PRIYADARSHINI ENGINEERING COLLEGE, VANIYAMBADI-635 751

ANALYSIS OF THE COVID-19 VACCINE DEVELOPMENT PROCESS

Abstract

The pace of the COVID-19 vaccine development process is unprecedented and is challenging the traditional paradigm of vaccinology science. The main pressure comes from the pandemic situation, but what makes it possible is a complex set of factors and innovative environments built along the times, which this manuscript aims to study.

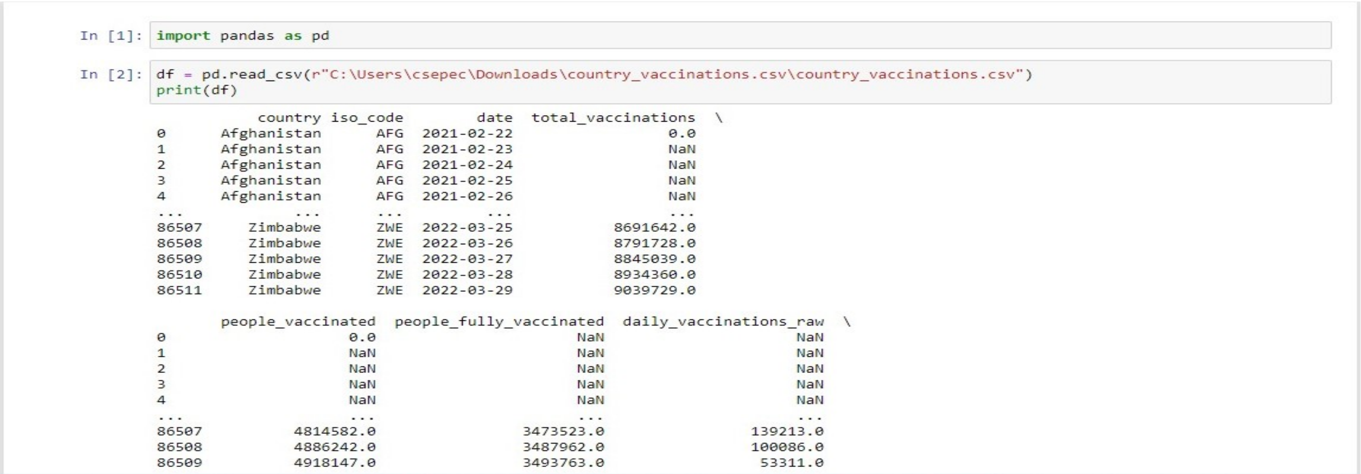
Introduction

The world has witnessed an unprecedented series of events triggered by the pandemic of COVID-19 (coronavirus disease), a disease caused by SARS-CoV-2, a new virus belonging to the Coronavideae family, of great impact on individual and collective health worldwide, and high impact implications for the global economy. On the other hand, it is possible to identify positive aspects in facing the pandemic, ranging from humanitarian solidarity aid actions to accelerating strategies for the development of vaccines, which assumes the position of main hope in solving this problem of global scope.

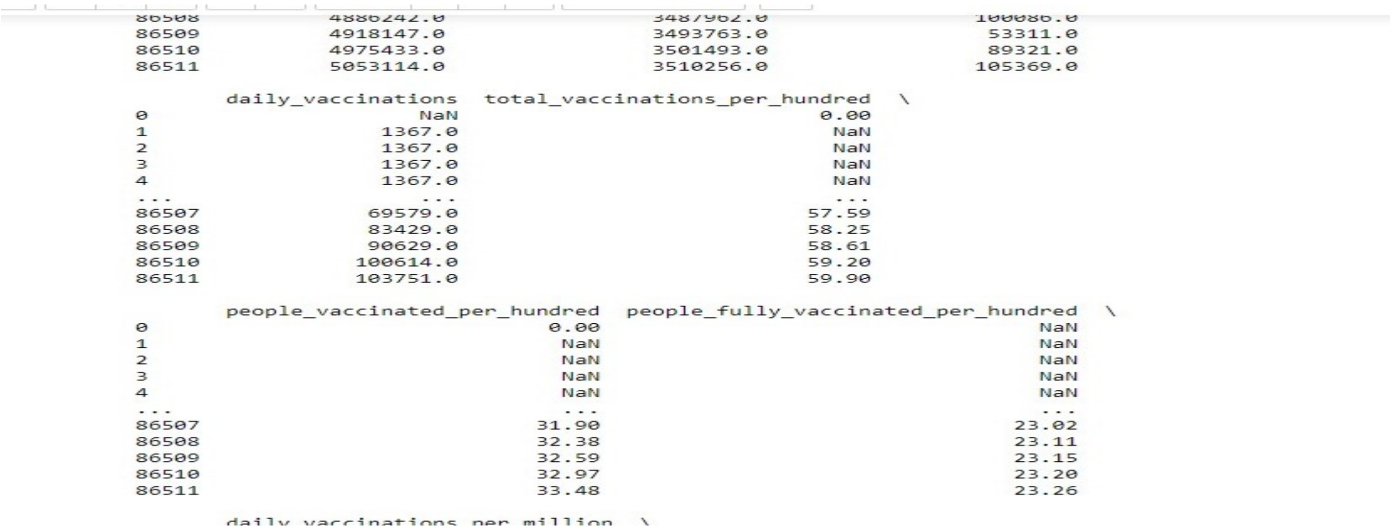
Exploratory Data Analysis

Importing The Datasets:

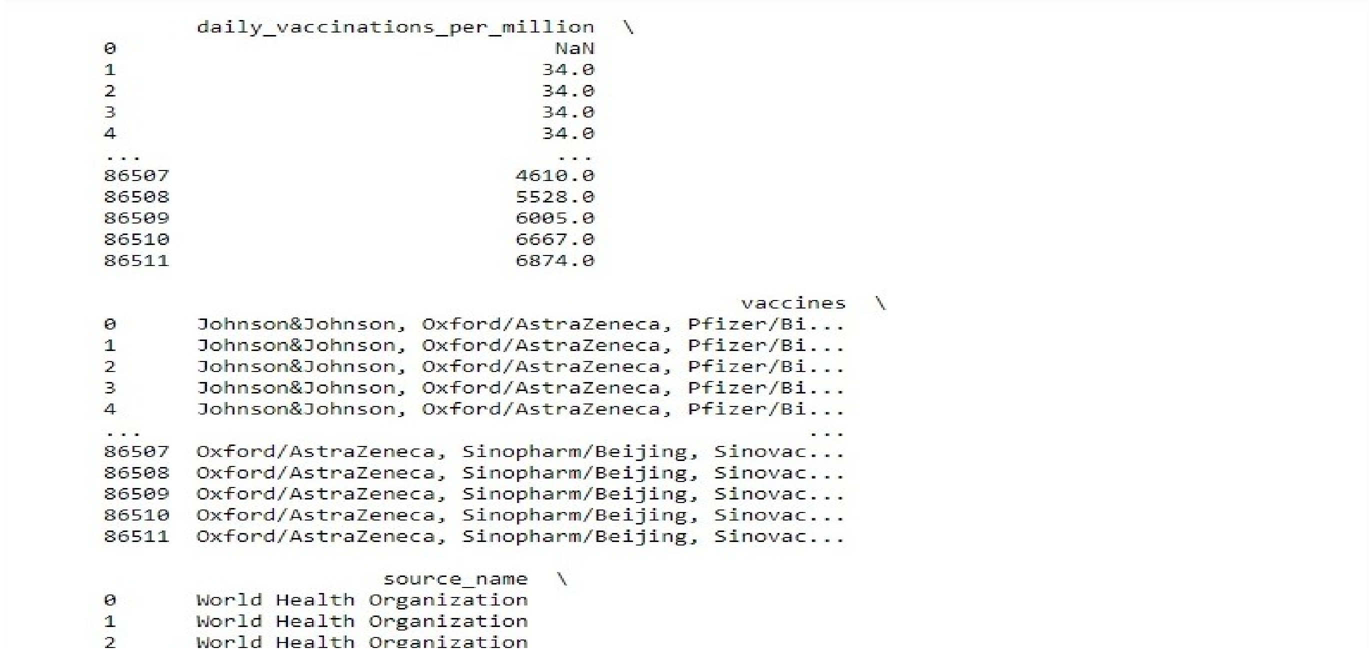
When running python programs, we need to use datasets for data analysis. Python has various modules which help us in importing the external data in various file formats to a python program. In this example we will see how to import data of various formats to a python program.



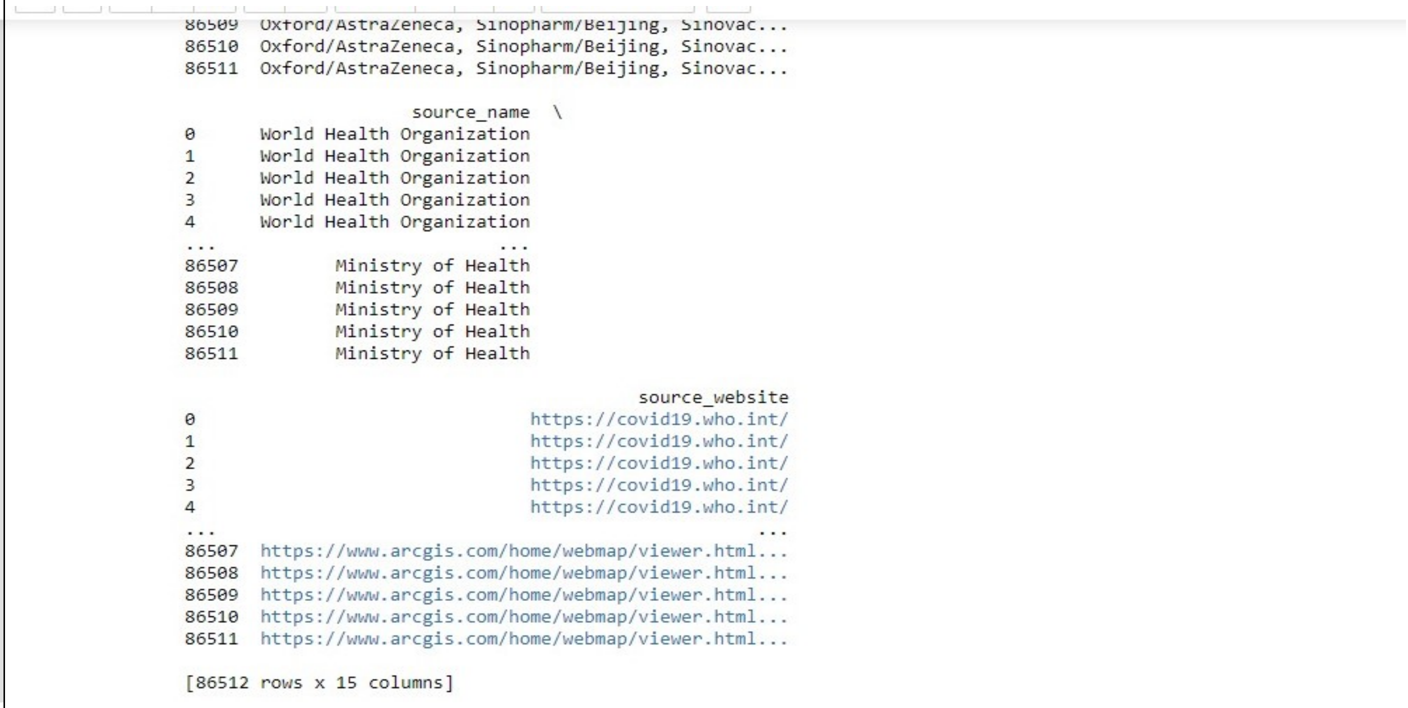
```
In [1]: import pandas as pd

In [2]: df = pd.read_csv(r"C:\Users\csepec\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
```



```
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
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  \
```



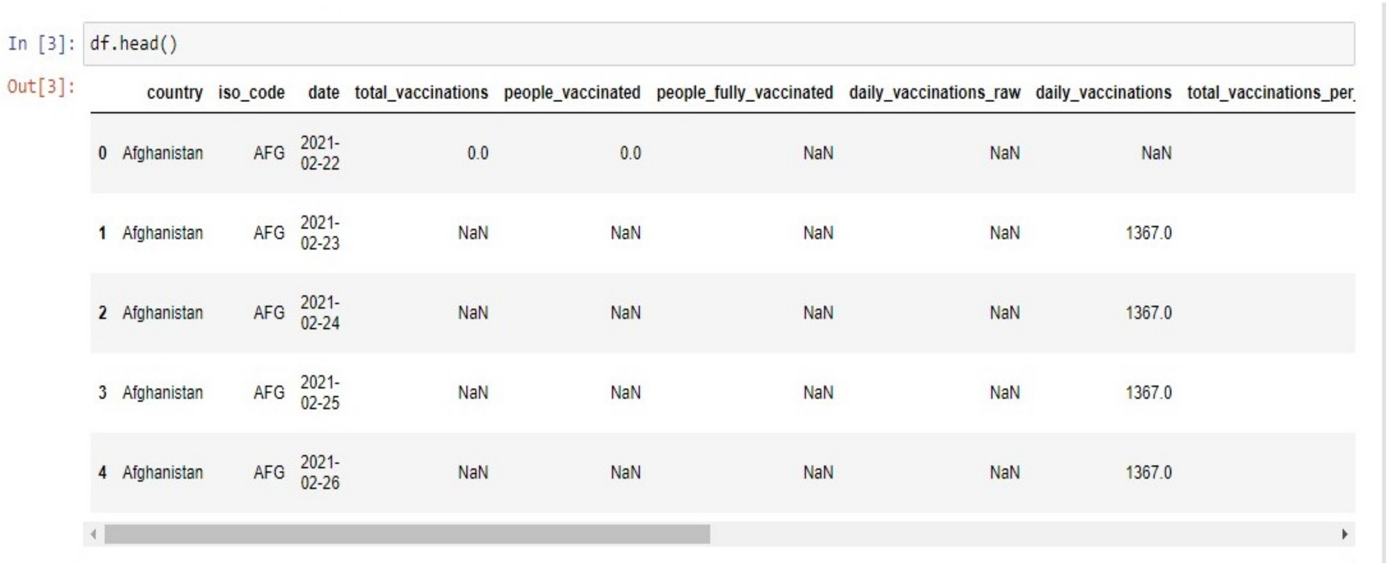
```
       daily_vaccinations_per_million  \
0                                 NaN
1                                34.0
2                                34.0
3                                34.0
4                                34.0
...                               ...
86507                          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
```



```
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
86511          Ministry of Health

                                          source_website
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]
```

Definition and Usage

The head() method returns a specified number of rows, string from the top. The head() method returns the first 5 rows if a number is not specified. ;] Note: The column names will also be returned, in addition to the specified rows.



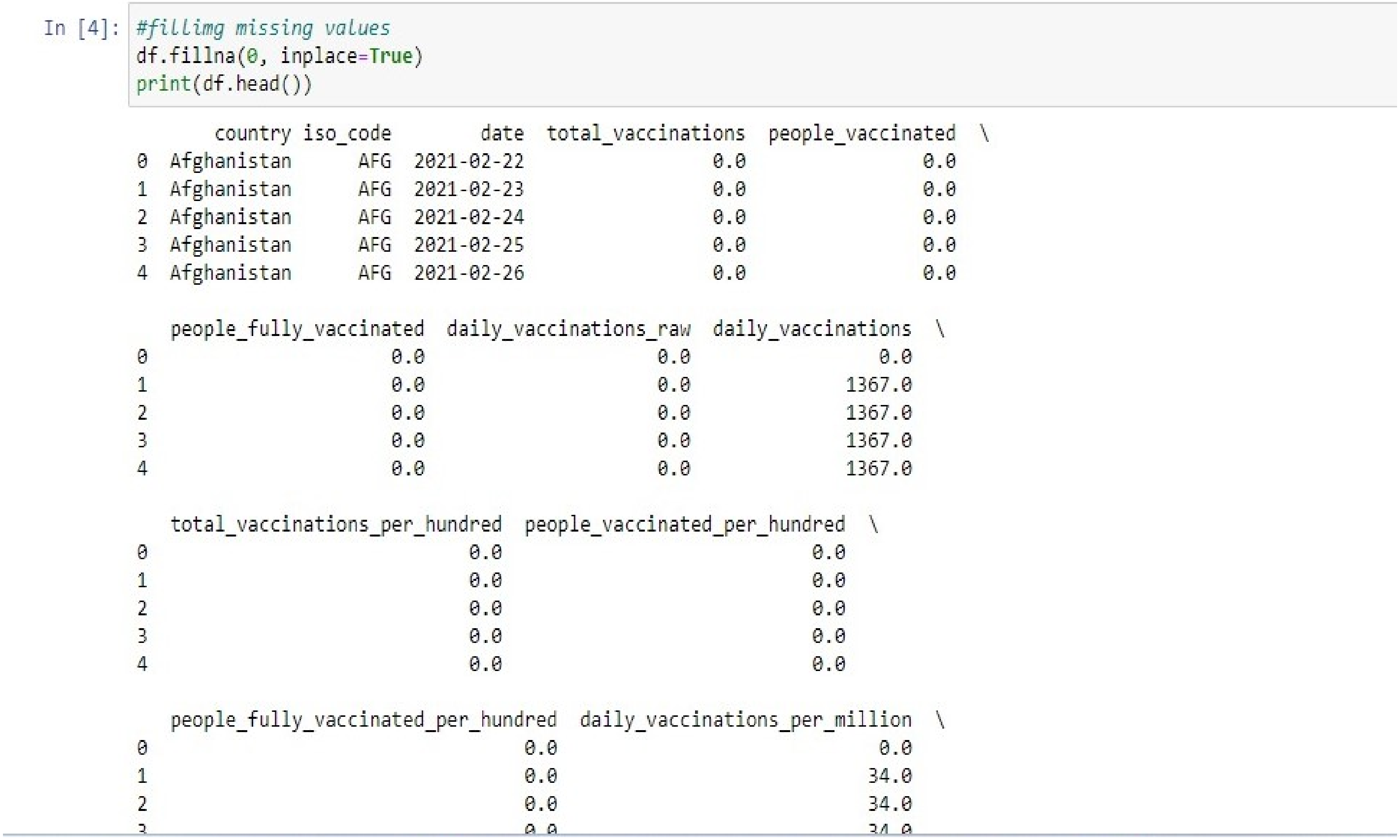
```
In [3]: df.head()
```

Out[3]:

	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	

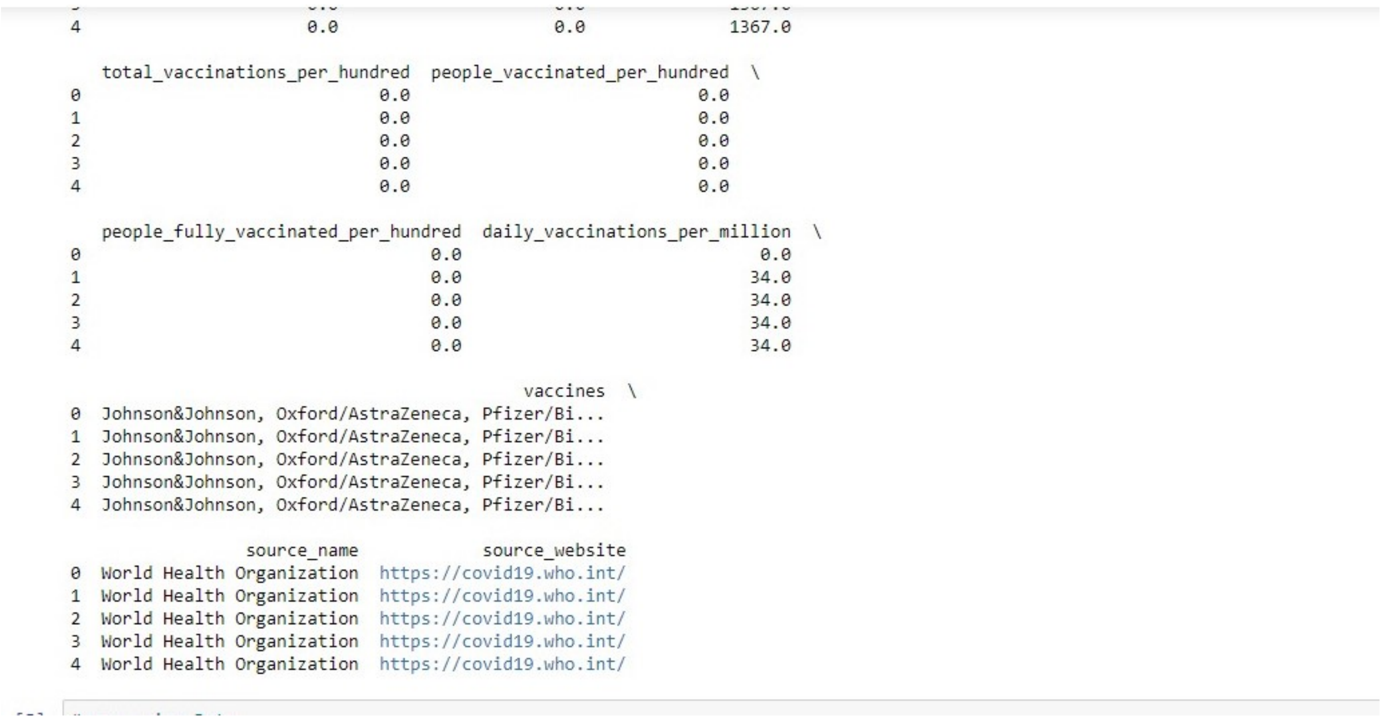
Filling Missing Values:

Missing Data can occur when no information is provided for one or more items or for a whole unit. Missing Data is a very big problem in a real-life scenarios. Missing Data can also refer to as NA(Not Available) values in pandas.



```
In [4]: #fillimg missing values
        df.fillna(0, inplace=True)
        print(df.head())
```

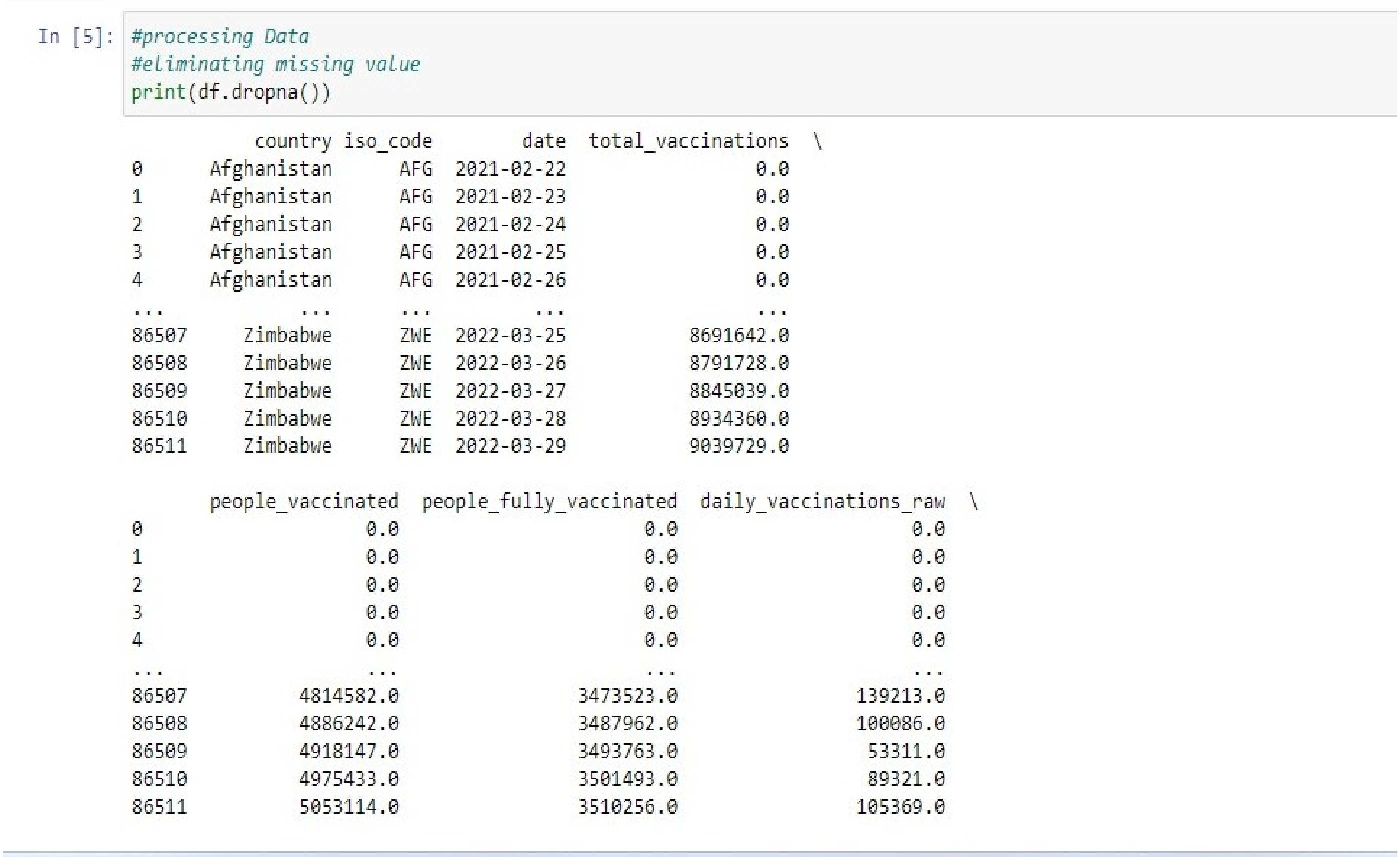
```
       country iso_code        date  total_vaccinations  people_vaccinated  \
0  Afghanistan      AFG  2021-02-22                 0.0                0.0
1  Afghanistan      AFG  2021-02-23                 0.0                0.0
2  Afghanistan      AFG  2021-02-24                 0.0                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
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                 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/
```

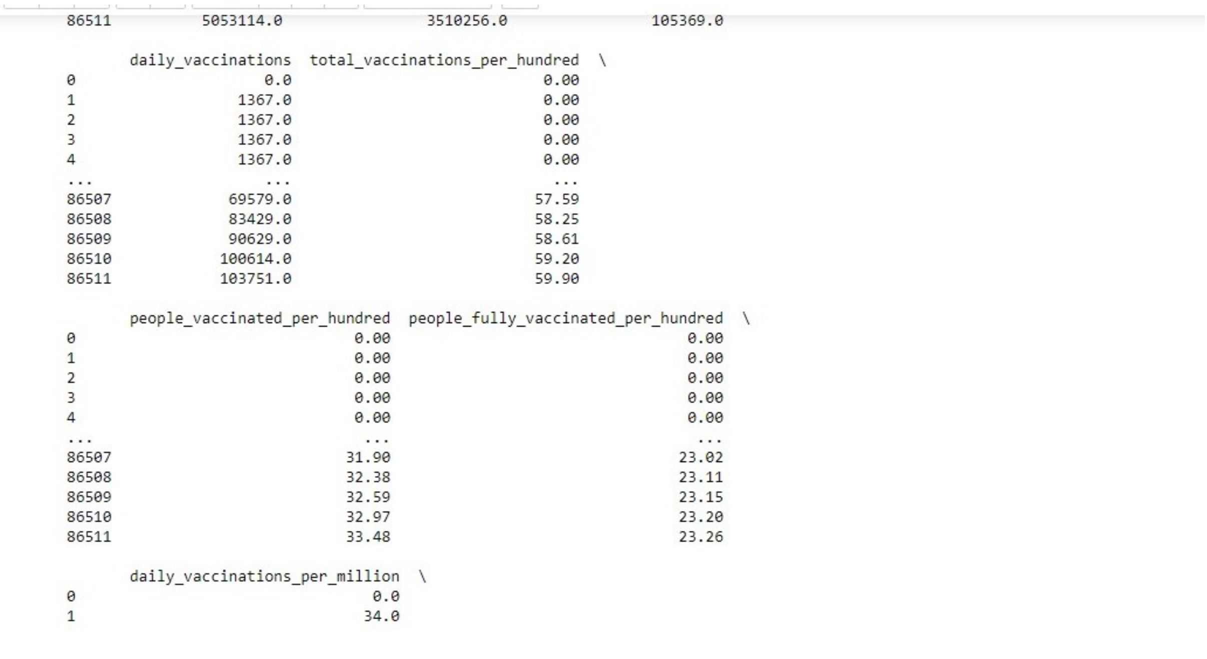
Preprocessing The Given Data:

Data preprocessing is the process of preparing data for analysis by cleaning, transforming, and selecting relevant features. It involves identifying and handling **missing** or **duplicate data**, **scaling features**, **encoding categorical data**, **reducing dimensionality**, and **splitting data** into training and testing sets.

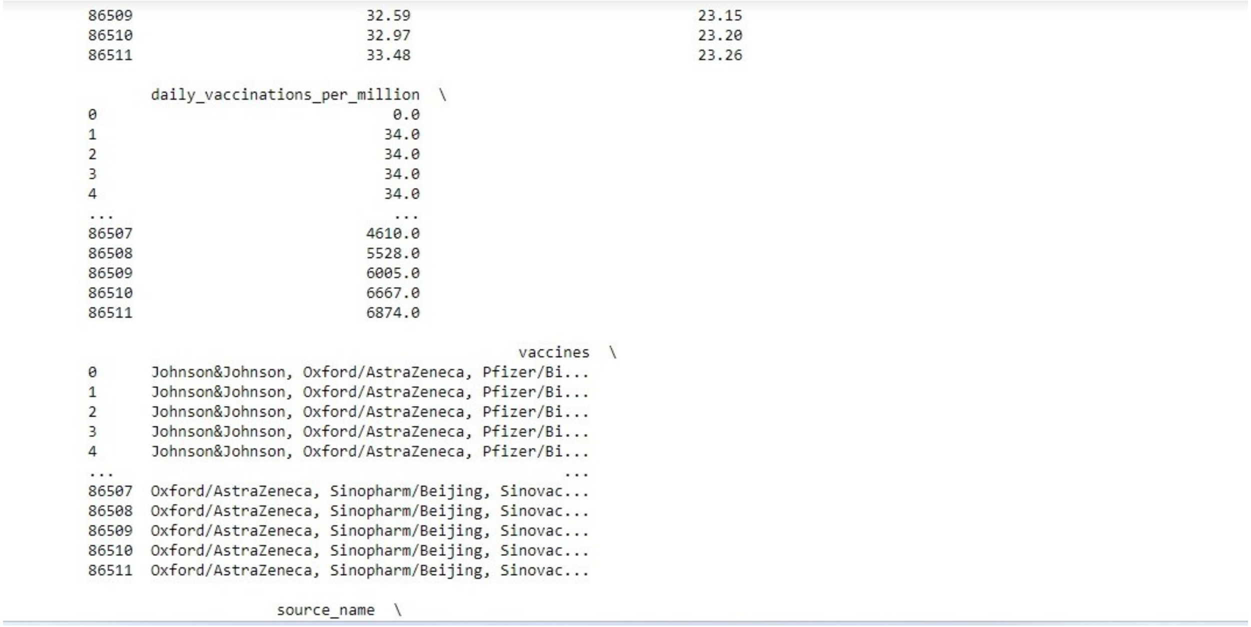


```
In [5]: #processing Data
        #eliminating missing value
        print(df.dropna())
```

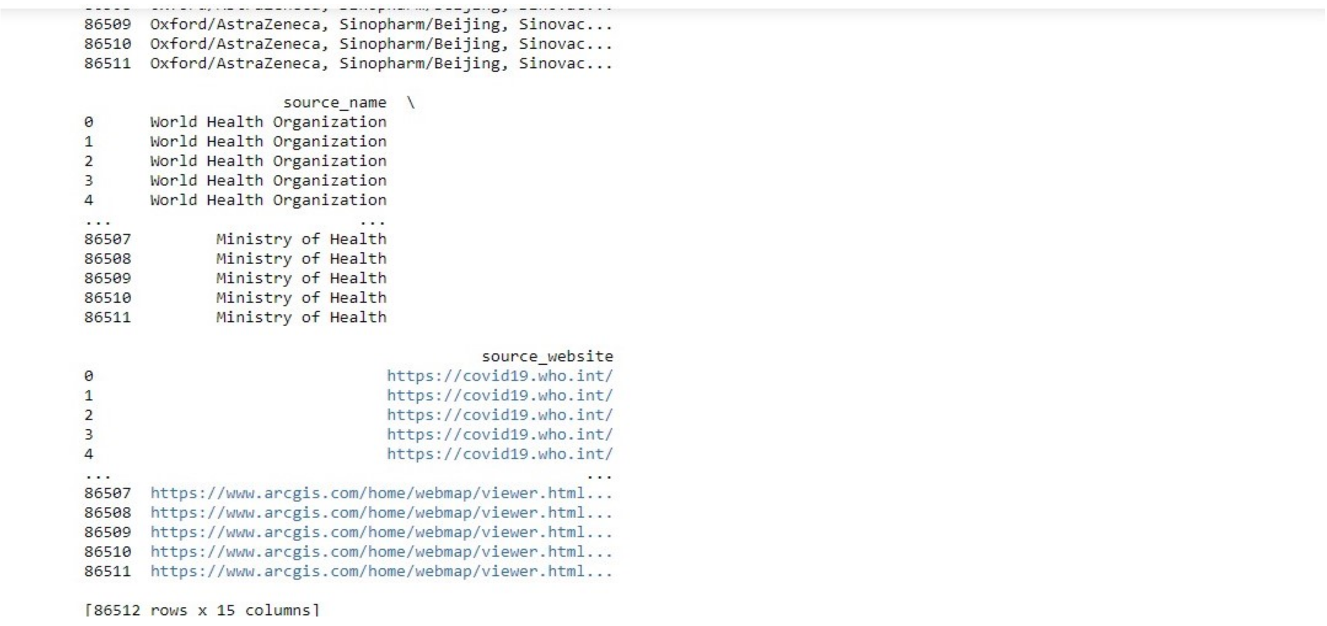
```
           country iso_code        date  total_vaccinations  \
0      Afghanistan      AFG  2021-02-22                 0.0
1      Afghanistan      AFG  2021-02-23                 0.0
2      Afghanistan      AFG  2021-02-24                 0.0
3      Afghanistan      AFG  2021-02-25                 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
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                      0.0                     0.0
1                    0.0                      0.0                     0.0
2                    0.0                      0.0                     0.0
3                    0.0                      0.0                     0.0
4                    0.0                      0.0                     0.0
...                  ...                      ...                     ...
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
```



```
86511       5053114.0             3510256.0             105369.0

       daily_vaccinations  total_vaccinations_per_hundred  \
0                     0.0                            0.00
1                  1367.0                            0.00
2                  1367.0                            0.00
3                  1367.0                            0.00
4                  1367.0                            0.00
...                   ...                             ...
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                                 0.00
1                               0.00                                 0.00
2                               0.00                                 0.00
3                               0.00                                 0.00
4                               0.00                                 0.00
...                              ...                                  ...
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                                 0.0
1                                34.0
```

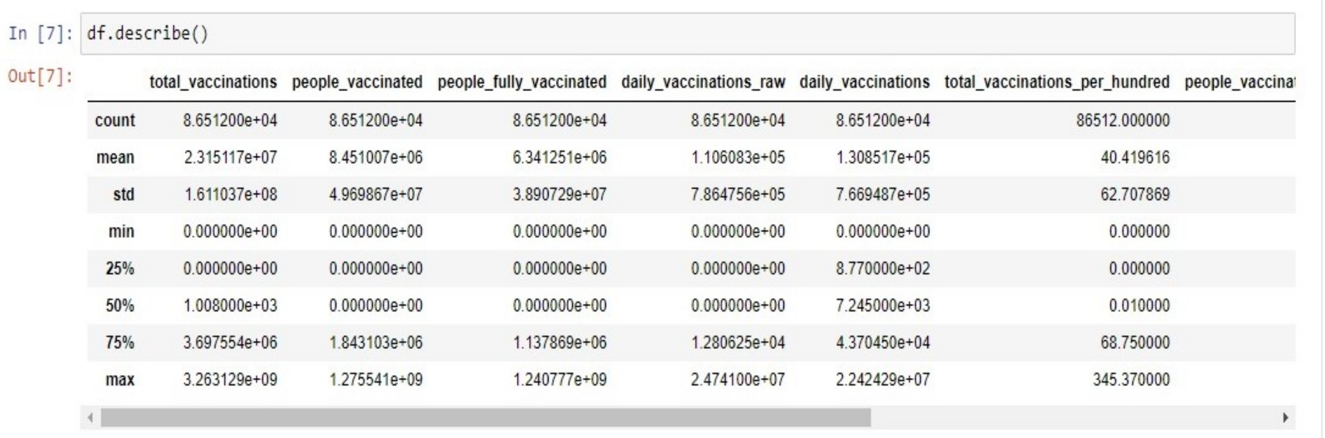


```
86509                          32.59                                23.15
86510                          32.97                                23.20
86511                          33.48                                23.26

       daily_vaccinations_per_million  \
0                                 0.0
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  \
```



```
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
86511         Ministry of Health

                                          source_website
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]
```

DataFrame describe()

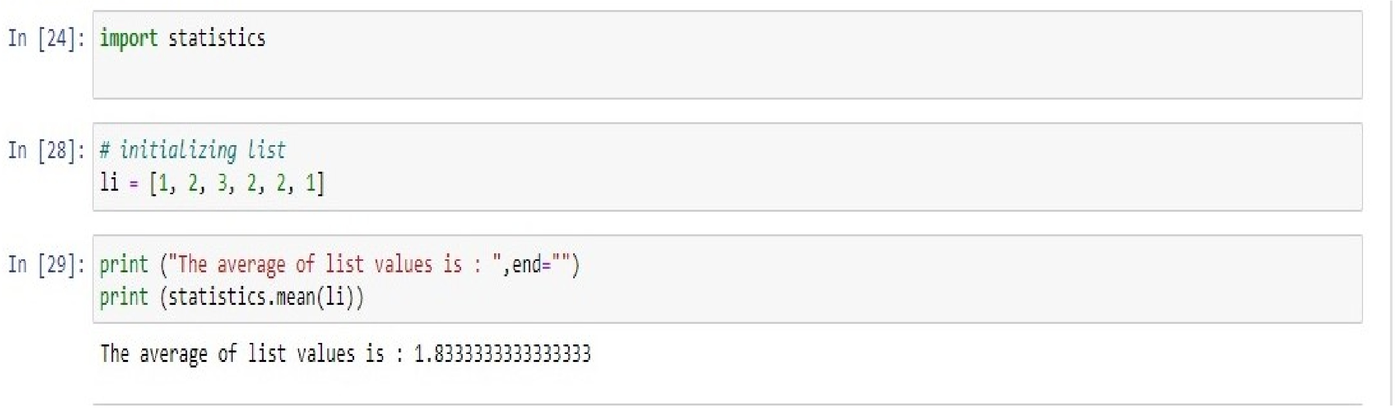


```
In [7]: df.describe()
Out[7]:
```

	total_vaccinations	people_vaccinated	people_fully_vaccinated	daily_vaccinations_raw	daily_vaccinations	total_vaccinations_per_hundred	people_vaccina
count	8.651200e+04	8.651200e+04	8.651200e+04	8.651200e+04	8.651200e+04	86512.000000	
mean	2.315117e+07	8.451007e+06	6.341251e+06	1.106083e+05	1.308517e+05	40.419616	
std	1.611037e+08	4.969867e+07	3.890729e+07	7.864756e+05	7.669487e+05	62.707869	
min	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	0.000000	
25%	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	8.770000e+02	0.000000	
50%	1.008000e+03	0.000000e+00	0.000000e+00	0.000000e+00	7.245000e+03	0.010000	
75%	3.697554e+06	1.843103e+06	1.137869e+06	1.280625e+04	4.370450e+04	68.750000	
max	3.263129e+09	1.275541e+09	1.240777e+09	2.474100e+07	2.242429e+07	345.370000	

Statistical Analysis

MEAN:

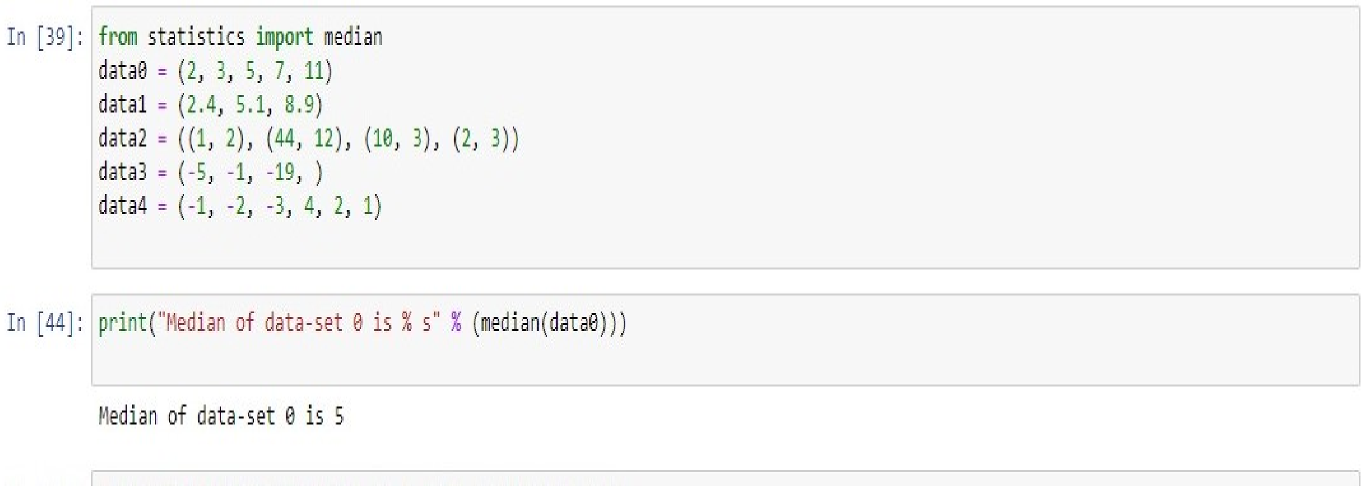


```
In [24]: import statistics
```

```
In [28]: # initializing list
         li = [1, 2, 3, 2, 2, 1]
```

```
In [29]: print ("The average of list values is : ",end="")
         print (statistics.mean(li))

         The average of list values is : 1.8333333333333333
```

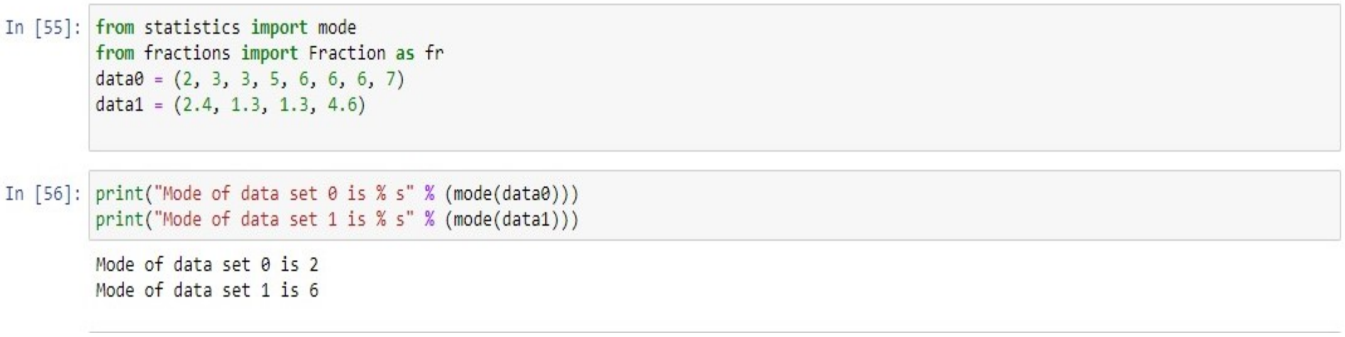
MEDIAN:



```
In [39]: from statistics import median
         data0 = (2, 3, 5, 7, 11)
         data1 = (2.4, 5.1, 8.9)
         data2 = ((1, 2), (44, 12), (10, 3), (2, 3))
         data3 = (-5, -1, -19, )
         data4 = (-1, -2, -3, 4, 2, 1)
```

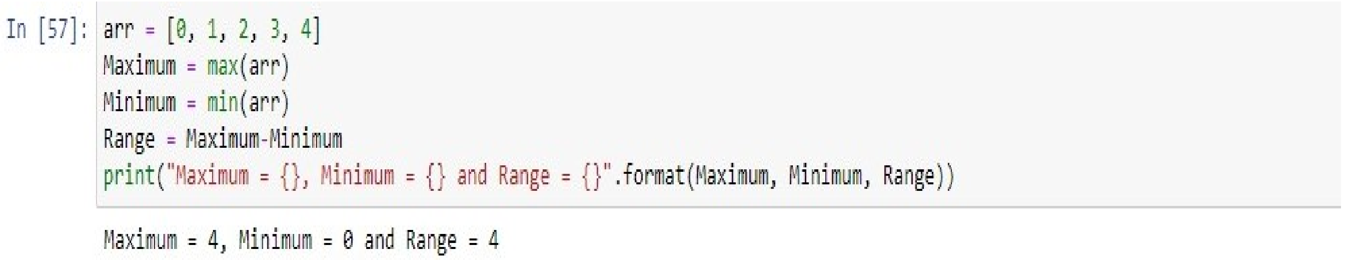
```
In [44]: print("Median of data-set 0 is % s" % (median(data0)))

         Median of data-set 0 is 5
```

MODE:



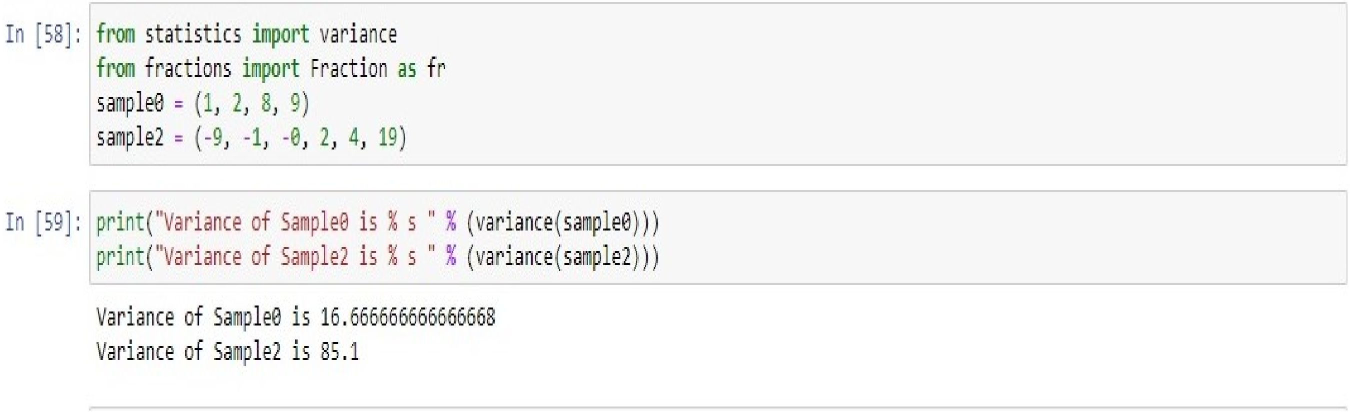
```
In [55]: from statistics import mode
         from fractions import Fraction as fr
         data0 = (2, 3, 3, 5, 6, 6, 6, 7)
         data1 = (2.4, 1.3, 1.3, 4.6)

In [56]: print("Mode of data set 0 is % s" % (mode(data0)))
         print("Mode of data set 1 is % s" % (mode(data1)))

         Mode of data set 0 is 2
         Mode of data set 1 is 6
```

RANGE:



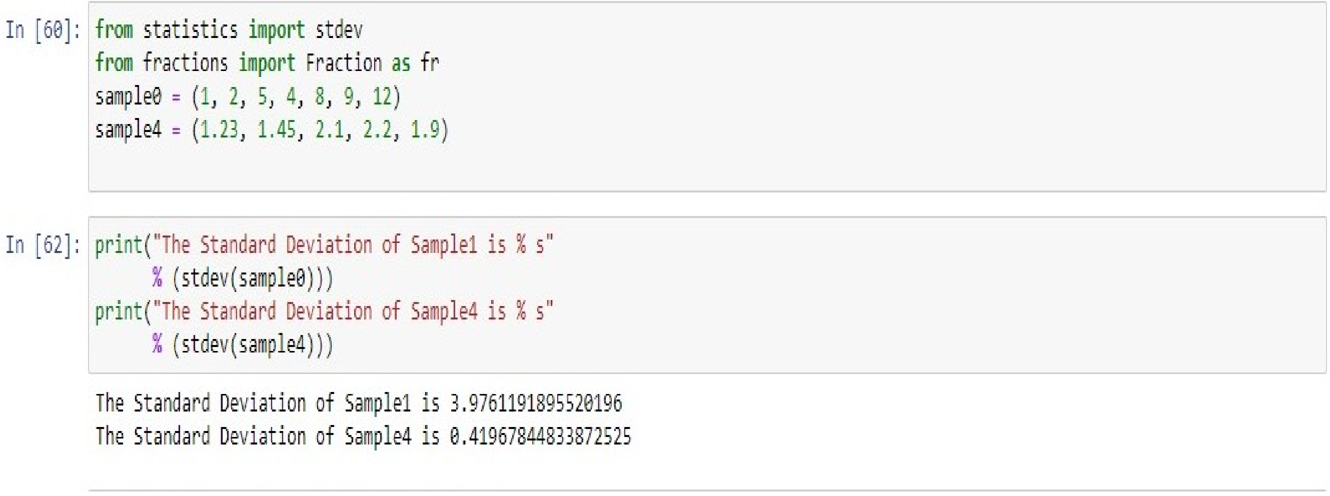
```
In [57]: arr = [0, 1, 2, 3, 4]
         Maximum = max(arr)
         Minimum = min(arr)
         Range = Maximum-Minimum
         print("Maximum = {}, Minimum = {} and Range = {}".format(Maximum, Minimum, Range))

         Maximum = 4, Minimum = 0 and Range = 4
```

VARIANCE:



```
In [58]: from statistics import variance
         from fractions import Fraction as fr
         sample0 = (1, 2, 8, 9)
         sample2 = (-9, -1, -0, 2, 4, 19)

In [59]: print("Variance of Sample0 is % s " % (variance(sample0)))
         print("Variance of Sample2 is % s " % (variance(sample2)))

         Variance of Sample0 is 16.666666666666668
         Variance of Sample2 is 85.1
```

STANDARD DEVIATION:

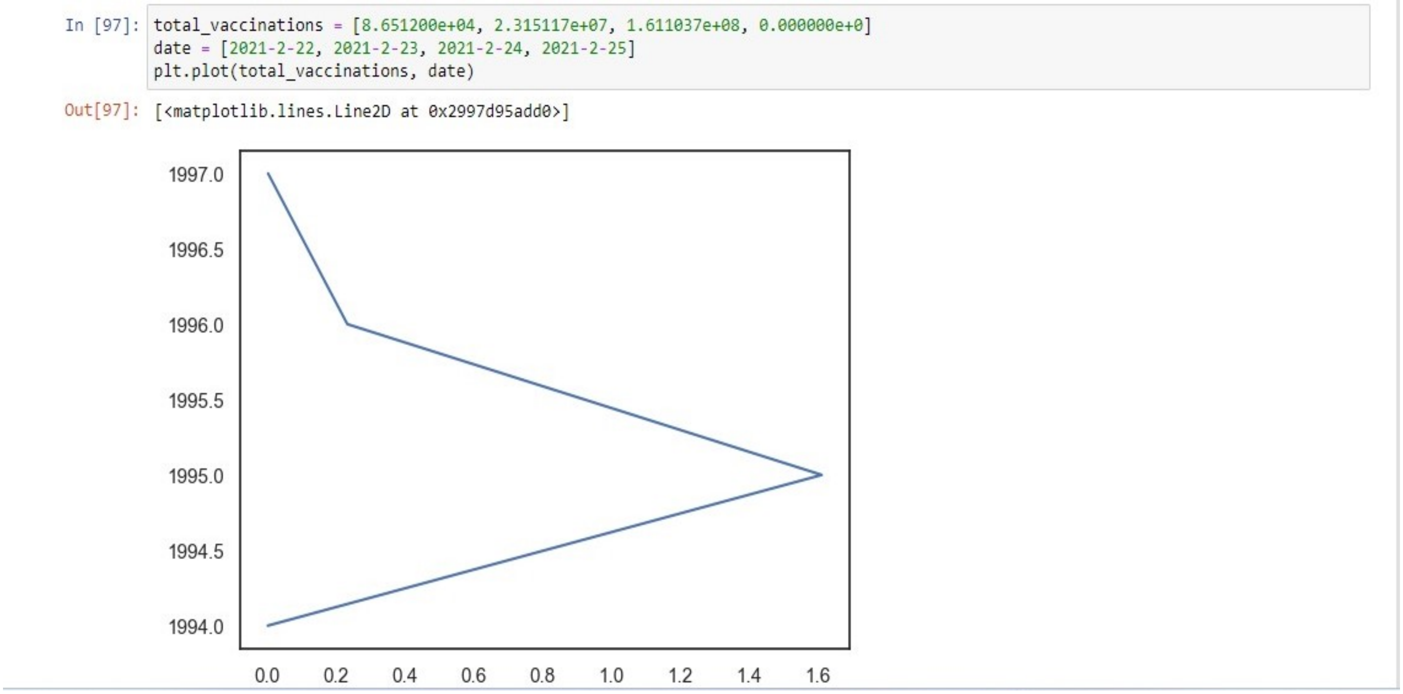


```
In [60]: from statistics import stdev
         from fractions import Fraction as fr
         sample0 = (1, 2, 5, 4, 8, 9, 12)
         sample4 = (1.23, 1.45, 2.1, 2.2, 1.9)

In [62]: print("The Standard Deviation of Sample1 is % s"
               % (stdev(sample0)))
         print("The Standard Deviation of Sample4 is % s"
               % (stdev(sample4)))

         The Standard Deviation of Sample1 is 3.9761191895520196
         The Standard Deviation of Sample4 is 0.41967844833872525
```

Visualization:

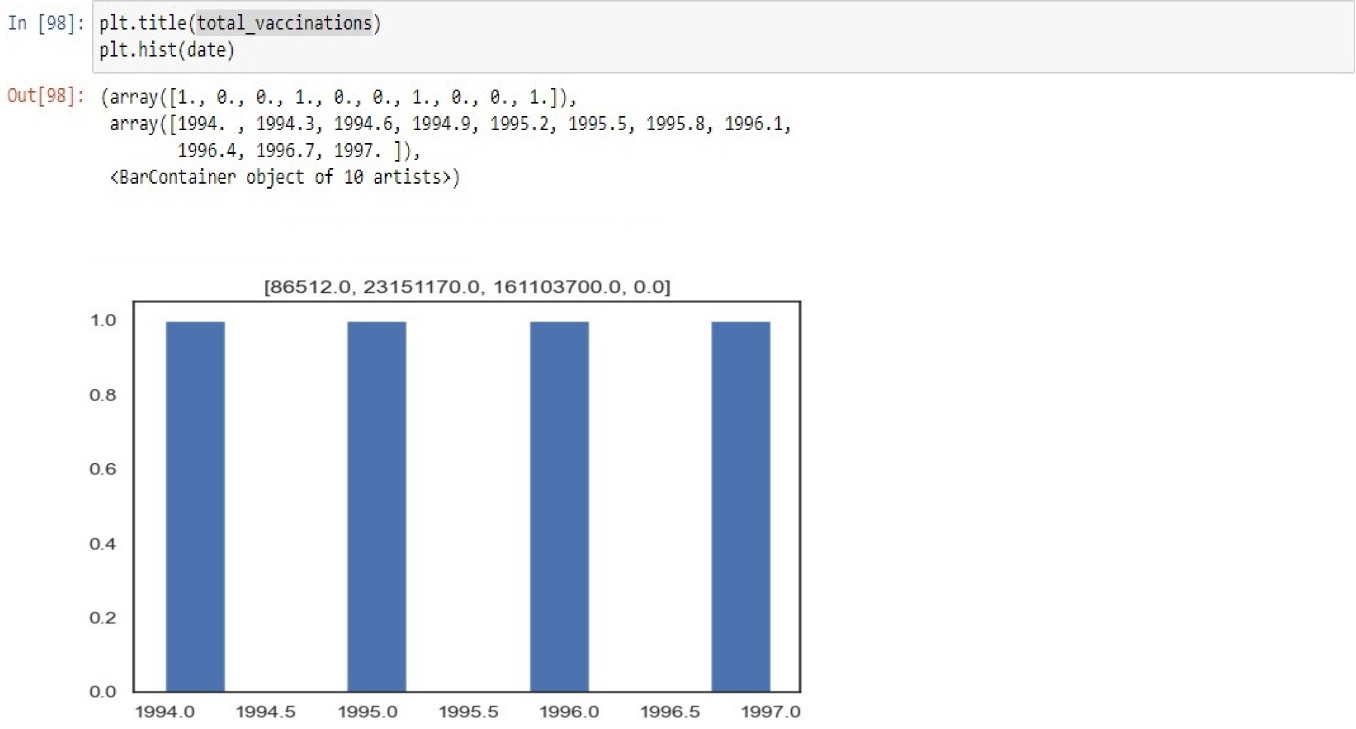
Pyplot:

matplotlib.pyplot is a collection of functions that make matplotlib work like MATLAB. Each pyplot function makes some change to a figure: e.g., creates a figure, creates a plotting area in a figure, plots some lines in a plotting area, decorates the plot with labels, etc.



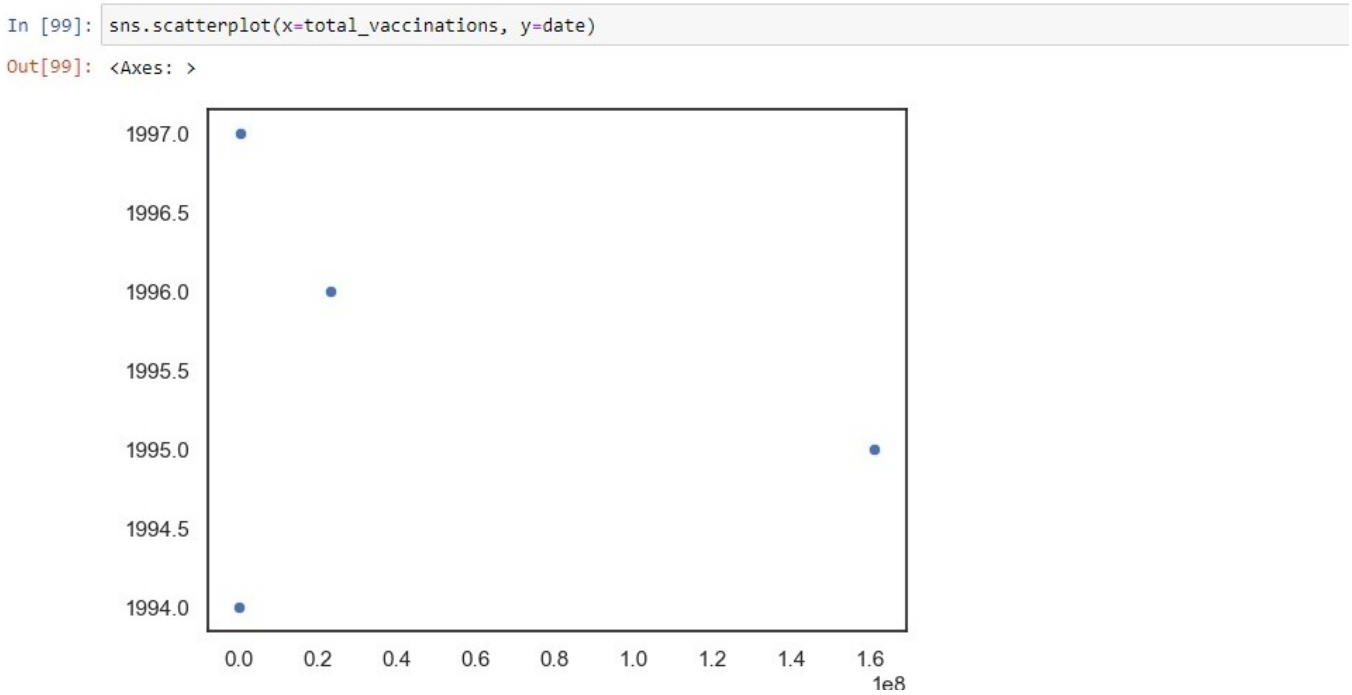
```
In [97]: total_vaccinations = [8.651200e+04, 2.315117e+07, 1.611037e+08, 0.000000e+0]
         date = [2021-2-22, 2021-2-23, 2021-2-24, 2021-2-25]
         plt.plot(total_vaccinations, date)

Out[97]: [<matplotlib.lines.Line2D at 0x2997d95add0>]
```

Histogram:



```
In [98]: plt.title(total_vaccinations)
         plt.hist(date)

Out[98]: (array([1., 0., 0., 1., 0., 0., 1., 0., 0., 1.]),
          array([1994. , 1994.3, 1994.6, 1994.9, 1995.2, 1995.5, 1995.8, 1996.1,
                 1996.4, 1996.7, 1997. ]),
          <BarContainer object of 10 artists>)
```

Scatter Plot:



```
In [99]: sns.scatterplot(x=total_vaccinations, y=date)
Out[99]: <Axes: >
```

Conclusions

The need to rapidly develop a vaccine against COVID-19 occurs at a time of great excitement in basic scientific understanding, as well as strategies learned in the past by industry and optimization of regulatory pathways. It is expected that these factors, arising from the global emergency, may redirect the R&D processes for new drugs, especially in times of pandemic.