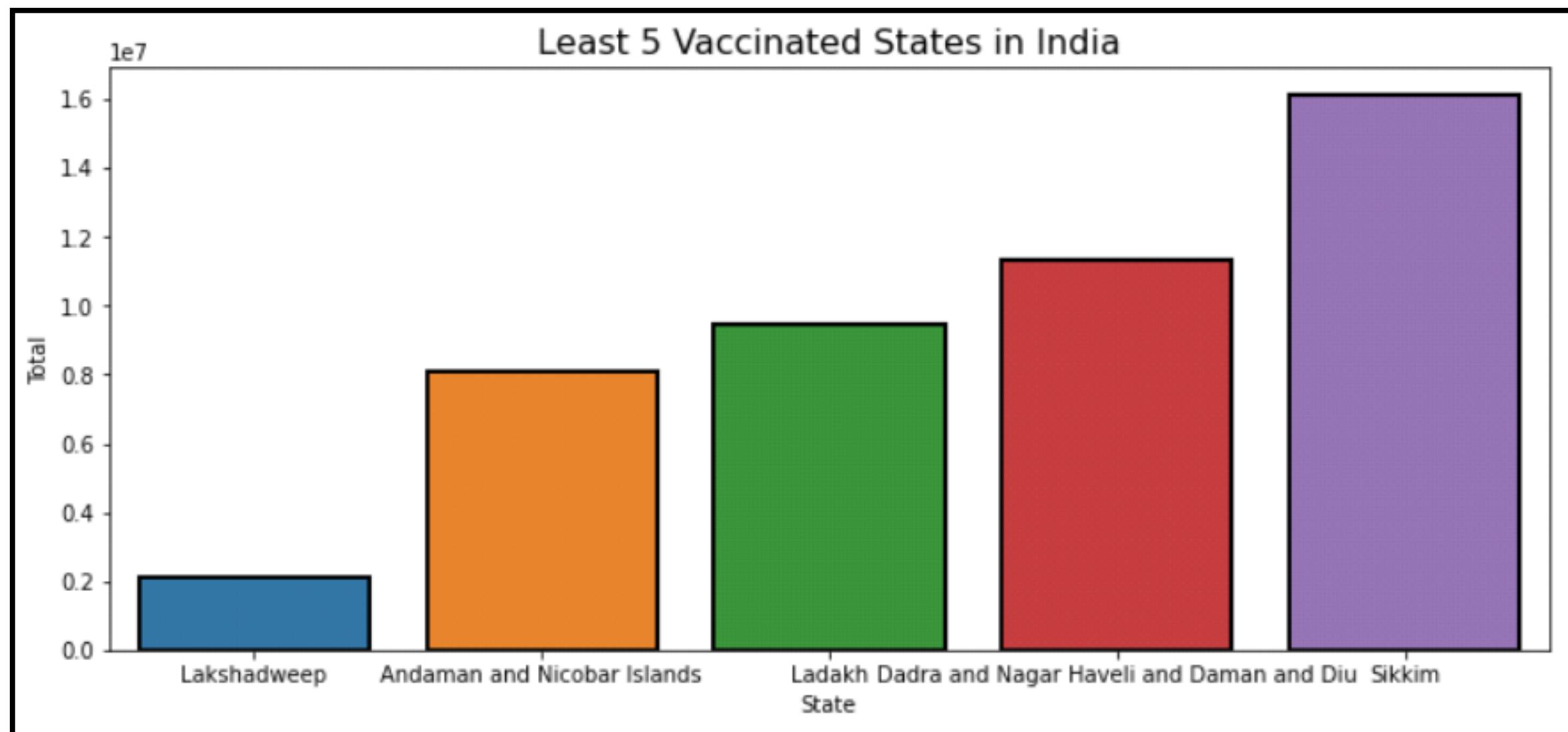
                           source_name \
0      World Health Organization
1      World Health Organization
2      World Health Organization
3      World Health Organization
4      World Health Organization
...
86507      ...
86508      ...
86509      ...
86510      ...
86511      ...

                           source_website
0          https://covid19.who.int/
1          https://covid19.who.int/
2          https://covid19.who.int/
3          https://covid19.who.int/
4          https://covid19.who.int/
...
86507  https://www.arcgis.com/home/webmap/viewer.html...
86508  https://www.arcgis.com/home/webmap/viewer.html...
86509  https://www.arcgis.com/home/webmap/viewer.html...
86510  https://www.arcgis.com/home/webmap/viewer.html...
86511  https://www.arcgis.com/home/webmap/viewer.html...
```

```

      people_vaccinated_per_hundred  people_fully_vaccinated_per_hundred \
0                           0.00                               NaN
1                           NaN                                NaN
2                           NaN                                NaN
3                           NaN                                NaN
4                           NaN                                NaN
...
86507                      ...
86508                      31.90                            23.02
86509                      32.38                            23.11
86510                      32.59                            23.15
86511                      32.97                            23.20
                                         ...
86507          daily_vaccinations_per_million \
0                           NaN
1                           34.0
2                           34.0
3                           34.0
4                           34.0
...
86507
86508
86509
86510
86511
                                         ...
86507          vaccines \
0   Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
1   Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
2   Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
3   Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
4   Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
...
86507  Oxford/AstraZeneca, Sinopharm/Beijing, Sinovac...
86508  Oxford/AstraZeneca, Sinopharm/Beijing, Sinovac...
86509  Oxford/AstraZeneca, Sinopharm/Beijing, Sinovac...
86510  Oxford/AstraZeneca, Sinopharm/Beijing, Sinovac...
86511  Oxford/AstraZeneca, Sinopharm/Beijing, Sinovac...
                                         ...
86507          source_name \
0   World Health Organization
1   World Health Organization
2   World Health Organization
3   World Health Organization
4   World Health Organization
...
86507
86508
86509
86510
86511
                                         ...
86507
86508
86509
86510
86511

```



```
df.people_vaccinated = df.people_vaccinated.astype(int)
df.people_fully_vaccinated = df.people_fully_vaccinated.astype(int)
df.daily_vaccinations_raw = df.daily_vaccinations_raw.astype(int)
df.daily_vaccinations = df.daily_vaccinations.astype(int)
df.total_vaccinations_per_hundred = df.total_vaccinations_per_hundred.astype(int)
df.people_fully_vaccinated_per_hundred =
df.people_fully_vaccinated_per_hundred.astype(int)
df.daily_vaccinations_per_million = df.daily_vaccinations_per_million.astype(int)
df.people_vaccinated_per_hundred = df.people_vaccinated_per_hundred.astype(int)
date = df.date.str.split('-', expand =True)
date
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4568 entries, 0 to 4567
Data columns (total 15 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   country          4568 non-null    object  
 1   iso_code          4260 non-null    object  
 2   date              4568 non-null    object  
 3   total_vaccinations 2988 non-null    float64
 4   people_vaccinated 2541 non-null    float64
 5   people_fully_vaccinated 1702 non-null    float64
 6   daily_vaccinations_raw 2523 non-null    float64
 7   daily_vaccinations 4409 non-null    float64
 8   total_vaccinations_per_hundred 2988 non-null    float64
 9   people_vaccinated_per_hundred 2541 non-null    float64
 10  people_fully_vaccinated_per_hundred 1702 non-null    float64
 11  daily_vaccinations_per_million 4409 non-null    float64
 12  vaccines          4568 non-null    object  
 13  source_name        4568 non-null    object  
 14  source_website     4568 non-null    object  
dtypes: float64(9), object(6)
memory usage: 535.4+ KB
```

Incubation period of covid-19 from exposure start to symptoms onset [20]:

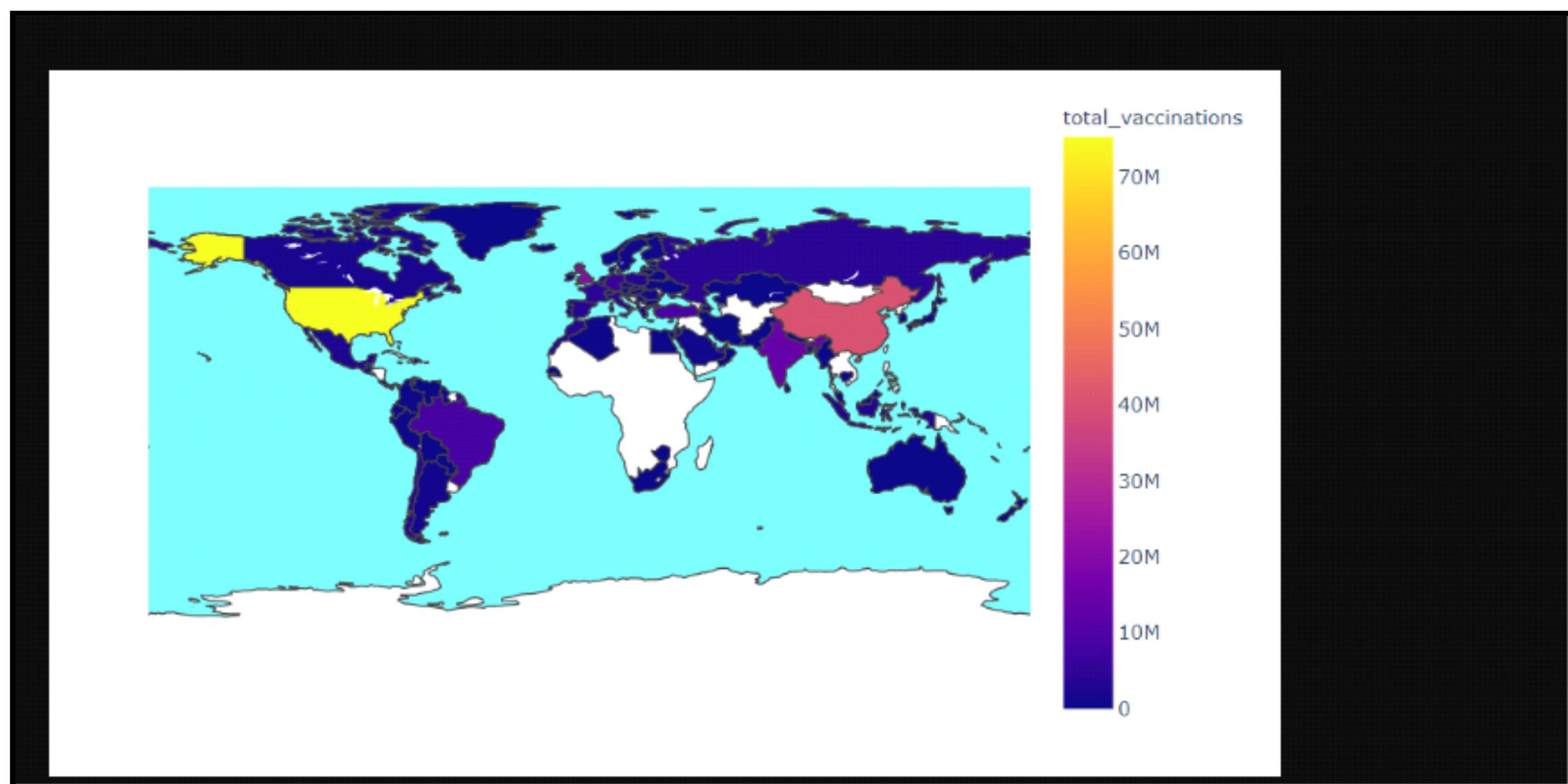
Days	Cases
1	4
2	5
3	3
4	11
5	11
6	9
7	5
8	8
9	7
10	4
11	2
12	2
13	2
14	0
Total	73

Incubation period of covid-19 from symptom onset to hospital visit [20]:

Days	Cases
1	221
2	48
3	41
4	30
5	27
6	17
7	20
8	17
9	2
10	3
11	2
12	18
13	1
14	5
Total	452

Infected people of different countries according to temperature domain [12,23,24]:

Country	Temperature domain	Temperature range (°C)	Average yearly temperature (°C)	Confirmed cases
Switzerland	Cold Temperate	(0-10)	5.50	24,551
China	Mix of Warm, Cold and Polar Temperate	(10-18) or (0-10)	6.95	81,953
UK	Cold Temperate	(0-10)	8.45	73,758
Germany	Cold Temperate	(0-10)	8.50	122,171
USA	Mostly Warm Temperate, and Cold Temperate	(10-18) or (0-10)	8.55	502,876
Belgium	Cold Temperate	(0-10)	9.55	26,667
Spain	Warm Temperate	(10-18)	13.30	158,273
Italy	Warm Temperate	(10-18)	13.45	147,577
Iran	Sub Tropical, and Warm temperate	(18-24) or (10-18)	17.25	68,192
South Africa	Mostly Warm Temperate	(10-18)	17.75	2003
India	Tropical Temperate	(24-34)	23.65	7,600
Saudi Arabia	Tropical, and Sub Tropical Temperate	(24-34) or (18-24)	24.65	3651
Oman	Tropical Temperate	(24-34)	25.60	484
Sudan	Tropical, and Sub Tropical Temperate	(24-34) or (18-24)	26.90	17



`describe()` function in pandas used to get the statistics of each feature present in our dataset. Some of the information we get include count, max, min, standard deviation, median, etc.

```
df.describe()
```

Covid Vaccination Progress | describe dataset

`unique()` function in pandas helps to get unique values present in the feature.

```
df.country.unique()
```

Unique country values

```
def size(m,n):
```

```
    fig = plt.gcf();
```

```
    fig.set_size_inches(m,n);
```

Word Art of Countries

Word Cloud is a unique way to get information from our dataset. The words are shown in the form of art where the size proportional depends on how much the particular word repeated in the dataset. This is made by using the WordCloud library. Check the below code on how to draw word cloud

```
wordCloud = WordCloud(
```

```
background_color='white',
max_font_size = 50).generate(' '.join(df.country))

plt.figure(figsize=(15,7))

plt.axis('off')

plt.imshow(wordCloud)

plt.show()
```

Covid Vaccination Progress | wordart countries

Total Vaccinated Till Date

In this section, we are going to see how many total vaccines have been used in each country. Check the below code for more information. The data shows the United States has administrated most vaccines in the world followed by China, United Kingdom, England, India and at the last some countries includes Saint Helena, San Marino has 0 vaccination.

```
country_wise_total_vaccinated = {}

for country in df.country.unique():

    vaccinated = 0

    for i in range(len(df)):

        if df.country[i] == country:

            vaccinated += df.daily_vaccinations[i]

    country_wise_total_vaccinated[country] = vaccinated

# made a seperate dict from the df

country_wise_total_vaccinated_df = 

pd.DataFrame.from_dict(country_wise_total_vaccinated,
orient='index',

columns = ['total_vaccinated_till_date'])

# converted dict to df

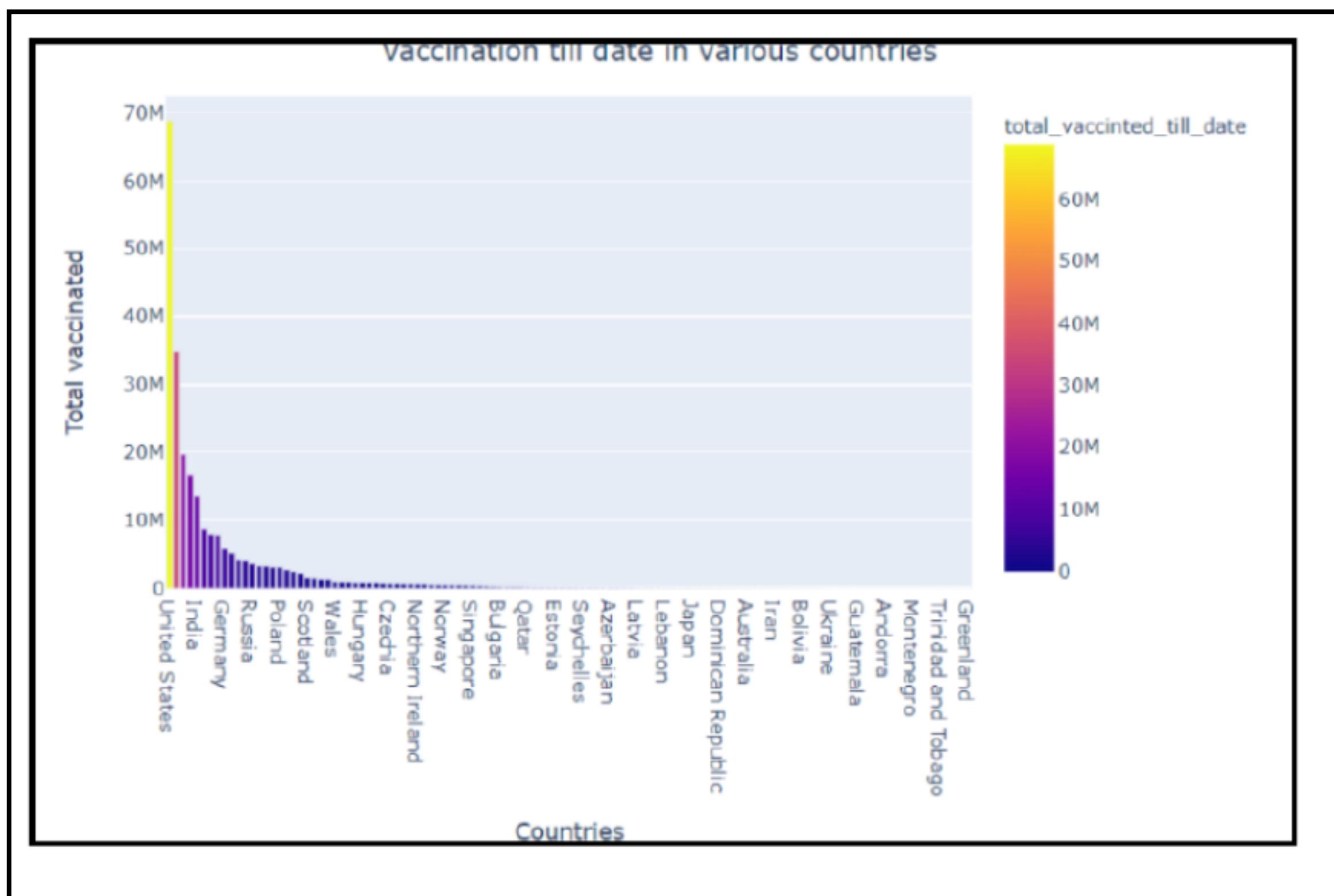
country_wise_total_vaccinated_df.sort_values(by = 'total_vaccinated_till_date',
```

```
ascending = False, inplace = True)

country_wise_total_vaccinated_df

Total vaccinated

fig = px.bar(country_wise_total_vaccinated_df,
y = 'total_vaccinted_till_date',
x = country_wise_total_vaccinated_df.index,
color = 'total_vaccinted_till_date',
color_discrete_sequence= px.colors.sequential.Viridis_r
)
fig.update_layout(
title={
'text': "Vaccination till date in various countries",
'y':0.95,
'x':0.5
},
xaxis_title="Countries",
yaxis_title="Total vaccinated",
legend_title="Total vaccinated"
)
fig.show()
```



Recommendations:-

- We can collect good datasets for analysis.
- We can analyze the data of the overall world.
- Like this dataset we can perform operations with various categories, city-wise or region-wise.

Conclusion:-

In conclusion, we can take a look at the final Report for further analysis.

As we can see in the Report the top countries vaccination-wise, Top Country by the vaccination in daily and best top 10 countries with our vaccines. Vaccination peaks in 2022. The Top 2 vaccine sources are India, and China.

