

Project Title: Covid-19 Vaccine Analysis

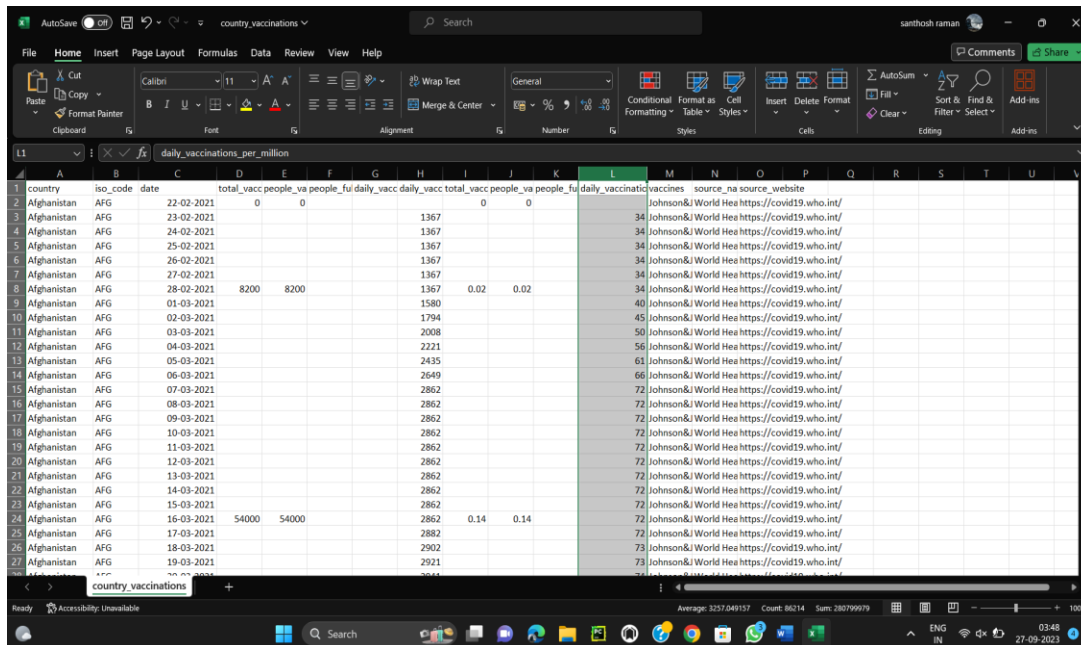
Project Definition:

This Project contains an In-depth Analysis of Covid-19 Vaccine data, Focusing on Vaccine efficacy, distribution and adverse effects. It involves data Collection, Data Preprocessing, Exploratory data Analysis, Statistical Analysis, and Virtualization.

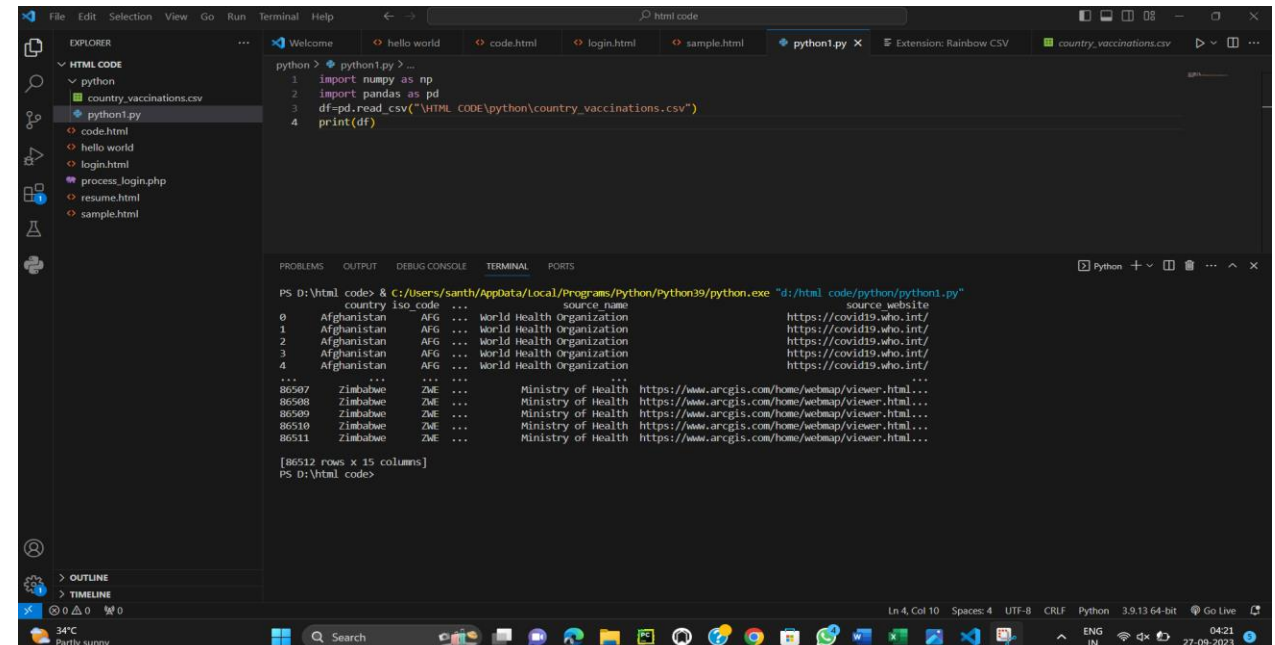
Data Collection:

By making use of the link given below we can download the dataset for our project.

<https://www.kaggle.com/datasets/gpreda/covid-world-vaccination-progress>



country	iso_code	date	total_vacc	people_vaccinated	daily_vacc	total_vacc_people	people_vaccinated	daily_vacc	vaccines	source	source_website
Afghanistan	AFG	22-02-2021	0	0					Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	23-02-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	24-02-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	25-02-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	26-02-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	27-02-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	28-02-2021	8200	8200					Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	01-03-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	02-03-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	03-03-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	04-03-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	05-03-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	06-03-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	07-03-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	08-03-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	09-03-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	10-03-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	11-03-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	12-03-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	13-03-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	14-03-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	15-03-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	16-03-2021	54000	54000					Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	17-03-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	18-03-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/
Afghanistan	AFG	19-03-2021							Johnson & Johnson	World Health Organization	https://covid19.who.int/



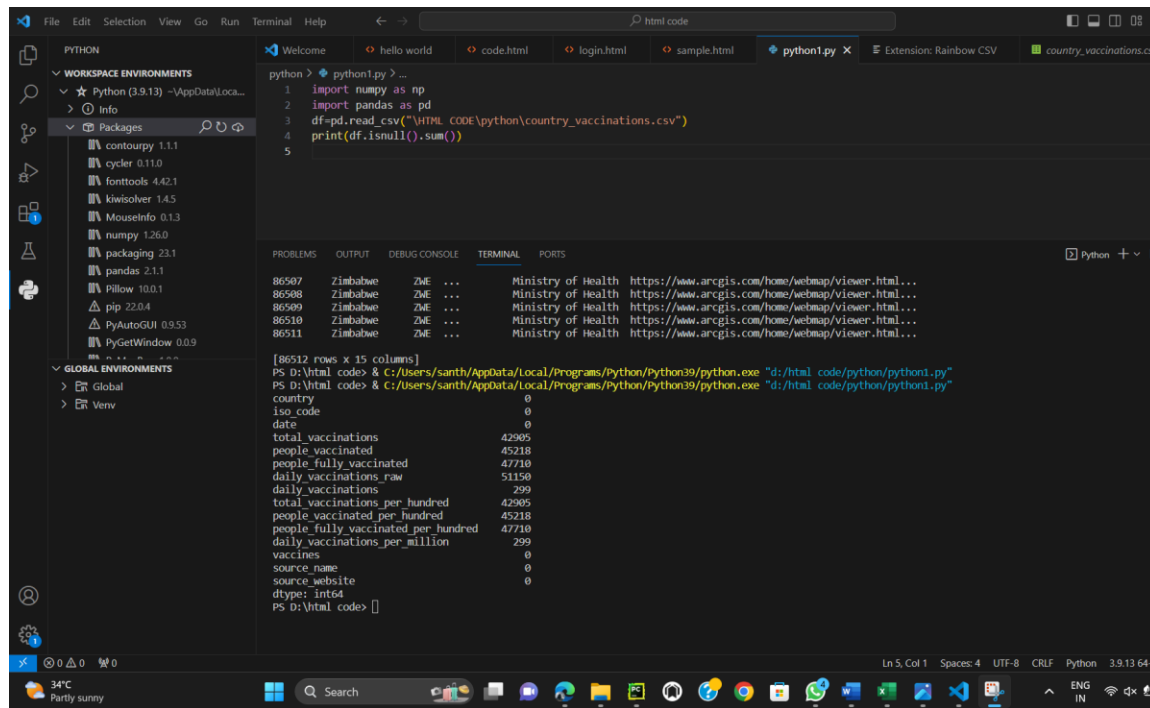
```
python> python1.py ...
1 import numpy as np
2 import pandas as pd
3 df=pd.read_csv("html code\python\country_vaccinations.csv")
4 print(df)
```

```
PS D:\html code> & C:\Users\santhos\AppData\Local\Programs\Python\Python39\python.exe "d:\html code\python\python1.py"
country iso_code date total_vacc people_vaccinated daily_vacc total_vacc_people people_vaccinated daily_vacc vaccines source source_website
0 Afghanistan AFG ... World Health Organization https://covid19.who.int/
1 Afghanistan AFG ... World Health Organization https://covid19.who.int/
2 Afghanistan AFG ... World Health Organization https://covid19.who.int/
3 Afghanistan AFG ... World Health Organization https://covid19.who.int/
4 Afghanistan AFG ... World Health Organization https://covid19.who.int/
... ..
86507 Zimbabwe ZWE ... Ministry of Health https://www.arcgis.com/home/webmap/viewer.html...
86508 Zimbabwe ZWE ... Ministry of Health https://www.arcgis.com/home/webmap/viewer.html...
86509 Zimbabwe ZWE ... Ministry of Health https://www.arcgis.com/home/webmap/viewer.html...
86510 Zimbabwe ZWE ... Ministry of Health https://www.arcgis.com/home/webmap/viewer.html...
86511 Zimbabwe ZWE ... Ministry of Health https://www.arcgis.com/home/webmap/viewer.html...

[86512 rows x 11 columns]
PS D:\html code>
```

Data Preprocessing:

- In Case of this part we will be finding the Total number of missing values in the given dataset and Handling it by plotting a Heatmap

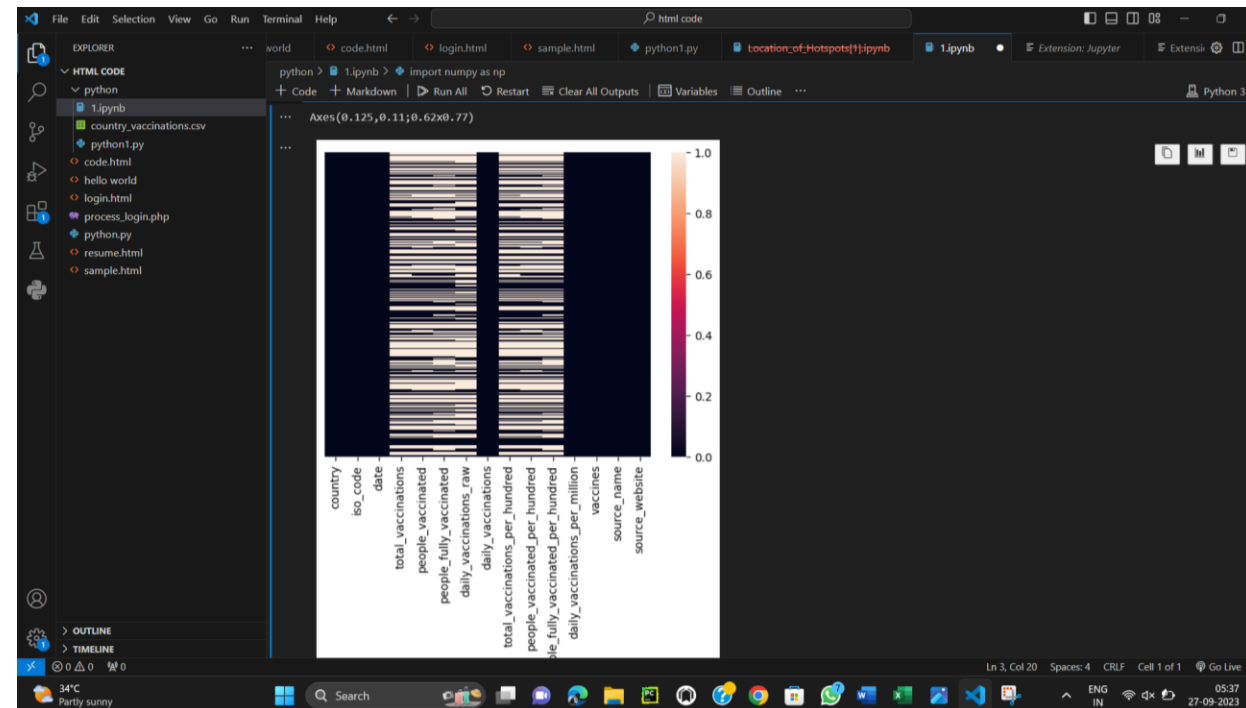


The screenshot shows a VS Code editor with a Python file named `python1.py`. The code reads a CSV file `country_vaccinations.csv` and prints the sum of missing values for each column. The output in the terminal shows that there are 0 missing values for most columns, but 299 missing values for the `daily_vaccinations` column.

```
python > python1.py > ...
1 import numpy as np
2 import pandas as pd
3 df=pd.read_csv("html code\python\country_vaccinations.csv")
4 print(df.isnull().sum())
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	people_fully_vaccinated_per_hundred	daily_vaccinations_per_million	vaccines	source_name	source_website
86507	Zimbabwe	ZME	...	Ministry of Health	https://www.arcgis.com/home/webmap/viewer.html...	0	0	42905	45218	47710	51150	299

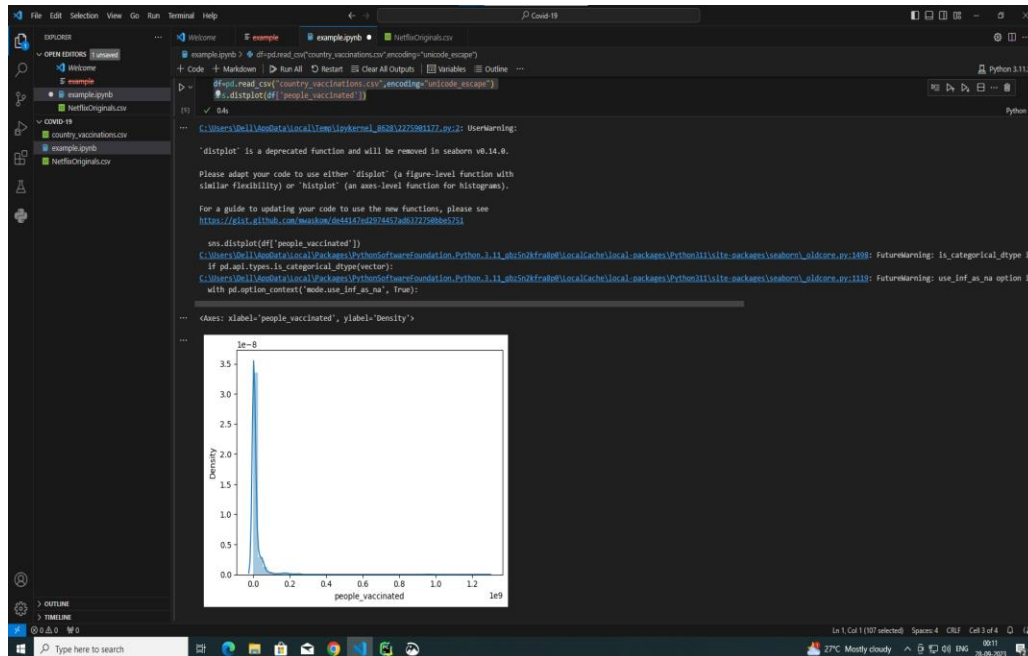
[86512 rows x 15 columns]
PS D:\html code> & C:/Users/santh/AppData/Local/Programs/Python/Python39/python.exe "d:/html code/python/python1.py"
PS D:\html code> & C:/Users/santh/AppData/Local/Programs/Python/Python39/python.exe "d:/html code/python/python1.py"



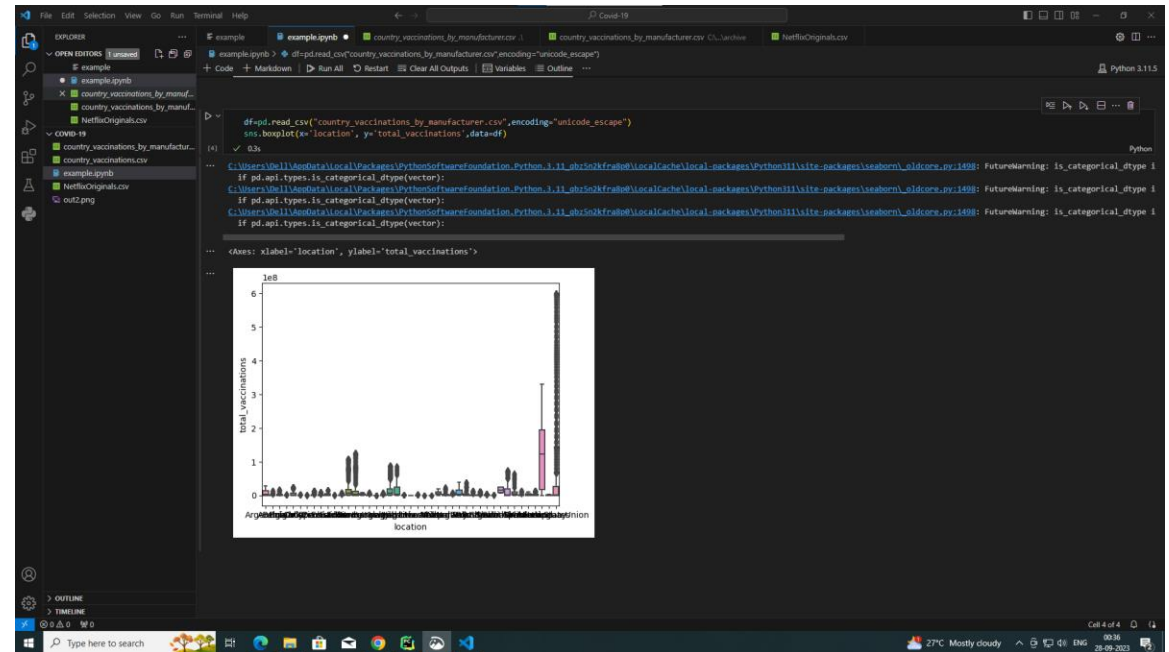
Exploratory Data Analysis:

In case of Exploratory Data Analysis(EDA) we have gone through the concept of identifying the outliers and plotting them in various type of Graphs.

1. As a very first step of this case we have plotted a distribution chart for the downloaded dataset by understanding it's characteristics.



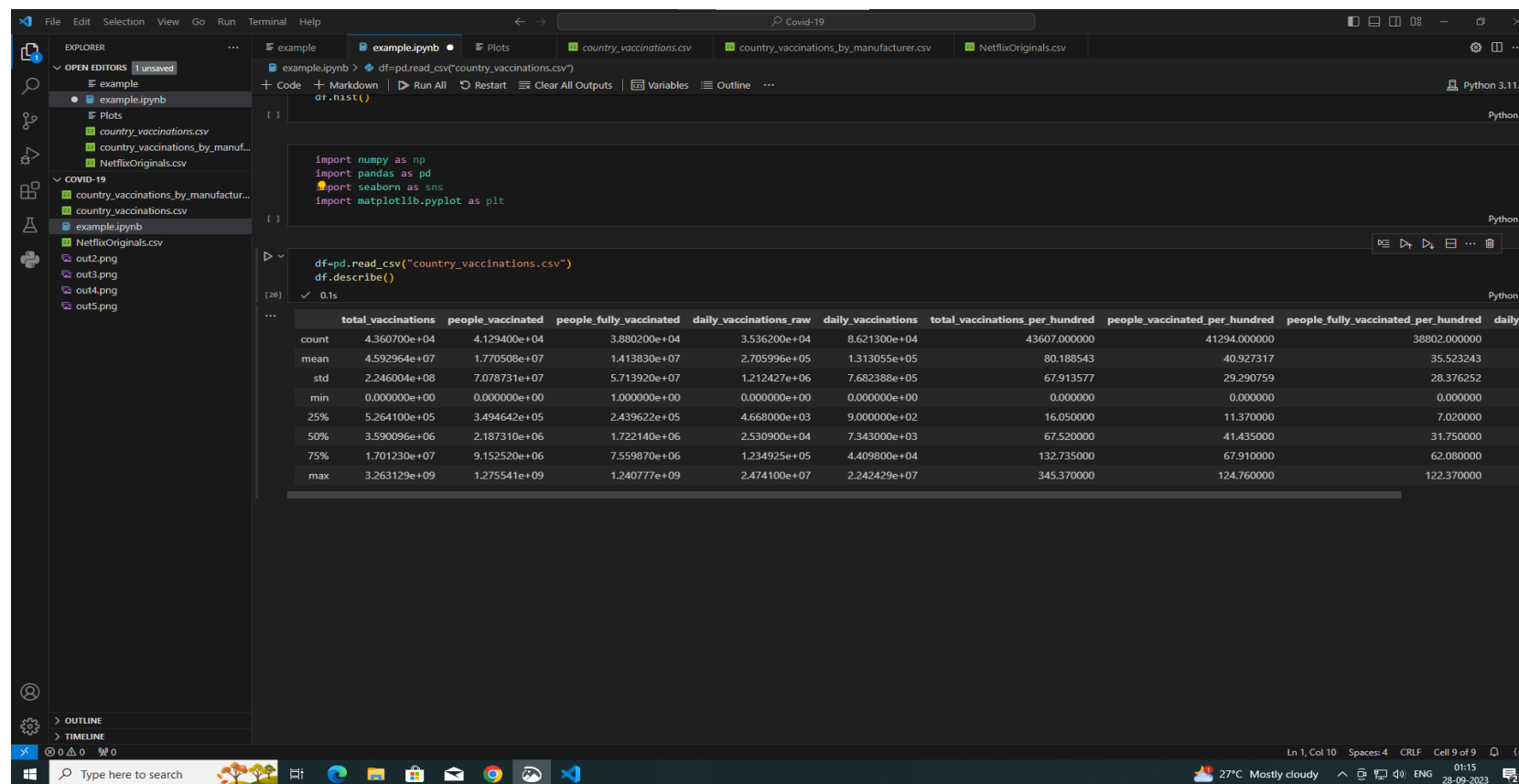
2. Secondly we have plotted the boxplot for given dataset by Understanding about Outliers .



Statistical Analysis:

- It simply describes the basic Statistics for all continuous variables And Nan values are automatically skipped in these statistics. It Indicates the Count of a variable, Mean, Standard deviation, Minimum and Maximum value

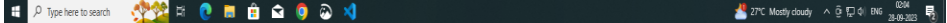
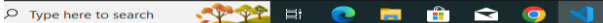
Statistical distribution contains the dataset variables that has either Integers or Floating Point type.



The screenshot shows a Jupyter Notebook interface with a file explorer on the left, a code editor in the center, and a terminal at the bottom. The code in the notebook reads a CSV file named 'country_vaccinations.csv' and displays its statistical distribution using the `df.describe()` method. The output is a table with 11 columns and 10 rows of statistics.

	total_vaccinations	people_vaccinated	people_fully_vaccinated	daily_vaccinations_raw	daily_vaccinations	total_vaccinations_per_hundred	people_vaccinated_per_hundred	people_fully_vaccinated_per_hundred	daily_vaccinations_per_hundred	daily_vaccinations_smoothed
count	4.360700e+04	4.129400e+04	3.880200e+04	3.536200e+04	8.621300e+04	43607.000000	41294.000000	38802.000000	35362.000000	35362.000000
mean	4.592964e+07	1.770508e+07	1.413830e+07	2.705996e+05	1.313055e+05	80.188543	40.927317	35.523243	28.376252	28.376252
std	2.246004e+08	7.078731e+07	5.713920e+07	1.212427e+06	7.682388e+05	67.913577	29.290759	28.376252	28.376252	28.376252
min	0.000000e+00	0.000000e+00	1.000000e+00	0.000000e+00	0.000000e+00	0.000000	0.000000	0.000000	0.000000	0.000000
25%	5.264100e+05	3.494642e+05	2.439622e+05	4.668000e+03	9.000000e+02	16.050000	11.370000	7.020000	7.020000	7.020000
50%	3.590096e+06	2.187310e+06	1.722140e+06	2.530900e+04	7.343000e+03	67.520000	41.435000	31.750000	31.750000	31.750000
75%	1.701230e+07	9.152520e+06	7.559870e+06	1.234925e+05	4.409800e+04	132.735000	67.910000	62.080000	62.080000	62.080000
max	3.263129e+09	1.275541e+09	1.240777e+09	2.474100e+07	2.242429e+07	345.370000	124.760000	122.370000	122.370000	122.370000

100




Insights and Recommendations:


Variants: Monitoring the impact of COVID-19 variants on vaccine effectiveness is essential. New variants may require booster shots or updated vaccines to maintain protection.



Vaccine Distribution: Ensuring equitable distribution of vaccines globally is critical to achieving widespread immunity. Disparities in vaccine access can prolong the pandemic.



Booster Shots: Research indicated that booster shots might be necessary to maintain immunity, especially for certain populations and in response to new variants.



Long-term Effects: Continuously monitor and research the long-term effects of COVID-19 vaccines, including any potential rare side effects.



Global Collaboration: Collaboration between countries, organizations, and pharmaceutical companies is essential for effective vaccine research, production, and distribution