



Project-1

COVID-19 Analysis

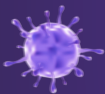
BY:Kariem Abdelmoniem

TABLE OF CONTENTS



01

INTRODUCTION

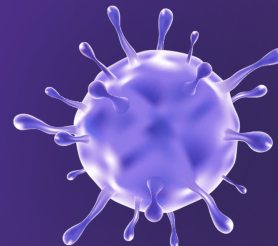


02

DATASET
& DESCRIPTION

03

Q&A





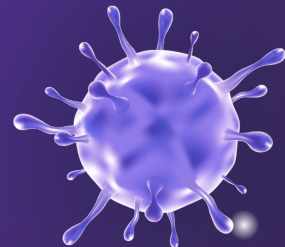
01

INTRODUCTION



Introduction

This project focuses on the analysis of COVID-19 epidemic has had significant impact on public health and has created a global need for data-driven insights to understand the spread of the virus.





02

DATASET & DESCRIPTION



dbo.Corona

Columns

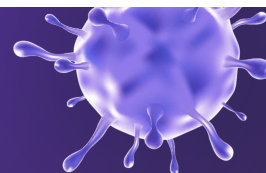
- Province (nvarchar(50), null)
- Country_Region (nvarchar(50), null)
- Latitude (varchar(50), null)
- Longitude (varchar(50), null)
- Date (date, null)
- Confirmed (int, null)
- Deaths (int, null)
- Recovered (int, null)

Keys

Dataset:

Description of each column in dataset:

- Province:** Geographic subdivision within a country/region.
- Country/Region:** Geographic entity where data is recorded.
- Latitude:** North-south position on Earth's surface.
- Longitude:** East-west position on Earth's surface.
- Date:** Recorded date of CORONA VIRUS data.
- Confirmed:** Number of diagnosed CORONA VIRUS cases.
- Deaths:** Number of CORONA VIRUS related deaths.
- Recovered:** Number of recovered CORONA VIRUS cases.



03


Q&A

-- Q1. Write a code to check NULL values

```
select  
  count(case when [Latitude] is null then 1 end) as nulls_in_Latitude,  
  count(case when [Longitude] is null then 1 end) as nulls_in_Longitude  
from corona
```

Results		Messages
	nulls_in_Latitude	nulls_in_Longitude
1	0	0

--Q2. If NULL values are present, update them with zeros for all columns.

```
 update Corona  
set  
[Latitude]=0 ,  
[Longitude]=0  
where  
[Latitude] is null or  
[Longitude] is null |
```

-- Q3. check total number of rows

```
select count(*) count_of_rows from Corona
```

	count_of_rows
1	78386

-- Q4. Check what is start_date and end_date

```
select min([date]) as start_date, max([date]) as end_date  
from Corona
```

	start_date	end_date
1	2020-01-22	2021-06-13

-- Q5. Number of month present in dataset

```
select year([date]) as year, count(distinct month([date])) as count_of_months  
from Corona  
group by year([date])
```

	year	count_of_months
1	2020	12
2	2021	6

-- Q6. Find monthly average for confirmed, deaths, recovered

```
-- select year([date]) year,  
month([date]) month,  
avg(confirmed) average_confirmed,  
avg(deaths) average_deaths,  
avg(recovered) average_recovered  
from corona  
group by year([date]), month([date])  
order by year([date]), month([date])
```

	year	month	average_confirmed	average_deaths	average_recovered
1	2020	1	4	0	0
2	2020	2	15	0	7
3	2020	3	161	8	27
4	2020	4	505	41	171
5	2020	5	574	30	318
6	2020	6	859	29	548
7	2020	7	1432	35	983
8	2020	8	1611	37	1299
9	2020	9	1784	34	1438
10	2020	10	2412	36	1420
11	2020	11	3592	56	1985
12	2020	12	4050	71	2497
13	2021	1	3911	84	1919
14	2021	2	2433	69	1558
15	2021	3	2916	59	1652
16	2021	4	4699	78	3074
17	2021	5	4005	76	4007
18	2021	6	2508	66	2769

-- Q8. Find minimum values for confirmed, deaths, recovered per year

```
select  year([date]) year,  
min(confirmed)min_confirmed,  
min(deaths)min_deaths,  
min(recovered) min_recovered  
from Corona  
  
group by year([date])  
order by year
```

	year	min_confirmed	min_deaths	min_recovered
1	2020	0	0	0
2	2021	0	0	0

-- Q9. Find maximum values of confirmed, deaths, recovered per year

```
select year([date]) year,  
max(confirmed) max_confirmed,  
max(deaths) max_deaths,  
max(recovered) max_recovered  
from Corona  
group by year([date])  
order by year
```

	year	max_confirmed	max_deaths	max_recovered
1	2020	823225	3752	1123456
2	2021	414188	7374	422436

-- Q10. The total number of case of confirmed, deaths, recovered each month

```
select year(date) year,  
month([date]) month, sum(confirmed) count_confirmed,  
sum(deaths) count_deaths ,  
sum(recovered) count_recovered  
from Corona  
  
group by year(date), month([date])  
order by year(date), month([date])
```

	year	month	count_confirmed	count_deaths	count_recovered
1	2020	1	6384	190	143
2	2020	2	68312	2651	31405
3	2020	3	769236	41346	133070
4	2020	4	2336798	191833	792987
5	2020	5	2744333	144561	1519547
6	2020	6	3969634	137757	2535417
7	2020	7	6838092	167613	4693120
8	2020	8	7694938	179200	6202833
9	2020	9	8244794	160671	6647749
10	2020	10	11515841	175484	6782150
11	2020	11	16595938	262247	9172292
12	2020	12	19336799	339996	11924903
13	2021	1	18672205	401893	9164347
14	2021	2	10492664	298239	6719785
15	2021	3	13924790	282620	7888013
16	2021	4	21711021	362387	14205507
17	2021	5	19121083	366549	19131842
18	2021	6	5022282	132657	5544438

–Q11. Check how coronavirus spread out with respect to confirmed case

```
select year([date]) year,  
month([date]) month,  
sum(confirmed) sum_confirmed,  
avg(confirmed) average_confirmed,  
STDEV(confirmed) stdev_confirmed  
from corona  
  
group by year([date]), month([date])  
order by year([date]), month([date])
```

	year	month	sum_confirmed	average_confirmed	stdev_confirmed
1	2020	1	6384	4	69.5417256311408
2	2020	2	68312	15	280.191050909673
3	2020	3	769236	161	1013.22713376042
4	2020	4	2336798	505	2648.31670377589
5	2020	5	2744333	574	2462.6917653358
6	2020	6	3969634	859	3712.43784177769
7	2020	7	6838092	1432	6850.09868045691
8	2020	8	7694938	1611	7376.99006391748
9	2020	9	8244794	1784	8326.44612252557
10	2020	10	11515841	2412	8306.78113850618
11	2020	11	16595938	3592	13994.9373480842
12	2020	12	19336799	4050	21447.1862515374
13	2021	1	18672205	3911	17786.8199440871
14	2021	2	10492664	2433	8922.24092007789
15	2021	3	13924790	2916	9151.10960032634
16	2021	4	21711021	4699	22385.7471236031
17	2021	5	19121083	4005	25075.4724471398
18	2021	6	5022282	2508	10535.0944628437

-- Q12. Check how corona virus spread out with respect to death case per month

```
select year([date]) year,  
month([date]) month,  
sum(Deaths) sum_Deaths,  
avg(Deaths) average_Deaths,  
STDEV(Deaths) stdev_Deaths  
from corona  
group by year([date]), month([date])  
order by year([date]), month([date])
```

	year	month	sum_Deaths	average_Deaths	stdev_Deaths
1	2020	1	190	0	2.06120740960683
2	2020	2	2651	0	8.26662872455985
3	2020	3	41346	8	62.4628651839385
4	2020	4	191833	41	201.278506486273
5	2020	5	144561	30	143.837566042174
6	2020	6	137757	29	130.127287243856
7	2020	7	167613	35	145.41177413497
8	2020	8	179200	37	152.570876726552
9	2020	9	160671	34	141.799581855918
10	2020	10	175484	36	132.60374901453
11	2020	11	262247	56	166.672752848512
12	2020	12	339996	71	255.654180153028
13	2021	1	401893	84	320.593140018966
14	2021	2	298239	69	261.715028514503
15	2021	3	282620	59	233.232425290674
16	2021	4	362387	78	307.623071355242
17	2021	5	366549	76	363.03867091102
18	2021	6	132657	66	336.184661457491

-- Q13. Check how corona virus spread out with respect to recovered case

```
select year([date]) year,  
month([date]) month,  
sum(recovered) sum_recovered,  
avg(recovered) average_recovered,  
STDEV(recovered) stdev_recovered  
from corona  
group by year([date]), month([date])  
order by year([date]), month([date])
```

	year	month	sum_recovered	average_recovered	stdev_recovered
1	2020	1	143	0	1.62335965662889
2	2020	2	31405	7	111.577101550499
3	2020	3	133070	27	200.303754294549
4	2020	4	792987	171	877.530461883054
5	2020	5	1519547	318	1406.63459194499
6	2020	6	2535417	548	2555.69682403668
7	2020	7	4693120	983	4984.88544901792
8	2020	8	6202833	1299	6338.67796758684
9	2020	9	6647749	1438	7552.21238309451
10	2020	10	6782150	1420	8587.61609332342
11	2020	11	9172292	1985	7123.10334437809
12	2020	12	11924903	2497	18076.5917837348
13	2021	1	9164347	1919	5612.51266537584
14	2021	2	6719785	1558	4942.98269296028
15	2021	3	7888013	1652	5908.0202316652
16	2021	4	14205507	3074	14982.2618898092
17	2021	5	19131842	4007	27483.3358595653
18	2021	6	5544438	2769	15269.2785148651

-- Q14. Find Country having highest number of the Confirmed case

```
select country_region, confirmed  
from Corona  
where Confirmed=(select max(Confirmed) from corona)
```

	country_region	confirmed
1	Turkey	823225

-- Q15. Find Country having lowest number of the death case

```
select top 1 Country_Region,min (Deaths)
from Corona
group by Country_Region
order by min(Deaths)|
```

	Country_Region	death
1	Gambia	0

-- Q16. Find top 5 countries having highest recovered case

```
SELECT TOP 5 Country_Region, SUM(recovered) as Total_Reccvered_Cases  
FROM Corona  
GROUP BY Country_Region  
ORDER BY SUM(Recovered) DESC
```

	Country_Region	Total_Reccvered_Cases
1	India	28089649
2	Brazil	15400169
3	US	6303715
4	Turkey	5202251
5	Russia	4745756