Covid 19 Vaccine Analysis Project

BATCH MEMBERS:

M.POTHINI

S.SHANOFAR BEGUM

P.SUBAMOZHI

R.VIJAYASHREE

Aim

The aim of this project is to analyze the data in detail on vaccinated people and fully vaccinated people by country in 2021-2022

Introduction

Since the start of the COVID-19 outbreak, WHO has worked with countries and areas in Different regions on public health measures to slow or stop the spread of the virus.

Safe and effective vaccines are an important tool, in combination with other measures, to protect people against COVID-19, save lives and reduce widescale social disruption.

The main objective of this project to analyze Covid – 19 Vaccination data and find out some important insights.

Problem Statement.

In this project, we find out in Different Countries how many people are Vaccinated or fully Vaccinated in 2021 and 2022 and also find out how many people in different countries have completed Daily Vaccinations, Total Vaccination per hundred, People Vaccination per hundred, and Daily vaccinations per Million.

Methodology

- In the Country, Vaccination data set the first column is that Country in country column presents different countries.
- In the iso code column containing the iso code for the country. and In the data set contain a date for Vaccination.
- Another column is total Vaccination, People Vaccination, People fully Vaccination, Daily Vaccinated raw, Daily Vaccinations, Total Vaccinations per hundred, people Vaccination

per hundred, People fully Vaccinated per Hundred, and Daily Vaccination per million in all columns Show how many people are Vaccinated in different criteria.

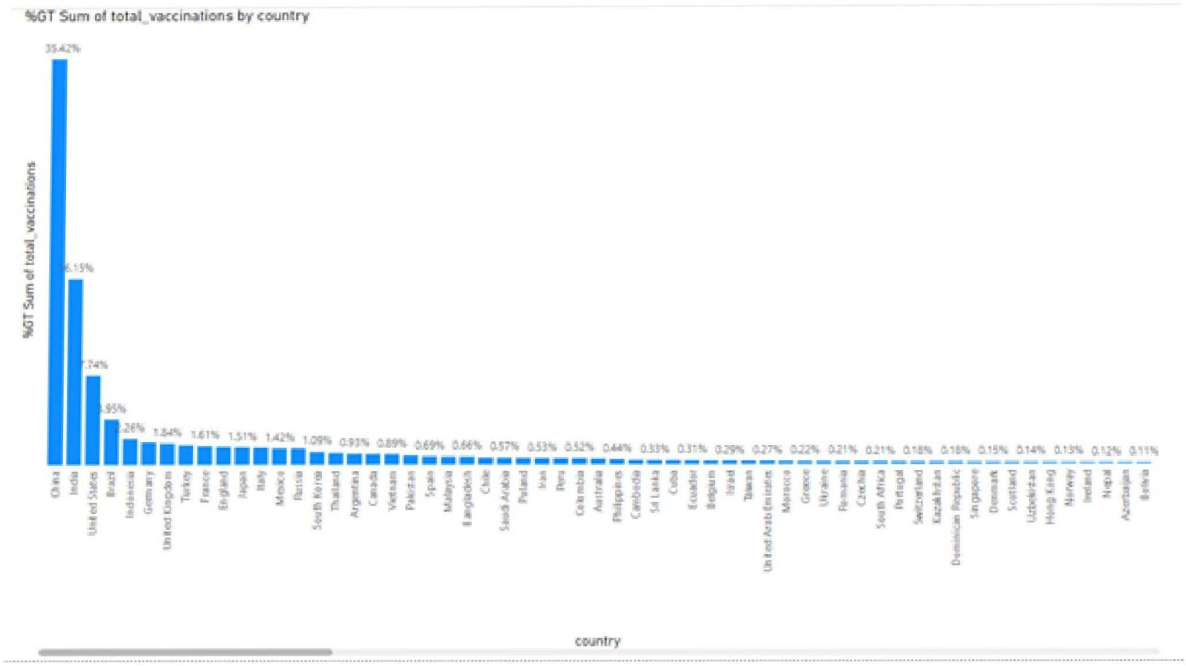
- The first step is cleaning the data. In data cleaning I cleaned the data with the transform data option present in Power BI because the dataset was not cleaned and included some missing values, so with the help of the duplicates function in Power BI, I cleaned the data and replace the null values with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

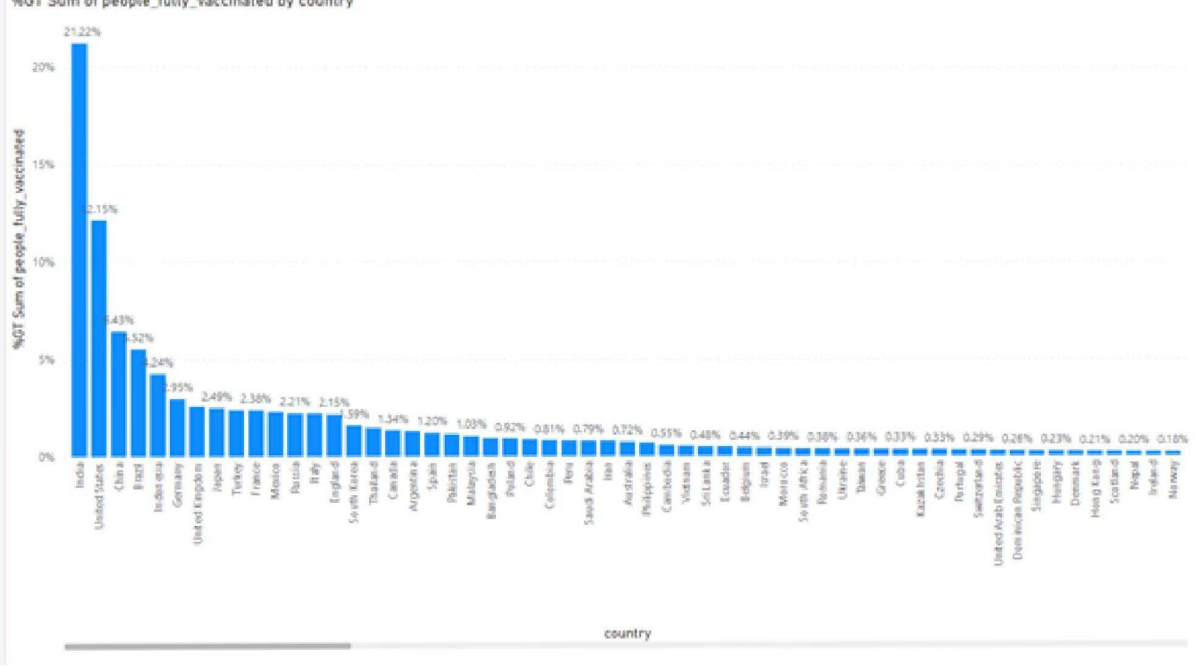
- After Data cleaning performs the Data interpretation step. In Data interpretation found some important information.

- The data set was almost clean and completed.
- Replace the Null values with 0
- There is a total of 8513 rows and 15 columns is present.

Visualization

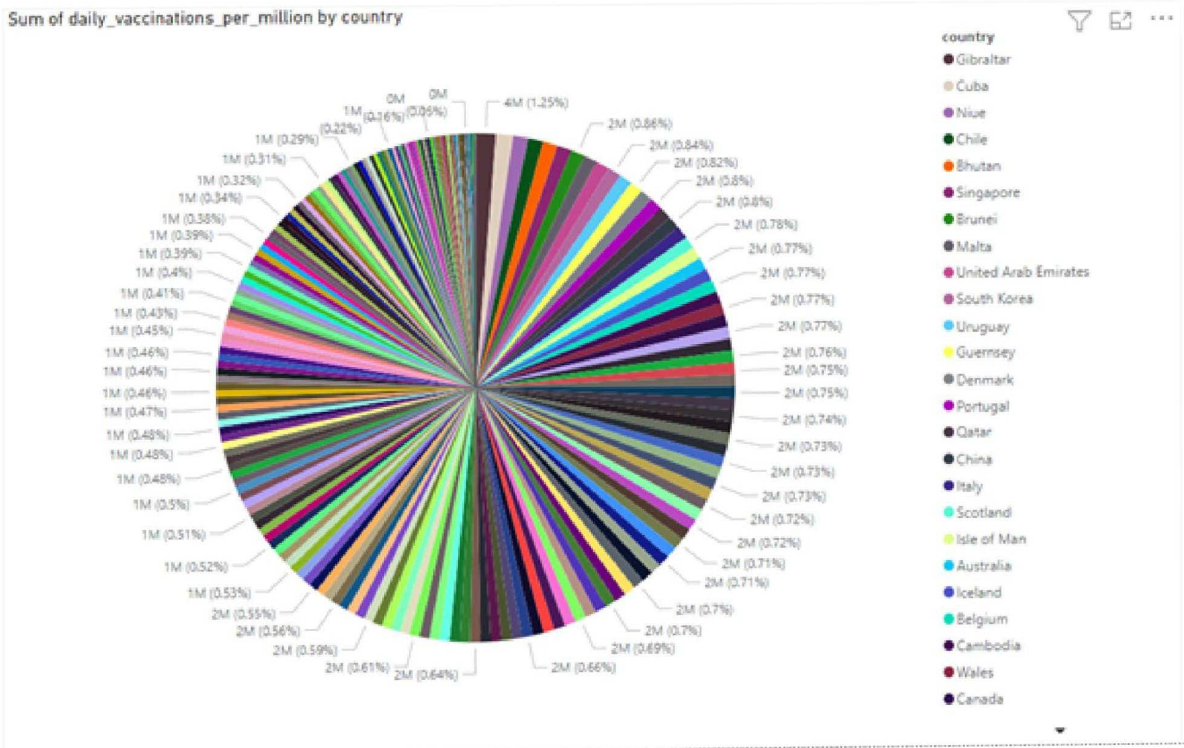
In visualization, I took the help of Power BI Desktop software to make graphs and charts here some relevant graphs and charts are attached.




%GT Sum of people_fully_vaccinated by country
%GT Sum of people_fully_vaccinated
21.22%
12.15%
6.43%
5.52%
4.24%
2.95%
2.49%
2.38%
2.21%
2.15%
1.59%
1.34%
1.20%
1.03%
0.92%
0.81%
0.79%
0.72%
0.55%
0.48%
0.44%
0.39%
0.38%
0.36%
0.33%
0.33%
0.29%
0.26%
0.23%
0.21%
0.20%
0.16%
20%
15%
10%
5%
0%
India
United States
China
Brazil
Indonesia
Germany
United Kingdom
Japan
Turkey
France
Mexico
Russia
Italy
England
South Korea
Thailand
Canada
Argentina
Spain
Pakistan
Malaysia
Bangladesh
Poland
Chile
Colombia
Peru
Saudi Arabia
Iran
Australia
Philippines
Cambodia
Vietnam
Sri Lanka
Ecuador
Belgium
Israel
Morocco
South Africa
Romania
Ukraine
Taiwan
Greece
Cuba
Kazakhstan
Czechia
Portugal
Switzerland
United Arab Emirates
Dominican Republic
Singapore
Hungary
Denmark
Hong Kong
Scotland
Nepal
Ireland
Norway
country


country
Afghanistan
Albania
Algeria
Andorra
Angola
Anguilla
Antigua and Barbuda
Argentina
Armenia
Aruba
Australia
Austria
Azerbaijan
Bahamas
Bahrain
Bangladesh
Barbados
Belarus
Belgium
Belize
Benin
Bermuda
Bhutan
Bolivia
Bonaire Sint Eustatius and Saba
Bosnia and Herzegovina
Botswana
Brazil


223
Count of country


Sum of daily_vaccinations_per_million by country
country
Gibraltar
Cuba
Niue
Chile
Bhutan
Singapore
Brunei
Malta
United Arab Emirates
South Korea
Uruguay
Guernsey
Denmark
Portugal
Qatar
China
Italy
Scotland
Isle of Man
Australia
Iceland
Belgium
Cambodia
Wales
Canada

```
data.vaccines.value_counts()
```

```
Moderna, Oxford/AstraZeneca, Pfizer/BioNTech
2587
Oxford/AstraZeneca
1673
Pfizer/BioNTech
1416
Oxford/AstraZeneca, Pfizer/BioNTech
845
Pfizer/BioNTech, Sinovac
475
Moderna, Pfizer/BioNTech
407
Sputnik V
351
Oxford/AstraZeneca, Sinovac
301
Oxford/AstraZeneca, Sinopharm/Beijing
268
Oxford/AstraZeneca, Sinopharm/Beijing, Sputnik V
235
Pfizer/BioNTech, Sinopharm/Beijing
208
Oxford/AstraZeneca, Pfizer/BioNTech, Sinopharm/Beijing, Sputnik V
202
Sinopharm/Beijing
186
Oxford/AstraZeneca, Pfizer/BioNTech, Sinovac
123
Sinopharm/Beijing, Sinopharm/Wuhan, Sinovac
117
EpiVacCorona, Sputnik V
117
Johnson&Johnson, Moderna, Pfizer/BioNTech
112
Oxford/AstraZeneca, Pfizer/BioNTech, Sinovac, Sputnik V
108
Moderna, Oxford/AstraZeneca, Pfizer/BioNTech, Sinopharm/Beijing, Sputnik V
104
Oxford/AstraZeneca, Pfizer/BioNTech, Sinopharm/Beijing, Sinopharm/Wuhan, Sputnik V
96
Covaxin, Oxford/AstraZeneca
86
Sinovac
84
Moderna, Oxford/AstraZeneca
79
Oxford/AstraZeneca, Sinopharm/Beijing, Sinovac
61
Johnson&Johnson
54
Sinopharm/Beijing, Sputnik V
51
Pfizer/BioNTech, Sputnik V
42
```

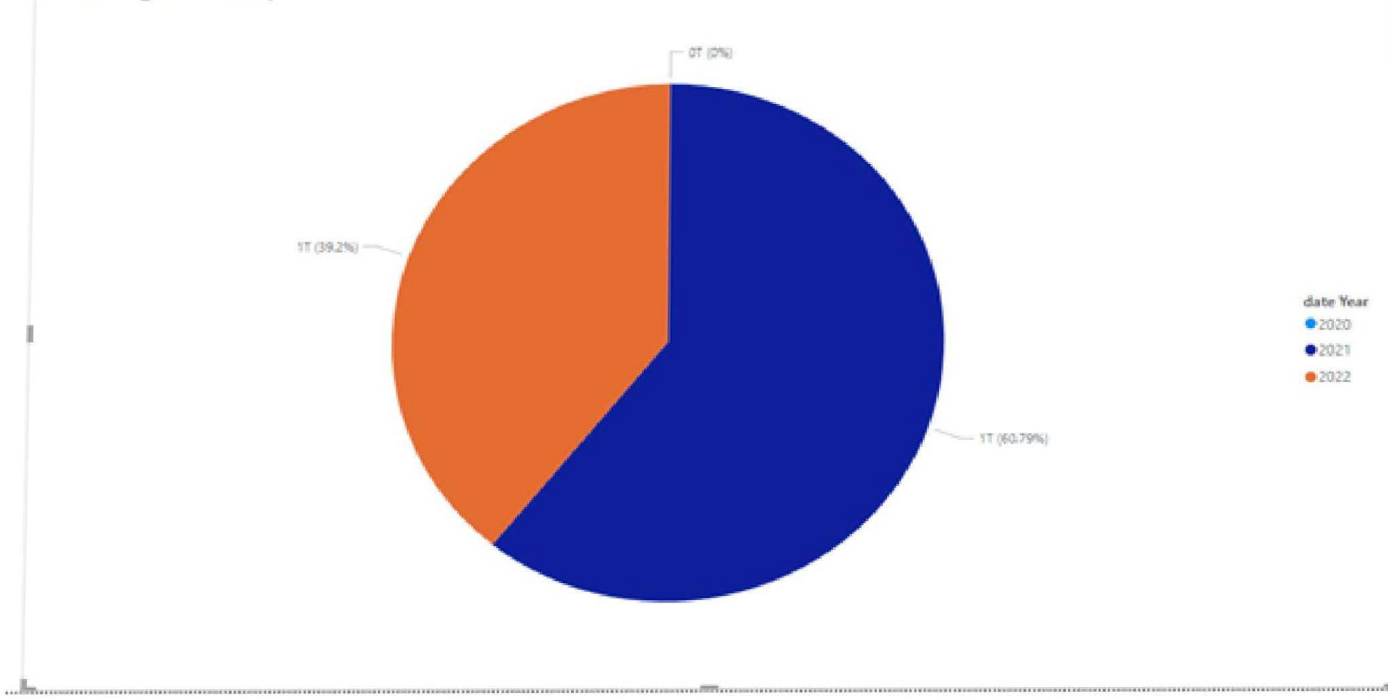
```
pd.to_datetime(data.date)
data.country.value_counts()
```

```
United Kingdom      118
Northern Ireland    118
Wales               118
England             118
Canada              118
                   ...
Mali                  4
Bahamas               2
Brunei                2
Laos                  1
Armenia               1
Name: country, Length: 175, dtype: int64
```

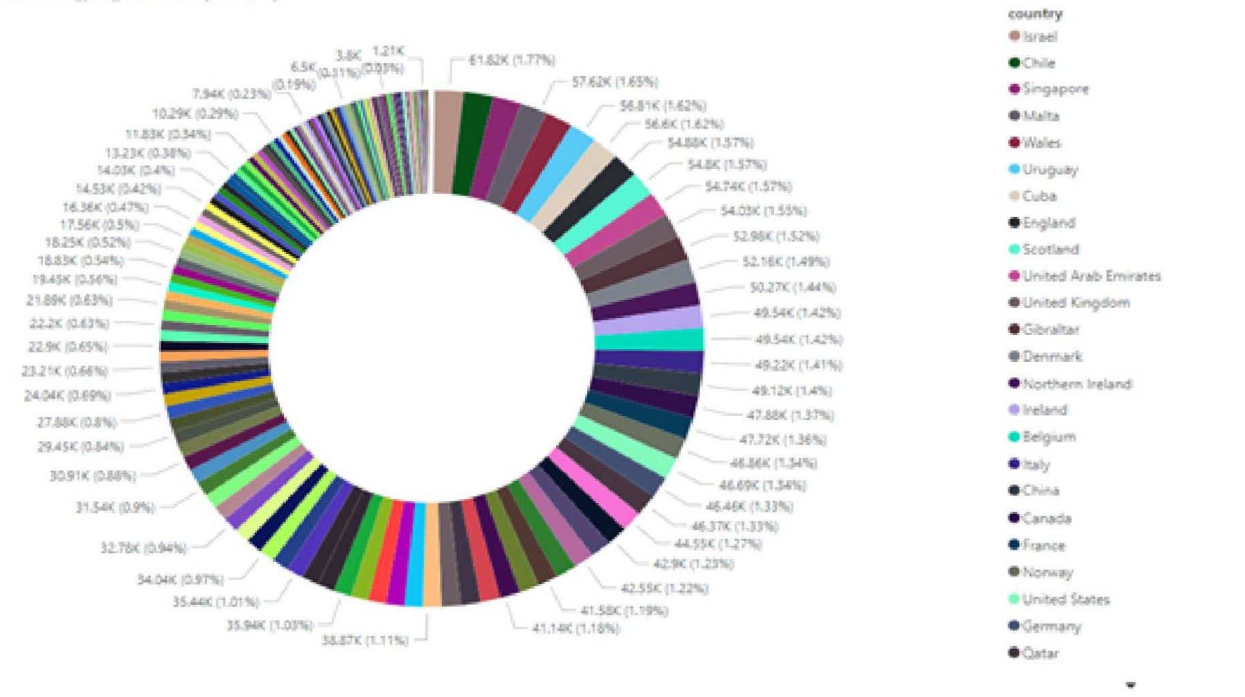
The United Kingdom is made up of England, Scotland, Wales, and Northern Ireland. But in the above data, these countries are mentioned separately with the same values as in the United Kingdom. So this may be an error while recording this data. So let's see how we can fix this error:

```
data = data[data.country.apply(lambda x: x not in ["England", "Sc
data.country.value_counts()
```

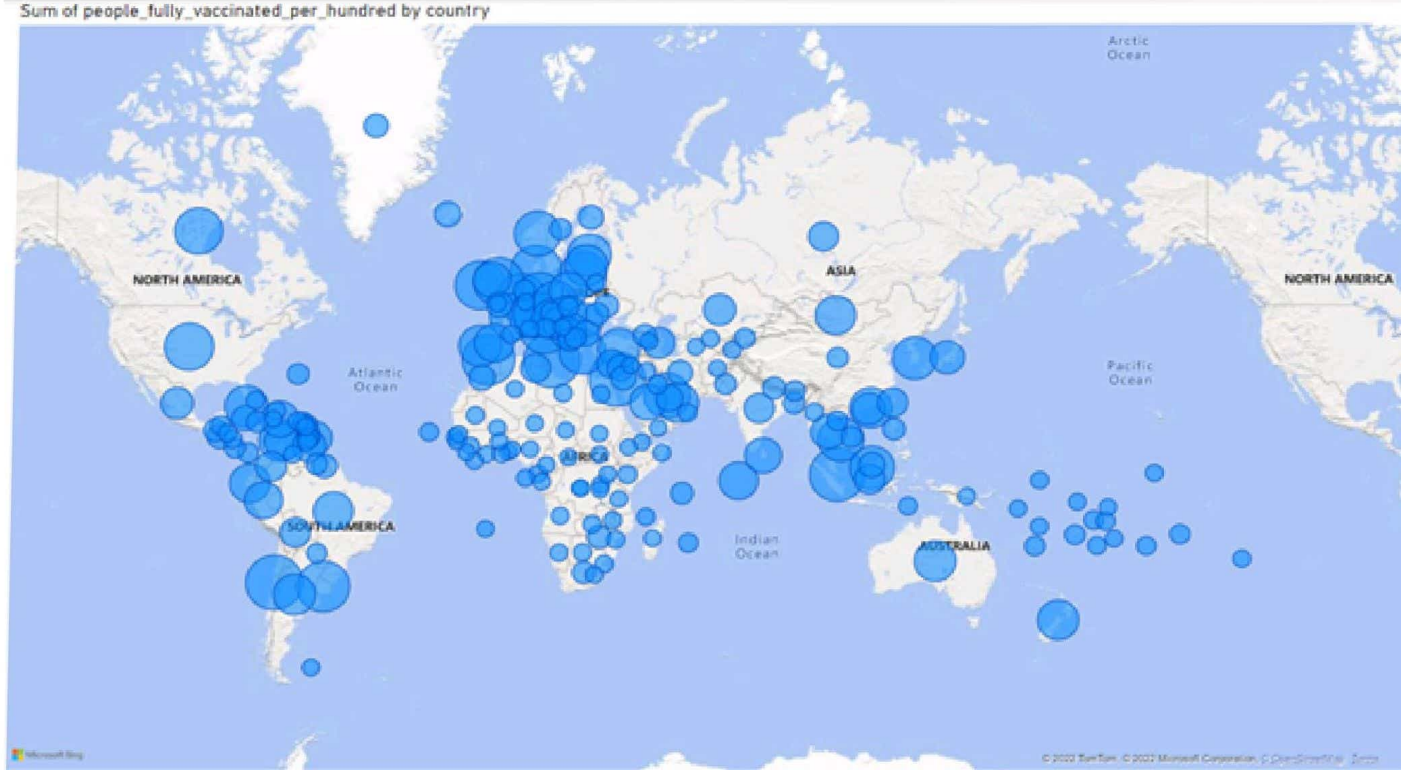
```
Canada              118
United Kingdom      118
China               117
Russia              117
Israel              113
                   ...
Mali                  4
Bahamas               2
Brunei                2
Laos                  1
Armenia               1
Name: country, Length: 171, dtype: int64
```



Sum of total_vaccinations_per_hundred by country



Sum of people_fully_vaccinated_per_hundred by country



Analysis

From the above Visualization find some interesting points that are mentioned below

- China is the number 1 Total Vaccinated people are present
- The Dataset is available for 2021 and 2022.
- The Total of 223 Countries participated in Vaccination.
- IN 2021 Total Vaccination is 60.79% completed.
- In 2022 Total Vaccination is 39.2% is completed.
- China, India, the United States, Brazil, Indonesia, Germany, United States, Turkey, France, and England There are the top 10 countries is completed the full Vaccinations.

Insights

- In conclusion, we can take a look at the final Dashboard for further Analysis.
- We can see the top 10 Sources

Recommendations

- We can analyze the data of the overall world.
- Like this dataset, we can perform operations with various categories, city-wise, or Region wise.
- We can analyze the data with the proper format
- We can collect good datasets for more effective analysis by using charts.

Conclusion

In Conclusion, we can take look at the Dashboard for further Analysis.

- In China and India in these two countries, most people are Vaccinated.
- In 2021 60.79% of people are fully Vaccinated and in 2020 only 39.2 % of people are fully Vaccinated.

- China, India, the United States, Brazil, Indonesia, Germany, the United States, Turkey, France, and England There are the top 10 countries is completed the full Vaccinations.