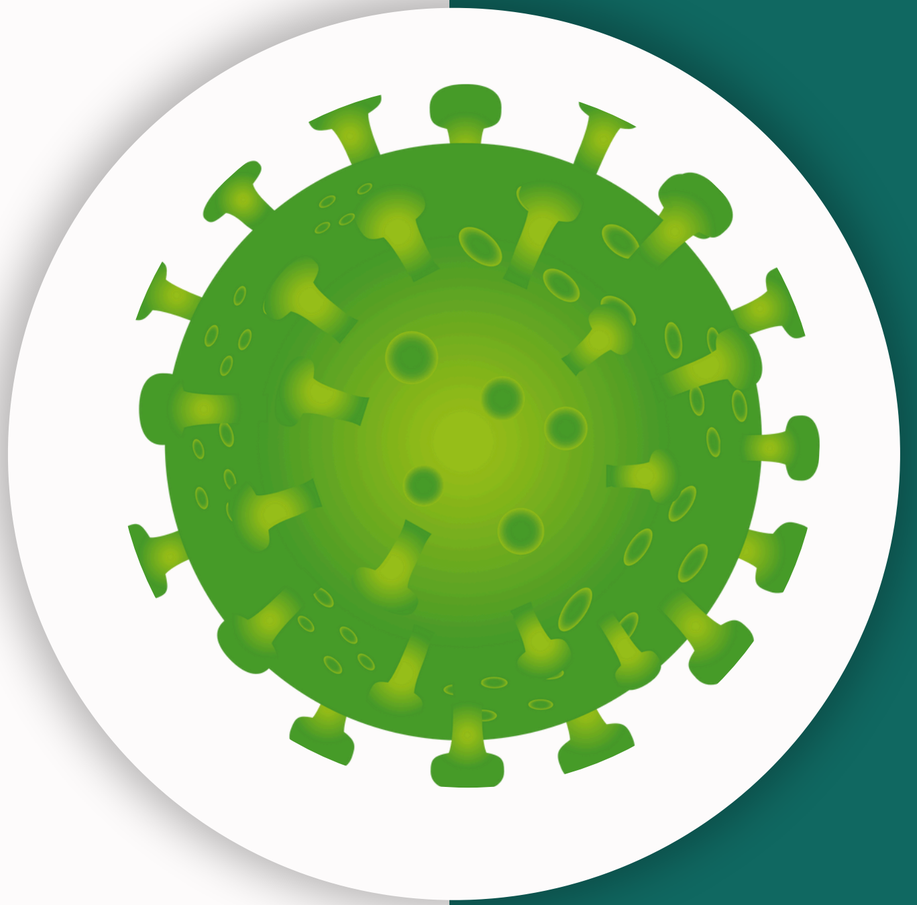
Corona Virus Analysis



About Dataset

- **Province:** Name of the province or region.
- **Country:** Name of the country.
- **Latitude:** The geographic coordinate that specifies the north–south position of a point on the Earth's surface.
- **Longitude:** The geographic coordinate that specifies the east-west position of a point on the Earth's surface.
- **C_Date:** Date of the data entry or record.
- **Confirmed:** Number of confirmed COVID-19 cases.
- **Deaths:** Number of deaths due to COVID-19.
- **Recovered:** Number of recovered COVID-19 cases.

Summary

Using SQL queries, I analyzed COVID-19 data, including total confirmed cases, deaths, and recoveries. Additionally, I calculated averages, standard deviations, and variances of death cases. I determined the number of unique months in the dataset. Each query provided insights into the pandemic's progression and statistical analysis.

```sql
--1.What is the total number of confirmed cases across all provinces and countries?
SELECT
    SUM([Confirmed]) AS TOTAL_CONFIRMED_CASES
FROM CovidData
```

100 %

Results | Messages

	TOTAL_CONFIRMED_CASES
1	169065144

```sql
--2.What is the total number of deaths across all provinces and countries?
SELECT
    SUM([Deaths]) AS TOTAL_DEATHS
FROM CovidData
```

03

100 %

Results | Messages

	TOTAL_DEATHS
1	3647894

```sql
--3.What is the total number of recovered cases across all provinces and countries?
SELECT
    SUM([Recovered]) AS TOTAL_RECOVER_CASES
FROM CovidData
```

03

100 %

Results | Messages

	TOTAL_RECOVER_CASES
1	113089548

```sql
--4.What is the average number of confirmed cases per province?
SELECT [Province],
        AVG([Confirmed]) AS AverageConfirmedCases
FROM CovidData
GROUP BY [Province]
```

100 %

Results | Messages

	Province	AverageConfirmedCases
1	United Arab Emirates	1174
2	Ghana	185
3	Belarus	796
4	Peru	3936
5	Western Australia	2
6	Mozambique	140
7	Sri Lanka	439
8	Iraq	2464
9	Mongolia	149
10	Barbados	7
11	Uganda	121
12	Sint Maarten	4
13	Turkey	10472
14	Panama	761
15	Yemen	13
16	Mauritania	39
17	Jamaica	96
18	Taiwan*	25
19	Portugal	1684

```sql
--5.What is the average number of deaths per province?
SELECT Province,
        AVG([Deaths])AS AverageDeathsCases
FROM CovidData
GROUP BY Province
```

100 %

Results | Messages

	Province	AverageDeathsCases
1	United Arab Emirates	3
2	Ghana	1
3	Belarus	5
4	Peru	370
5	Western Australia	0
6	Mozambique	1
7	Sri Lanka	4
8	Iraq	32
9	Mongolia	0
10	Barbados	0
11	Uganda	0
12	Sint Maarten	0
13	Turkey	95
14	Panama	12

```sql
--6.What is the average number of recovered cases per province?
 SELECT Province,
        AVG([Recovered]) AS AverageRecoverCase
FROM CovidData
GROUP BY Province
```

100 %

Results | Messages

	Province	AverageRecoverCase
1	United Arab Emirates	1134
2	Ghana	181
3	Belarus	783
4	Peru	3860
5	Western Australia	2
6	Mozambique	137
7	Sri Lanka	370
8	Iraq	2303
9	Mongolia	111
10	Barbados	7
11	Uganda	94
12	Sint Maarten	4
13	Turkey	10220
14	Panama	730
15	Yemen	7
16	Mauritania	37
17	Jamaica	54
18	Taiwan*	2
19	Portugal	1601

```sql
--7.Which province has the highest number of confirmed cases?
SELECT TOP 1 [Province],
            SUM(Confirmed) AS TotalConfirmedCases
FROM CovidData
GROUP BY [Province]
ORDER BY TotalConfirmedCases DESC;

```

100 %

Results | Messages

	Province	TotalConfirmedCases
1	US	33461982

```sql
--8.Which province has the highest number of deaths?
SELECT TOP 1
        [Province],
        SUM([Deaths]) AS TotalDeathCase
FROM CovidData
GROUP BY Province
ORDER BY TotalDeathCase DESC;
```

03

100 %

Results | Messages

	Province	TotalDeathCase
1	US	599769

```sql
--9.Which province has the highest recovery rate (recovered cases/confirmed cases)?
SELECT TOP 1
    [Province],
    SUM(Recovered) AS TotalRecoveredCases,
    SUM(Confirmed) AS TotalConfirmedCases,
    ROUND(SUM(Recovered) * 1.00 / SUM(Confirmed), 2) AS RecoveryRate
FROM
    CovidData
GROUP BY
    [Province]
ORDER BY
    RecoveryRate DESC;
```

100 %

Results | Messages

	Province	TotalRecoveredCases	TotalConfirmedCases	RecoveryRate
1	Lithuania	296084	277746	1.0700000000000

```sql
--10.Which country has the highest number of confirmed cases?
SELECT TOP 1
    [Country],
    SUM([Confirmed]) AS CONFIRM_CASE
FROM
    CovidData
GROUP BY
    [Country]
ORDER BY
    CONFIRM_CASE DESC;
```

100 %

Results | Messages

	Country	CONFIRM_CASE
1	US	33461982

```sql
--11.Which country has the highest number of deaths?
SELECT TOP 1
    Country,
    SUM([Deaths]) AS TOTALDEATH
FROM
    CovidData
GROUP BY
    Country
ORDER BY
    TOTALDEATH DESC;
```

100 %

Results | Messages

	Country	TOTALDEATH
1	US	599769

```sql
--12.What is the timeline of confirmed cases for a specific province/country?
SELECT
    CONVERT(varchar, [C_Date], 105) AS Date,
    SUM([Confirmed]) AS ConfirmedCases
FROM
    CovidData
WHERE
    [Province] = 'INDIA'
GROUP BY
    CONVERT(varchar, [C_Date], 105)
ORDER BY
    CONVERT(varchar, [C_Date], 105);
```

100 %

Results | Messages

	Date	ConfirmedCases
1	2020-01-22	0
2	2020-01-23	0
3	2020-01-24	0
4	2020-01-25	0
5	2020-01-26	0
6	2020-01-27	0
7	2020-01-28	0
8	2020-01-29	0
9	2020-01-30	1
10	2020-01-31	0
11	2020-02-01	0

```sql
--13.What is the mortality rate (deaths/confirmed cases) for each province/country?
SELECT
    [Country],
    [Province],
    SUM([Deaths]) AS TotalDeaths,
    SUM([Confirmed]) AS TotalConfirmedCases,
    SUM([Deaths]) * 100.0 / SUM([Confirmed]) AS MortalityRate
FROM
    CovidData
GROUP BY
    [Country], [Province];
```

100 %

Results | Messages

	Country	Province	TotalDeaths	TotalConfirmedCases	MortalityRate
1	China	Fujian	1	636	0.157232704402
2	Zimbabwe	Zimbabwe	1632	39965	4.083573126485
3	Mexico	Mexico	230150	2454176	9.377893028046
4	Egypt	Egypt	15623	273182	5.718898024027
5	Ukraine	Ukraine	53795	2283746	2.355559681330
6	Canada	New Brunswick	45	2302	1.954821894005
7	Dominican Republic	Dominican Republic	3707	308650	1.201036773043
8	Syria	Syria	1810	24814	7.294269364068
9	China	Hubei	4495	67715	6.638115631691
10	Mauritius	Mauritius	18	1675	1.074626865671
11	United Kingdom	Bermuda	33	2497	1.321585903083
12	Central African Republic	Central African Republic	98	7101	1.380087311646
13	West Bank and Gaza	West Bank and Gaza	3533	311534	1.134065623655

```sql
--14.How many new cases were confirmed each day globally?
SELECT
    [C_Date] AS Date,
    SUM([Confirmed]) AS ConfirmedCases,
    SUM([Confirmed]) - LAG(SUM([Confirmed]), 1, 0) OVER (ORDER BY [C_Date]) AS NewCases
FROM
    CovidData
GROUP BY
    [C_Date]
ORDER BY
    [C_Date];
```

100 %

Results | Messages

	Date	ConfirmedCases	NewCases
1	2020-01-22	9	9
2	2020-01-23	35	26
3	2020-01-24	153	118
4	2020-01-25	293	140
5	2020-01-26	410	117
6	2020-01-27	488	78
7	2020-01-28	2283	1795
8	2020-01-29	142	-2141
9	2020-01-30	1479	1337
10	2020-01-31	1092	-387
11	2020-02-01	1529	437
12	2020-02-02	4174	2645
13	2020-02-03	2501	-1673

```sql
--15.Find Monthly Average For Confirmed,Deaths And Recovered.
SELECT
    DATEPART(YEAR, [C_Date]) AS Year,
    DATEPART(MONTH, [C_Date]) AS Month,
    AVG([Confirmed]) AS AvgConfirmed,
    AVG([Deaths]) AS AvgDeaths,
    AVG([Recovered]) AS AvgRecovered
FROM
    CovidData
GROUP BY
    DATEPART(YEAR, [C_Date]), DATEPART(MONTH, [C_Date])
ORDER BY
    Year, Month;
```

100 %

Results | Messages

	Year	Month	AvgConfirmed	AvgDeaths	AvgRecovered
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

```sql
--16.Find Most Frequent Value for Confirmed,Death,Recover.
SELECT
    SUBSTRING([C_Date], 1, 4) AS YEARS,
    SUBSTRING([C_Date], 6, 2) AS MONTHS,
    MAX(CAST([Confirmed] AS INT)) AS FREQUENT_CONFIRM,
    MAX(CAST([Deaths] AS INT)) AS FREQUENT_DEATH,
    MAX(CAST([Recovered] AS INT)) AS FREQUENT_RECOVER
FROM
    CovidData
GROUP BY
    SUBSTRING([C_Date], 1, 4), SUBSTRING([C_Date], 6, 2)
ORDER BY
    YEARS, MONTHS ASC;
```

100 %

Results | Messages

	YEARS	MONTHS	FREQUENT_CONFIRM	FREQUENT_DEATH	FREQUENT_RECOVER
1	2020	01	2131	49	51
2	2020	02	14840	242	3418
3	2020	03	26314	1085	4289
4	2020	04	50740	2607	33227
5	2020	05	34907	2309	51717
6	2020	06	54771	2003	94305
7	2020	07	75866	1595	140050
8	2020	08	85687	1505	95881
9	2020	09	97894	1703	101468
10	2020	10	99264	3351	388340
11	2020	11	207933	2259	139292

```sql
--17.Find the maximum values of confirmed,death,and recover per year
SELECT
    YEAR(CONVERT(date, [C_Date], 23)) AS Year,
    MAX(CAST([Confirmed] AS INT)) AS MaxConfirmed,
    MAX(CAST([Deaths] AS INT)) AS MaxDeaths,
    MAX(CAST([Recovered] AS INT)) AS MaxRecovered
FROM
    CovidData
GROUP BY
    YEAR(CONVERT(date, [C_Date], 23));
```

100 %

Results | Messages

	Year	MaxConfirmed	MaxDeaths	MaxRecovered
1	2021	414188	7374	422436
2	2020	823225	3752	1123456

```sql
--18.Find the minimum values of confirmed,death,and recover per year
SELECT
    YEAR(CONVERT(date, [C_Date], 23)) AS Year,
    Min(CAST([Confirmed] AS INT)) AS MaxConfirmed,
    Min(CAST([Deaths] AS INT)) AS MaxDeaths,
    Min(CAST([Recovered] AS INT)) AS MaxRecovered
FROM
    CovidData
GROUP BY
    YEAR(CONVERT(date, [C_Date], 23));
```

100 %

Results | Messages

	Year	MaxConfirmed	MaxDeaths	MaxRecovered
1	2021	0	0	0
2	2020	0	0	0

```sql
--19.Find the total number of case of confirmed,death and recover for each months
SELECT
    MONTH(CONVERT(date, [C_Date], 23)) AS Month,
    SUM(CAST([Confirmed] AS INT)) AS TotalConfirmed,
    SUM(CAST([Deaths] AS INT)) AS TotalDeaths,
    SUM(CAST([Recovered] AS INT)) AS TotalRecovered
FROM
    CovidData
GROUP BY
    MONTH(CONVERT(date, [C_Date], 23))
ORDER BY
    Month;
```

100 %

Results | Messages

	Month	TotalConfirmed	TotalDeaths	TotalRecovered
1	1	18678589	402083	9164490
2	2	10560976	300890	6751190
3	3	14694026	323966	8021083
4	4	24047819	554220	14998494
5	5	21865416	511110	20651389
6	6	8991916	270414	8079855
7	7	6838092	167613	4693120
8	8	7694938	179200	6202833
9	9	8244794	160671	6647749
10	10	11515841	175484	6782150
11	11	16595938	262247	9172292
12	12	19336799	339996	11924903

```sql
--20.Check How Corona virus spread out respect to confirned case.
SELECT
    DATEPART(YEAR, CONVERT(date, [C_Date])) AS years,
    DATEPART(MONTH, CONVERT(date, [C_Date])) AS months,
    SUM(CAST([Confirmed] AS INT)) AS TotalConfirmed,
    SUM(CAST([Deaths] AS INT)) AS TotalDeaths,
    SUM(CAST([Recovered] AS INT)) AS TotalRecovered
FROM
    CovidData
GROUP BY
    DATEPART(YEAR, CONVERT(date, [C_Date])),
    DATEPART(MONTH, CONVERT(date, [C_Date]))
ORDER BY
    years, months;
```

100 %

Results | Messages

	years	months	TotalConfirmed	TotalDeaths	TotalRecovered
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

```sql
--21.Check How Corona virus spread out respect to death case per month.
SELECT
    DATEPART(YEAR, CONVERT(date, [C_Date])) AS years,
    DATEPART(MONTH, CONVERT(date, [C_Date])) AS months,
    SUM(CAST([Deaths] AS INT)) AS TotalDeaths,
    ROUND(AVG(CAST([Deaths] AS FLOAT)), 2) AS AVG_DEATH,
    ROUND(STDEV(CAST([Deaths] AS FLOAT)), 2) AS STDDEV_DEATH,
    ROUND(VAR(CAST([Deaths] AS FLOAT)), 2) AS VARIANCE_DEATH
FROM
    CovidData
GROUP BY
    DATEPART(YEAR, CONVERT(date, [C_Date])),
    DATEPART(MONTH, CONVERT(date, [C_Date]))
ORDER BY
    years, months ASC;
```

100 %

Results | Messages

	years	months	TotalDeaths	AVG_DEATH	STDDEV_DEATH	VARIANCE_DEATH
1	2020	1	190	0.12	2.06	4.25
2	2020	2	2651	0.59	8.27	68.34
3	2020	3	41346	8.66	62.46	3901.61
4	2020	4	191833	41.52	201.28	40513.04
5	2020	5	144561	30.28	143.84	20689.25
6	2020	6	137757	29.82	130.13	16933.11
7	2020	7	167613	35.11	145.41	21144.58
8	2020	8	179200	37.54	152.57	23277.87
9	2020	9	160671	34.78	141.8	20107.12

```sql
--22.Check How Corona virus spread out respect to RECOVER case per month
SELECT
    DATEPART(YEAR, CONVERT(date, [C_Date])) AS years,
    DATEPART(MONTH, CONVERT(date, [C_Date])) AS months,
    SUM(CAST([Recovered] AS INT)) AS TotalRecover,
    ROUND(AVG(CAST([Recovered] AS FLOAT)), 2) AS AVG_RECOVER,
    ROUND(STDEV(CAST([Recovered] AS FLOAT)), 2) AS STDDEV_RECOVER,
    ROUND(VAR(CAST([Recovered] AS FLOAT)), 2) AS VARIANCE_RECOVER
FROM
    CovidData
GROUP BY
    DATEPART(YEAR, CONVERT(date, [C_Date])),
    DATEPART(MONTH, CONVERT(date, [C_Date]))
ORDER BY
    years, months ASC;
```

100 %

Results | Messages

	years	months	TotalRecover	AVG_RECOVER	STDDEV_RECOVER	VARIANCE_RECOVER
1	2020	1	143	0.09	1.62	2.64
2	2020	2	31405	7.03	111.58	12449.45
3	2020	3	133070	27.87	200.3	40121.59
4	2020	4	792987	171.64	877.53	770059.71
5	2020	5	1519547	318.3	1406.63	1978620.88
6	2020	6	2535417	548.79	2555.7	6531586.26
7	2020	7	4693120	983.06	4984.89	24849082.94
8	2020	8	6202833	1299.29	6338.68	40178838.38

```sql
--23.Find The Country Having Lowest Number Of The Death case.
SELECT [Country],
        MIN([Deaths]) AS lOW_Deaths
FROM CovidData
GROUP BY Country
ORDER BY lOW_Deaths ASC
```

100 %

Results | Messages

	Country	lOW_Deaths
1	United Arab Emirates	0
2	Ghana	0
3	Belarus	0
4	Peru	0
5	Mozambique	0
6	Sri Lanka	0
7	Iraq	0
8	Mongolia	0
9	Barbados	0
10	Uganda	0
11	Turkey	0
12	Panama	0
13	Yemen	0
14	Mauritania	0
15	Jamaica	0
16	Taiwan*	0
17	Portugal	0
18	Greece	0
19	Saint Lucia	0

```sql
--24.Find The Top 5 Country Having Highest Number Of The Recover case.
SELECT Top 5
    [Country],
    SUM([Recovered]) AS Highest_Recover
FROM CovidData
GROUP BY [Country]
ORDER BY Highest_Recover DESC
```

100 %

Results | Messages

	Country	Highest_Recover
1	India	28089649
2	Brazil	15400169
3	US	6303715
4	Turkey	5202251
5	Russia	4745756

03

Thank You