NAAN MUDHULVAN PHASE-4

DEVELOPMENT PART 2

Project name: COVID Vaccines Analysis

DATA ANALYTICS OF COVID VACCINES ANALYSIS

Program implementation:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
data = pd.read_csv("C:\Users\student\Documents\country_vaccinations.csv")
data.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	NaN	NaN	
2	Afghanistan	AFG	2021-02-24	NaN	NaN	
3	Afghanistan	AFG	2021-02-25	NaN	NaN	
4	Afghanistan	AFG	2021-02-26	NaN	NaN	

	people_fully_vaccinated	daily_vaccinations_raw	daily_vaccinations	\
0	NaN	NaN	NaN	
1	NaN	NaN	1367.0	
2	NaN	NaN	1367.0	
3	NaN	NaN	1367.0	
4	NaN	NaN	1367.0	

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

	people_fully_vaccinated_per_hundred	daily_vaccinations_per_million	\
0	NaN	NaN	
1	NaN	34.0	
2	NaN	34.0	
3	NaN	34.0	

```
4                                    NaN                              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/


data.describe()

       total_vaccinations  people_vaccinated  people_fully_vaccinated  \
count        4.360700e+04       4.129400e+04             3.880200e+04
mean         4.592964e+07       1.770508e+07             1.413830e+07
std          2.246004e+08       7.078731e+07             5.713920e+07
min          0.000000e+00       0.000000e+00             1.000000e+00
25%          5.264100e+05       3.494642e+05             2.439622e+05
50%          3.590096e+06       2.187310e+06             1.722140e+06
75%          1.701230e+07       9.152520e+06             7.559870e+06
max          3.263129e+09       1.275541e+09             1.240777e+09


       daily_vaccinations_raw  daily_vaccinations  \
count            3.536200e+04        8.621300e+04
mean             2.705996e+05        1.313055e+05
std              1.212427e+06        7.682388e+05
min              0.000000e+00        0.000000e+00
25%              4.668000e+03        9.000000e+02
50%              2.530900e+04        7.343000e+03
75%              1.234925e+05        4.409800e+04
max              2.474100e+07        2.242429e+07


       total_vaccinations_per_hundred  people_vaccinated_per_hundred  \
count                    43607.000000                   41294.000000
mean                        80.188543                      40.927317
std                         67.913577                      29.290759
min                          0.000000                       0.000000
25%                         16.050000                      11.370000
```

```
50%                                         67.520000                        41.435000
75%                                        132.735000                        67.910000
max                                        345.370000                       124.760000


        people_fully_vaccinated_per_hundred    daily_vaccinations_per_million
count                         38802.000000                      86213.000000
mean                             35.523243                       3257.049157
std                              28.376252                       3934.312440
min                               0.000000                          0.000000
25%                               7.020000                        636.000000
50%                              31.750000                       2050.000000
75%                              62.080000                       4682.000000
max                             122.370000                     117497.000000


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


country
Norway                               482
Latvia                               480
Denmark                              476
United States                        471
Russia                               470
                                     ...
Bonaire Sint Eustatius and Saba      146
Tokelau                              114
Saint Helena                          92
Pitcairn                              85
Falkland Islands                      67
Name: count, Length: 223, dtype: int 64
```
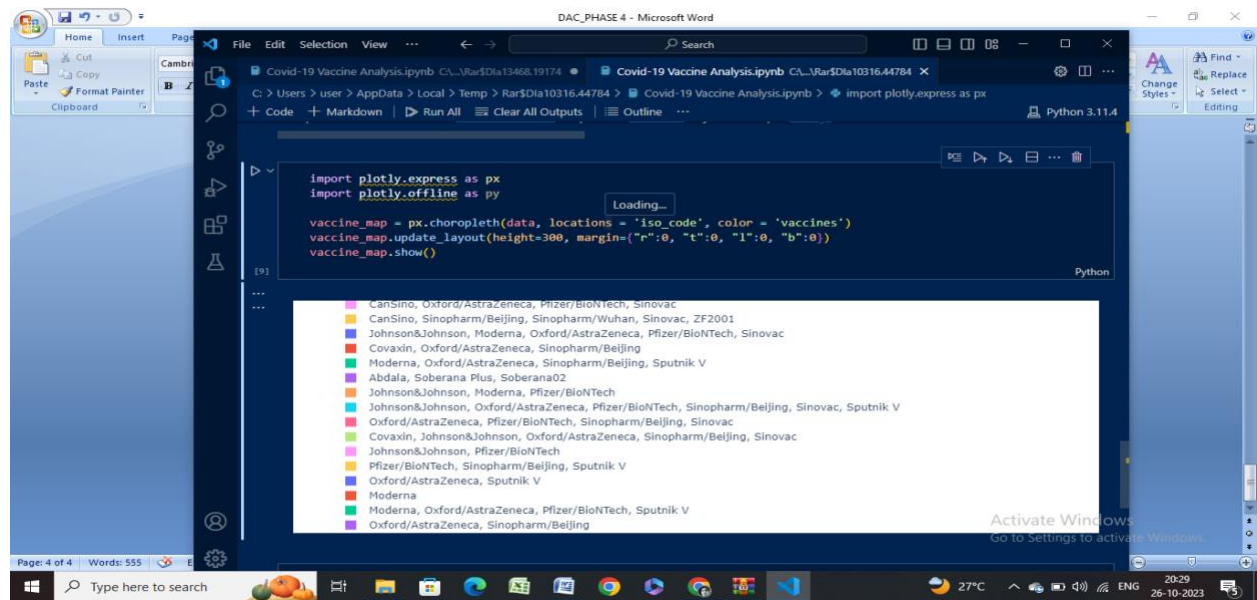
Statistical Analysis:

1. <u>Hypothesis Testing: -</u>

Perform hypothesis tests to determine if there are statistically significant differences in vaccination rates between different groups or regions. For example, you can use t-tests or ANOVA to compare vaccination rates by age groups or between different states.

2. <u>Regression Analysis:-</u>

Perform regression analysis to model the factors that influence vaccination rates. Multiple linear regression or logistic regression can help you understand which variables have the most significant impact on vaccination rates.

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

vaccines

Johnson&Johnson, Moderna, Oxford/AstraZeneca, Pfizer/BioNTech	7608
Moderna, Oxford/AstraZeneca, Pfizer/BioNTech	6263
Oxford/AstraZeneca	6022
Oxford/AstraZeneca, Pfizer/BioNTech	4629
Johnson&Johnson, Moderna, Novavax, Oxford/AstraZeneca, Pfizer/BioNTech	3564
...	
Johnson&Johnson, Oxford/AstraZeneca, Sinovac	312
Moderna, Oxford/AstraZeneca, Pfizer/BioNTech, Sinovac, Sputnik V	311
Johnson&Johnson, Moderna	251
Johnson&Johnson, Pfizer/BioNTech, Sinopharm/Beijing	228
EpiVacCorona, Oxford/AstraZeneca, QazVac, Sinopharm/Beijing, Sputnik V, ZF2001	190

Name: count, Length: 84, dtype: int64

```
df = data[["vaccines", "country"]]
df.head()
```

	vaccines	country
0	Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...	Afghanistan
1	Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...	Afghanistan
2	Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...	Afghanistan

3 Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi... Afghanistan

4 Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi... Afghanistan

Exploratory Data Analysis:

Data Virtualization:

It an approach to data management that allows an application to retrieve and manipulate data without requiring technical details about the data, such as how it is formatted at source, or where it is physically located, and can provide a single customer view of the overall data.