NAME: SAUMYAJEET SINGH BHATI

ROLL NO.: CS8-03

PRN: 202401110065

BATCH: C-81

DIVISION: CS8

DATASET USED: COVID-19

EMAIL: 202401110065@MITAOE.AC.IN

SUBJECT: ESSENTIAL OF DATASCIENCE

GOOGLE COL&B LINK:

HTTPS://COL&B.RESE&RCH.GOOGLE.CO
M/DRIVE/19Y1HBZMDFZHNQFOU7HVECC
IQIZHYYTCK?USP=DRIVE_LINK

D&T&SET:

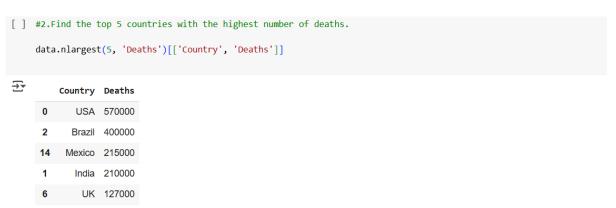
HTTPS://DOCS.GOOGLE.COM/SPREADSHEETS/D/
1AOTLPHLEPZ4KHIIEHRX1TWSROYBDIKOO/EDIT
?USP=DRIVE_LINK&OUID=100157696272359354
723&RTPOF=TRUE&SD=TRUE

```
import pandas as pd
import numpy as np

# Load the dataset
data = pd.read_csv('/content/covid19_dataset.csv') # Replace with your file path
```

1.FIND THE TOP 5 COUNTRIES WITH THE HIGHEST NUMBER OF CONFIRMED CASES.

2.FIND THE TOP 5 COUNTRIES WITH THE HIGHEST NUMBER OF DEATHS.



3.CALCULATE THE OVERALL GLOBAL VACCINATION RATE (AVERAGE VACCINATION %).

```
[ ] #3.Calculate the overall global vaccination rate (average vaccination %).
    data['Vaccination Rate (%)'].mean()

    np.float64(20.183333333333333)
```

4.LIST ALL COUNTRIES WITH VACCINATION RATE GREATER THAN 50%.

```
#4.List all countries with vaccination rate greater than 50%.

data[data['Vaccination Rate (%)'] > 50][['Country', 'Vaccination Rate (%)']]

Country Vaccination Rate (%)

6 UK 50.1

30 Israel 60.5

39 UAE 54.3
```

5.FIND THE COUNTRY WITH THE HIGHEST ACTIVE CASES.

```
[ ] #5.Find the country with the highest active cases.
   data.loc[data['Active Cases'].idxmax()][['Country', 'Active Cases']]
```



6.FIND THE COUNTRY WITH THE LOWEST ACTIVE CASES (EXCLUDING COUNTRIES WITH 0 CASES).

0			the lowest active cases (excluding countries with 0 cases). > 0].loc[data['Active Cases'].idxmin()][['Country', 'Active Cases']]
₹		59	
	Country	New Zealand	
	Active Cases	74	
	dtype: object		

7.CALCULATE THE DEATH RATE (%) FOR EACH COUNTRY AND FIND THE TOP 5 COUNTRIES WITH THE HIGHEST DEATH RATE.

8.CALCULATE THE RECOVERY RATE (%) FOR EACH COUNTRY AND FIND THE TOP 5 COUNTRIES WITH THE HIGHEST RECOVERY RATE.



9.FIND THE AVERAGE NUMBER OF CONFIRMED CASES AMONG ALL COUNTRIES.

[]	<pre>#9.Find the average number of confirmed cases among all countries. data['Confirmed Cases'].mean()</pre>
→	np.float64(2343560.0)

10. FIND HOW MANY COUNTRIES HAVE MORE THAN 1 MILLION CONFIRMED CASES.

```
[ ] #10.Find how many countries have more than 1 million confirmed cases.
   data[data['Confirmed Cases'] > 1_000_000].shape[0]
```

→ 24

11. FIND THE CORRELATION BETWEEN VACCINATION RATE AND ACTIVE CASES.

[] #11.Find the correlation between vaccination rate and active cases.
 data[['Vaccination Rate (%)', 'Active Cases']].corr()

 Vaccination Rate (%)
 Active Cases

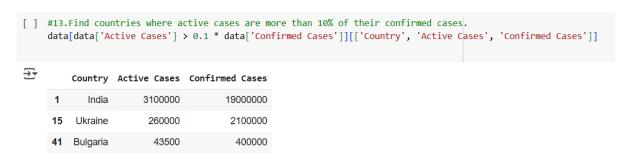
 Vaccination Rate (%)
 1.00000
 0.02588

 Active Cases
 0.02588
 1.00000

12. GROUP COUNTRIES BASED ON VACCINATION RATE INTO CATEGORIES: LOW (<20%), MEDIUM (20%-50%), HIGH (>50%) AND COUNT COUNTRIES IN EACH CATEGORY.



13. FIND COUNTRIES WHERE ACTIVE CASES ARE MORE THAN 10% OF THEIR CONFIRMED CASES.



14. WHICH COUNTRIES HAVE MORE DEATHS THAN ACTIVE CASES?

[] #14.Which countries have more deaths than active cases? data[data['Deaths'] > data['Active Cases']][['Country', 'Deaths', 'Active Cases']] ₹ Country Deaths Active Cases 6 70000 UK 127000 14 Mexico 215000 190000 19 South Africa 54000 40000 30 Israel 6300 1700 46 **Ecuador** 17000 13000 55 Panama 5700 4500

15. FIND THE TOTAL NUMBER OF CONFIRMED CASES, DEATHS, RECOVERED CASES, AND ACTIVE CASES GLOBALLY.

90

58

Australia

910



16. FIND THE COUNTRY WITH THE HIGHEST PERCENTAGE OF POPULATION STILL HAVING ACTIVE CASES (ACTIVE/CONFIRMED * 100).

[]	#16.Find the codata['Active Codata.loc[data[ase %'] = (
₹		1
	Country	India
	Active Case %	16.315789
	dtype: object	

17. LIST THE TOP 10 COUNTRIES WITH THE LOWEST VACCINATION RATES.

#17.List the top 10 countries with the lowest vaccination rates.
data.nsmallest(10, 'Vaccination Rate (%)')[['Country', 'Vaccination Rate (%)']]

Country Vaccination Rate (%)

19 South Africa
2.1

	•	` '
19	South Africa	2.1
28	Pakistan	2.1
49	Nepal	2.5
37	Japan	2.6
27	Philippines	2.9
25	Iraq	3.2
42	Malaysia	3.4
32	Bangladesh	3.5
53	Sri Lanka	4.0
17	Indonesia	4.3

18. FIND COUNTRIES WHERE RECOVERY RATE IS BELOW 85%.

19. CALCULATE THE STANDARD DEVIATION OF CONFIRMED CASES.

```
#19.Calculate the standard deviation of confirmed cases.
data['Confirmed Cases'].std()
```

4992234.341223869

20. FIND COUNTRIES WHERE CONFIRMED CASES ARE BETWEEN 500,000 AND 1,000,000.

ata	[(data['Con	firmed Cases'] >=	500_000) & (data['Confirmed Cases']	<= 1_000_0	00)][['Country', 'Con	firmed Cases']
	Country	Confirmed Cases				
24	Belgium	1000000				
25	Iraq	990000				
26	Sweden	960000				
27	Philippines	950000				
28	Pakistan	820000				
29	Portugal	830000				
30	Israel	840000				
31	Hungary	780000				
32	Bangladesh	760000				
3	Jordan	720000				
34	Serbia	700000				
35	Switzerland	690000				
36	Austria	650000				
37	Japan	630000				
38	Morocco	510000				
39	UAE	500000				

