

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 COLAB LINK:

[HTTPS://COLAB.RESEARCH.GOOGLE.COM/DRIVE/19Y1HBZMDFZHNQFOU7HVECCIQIZHYTCK?USP=DRIVE_LINK](https://colab.research.google.com/drive/19Y1HBZMDFZHNQFOU7HVECCIQIZHYTCK?USP=DRIVE_LINK)

DATASET:

[HTTPS://DOCS.GOOGLE.COM/SPREADSHEETS/D/1A0TLPHLEPZ4KHIIHRX1TWSROYBDIKOO/EDIT?USP=DRIVE_LINK&OUID=100157696272359354723&RTPOF=TRUE&SD=TRUE](https://docs.google.com/spreadsheets/d/1A0TLPHLEPZ4KHIIHRX1TWSROYBDIKOO/edit?USP=DRIVE_LINK&OUID=100157696272359354723&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.

```
[ ] #1.Find the top 5 countries with the highest number of confirmed cases.

data.nlargest(5, 'Confirmed Cases')[['Country', 'Confirmed Cases']]
```




	Country	Confirmed Cases
0	USA	32000000
1	India	19000000
2	Brazil	14500000
3	France	5600000
4	Turkey	4800000

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

```
[ ] #2.Find the top 5 countries with the highest number of deaths.

data.nlargest(5, 'Deaths')[['Country', 'Deaths']]
```



	Country	Deaths
0	USA	570000
2	Brazil	400000
14	Mexico	215000
1	India	210000
6	UK	127000

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

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

np.float64(20.183333333333334)

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']]
```

Country India

Active Cases	3100000
--------------	---------

dtype: object

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

```
#6.Find the country with the lowest active cases (excluding countries with 0 cases).
data[data['Active 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.

```
[ ] #7.Calculate the death rate (%) for each country and find the top 5 countries with the highest death rate.
data['Death Rate (%)'] = (data['Deaths'] / data['Confirmed Cases']) * 100
data.nlargest(5, 'Death Rate (%)')[['Country', 'Death Rate (%)']]
```



	Country	Death Rate (%)
14	Mexico	9.347826
46	Ecuador	5.000000
54	Bolivia	4.565217
41	Bulgaria	4.125000
48	Tunisia	3.548387

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

```
[ ] #8.Calculate the recovery rate (%) for each country and find the top 5 countries with the highest recovery rate.
data['Recovery Rate (%)'] = (data['Recovered'] / data['Confirmed Cases']) * 100
data.nlargest(5, 'Recovery Rate (%)')[['Country', 'Recovery Rate (%)']]
```



	Country	Recovery Rate (%)
30	Israel	98.809524
57	Singapore	98.360656
39	UAE	98.000000
44	Belarus	97.222222
58	Australia	96.666667

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

```
[ ] #9.Find the average number of confirmed cases among all countries.  
data['Confirmed Cases'].mean()
```

⇒ 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.

```
#12.Group countries based on vaccination rate into categories: Low (<20%), Medium (20%-50%), High (>50%) and count countries in each category.
def categorize(rate):
    if rate < 20:
        return 'Low'
    elif rate <= 50:
        return 'Medium'
    else:
        return 'High'

data['Vaccination Category'] = data['Vaccination Rate (%)'].apply(categorize)
data['Vaccination Category'].value_counts()
```

Vaccination Category	count
Low	34
Medium	23
High	3

dtype: int64

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

```
[ ] #13.Find countries where active cases are more than 10% of their confirmed cases.
data[data['Active Cases'] > 0.1 * data['Confirmed Cases']][['Country', 'Active Cases', 'Confirmed Cases']]
```

	Country	Active Cases	Confirmed Cases
1	India	3100000	19000000
15	Ukraine	260000	2100000
41	Bulgaria	43500	400000

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	UK	127000	70000
14	Mexico	215000	190000
19	South Africa	54000	40000
30	Israel	6300	1700
46	Ecuador	17000	13000
55	Panama	5700	4500
58	Australia	910	90

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

```
[ ] #15.Find the total number of confirmed cases, deaths, recovered cases, and active cases globally.  
data[['Confirmed Cases', 'Deaths', 'Recovered', 'Active Cases']].sum()
```



	0
Confirmed Cases	140613600
Deaths	2985367
Recovered	126076500
Active Cases	11528664

dtype: int64

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

```
[ ] #16.Find the country with the highest percentage of population still having active cases (Active/Confirmed * 100).
data['Active Case %'] = (data['Active Cases'] / data['Confirmed Cases']) * 100
data.loc[data['Active Case %'].idxmax()][['Country', 'Active Case %']]
```



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
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%.

```
[ ] #18.Find countries where recovery rate is below 85%.  
data[data['Recovery Rate (%)'] < 85][['Country', 'Recovery Rate (%)']]
```



	Country	Recovery Rate (%)
1	India	82.631579
14	Mexico	82.608696

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.

```
#20. Find countries where confirmed cases are between 500,000 and 1,000,000.  
data[(data['Confirmed Cases'] >= 500_000) & (data['Confirmed Cases'] <= 1_000_000)][['Country', 'Confirmed 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
33	Jordan	720000
34	Serbia	700000
35	Switzerland	690000
36	Austria	650000
37	Japan	630000
38	Morocco	510000
39	UAE	500000

A graphic featuring the text "Thank You!" in a bold, white, sans-serif font. The text is centered and has a slight 3D effect with a soft shadow. It is surrounded by a dense, overlapping cluster of autumn leaves in shades of red, orange, and yellow. The entire graphic is set against a plain white background.

Thank You!