Weekly Project 16 :

COVID Analysis - Visualisation :

In [1]:
```
import pandas as pd
import numpy as np
import datetime
```

In [2]:
```
df = pd.read_csv("country_vaccinations.csv")
```

In [3]:
```
df.isnull().sum()
```

Out[3]:
```
country                              0
iso_code                             0
date                                 0
total_vaccinations                42905
people_vaccinated                 45218
people_fully_vaccinated           47710
daily_vaccinations_raw            51150
daily_vaccinations                  299
total_vaccinations_per_hundred    42905
people_vaccinated_per_hundred     45218
people_fully_vaccinated_per_hundred    47710
daily_vaccinations_per_million      299
vaccines                              0
source_name                           0
source_website                        0
dtype: int64
```

In [4]:
```
df
```

Out[4]:

	country	iso_code	date	total_vaccinations	people_vaccinated	people_fully_vaccinated	daily_vaccinations_raw	daily_vaccinations	total_vaccinations_per_hundred	people_vaccinated_per_hundred
0	Afghanistan	AFG	2021-02-22	0.0	0.0	NaN	NaN	NaN	0.00	0.00
1	Afghanistan	AFG	2021-02-23	NaN	NaN	NaN	NaN	1367.0	NaN	NaN
2	Afghanistan	AFG	2021-02-24	NaN	NaN	NaN	NaN	1367.0	NaN	NaN
3	Afghanistan	AFG	2021-02-25	NaN	NaN	NaN	NaN	1367.0	NaN	NaN
4	Afghanistan	AFG	2021-02-26	NaN	NaN	NaN	NaN	1367.0	NaN	NaN
...	...	...	...	...	...	...	...	...	...	...
86507	Zimbabwe	ZWE	2022-03-25	8691642.0	4814582.0	3473523.0	139213.0	69579.0	57.59	31.90
86508	Zimbabwe	ZWE	2022-03-26	8791728.0	4886242.0	3487962.0	100086.0	83429.0	58.25	32.38
86509	Zimbabwe	ZWE	2022-03-27	8845039.0	4918147.0	3493763.0	53311.0	90629.0	58.61	32.59
86510	Zimbabwe	ZWE	2022-03-28	8934360.0	4975433.0	3501493.0	89321.0	100614.0	59.20	32.97
86511	Zimbabwe	ZWE	2022-03-29	9039729.0	5053114.0	3510256.0	105369.0	103751.0	59.90	33.48

86512 rows × 15 columns

In [5]:
```
df.head()
```

Out[5]:

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

Add Month Year and Date column :

In [6]:
```
date = []
month = []
year = []
for i in df["date"]:
    date.append(i.split("-")[2])
    month.append(i.split("-")[1])
    year.append(i.split("-")[0])
df["Year"] = year
df["Month"] = month
df["Date"] = date

df.head()
```

Out[6]:

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

Add Month_Name Column :

In [7]:
```
month_name = []
for i in df["date"]:
    date = int(i.split("-")[2])
    month = int(i.split("-")[1])
    year = int(i.split("-")[0])
    a = (datetime.date(year,month,date).strftime("%b"))
    month_name.append(a)
df["Month_Name"] = month_name
df.head()
```

Out[7]:

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

1.Total Number of Vaccinations in India in Year 2020,2021 and 2022 :

In [8]:
```
total_vaccination_2020 = df[(df['country'] == 'India') & (df['Year'] == '2020')]['total_vaccinations'].sum()
total_vaccination_2021 = df[(df['country'] == 'India') & (df['Year'] == '2021')]['total_vaccinations'].sum()
total_vaccination_2022 = df[(df['country'] == 'India') & (df['Year'] == '2022')]['total_vaccinations'].sum()

print("The Total Vaccinations in India in Year 2020 :", total_vaccination_2020)
print("The Total Vaccinations in India in Year 2021 :", total_vaccination_2021)
print("The Total Vaccinations in India in Year 2022 :", total_vaccination_2022)
```

```
The Total Vaccinations in India in Year 2020 : 0.0
The Total Vaccinations in India in Year 2021 : 174118546779.0
The Total Vaccinations in India in Year 2022 : 149321759019.0
```

2.Compare number of total vaccinations in year 2020 of India and USA :

In [9]:
```
total_vaccination_india_2020 = df[(df['country'] == 'India') & (df['Year'] == '2020')]['total_vaccinations'].sum()
total_vaccination_USA_2020   = df[(df['country'] == 'United States') & (df['Year'] == '2020')]['total_vaccinations'].sum()

print("The Total Vaccinations in India in Year 2020 :", total_vaccination_india_2020)
print("The Total Vaccinations in USA in Year 2020   :", total_vaccination_USA_2020)
```

```
The Total Vaccinations in India in Year 2020 : 0.0
The Total Vaccinations in USA in Year 2020   : 41094416.0
```

3.Compare number of total vaccinations in year 2021 of India and China :

In [10]:
```
total_vaccination_india_2021 = df[(df['country'] == 'India') & (df['Year'] == '2021')]['total_vaccinations'].sum()
total_vaccination_china_2021 = df[(df['country'] == 'china') & (df['Year'] == '2021')]['total_vaccinations'].sum()

print("The Total Vaccinations in India in Year 2021 :", total_vaccination_india_2021)
print("The Total Vaccinations in China in Year 2021 :", total_vaccination_india_2021)
```

```
The Total Vaccinations in India in Year 2021 : 174118546779.0
The Total Vaccinations in China in Year 2021 : 174118546779.0
```

4.Find the number of Vaccinations in each month in India in the year 2021:

In [11]:
```
df.head()
```

Out[11]:

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

In [12]:
```
df = df.dropna()
```

In [13]:
```
df.isnull().sum()
```

Out[13]:
```
country                              0
iso_code                             0
date                                 0
total_vaccinations                   0
people_vaccinated                    0
people_fully_vaccinated              0
daily_vaccinations_raw               0
daily_vaccinations                   0
total_vaccinations_per_hundred       0
people_vaccinated_per_hundred        0
people_fully_vaccinated_per_hundred    0
daily_vaccinations_per_million       0
vaccines                             0
source_name                          0
source_website                       0
Year                                 0
Month                                0
Date                                 0
Month_Name                           0
dtype: int64
```

In [14]:
```
n = len(df['country'])
l = ['Jan','Feb','Mar','Apr','May','Jun','Jly','Aug','Sep','Oct','Nov','Dec']
lst = []
for i in l:
    a = 0
    for j in range(n):
        try:
            if df['Month_Name'][j] == i and df['Year'][j] == '2021':
                a += df['total_vaccinations'][j].sum()
        except KeyError:
            pass
    lst.append(a)
```

In [15]:
```
df1 = pd.DataFrame(list(zip(l,lst)),columns = ["Months","Total_Vaccination_2021"])
print("The Number of Vaccination in each month in India in 2021 are : ")
df1.head(20)
```

```
The Number of Vaccination in each month in India in 2021 are :
```

Out[15]:

	Months	Total_Vaccination_2021
0	Jan	2.092826e+08
1	Feb	9.100231e+08
2	Mar	2.397304e+09
3	Apr	4.756036e+09
4	May	8.256727e+09
5	Jun	1.215447e+10
6	Jly	0.000000e+00
7	Aug	2.424152e+10
8	Sep	2.888268e+10
9	Oct	2.898878e+10
10	Nov	3.306481e+10
11	Dec	3.468745e+10

5.Which month has the most number of total vaccinations in India in 2021:

In [16]:
```
n = len(df1['Months'])
a = df1['Total_Vaccination_2021'].max()
for i in range(n):
    if (df1['Total_Vaccination_2021'][i]):
        b = df1['Months'][i]
print("This the Month that has the most Number of Total Vaccinations in India in 2021","-",b)
```

```
This the Month that has the most Number of Total Vaccinations in India in 2021 - Dec
```