```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.impute import KNNImputer
import warnings
warnings.filterwarnings("ignore")

data=pd.read_csv("/content/drive/MyDrive/country_vaccinations.csv")
data_manu=pd.read_csv("/content/drive/MyDrive/country_vaccinations_by_
manufacturer.csv")
print(data.shape)
print(data_manu.shape)
```

```
(86512, 15)
(35623, 4)
```

```
data.head(2)
```

```
       country iso_code        date  total_vaccinations
people_vaccinated  \
0  Afghanistan      AFG  2021-02-22                 0.0
0.0
1  Afghanistan      AFG  2021-02-23                 NaN
NaN

   people_fully_vaccinated  daily_vaccinations_raw  daily_vaccinations
\
0                      NaN                     NaN                 NaN

1                      NaN                     NaN              1367.0


   total_vaccinations_per_hundred  people_vaccinated_per_hundred  \
0                             0.0                            0.0
1                             NaN                            NaN

   people_fully_vaccinated_per_hundred  daily_vaccinations_per_million
\
0                                  NaN                            NaN

1                                  NaN                           34.0


                                        vaccines  \
0  Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
1  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/
```

```
data_manu.head(4)
```

```
    location        date              vaccine  total_vaccinations
0  Argentina  2020-12-29              Moderna                   2
1  Argentina  2020-12-29  Oxford/AstraZeneca                   3
2  Argentina  2020-12-29   Sinopharm/Beijing                   1
3  Argentina  2020-12-29           Sputnik V               20481
```

preprocessing the datasets

handle the missing values

```
print(data.isnull().sum())
print("----------------------------------------")
print(data_manu.isnull().sum())
```

```
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
----------------------------------------
location              0
date                  0
vaccine               0
total_vaccinations    0
dtype: int64
```

```
data.dropna(axis=0,inplace=True)
```

```
data.shape
```

```
(30847, 15)
```

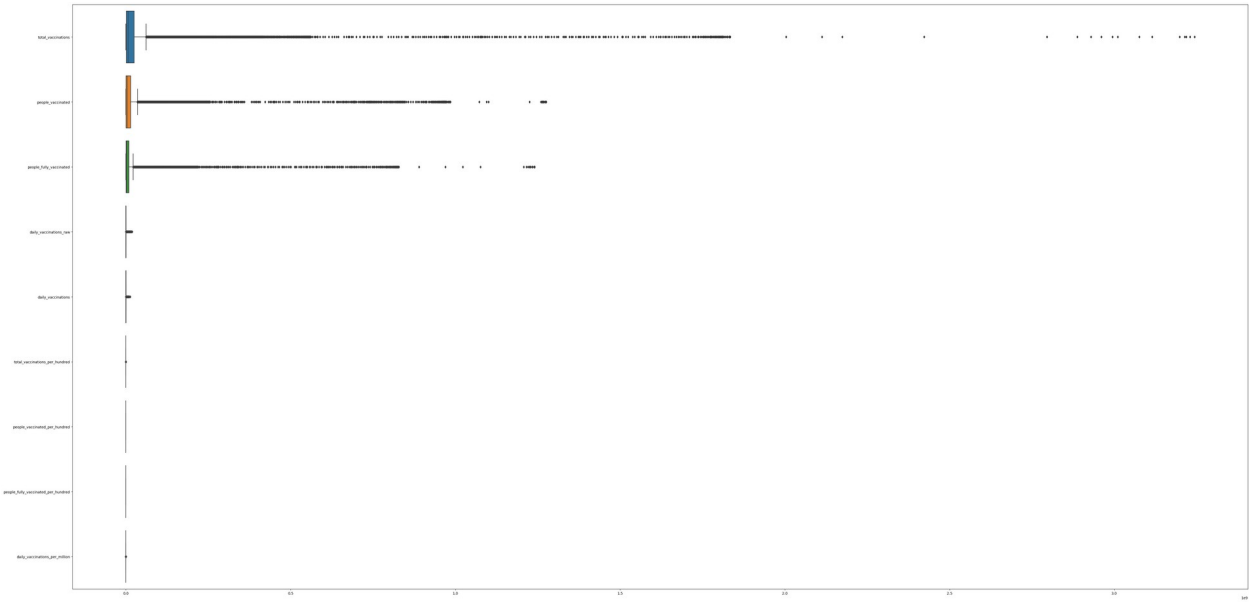
Outlierhandling

Finding outlier and its distribution

Covid vaccination

```
plt.figure(figsize=(60,30))
sns.boxplot(data,orient='h')
plt.show
```

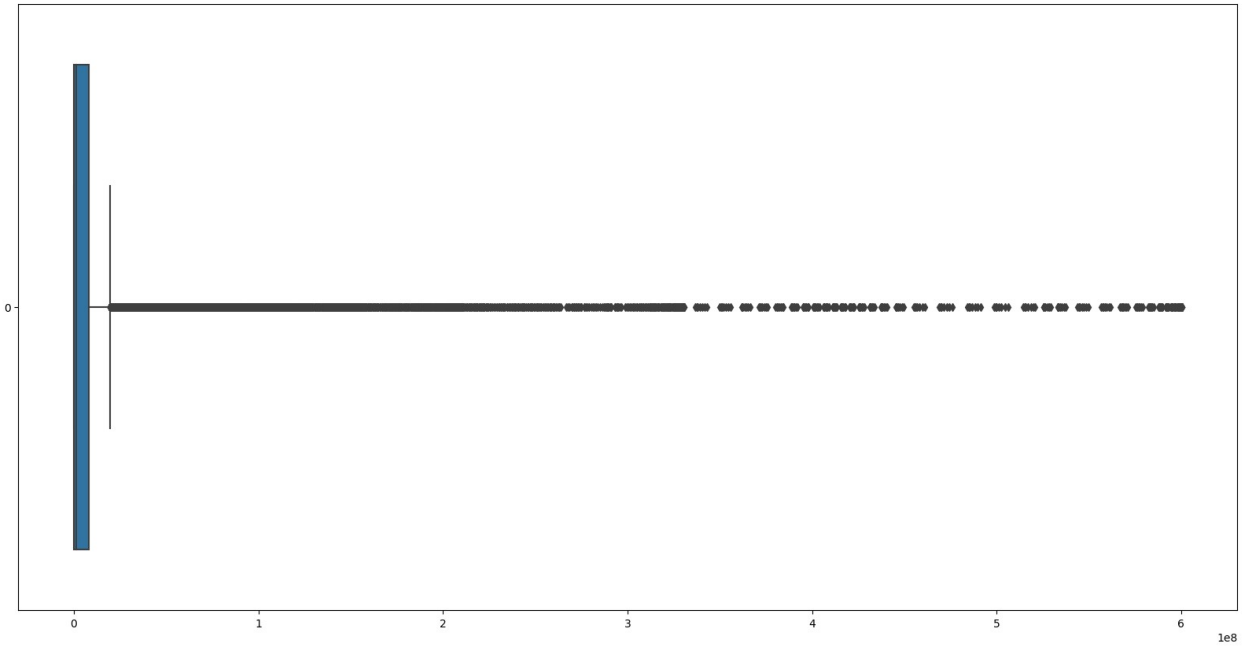
```
<function matplotlib.pyplot.show(close=None, block=None)>
```



In manufacture dataset

```
plt.figure(figsize=(20,10))
sns.boxplot(data_manu["total_vaccinations"],orient='h')
plt.show
```

```
<function matplotlib.pyplot.show(close=None, block=None)>
```




```
list1=['total_vaccinations','people_vaccinated','people_fully_vaccinat
ed','daily_vaccinations_raw','daily_vaccinations','total_vaccinations_
per_hundred','daily_vaccinations_per_million']

for c in list1:
  col=data[c]
  q1=col.quantile(0.25)
  q3=col.quantile(0.75)
  iqr=q3-q1
  lower=q1 - 1.5 *iqr
  upper=q3 + 1.5*iqr
  length=len(col[(col<lower) | (col>upper)])
  print(f"Outlier on  {c}    is {length}")
  print("                    ")
```

```
Outlier on  total_vaccinations    is 4407

Outlier on  people_vaccinated    is 4384

Outlier on  people_fully_vaccinated    is 4826

Outlier on  daily_vaccinations_raw    is 4091

Outlier on  daily_vaccinations    is 4004

Outlier on  total_vaccinations_per_hundred    is 26

Outlier on  daily_vaccinations_per_million    is 769
```

```
for c in list1:
  col=data[c]
  q1=col.quantile(0.25)
  q3=col.quantile(0.75)
  iqr=q3-q1
  lower=q1 - 1.5 *iqr
  upper=q3 + 1.5*iqr
  col[col<lower]=lower
  col[col>upper]=upper
  print("completed")
```

```
completed
completed
completed
completed
completed
completed
completed
```

```
  col=data_manu["total_vaccinations"]
  q1=col.quantile(0.25)
  q3=col.quantile(0.75)
  iqr=q3-q1
  lower=q1 - 1.5 *iqr
  upper=q3 + 1.5*iqr
  length=len(col[(col<lower) | (col>upper)])
  print(f"Outlier on  total_vaccinations    is {length}")
  print("                    ")
```

```
Outlier on  total_vaccinations    is 4544
```

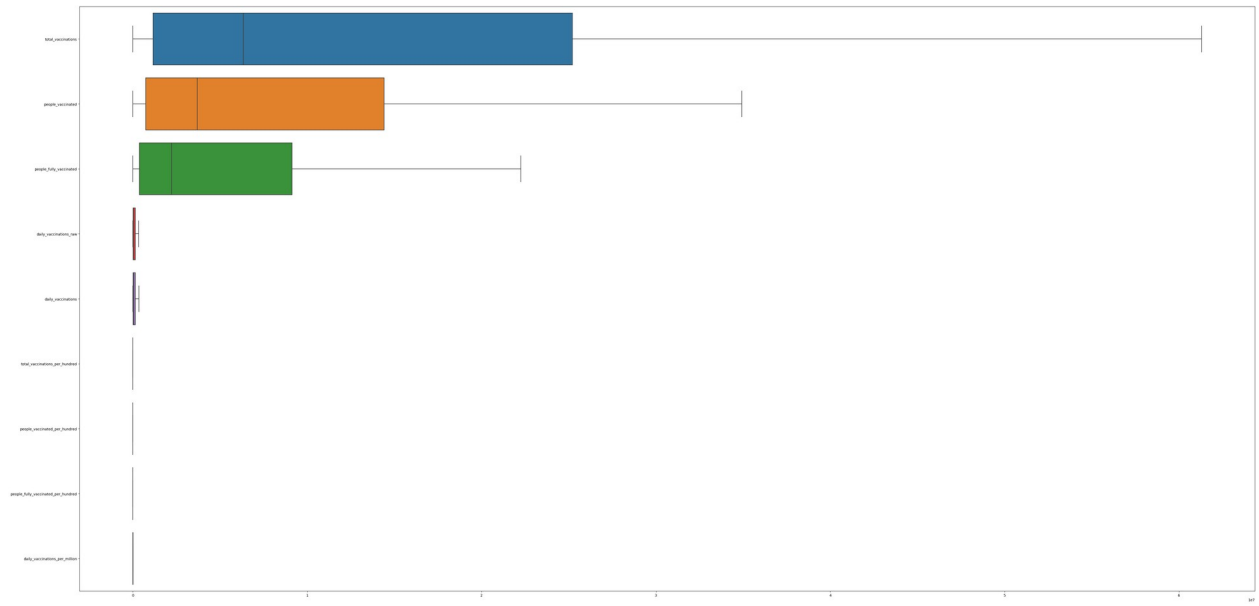
```
  col[col<lower]=lower
  col[col>upper]=upper

  length=len(col[(col<lower) | (col>upper)])
  print(f"Outlier on  total_vaccinations    is {length}")
```

```
Outlier on  total_vaccinations    is 0
```

After handling the outliers both dataset boxplot

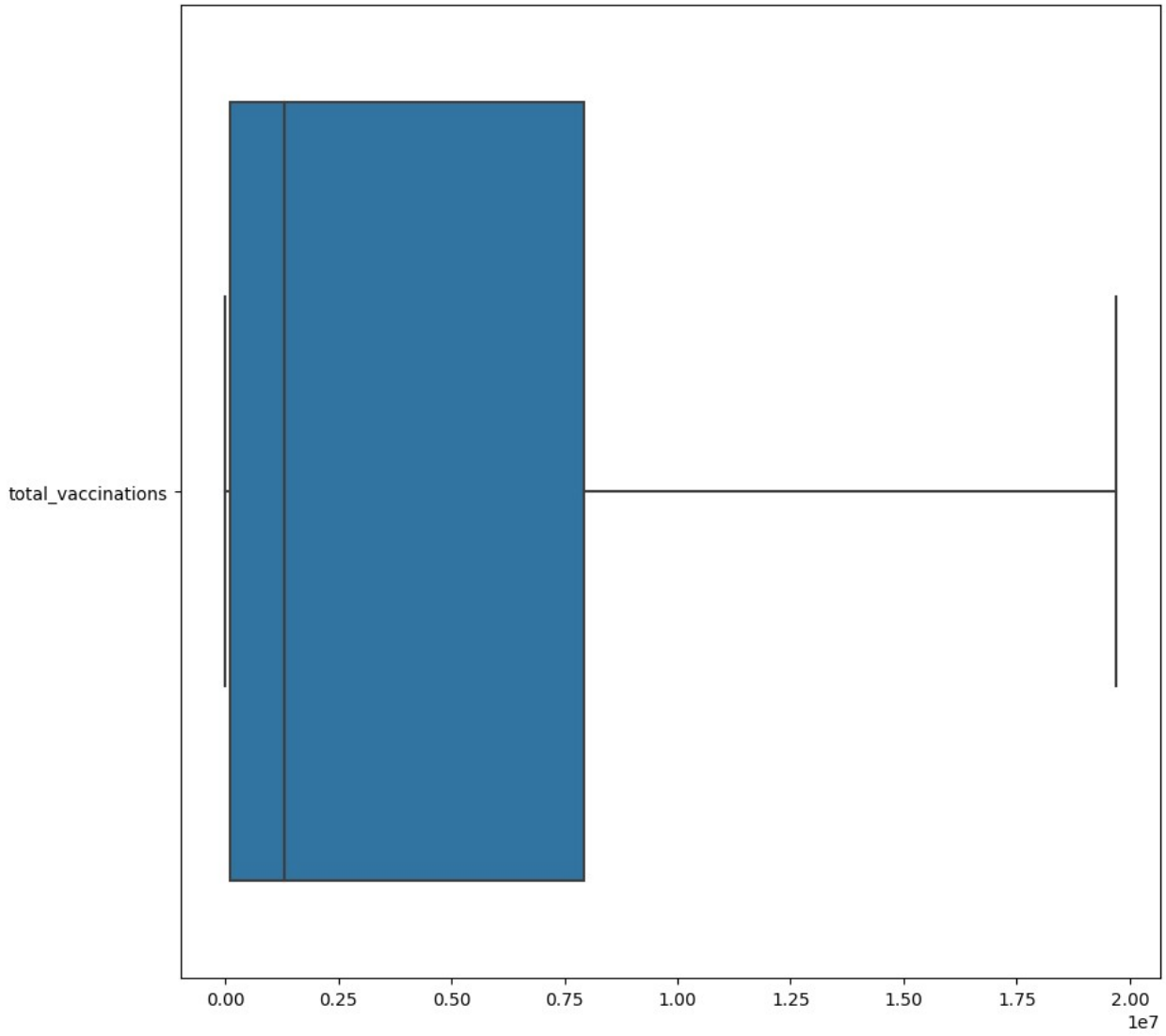
```
plt.figure(figsize=(60,30))
sns.boxplot(data,orient='h')
plt.show
```

```
<function matplotlib.pyplot.show(close=None, block=None)>
```




```
plt.figure(figsize=(10,10))
sns.boxplot(data_manu,orient='h')
plt.show
```

```
<function matplotlib.pyplot.show(close=None, block=None)>
```




```
data.describe()
```

	total_vaccinations	people_vaccinated	people_fully_vaccinated
count	3.084700e+04	3.084700e+04	3.084700e+04
mean	1.705935e+07	9.587147e+06	6.369341e+06
std	2.187500e+07	1.237931e+07	8.118668e+06
min	3.000000e+00	3.000000e+00	1.000000e+00
25%	1.153332e+06	7.339795e+05	3.704450e+05
50%	6.335305e+06	3.688092e+06	2.211035e+06
75%	2.520629e+07	1.440668e+07	9.121526e+06

```
max          6.128573e+07       3.491573e+07              2.224815e+07


       daily_vaccinations_raw  daily_vaccinations  \
count            30847.000000        30847.000000
mean             89622.957759        93837.217071
std             116020.390309       119257.938525
min                 0.000000            0.000000
25%              5498.000000         7329.500000
50%             29081.000000        32472.000000
75%            134458.000000       140291.500000
max            327898.000000       339734.500000

       total_vaccinations_per_hundred
people_vaccinated_per_hundred  \
count                    30847.000000                    30847.000000

mean                        88.595403                       44.793028

std                         67.444199                       28.464379

min                          0.000000                        0.000000

25%                         25.475000                       17.190000

50%                         81.470000                       48.160000

75%                        140.745000                       70.330000

max                        313.650000                      124.760000


       people_fully_vaccinated_per_hundred
daily_vaccinations_per_million
count                           30847.000000
30847.000000
mean                               36.563440
4144.458635
std                                28.532602
3217.135674
min                                 0.000000
0.000000
25%                                 7.400000
1567.500000
50%                                34.110000
3254.000000
75%                                63.570000
6069.500000
```

```
max                                   122.370000
12822.500000
```

on the manufactururas data we can only have total vaccination in the countries on day by day

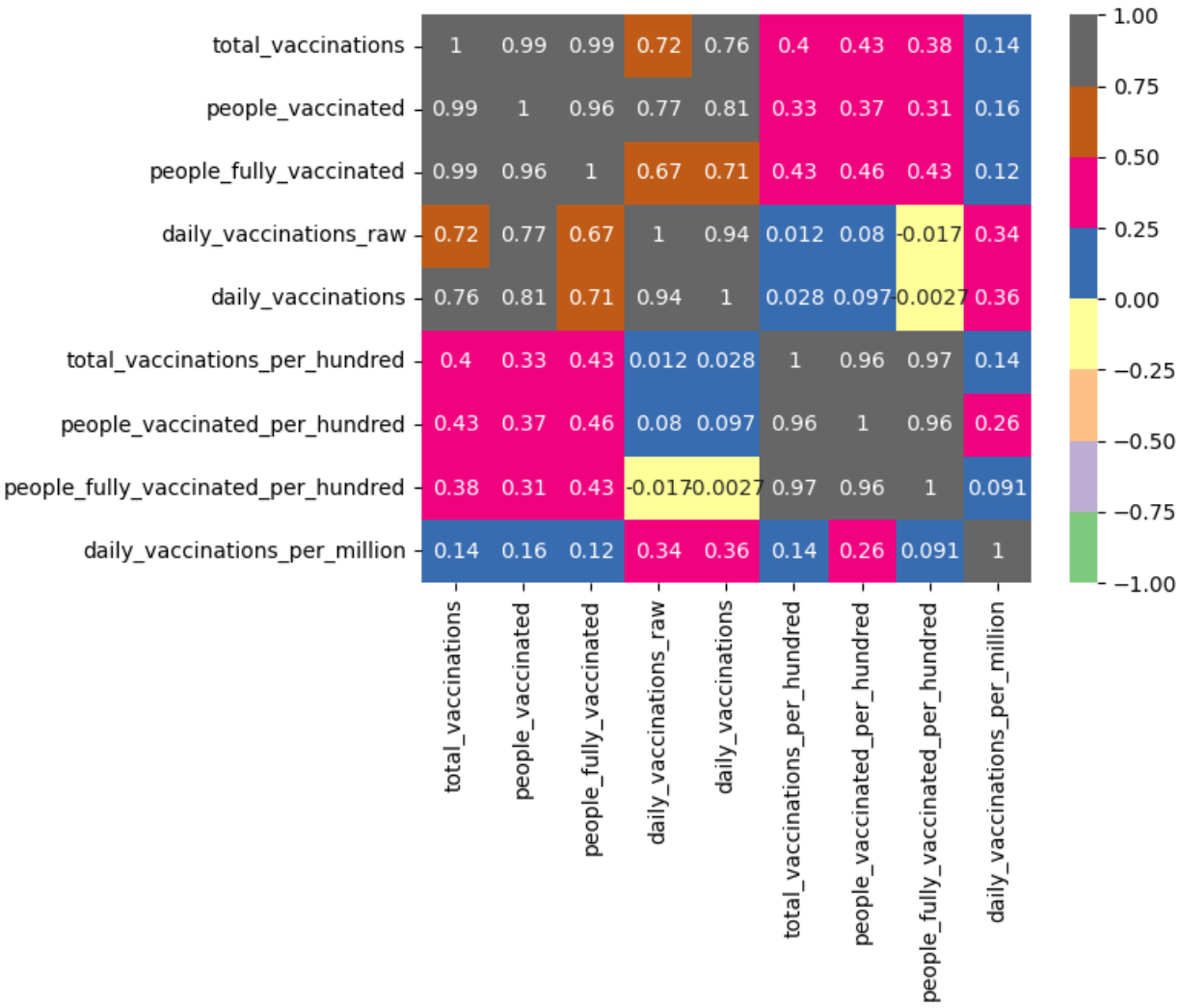
```
data_manu.describe()
```

```
       total_vaccinations
count        3.562300e+04
mean         5.137679e+06
std          7.003622e+06
min          0.000000e+00
25%          9.777600e+04
50%          1.305506e+06
75%          7.932423e+06
max          1.968439e+07
```

Find Correlation of the column

```
sns.heatmap(data.corr(),cmap="Accent",annot=True,vmin=-
1,vmax=1,center=0)
```

```
<Axes: >
```



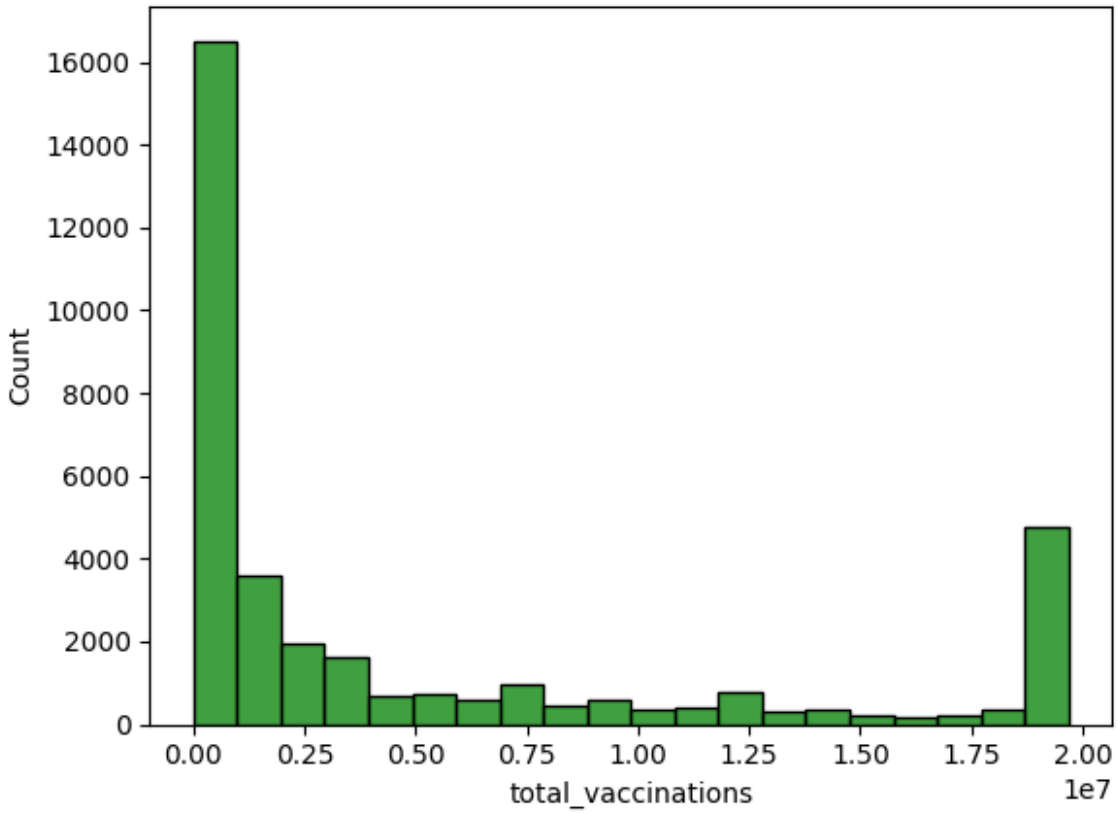
Number of Countries that are present in dataset

```
len(data["country"].unique())
```

```
169
```

```
sns.histplot(data=data_manu["total_vaccinations"],bins=20,color= 'g')
```

```
<Axes: xlabel='total_vaccinations', ylabel='Count'>
```



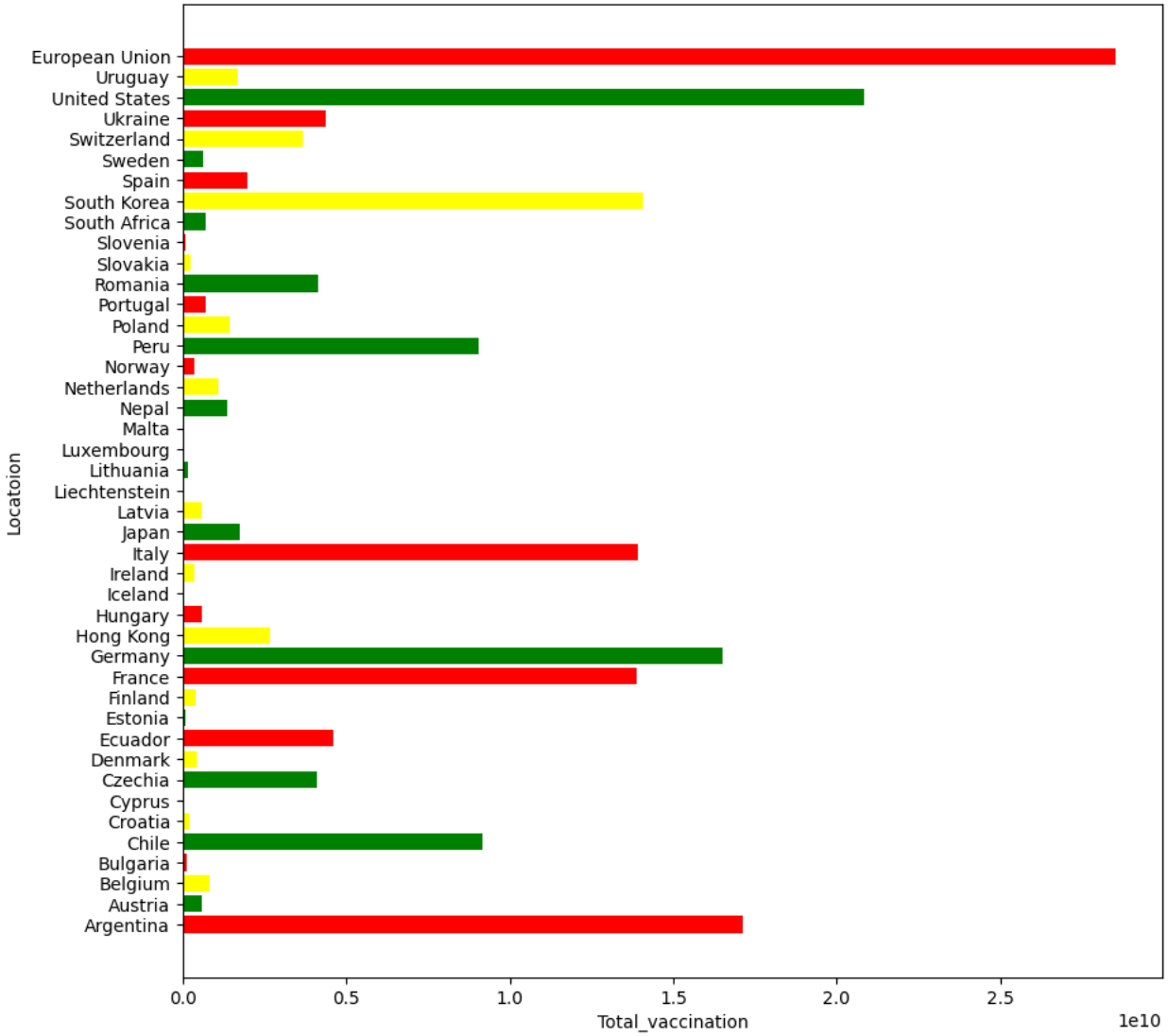
Calculate the Total vaccination on date wise

```
total_vac=[]
for x in data_manu["location"].unique():
     z=data_manu.loc[data_manu["location"]== x]
     total=0
     total=z["total_vaccinations"].sum()


     total_vac.append(total)

df=pd.DataFrame({'location':data_manu["location"].unique(),'total_vacc
inations':total_vac})
df.head()
```

	location	total_vaccinations
0	Argentina	1.711444e+10
1	Austria	5.965148e+08
2	Belgium	8.343959e+08
3	Bulgaria	1.342383e+08
4	Chile	9.170587e+09

Country wise vaccination

```
x=data_manu["location"].unique()
y=total_vac
plt.figure(figsize=(10,10))
plt.ylabel("Locatoion")
plt.xlabel("Total_vaccination")
plt.barh(x,y,color=["red",'green','yellow'])
plt.show()
```



In Country vaccination csv

```
total_vac1=[]
for x in data["country"].unique():
     z=data.loc[data["country"]== x]
     total=0
     total=z["total_vaccinations"].sum()
```

```
    total_vac1.append(total)
```

Calculate the peple_vaccinated and people_fully_vaccinated for countries

```
total_vac2=[]
for x in data["country"].unique():
    z=data.loc[data["country"]== x]
    total=0
    total=z["people_vaccinated"].sum()


    total_vac2.append(total)

total_vac3=[]
for x in data["country"].unique():
    z=data.loc[data["country"]== x]
    total=0
    total=z["people_fully_vaccinated"].sum()


    total_vac3.append(total)

data["country"].value_counts()
```

```
United States    470
Israel           465
Switzerland      462
Estonia          457
Germany          457
                ...
Ghana              1
Ethiopia           1
Mauritius          1
Gambia             1
Sierra Leone       1
Name: country, Length: 169, dtype: int64
```

Create dataframe for country wise vaccination distibution

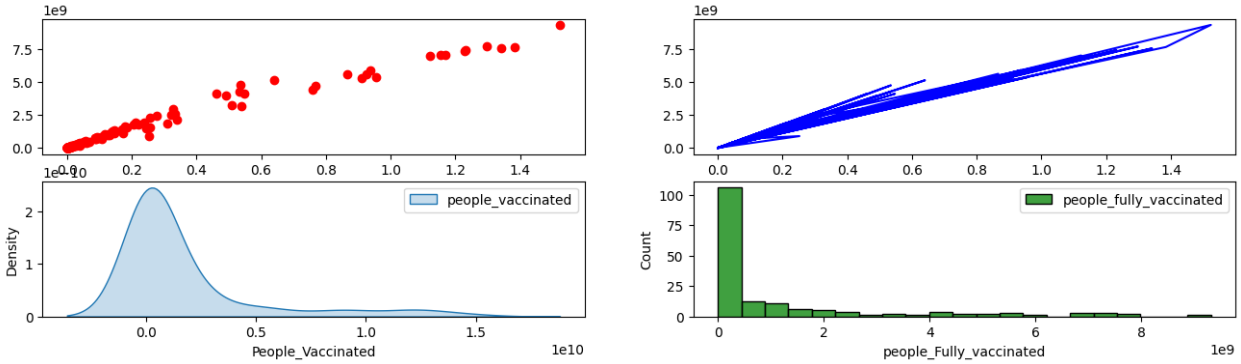
```
datafre={"country":data['country'].unique(),"total_vaccinations":total
_vac1,"People_Vaccinated":total_vac2,"people_Fully_vaccinated":total_v
ac3}
new_data=pd.DataFrame(datafre)
new_data
```

```
                country  total_vaccinations  People_Vaccinated  \
0           Afghanistan        6.304682e+06       5.478754e+06
1               Albania        1.748274e+08       9.592519e+07
2               Algeria        2.432556e+07       1.357837e+07
```

```
3                 Andorra        1.526900e+04       9.781000e+03
4    Antigua and Barbuda         6.160890e+05       3.551400e+05
..                   ...                  ...                ...
164           Uzbekistan         2.648153e+08       1.442560e+08
165              Vietnam         3.551691e+09       2.521270e+09
166                Wales         1.805534e+09       8.840849e+08
167               Zambia         1.662901e+07       1.146332e+07
168             Zimbabwe         1.534183e+09       8.799751e+08

     people_Fully_vaccinated
0                4.131076e+06
1                7.658774e+07
2                1.070525e+07
3                4.484000e+03
4                2.609490e+05
..                        ...
164              6.025293e+07
165              8.917782e+08
166              6.909660e+08
167              5.165692e+06
168              6.423882e+08

[169 rows x 4 columns]
```

```
plt.figure(figsize=(20,20))
x1=new_data["People_Vaccinated"]
y1=new_data["people_Fully_vaccinated"]
plt.figure(figsize=(16,4))
plt.subplot(2,2,1)
plt.scatter(x1,y1,color="red")
plt.subplot(2,2,2)
plt.plot(x1,y1,color='blue')
plt.subplot(2,2,3)
sns.kdeplot(x1,shade=True,label="people_vaccinated",fill=None)
plt.legend()
plt.subplot(2,2,4)
sns.histplot(y1,color='green',label='people_fully_vaccinated')
plt.legend()
plt.show()
```

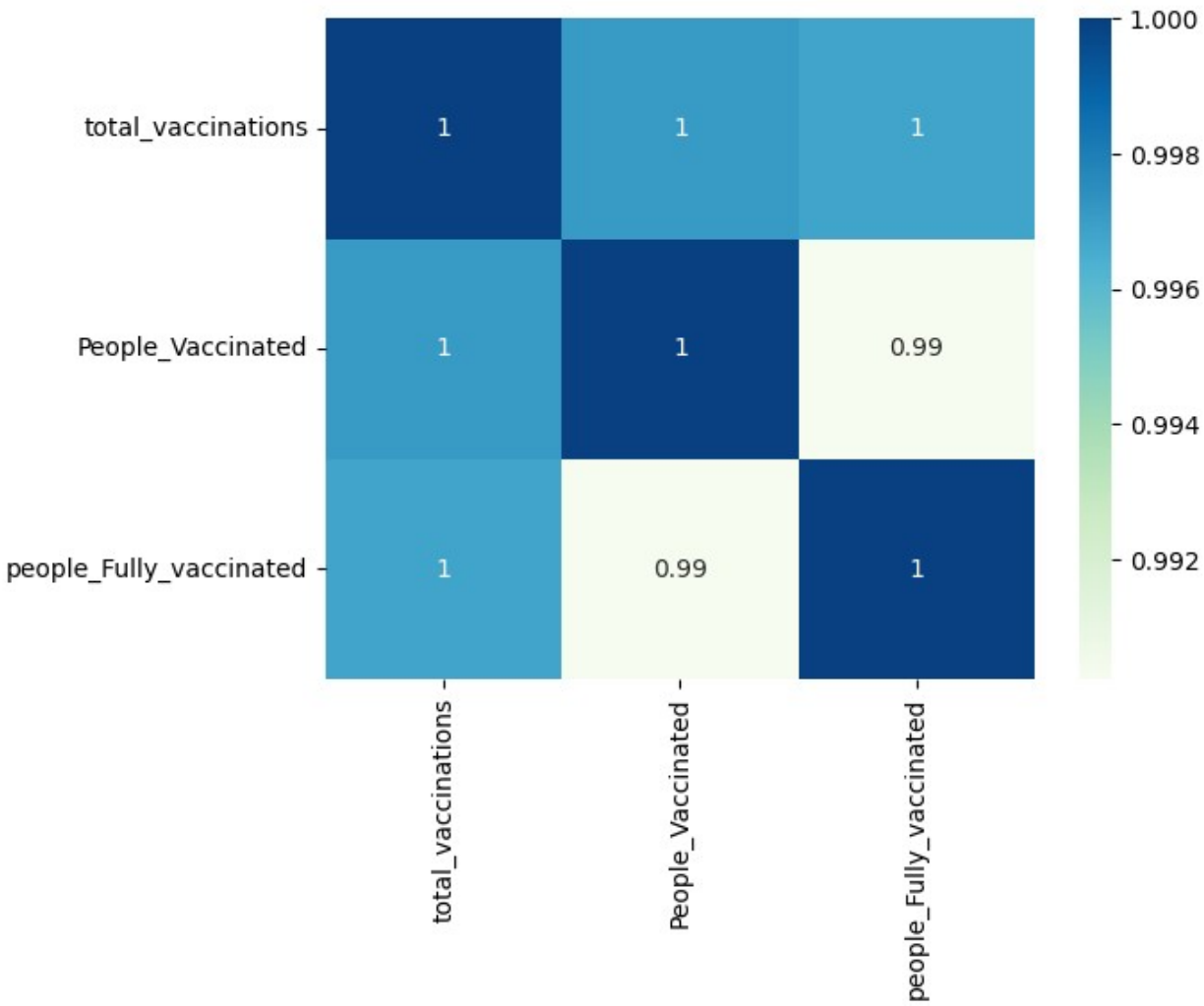
```
<Figure size 2000x2000 with 0 Axes>
```



From above graph we can say that these two fields are positively relatived

```
print(sns.heatmap(new_data.corr(),cmap='GnBu',annot=True))
```

```
Axes(0.125,0.11;0.62x0.77)
```



```
len(data_manu["vaccine"].unique())
```

```
10
```

On the vaccinations there are 10 type of vaccines used in the all over world according to the manufacturer dat

That vaccines

```
vaccines =data_manu["vaccine"].unique()
vaccines
```

```
array(['Moderna', 'Oxford/AstraZeneca', 'Sinopharm/Beijing', 'Sputnik
V',
       'CanSino', 'Pfizer/BioNTech', 'Johnson&Johnson', 'Novavax',
       'Sinovac', 'Covaxin'], dtype=object)
```

```
data.vaccines.unique()
```

```
array(['Johnson&Johnson, Oxford/AstraZeneca, Pfizer/BioNTech,
Sinopharm/Beijing',
       'Oxford/AstraZeneca, Pfizer/BioNTech, Sinovac, Sputnik V',
       'Oxford/AstraZeneca, Sinopharm/Beijing, Sinovac, Sputnik V',
       'Moderna, Oxford/AstraZeneca, Pfizer/BioNTech',
       'Oxford/AstraZeneca, Pfizer/BioNTech, Sputnik V',
       'CanSino, Moderna, Oxford/AstraZeneca, Pfizer/BioNTech,
Sinopharm/Beijing, Sputnik V',
       'Pfizer/BioNTech',
       'Johnson&Johnson, Oxford/AstraZeneca, Pfizer/BioNTech',
       'Johnson&Johnson, Moderna, Oxford/AstraZeneca, Pfizer/BioNTech,
Sinopharm/Beijing, Sputnik Light, Sputnik V',
       'Johnson&Johnson, Moderna, Oxford/AstraZeneca, Pfizer/BioNTech,
Sinopharm/Beijing, Sinovac',
       'Oxford/AstraZeneca, Pfizer/BioNTech, Sinopharm/Beijing',
       'Johnson&Johnson, Moderna, Oxford/AstraZeneca,
Pfizer/BioNTech',
       'Moderna, Oxford/AstraZeneca, Pfizer/BioNTech,
Sinopharm/Beijing',
       'Johnson&Johnson, Oxford/AstraZeneca, Pfizer/BioNTech,
Sinopharm/Beijing, Sputnik V',
       'Covaxin, Johnson&Johnson, Moderna, Oxford/AstraZeneca,
Pfizer/BioNTech, Sinovac',
       'Johnson&Johnson, Oxford/AstraZeneca, Pfizer/BioNTech,
Sinovac',
       'Johnson&Johnson, Moderna, Oxford/AstraZeneca, Pfizer/BioNTech,
Sinopharm/Beijing',
       'Sinopharm/Beijing',
       'Johnson&Johnson, Oxford/AstraZeneca, Sinopharm/Beijing,
Sinovac',
       'Oxford/AstraZeneca, Pfizer/BioNTech',
```

```
       'CanSino, Oxford/AstraZeneca, Pfizer/BioNTech, Sinovac',
       'CanSino, Sinopharm/Beijing, Sinopharm/Wuhan, Sinovac, ZF2001',
       'Johnson&Johnson, Moderna, Oxford/AstraZeneca, Pfizer/BioNTech,
Sinovac',
       'Abdala, Soberana Plus, Soberana02', 'Moderna,
Pfizer/BioNTech',
       'Johnson&Johnson, Moderna, Novavax, Oxford/AstraZeneca,
Pfizer/BioNTech',
       'Oxford/AstraZeneca', 'Johnson&Johnson, Moderna,
Pfizer/BioNTech',
       'Oxford/AstraZeneca, Pfizer/BioNTech, Sinopharm/Beijing,
Sinovac',
       'Johnson&Johnson, Oxford/AstraZeneca, Pfizer/BioNTech,
Sinopharm/Beijing, Sinovac, Sputnik V',
       'Covaxin, Johnson&Johnson, Oxford/AstraZeneca,
Sinopharm/Beijing, Sinovac',
       'Johnson&Johnson, Pfizer/BioNTech',
       'Pfizer/BioNTech, Sinopharm/Beijing, Sputnik V',
       'Johnson&Johnson, Oxford/AstraZeneca, Sinopharm/Beijing',
       'Oxford/AstraZeneca, Sputnik V', 'Moderna',
       'Moderna, Oxford/AstraZeneca, Pfizer/BioNTech, Sputnik V',
       'Oxford/AstraZeneca, Sinopharm/Beijing',
       'Moderna, Oxford/AstraZeneca, Pfizer/BioNTech,
Sinopharm/Beijing, Sputnik V',
       'Johnson&Johnson, Moderna',
       'Johnson&Johnson, Moderna, Oxford/AstraZeneca, Pfizer/BioNTech,
Sputnik V',
       'Pfizer/BioNTech, Sinovac',
       'Johnson&Johnson, Moderna, Oxford/AstraZeneca, Pfizer/BioNTech,
Sinopharm/Beijing, Sputnik V',
       'Covaxin, Oxford/AstraZeneca, Sputnik V',
       'Johnson&Johnson, Moderna, Novavax, Oxford/AstraZeneca,
Pfizer/BioNTech, Sinopharm/Beijing, Sinovac',
       'COVIran Barekat, Covaxin, FAKHRAVAC, Oxford/AstraZeneca, Razi
Cov Pars, Sinopharm/Beijing, Soberana02, SpikoGen, Sputnik V',
       'QazVac, Sinopharm/Beijing, Sputnik V',
       'Sinopharm/Beijing, Sputnik V',
       'Johnson&Johnson, Oxford/AstraZeneca, Pfizer/BioNTech,
Sinopharm/Beijing, Sinovac, Sputnik Light, Sputnik V',
       'Johnson&Johnson, Moderna, Novavax, Pfizer/BioNTech',
       'Oxford/AstraZeneca, Pfizer/BioNTech, Sinopharm/Beijing,
Sputnik V',
       'Oxford/AstraZeneca, Pfizer/BioNTech, Sinopharm/Beijing,
Sinovac, Sputnik V',
       'Pfizer/BioNTech, Sinopharm/Beijing',
       'Johnson&Johnson, Oxford/AstraZeneca',
       'CanSino, Oxford/AstraZeneca, Pfizer/BioNTech,
Sinopharm/Beijing, Sinovac',
       'Covaxin, Oxford/AstraZeneca, Sinopharm/Beijing',
```

```
       'CanSino, Johnson&Johnson, Moderna, Oxford/AstraZeneca,
Pfizer/BioNTech, Sinovac, Sputnik V',
       'Oxford/AstraZeneca, Pfizer/BioNTech, Sinovac',
       'CanSino, Covaxin, Moderna, Oxford/AstraZeneca,
Pfizer/BioNTech, Sinopharm/Beijing, Sinovac, Sputnik V',
       'Johnson&Johnson, Moderna, Oxford/AstraZeneca, Pfizer/BioNTech,
Sinopharm/Beijing, Sinovac, Sputnik Light, Sputnik V',
       'Covaxin, Moderna, Oxford/AstraZeneca, Pfizer/BioNTech,
Sinopharm/Beijing, Sinovac, Sputnik V',
       'EpiVacCorona, Sputnik V',
       'Johnson&Johnson, Moderna, Oxford/AstraZeneca, Pfizer/BioNTech,
Sinopharm/Beijing, Sinovac, Sputnik V',
       'Pfizer/BioNTech, Sputnik V',
       'Oxford/AstraZeneca, Sinopharm/Beijing, Sputnik V',
       'Moderna, Pfizer/BioNTech, Sinopharm/Beijing, Sinovac',
       'Johnson&Johnson, Oxford/AstraZeneca, Pfizer/BioNTech,
Sinopharm/Beijing, Sinovac',
       'Medigen, Moderna, Oxford/AstraZeneca, Pfizer/BioNTech',
       'Moderna, Oxford/AstraZeneca, Pfizer/BioNTech, Sinovac, Sputnik
V',
       'Moderna, Oxford/AstraZeneca, Pfizer/BioNTech,
Sinopharm/Beijing, Sinovac',
       'Pfizer/BioNTech, Sinovac, Turkovac',
       'Oxford/AstraZeneca, Pfizer/BioNTech, Sinopharm/Beijing,
Sinopharm/Wuhan, Sputnik V',
       'Moderna, Oxford/AstraZeneca, Pfizer/BioNTech, Sinovac, Sputnik
Light, Sputnik V, ZF2001',
       'Abdala, Moderna, Oxford/AstraZeneca, Pfizer/BioNTech,
Sinopharm/Beijing, Sputnik V'],
      dtype=object)
```

according to the covid vaccinations data each and every day different type vaccines used in each every country

74 combinations used all over the countries according to the data

Exploratory Data Analysis on covid_vaccination by manufacturer

```
data_manu.head(3)
```

```
    location        date             vaccine  total_vaccinations
0  Argentina  2020-12-29             Moderna                 2.0
1  Argentina  2020-12-29  Oxford/AstraZeneca                 3.0
2  Argentina  2020-12-29   Sinopharm/Beijing                 1.0
```

```
vac=[]
for x in data_manu["location"].unique():
     z=data_manu.loc[data_manu["location"]== x]
     vac1=[]
     for y in z["vaccine"]:
```

```
      vac1.append(y)
     vac.append(vac1)
```

A new Dataset with country and its used vaccine

```
vac_dataframe=pd.DataFrame({"country":data_manu["location"].unique(),"
vaccines":vac})
vac_dataframe.head()
```

```
     country                                               vaccines
0  Argentina  [Moderna, Oxford/AstraZeneca, Sinopharm/Beijin...
1    Austria  [Johnson&Johnson, Moderna, Novavax, Oxford/Ast...
2    Belgium  [Johnson&Johnson, Moderna, Novavax, Oxford/Ast...
3   Bulgaria  [Pfizer/BioNTech, Moderna, Pfizer/BioNTech, Mo...
4      Chile  [Pfizer/BioNTech, Pfizer/BioNTech, Pfizer/BioN...
```

```
dt=[]
for i in vac_dataframe["vaccines"]:
  dat=[]
  dat=list(dict.fromkeys(i))
  dt.append(dat)
vac_dataframe["vaccines"]=dt
```

After remove duplicates

```
vac_dataframe.head()
```

```
     country                                               vaccines
0  Argentina  [Moderna, Oxford/AstraZeneca, Sinopharm/Beijin...
1    Austria  [Johnson&Johnson, Moderna, Novavax, Oxford/Ast...
2    Belgium  [Johnson&Johnson, Moderna, Novavax, Oxford/Ast...
3   Bulgaria  [Pfizer/BioNTech, Moderna, Oxford/AstraZeneca,...
4      Chile  [Pfizer/BioNTech, Sinovac, Oxford/AstraZeneca,...
```

Find the number of countries per vaccine used

```
count=[]
for i in data_manu["vaccine"].unique():
  a=0
  for j in vac_dataframe["vaccines"]:
    for k in j:
      if i == k:
        a=a+1
  count.append(a)


count
```

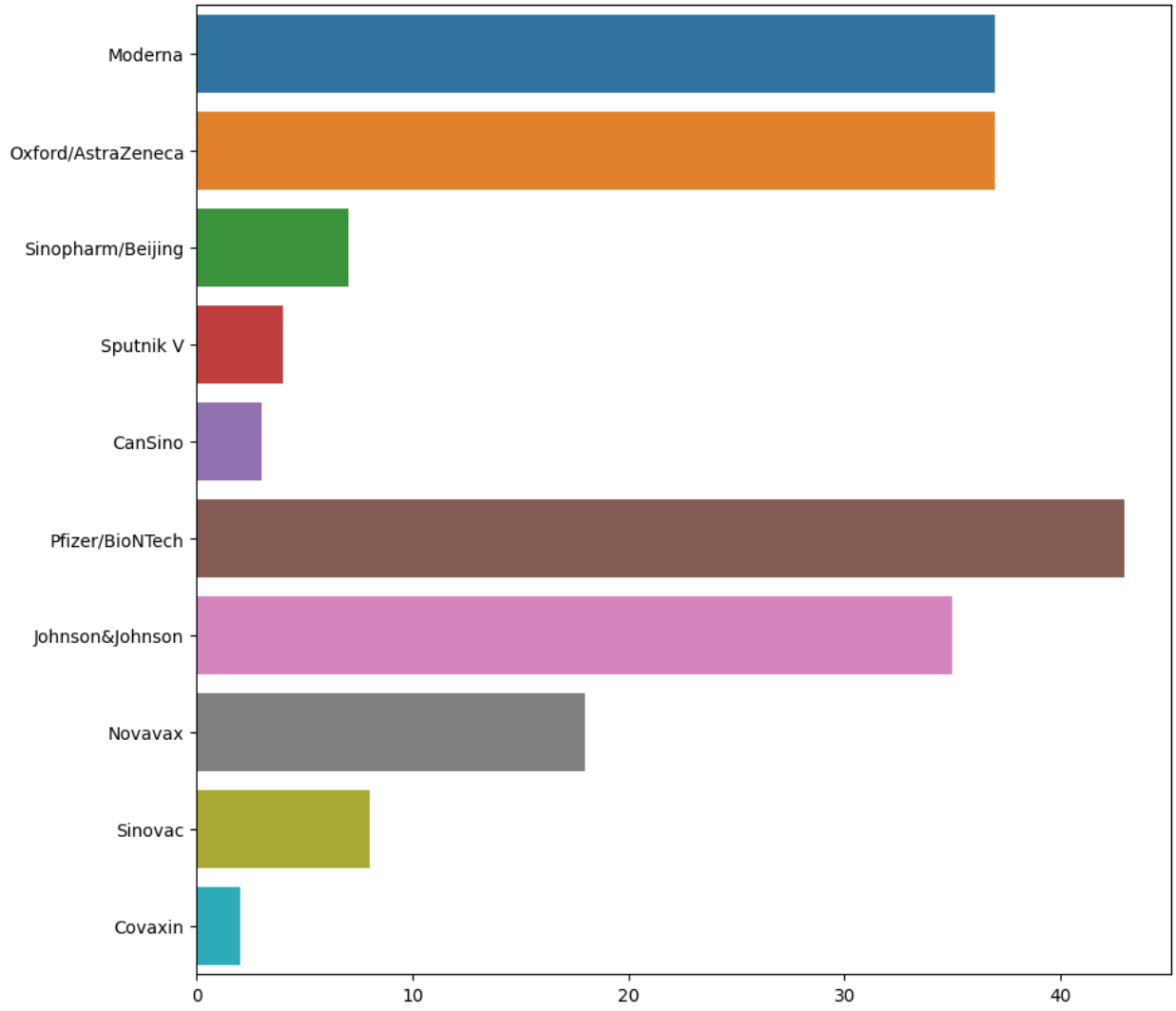
```
[37, 37, 7, 4, 3, 43, 35, 18, 8, 2]
```

```
new_vac=pd.DataFrame({"vaccine":data_manu["vaccine"].unique(),"No fo
country":count})
new_vac
```

	vaccine	No fo country
0	Moderna	37
1	Oxford/AstraZeneca	37
2	Sinopharm/Beijing	7
3	Sputnik V	4
4	CanSino	3
5	Pfizer/BioNTech	43
6	Johnson&Johnson	35
7	Novavax	18
8	Sinovac	8
9	Covaxin	2

Visualize the data

```
plt.figure(figsize=(10,10))
x=data_manu["vaccine"].unique()
y=count
sns.barplot(x=y,y=x,orient='h')
plt.show()
```



Exploratory Data Analysis on the covid_vaccination data

```
data.head()
```

```
         country iso_code        date  total_vaccinations
people_vaccinated  \
94   Afghanistan      AFG  2021-05-27            593313.0
479574.0
101  Afghanistan      AFG  2021-06-03            630305.0
481800.0
339  Afghanistan      AFG  2022-01-27           5081064.0
4517380.0
433      Albania      ALB  2021-02-18              3049.0
2438.0
515      Albania      ALB  2021-05-11            622507.0
440921.0

     people_fully_vaccinated  daily_vaccinations_raw
```

```
daily_vaccinations  \
94                   113739.0                    2859.0
6487.0
101                  148505.0                    4015.0
5285.0
339                 3868832.0                    6868.0
9802.0
433                     611.0                    1348.0
254.0
515                  181586.0                    9548.0
12160.0

     total_vaccinations_per_hundred  people_vaccinated_per_hundred  \
94                             1.49                           1.20
101                            1.58                           1.21
339                           12.76                          11.34
433                            0.11                           0.08
515                           21.67                          15.35

     people_fully_vaccinated_per_hundred
daily_vaccinations_per_million  \
94                                  0.29
163.0
101                                 0.37
133.0
339                                 9.71
246.0
433                                 0.02
88.0
515                                 6.32
4233.0

                                                vaccines  \
94   Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
101  Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
339  Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
433  Oxford/AstraZeneca, Pfizer/BioNTech, Sinovac, ...
515  Oxford/AstraZeneca, Pfizer/BioNTech, Sinovac, ...

                    source_name  \
94   World Health Organization
101  World Health Organization
339  World Health Organization
433         Ministry of Health
515         Ministry of Health

                                      source_website
94                           https://covid19.who.int/
101                          https://covid19.who.int/
339                          https://covid19.who.int/
```

```
433  https://shendetesia.gov.al/vaksinimi-anticovid...
515  https://shendetesia.gov.al/vaksinimi-anticovid...
```

```
new_data.head()
```

```
                country  total_vaccinations  People_Vaccinated  \
0           Afghanistan           6304682.0          5478754.0
1               Albania         174827362.0         95925186.0
2               Algeria          24325560.0         13578367.0
3               Andorra             15269.0             9781.0
4  Antigua and Barbuda            616089.0           355140.0

   people_Fully_vaccinated
0                4131076.0
1               76587739.0
2               10705248.0
3                   4484.0
4                 260949.0
```