

Cv analysis with SQL

Presented by:- J.shirisha
batch no :- MIP-DA-09
profile name :- Data analysis







carona virus Data set

Province	Country/Region	Latitude	Longitude	Date	Confirmed	Deaths	Recovered
Afghanistan	Afghanistan	33.93911	67.709953	22-01-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	23-01-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	24-01-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	25-01-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	26-01-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	27-01-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	28-01-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	29-01-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	30-01-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	31-01-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	01-02-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	02-02-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	03-02-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	04-02-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	05-02-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	06-02-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	07-02-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	08-02-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	09-02-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	10-02-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	11-02-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	12-02-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	13-02-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	14-02-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	15-02-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	16-02-2020	0	0	0
Afghanistan	Afghanistan	33.93911	67.709953	17-02-2020	0	0	0




1Q:- write a code to check null values ?

```
• SELECT
    SUM(CASE WHEN Province IS NULL THEN 1 ELSE 0 END) AS NullCount_Province,
    SUM(CASE WHEN `Country/Region` IS NULL THEN 1 ELSE 0 END) AS NullCount_Country_Region,
    SUM(CASE WHEN Latitude IS NULL THEN 1 ELSE 0 END) AS NullCount_Latitude,
    SUM(CASE WHEN Longitude IS NULL THEN 1 ELSE 0 END) AS NullCount_Longitude,
    SUM(CASE WHEN Date IS NULL THEN 1 ELSE 0 END) AS NullCount_Date,
    SUM(CASE WHEN Confirmed IS NULL THEN 1 ELSE 0 END) AS NullCount_Confirmed,
    SUM(CASE WHEN Deaths IS NULL THEN 1 ELSE 0 END) AS NullCount_Deaths,
    SUM(CASE WHEN Recovered IS NULL THEN 1 ELSE 0 END) AS NullCount_Recovered
FROM
    covid.corona;
```

ult Grid   Filter Rows: <input type="text"/> Export:  Wrap Cell Content: 							
NullCount_Province	NullCount_Country_Region	NullCount_Latitude	NullCount_Longitude	NullCount_Date	NullCount_Confirmed	NullCount_Deaths	NullC
0	0	0	0	0	0	0	0




2Q:- check total number of rows ?

```
1 • SELECT
2     COUNT(*) AS TotalRows from covid.corona
```

Result Grid 		 Filter Rows: <input type="text"/>	Export: 	Wrap Cell Content:
	TotalRows			
▶	78386			

3Q:- check what is start, date & end time ?





```
1 • SELECT
2     MIN(Date) AS StartDate,
3     MAX(Date) AS EndDate
4 FROM
5     covid.corona;
```

Result Grid |  Filter Rows: | Export:  | Wrap Cell Content: 

	StartDate	EndDate
▶	01-01-2021	31-12-2020

4Q:- Number of month present in dataset ?

```
1 • SELECT
2     COUNT(DISTINCT YEAR(Date) * 100 + MONTH(Date)) AS NumberOfMonths
3 FROM
4     covid.corona;
5
```

Result Grid				Filter Rows:	Export:	Wrap Cell Content:
	NumberOfMonths					
▶	0					



5Q:- find monthly average for confirmed death's, recovered ?

```
1 • SELECT
2     YEAR(Date) AS Year,
3     MONTH(Date) AS Month,
4     AVG(Confirmed) AS Avg_Confirmed,
5     AVG(Deaths) AS Avg_Deaths,
6     AVG(Recovered) AS Avg_Recovered
7 FROM
8     covid.corona
9 GROUP BY
10    Year, Month
11 ORDER BY
12    Year, Month;
```

Result Grid					
		Filter Rows:	Export:		Wrap Cell Content:
Year	Month	Avg_Confirmed	Avg_Deaths	Avg_Recovered	
NULL	NULL	2156.8283	46.5376	1442.7264	

6Q:- find minimum values for confirmed death's, recovered per year ?

```
1 • SELECT
2     YEAR(Date) AS Year,
3     MIN(Confirmed) AS Min_Confirmed,
4     MIN(Deaths) AS Min_Deaths,
5     MIN(Recovered) AS Min_Recovered
6 FROM
7     covid.corona
8 GROUP BY
9     Year;
```

Result Grid				
Filter Rows: <input type="text"/>				
Export:  Wrap Cell Content: 				
	Year	Min_Confirmed	Min_Deaths	Min_Recovered
▶	NULL	0	0	0

7Q:- find maximum values of confirmed death's, recovered per year ?




```
1 • SELECT
2     YEAR(Date) AS Year,
3     MAX(Confirmed) AS Max_Confirmed,
4     MAX(Deaths) AS Max_Deaths,
5     MAX(Recovered) AS Max_Recovered
6 FROM
7     covid.corona
8 GROUP BY
9     Year;
```

10

Result Grid			Filter Rows: <input type="text"/>	Export: 	Wrap Cell Content: 
	Year	Max_Confirmed	Max_Deaths	Max_Recovered	
▶	NULL	823225	7374	1123456	

8Q:- The total number of cases confirmed death's, recovered each month?





```
1 • SELECT
2     YEAR(Date) AS Year,
3     MONTH(Date) AS Month,
4     SUM(Confirmed) AS Total_Confirmed,
5     SUM(Deaths) AS Total_Deaths,
6     SUM(Recovered) AS Total_Recovered
7 FROM
8     covid.corona
9 GROUP BY
10    Year, Month
11 ORDER BY
12    Year, Month;
```

Result Grid |  Filter Rows: | Export:  | Wrap Cell Content: 

	Year	Month	Total_Confirmed	Total_Deaths	Total_Recovered
▶	NULL	NULL	169065144	3647894	113089548

9Q:-check how carona virus spread out with respect to confirmed case ?





```
1 • SELECT
2     SUM(Confirmed) AS Total_Confirmed,
3     AVG(Confirmed) AS Avg_Confirmed,
4     VARIANCE(Confirmed) AS Variance_Confirmed,
5     STDDEV(Confirmed) AS Stdev_Confirmed
6 FROM
7     covid.corona;
```

Result Grid   Filter Rows: <input type="text"/> Export:  Wrap Cell Content: 				
	Total_Confirmed	Avg_Confirmed	Variance_Confirmed	Stdev_Confirmed
▶	169065144	2156.8283	157288925.07796532	12541.488152446875

10Q:- check how carona virus spread out with respect to death cases per month ?

```
1 • SELECT
2     YEAR(Date) AS Year,
3     MONTH(Date) AS Month,
4     SUM(Deaths) AS Total_Deaths,
5     AVG(Deaths) AS Avg_Deaths,
6     VARIANCE(Deaths) AS Variance_Deaths,
7     STDDEV(Deaths) AS Stdev_Deaths
8 FROM
9     covid.corona
10 GROUP BY
11     Year, Month
12 ORDER BY
13     Year, Month;
```

1/1

Result Grid			Filter Rows: <input type="text"/>	Export: 	Wrap Cell Content: 
Year	Month	Total_Deaths	Avg_Deaths	Variance_Deaths	Stdev_Deaths
NULL	NULL	3647894	46.5376	45892.01885355753	214.22422564583476

11Q:-check how carona virus spread out with respect to recovered case ?

1 • SELECT

2 YEAR(Date) AS Year,

3 MONTH(Date) AS Month,

4 SUM(Deaths) AS Total_Deaths,

5 AVG(Deaths) AS Avg_Deaths,

6 VARIANCE(Deaths) AS Variance_Deaths,

7 STDDEV(Deaths) AS Stdev_Deaths

8 FROM

9 covid.corona

10 GROUP BY


11 Year, Month


12 ORDER BY



13 Year, Month;

14

Result Grid

 Filter Rows:






 Export:

 Wrap Cell Content: 

	Year	Month	Total_Deaths	Avg_Deaths	Variance_Deaths	Stdev_Deaths
▶	NULL	NULL	3647894	46.5376	45892.01885355753	214.22422564583476

12Q:- check how carona virus spread out with respect to death cases per month?

```
1 • SELECT
2     `Country/Region`,
3     SUM(Confirmed) AS Total_Confirmed
4 FROM
5     covid.corona
6 GROUP BY
7     `Country/Region`
8 ORDER BY
9     Total_Confirmed DESC
10 LIMIT 1;
11
```

Result Grid			Filter Rows: <input type="text"/>	Export: 	Wrap Cell Content: 	Fetch rows: 
	Country/Region	Total_Confirmed				
▶	US	33461982				

13Q:- find country having lowest number of the confirmed cases?

```
1 • SELECT
2     `Country/Region`,
3     SUM(Deaths) AS Total_Deaths
4 FROM
5     covid.corona
6 GROUP BY
7     `Country/Region`
8 ORDER BY
9     Total_Deaths ASC
10 LIMIT 1;
```

Result Grid



Filter Rows:

Export:



Wrap Cell Content:



Fetch rows:



	Country/Region	Total_Deaths
▶	Dominica	0

14Q:- find top countries having highest recovered cases

```
1 • SELECT
2     `Country/Region`,
3     SUM(Recovered) AS Total_Recovered
4 FROM
5     covid.corona
6 GROUP BY
7     `Country/Region`
8 ORDER BY
9     Total_Recovered DESC
10 LIMIT 5;
11
```

Result Grid			Filter Rows:	Export:	Wrap Cell Content:	Fetch rows:
	Country/Region	Total_Recovered				
▶	India	28089649				
	Brazil	15400169				
	US	6303715				
	Turkey	5202251				
	Russia	4745756				

Thank you