

COVID-19

**REPORT** 

LET'S TALK ABOUT IT

**CLICK HERE** 



### **Checking NULL values**

```
SELECT
FROM
    covid
WHERE
    Province IS NULL
        OR Country_Region IS NULL
        OR Latitude IS NULL
        OR Longitude IS NULL
        OR Date IS NULL
        OR Confirmed IS NULL
        OR Deaths IS NULL
        OR Recovered IS NULL;
```

	Province	Country_Region	Latitude	Longitude	Date	Confirmed	Deaths	Recovered	]
--	----------	----------------	----------	-----------	------	-----------	--------	-----------	---

### If NULL values are present, update them with zeros for all columns

```
UPDATE covid
SET
    Confirmed = COALESCE(Confirmed, 0),
    Deaths = COALESCE(Deaths, 0),
    Recovered = COALESCE(Recovered, 0)
WHERE
    Confirmed IS NULL AND DEATHS IS NULL
        AND Recovered IS NULL;
```

### What is Start\_date and End\_date

```
SELECT
    MIN(Date) AS start_date, MAX(Date) AS end_date
FROM
    covid;
```

### **OUTPUT**

	start_date	end_date
<b>&gt;</b>	2020-01-22	2021-06-13

### **Total Number Of Rows**

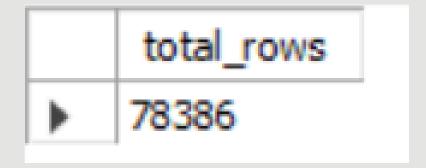
```
SELECT

COUNT(*) AS total_rows

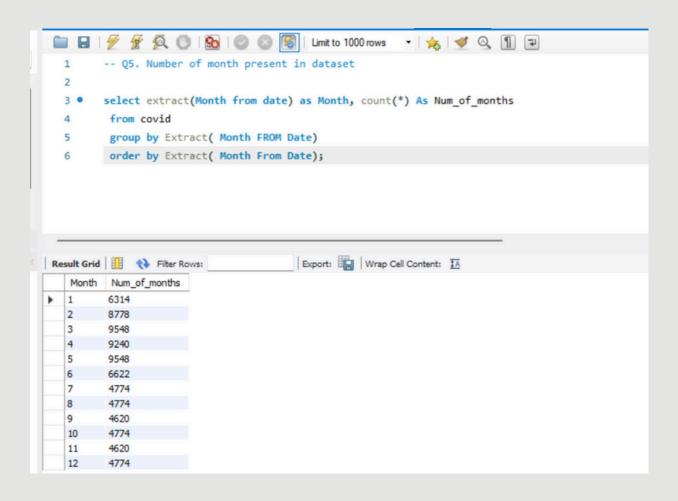
FROM

covid
```

### Output:-



### Number of month present in dataset



### Find monthly average for confirmed, deaths, recovered

```
select
extract(Month from date) as Month,
extract(Year from date) as Year,
Round(avg(confirmed),3) AS AVG_confirmed,
Round(avg(Deaths),3) AS AVG_deaths,
Round(avg(recovered),3) AS AVG_Recovered
From covid
group by
```

extract(YEAR FROM date) ,
extract(month from date)
ORDER BY
extract(YEAR FROM date) ,

extract(month from date)

				1
Month	Year	AVG_confirmed	AVG_deaths	AVG_Recovered
1	2020	4.145	0.123	0.093
2	2020	15.296	0.594	7.032
3	2020	161.130	8.661	27.874
4	2020	505.800	41.522	171.642
5	2020	574.850	30.281	318.296
6	2020	859.228	29.818	548.792
7	2020	1432.361	35.110	983.058
8	2020	1611.843	37.537	1299.295
9	2020	1784.587	34.777	1438.907
10	2020	2412.200	36.758	1420.643
11	2020	3592.194	56.763	1985.345
12	2020	4050.440	71.218	2497.885
1	2021	3911.229	84.184	1919.637
2	2021	2433.364	69.165	1558.392
3	2021	2916.797	59.200	1652.286
4	2021	4699.355	78.439	3074.785
_				

### Find most frequent value for confirmed, deaths, recovered each month

```
WITH AggregatedData As (

Select
    extract(Year from date) As Year,
    extract(Month from date) as Month,

    GROUP_CONCAT(Confirmed ORDER BY Confirmed DESC SEPARATOR ',') AS AggConfirmed,
    GROUP_CONCAT(Deaths ORDER BY Deaths DESC SEPARATOR ',') AS AggDeaths,
    GROUP_CONCAT(Recovered ORDER BY Recovered DESC SEPARATOR ',') AS AggRecovered

From Covid

group by Year, Month
)

Select Year, Month,

SUBSTRING_INDEX(AggConfirmed,',',1) AS MostFrequentConfirmed,

SUBSTRING_INDEX(AggDeaths,',',1) AS MostFrequentDeaths,

SUBSTRING_INDEX(AggRecovered,',',1) AS MostFrequentRecovered

From AggregatedData
Order by Year, Month;
```

#### **OUTPUT**

	Year	Month	MostFrequentConfirmed	MostFrequentDeaths	MostFrequentRecovered
•	2020	1	2131	49	51
	2020	2	14840	242	3418
	2020	3	26314	1085	4289
	2020	4	50740	2607	33227
	2020	5	34907	2309	51717
	2020	6	54771	2003	94305
	2020	7	75866	1595	140050
	2020	8	85687	1505	95881
	2020	9	97894	1703	101468
	2020	10	99264	3351	388340
	2020	11	207933	2259	139292
	2020	12	823225	3752	1123456
	2021	1	300462	4475	87090
	2021	2	134975	3907	98389
	2021	3	100158	3869	102138
	2021	4	401993	4249	299988
	2021	5	414188	4529	422436
	2021	6	134154	7374	231456

### Find minimum values for confirmed, deaths, recovered per year

```
select
extract(YEAR FROM DATE) AS YEAR,
MIN(Confirmed) AS Min_confirmed,
MIN(deaths) AS Min_deaths,
MIN(recovered) AS Min_recovered
from
covid
group by
extract(YEAR FROM DATE)
order by
year;
```

Re	sult Grid	III 🙌 Filter	Rows:	Exp	ort: ≝🔚
	YEAR	Min_confirmed	Min_deaths	Min_recovered	
•	2020	0	0	0	
	2021	0	0	0	

### Find maximum values of confirmed, deaths, recovered per year

```
select
extract(Year from date) as Year,
MAX(confirmed) as Max_confirmed,
MAX(deaths) as Max_deaths,
MAX(recovered) as Max_recovered
from
covid
group by
extract(Year from Date)
order by
year;
```

Y	ear M	lax_confirmed	Max_deaths	Max_recovered
<b>)</b> 20	20 82	23225	3752	1123456
20	21 41	14188	7374	422436

### Total Number of case of confirmed, deaths, recovered each month

```
select
extract(year from date) as Year,
extract(Month from date) as Month,
sum(confirmed) as Total_confirmed,
sum(deaths) as Total_deaths,
sum(recovered) as Total_recovered
from
covid
group by
extract(Month from date),
extract(year from date)
order by
Year,
Month;

Year Month Total confirmed Total deaths Total recovered
```

Year Month Total_confirmed Total_deaths Total_recov	ered
2020 1 6384 190 143	
2020 2 68312 2651 31405	
2020 3 769236 41346 133070	
2020 4 2336798 191833 792987	
2020 5 2744333 144561 1519547	
2020 6 3969634 137757 2535417	
2020 7 6838092 167613 4693120	
2020 8 7694938 179200 6202833	
2020 9 8244794 160671 6647749	
2020 10 11515841 175484 6782150	
2020 11 16595938 262247 9172292	
2020 12 19336799 339996 11924903	
2021 1 18672205 401893 9164347	
2021 2 10492664 298239 6719785	
2021 3 13924790 282620 7888013	
2021 4 21711021 362387 14205507	
2021 5 19121083 366549 19131842	
2021 6 5022282 132657 5544438	

## Check how corona virus spread out with respect to confirmed case (Eg.: total confirmed cases, their average, variance & STDEV)

```
select
sum(Confirmed) as total_Confirmed,
Round(avg(Confirmed),3) as Avg_Confirmed,
variance(Confirmed) as Var_Confirmed,
stddev(Confirmed) as Std_Confirmed
from
covid
```

	total_Confirmed	Avg_Confirmed	Var_Confirmed	Std_Confirmed
•	169065144	2156.828	157288925.07796532	12541.488152446875

Check how corona virus spread out with respect to death case per month(Eg.: total confirmed cases, their average, variance & STDEV)

```
extract(month from date) as Month,
sum(Confirmed) as Total_confirmed,
Round(avg(Confirmed),2) as Avg_Confirmed,
Round(variance(Confirmed),2) as Var_Confirmed,
Round(stddev(Confirmed),2) as Std_Confirmed
From
covid
group by
extract(month from date)
order by
Month;
```

	Month	Total_confirmed	Avg_Confirmed	Var_Confirmed	Std_Confirmed
١	1	18678589	2958.28	241973521.82	15555.5
	2	10560976	1203.12	40597067.13	6371.58
	3	14694026	1538.96	44274264.91	6653.89
	4	24047819	2602.58	258409110.21	16075.11
	5	21865416	2290.05	320297513.2	17896.86
	6	8991916	1357.89	43725047.85	6612.49
	7	6838092	1432.36	46914022.89	6849.38
	8	7694938	1611.84	54408583.16	7376.22
	9	8244794	1784.59	69314698.6	8325.54
	10	11515841	2412.20	68988159.05	8305.91
	11	16595938	3592.19	195815877.81	13993.42
	12	19336799	4050.44	459885446.66	21444.94

### Check how corona virus spread out with respect to recovered case

```
select
sum(Confirmed) as Total_Confirmed,
Round(avg(Confirmed),2) as Avg_Confirmed,
variance(Confirmed) as Var_Confirmed,
stddev(Confirmed) as std_confirmed
from
covid
```

	Total_Confirmed	Avg_Confirmed	Var_Confirmed	std_confirmed
<b>&gt;</b>	169065144	2156.83	157288925.07796532	12541.488152446875

### Find Country having highest number of the Confirmed case

```
select
Country_Region,
sum(Confirmed) as Total_confirmed
FROM
Covid
Group By
Country_Region
Order by
Total_confirmed DESC
LIMIT 5;
```

	Country_Region	Total_confirmed
•	US	33461982
	India	29460523
	Brazil	17412766
	France	6106009
	Turkey	5330447

### Find Country having lowest number of the death case

```
Country_Region,
Min(Deaths) as Min_Death
From
covid
group by Country_Region
order by
MIN_Death DESC
limit 5;
```

		Country_Region	Min_Death
	•	Afghanistan	0
L		Algeria	0
		Argentina	0
		Australia	0
L		Austria	0
Г			

### Find top 5 countries having highest recovered case

```
select
Country_Region,
sum(Recovered) as Highest_Recovered
FROM
covid
group by
Country_region
order by
Highest_Recovered DESC
limit 5;
```

### **OUTPUT**;-

	Country_Region	Highest_Recovered
•	India	28089649
	Brazil	15400169
	US	6303715
	Turkey	5202251
	Russia	4745756

# THANK