Analyze Covid Vaccination Progress Using Python

Introduction

Covid-19 has affected our lives very much in very accepts it could be economical, mentally, etc. In this blog, we are going to explore how the vaccination drive is going around the world. For the past 1 year, we have been hoping for vaccines so that we can enjoy our life as we were doing before.

Hope this vaccination drive will help millions of people and save them. We are going to first read the dataset, then clean and draw some beautiful visuals.

Dataset

IMPORT LIBRARIES

For analyzing data, we need some libraries. In this section, we are importing all the required libraries like pandas, NumPy, matplotlib, plotly, seaborn, and word cloud that are required for data analysis. Check the below code to import all the required libraries.

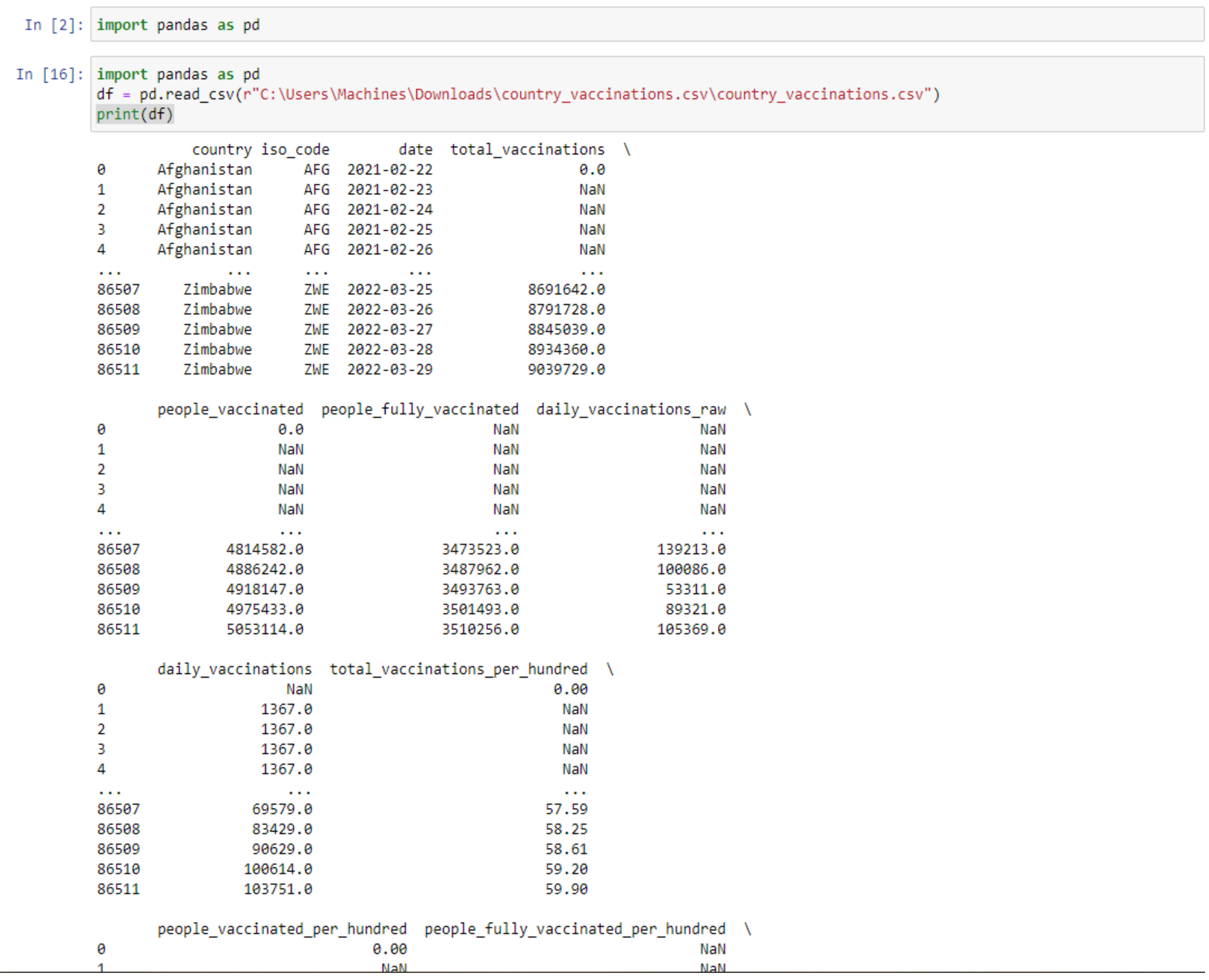
READ DATA AND BASIC INFORMATION

Read the CSV file using pandas read_csv() function and show the output using head() function.

Observation:

Dataset has columns like country, iso_code, date, total_vaccinations, people_vaccinated, people_fully vaccinated, etc. An initial look at the above table shows that data has null values too. We will deal with null values later.

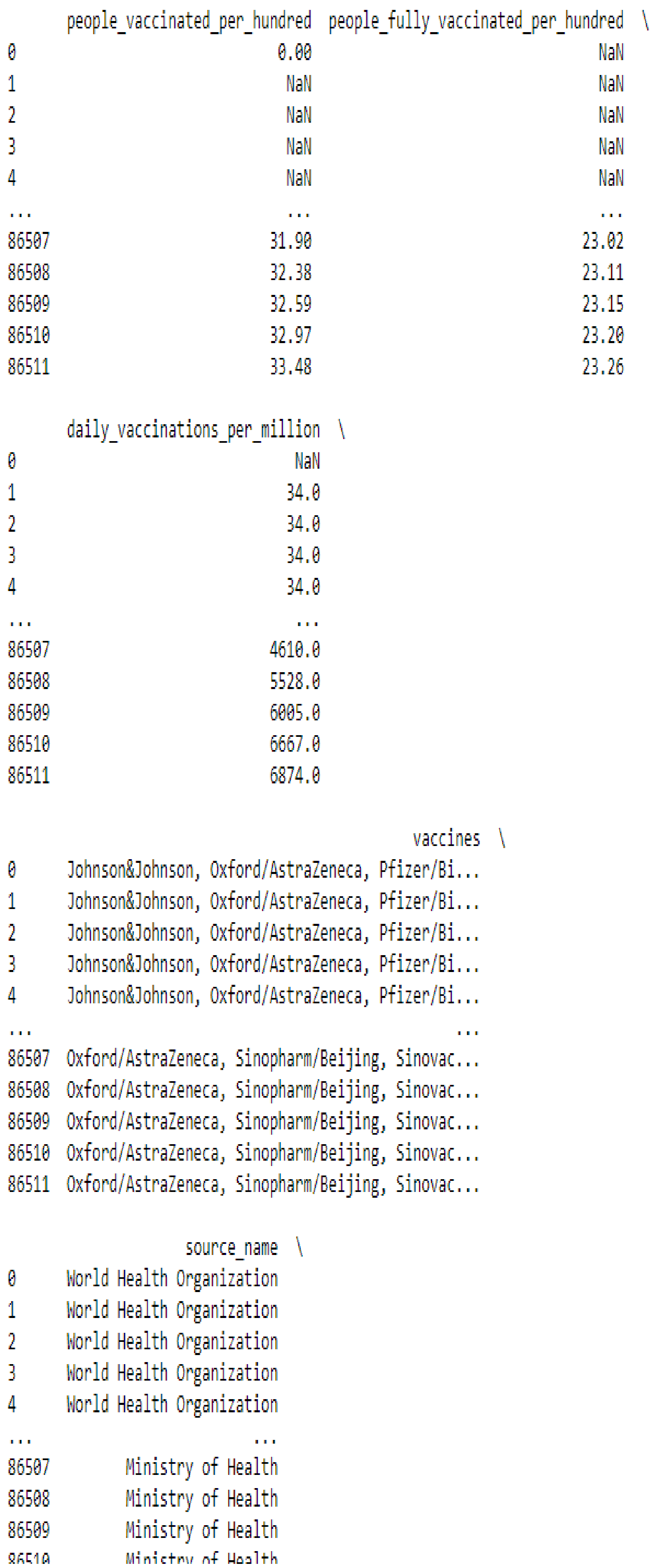
The below picture shows tables like country, date, vaccines, source_name has 0 null values. Features like people_fully_vaccinated have a maximum of 2866 null values.



```
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)
```

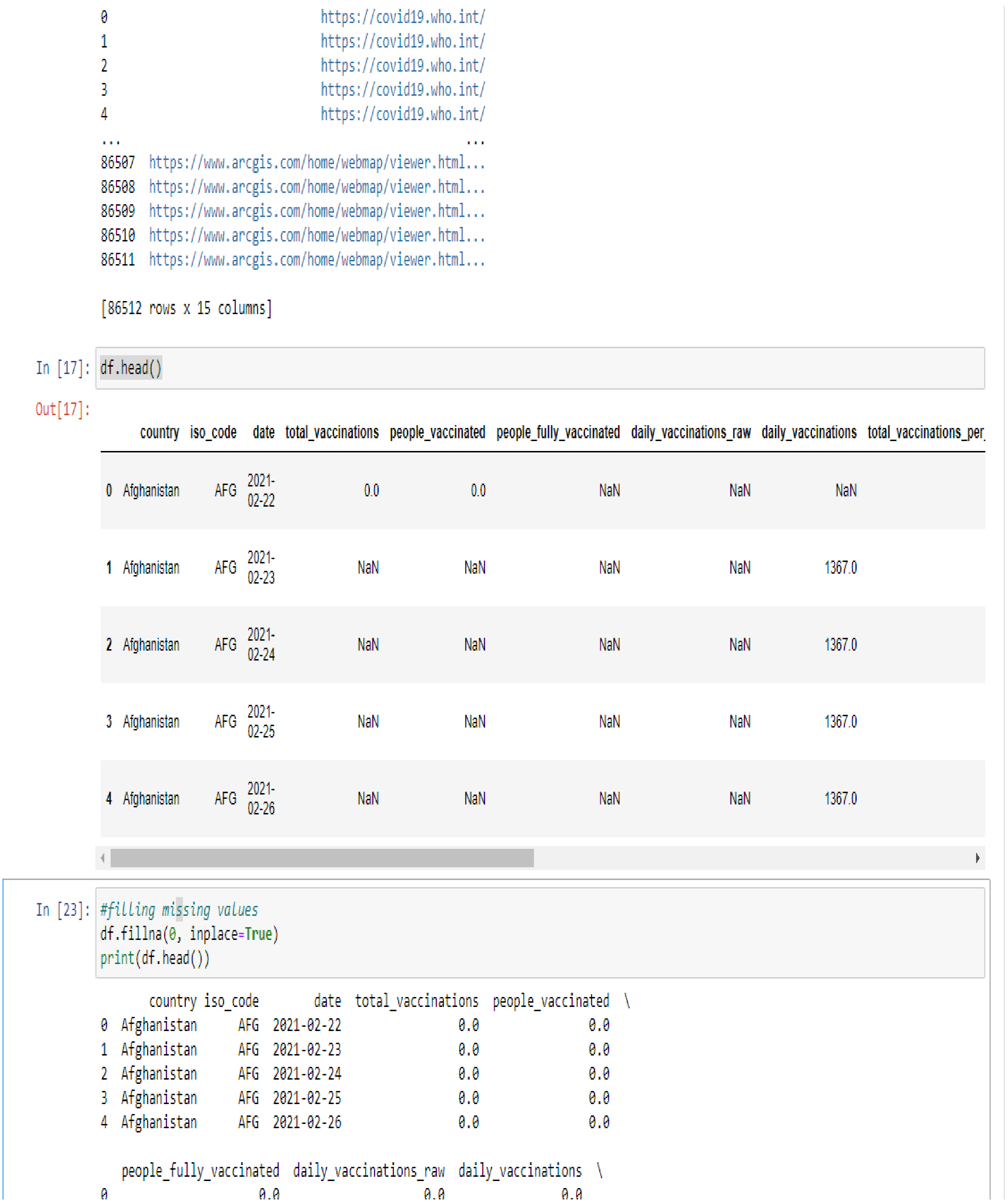
```
           country iso_code        date  total_vaccinations  \
0      Afghanistan      AFG  2021-02-22                 0.0   
1      Afghanistan      AFG  2021-02-23                 NaN   
2      Afghanistan      AFG  2021-02-24                 NaN   
3      Afghanistan      AFG  2021-02-25     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



```
       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                          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                                 NaN
1                                34.0
2                                34.0
3                                34.0
4                                34.0
...                               ...
86507                          4610.0
86508                          5528.0
86509                          6005.0
86510                          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         Ministry of Health
86508         Ministry of Health
86509         Ministry of Health
86510         Ministry of Health
```

info() function is used to get the overview of data like data type of feature, a number of null values in each column, and many more.

```
df.info()
```



```
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 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_
0	Afghanistan	AFG	2021-02-22	0.0	0.0	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	

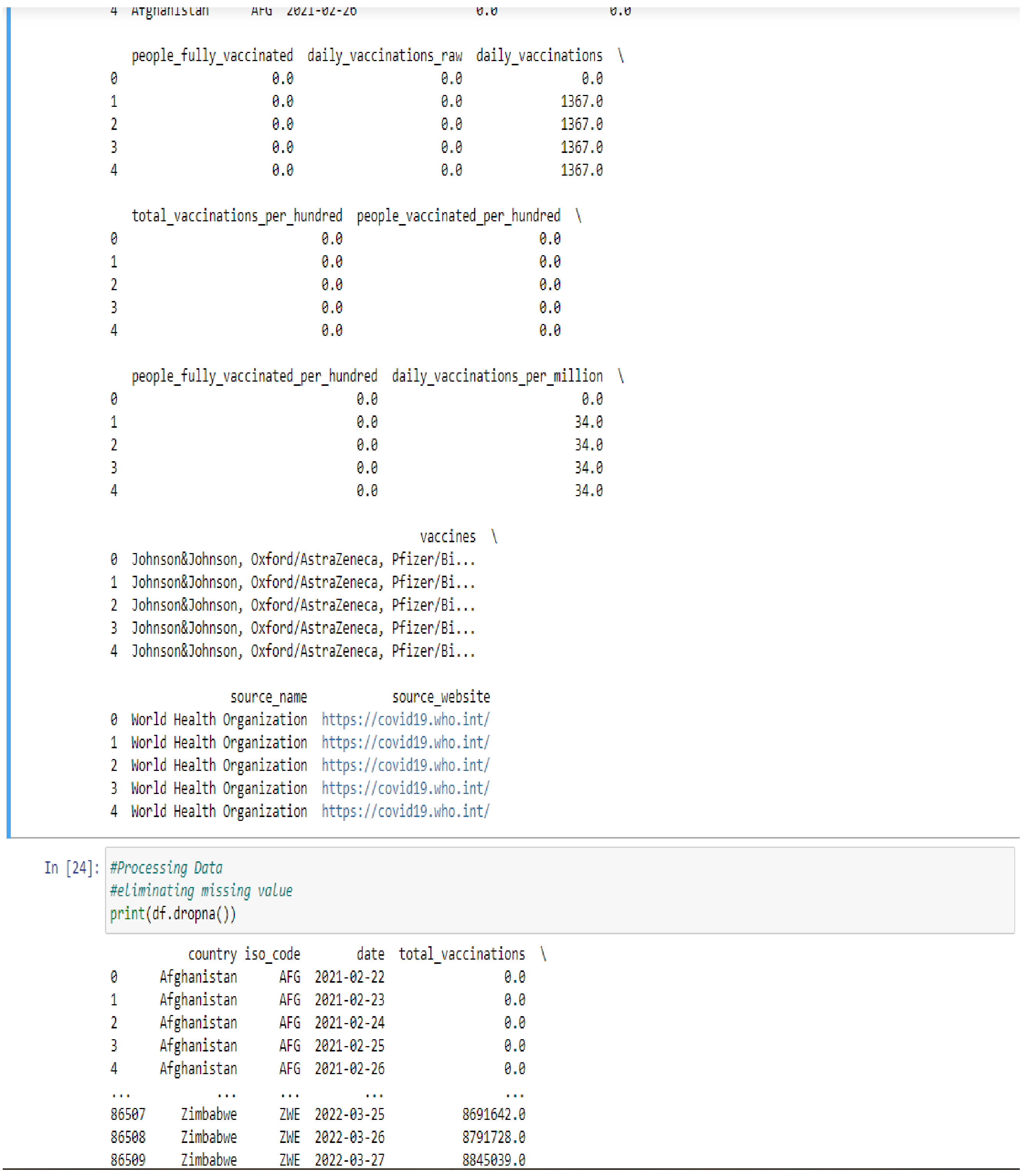
```
In [23]: #filling missing values
         df.fillna(0, inplace=True)
         print(df.head())
```

```
      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
```

Observation:

The above picture shows that there are many null values in our dataset. We will deal with these null values later in this blog. There are two data types as seen from the table object means string and float.

We Use this to fill null value with any value say 0: train_data.fillna(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
```

Observation:

The above picture shows that In order to check missing values in Pandas DataFrame, we use a function **isnull() and notnull().** Both function help in checking whether a value is NaN or not. These function can also be used in Pandas Series in order to find null values in a series.

The **read_csv** function of the pandas library can also be used to read some specific columns and a range of rows.

Use this to drop the rows that contains null values from dataset: train_data.dropna()

Accuracy:

In multilabel classification, this function computes subset **accuracy**: the set of labels predicted for a sample must exactly match the corresponding set of labels in y_true. Read more in the User Guide. Parameters: y_true1d array.

CONCLUSION

We just listed some basics to medium-advanced analysis over here, to give you an idea of how to use the data …