CORONA DATASET DATA ANALYSIS

**Q1. Write a code to check NULL values**

SELECT \*

FROM corona.dbo.Covid19Data

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;

Ans:No Columns has any empty/Null Values

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

UPDATE corona.dbo.Covid19Data

SET

Province=COALESCE(Province, 'Unknown'),

Country\_Region=ISNULL(Country\_Region, 'Unknown'),

Latitude=ISNULL(Latitude, 0),

Longitude=ISNULL(Longitude, 0),

Date=ISNULL(Date, 'Unknown'),

Confirmed=ISNULL(Confirmed, 0),

Deaths=ISNULL(Deaths, 0),

Recovered=ISNULL(Recovered, 0)

USE corona

--Checking the Table for updates made in Q2.

SELECT \* FROM Covid19Data

WHERE Province = 'Unknown' OR Date IS Null

**Q3. check total number of rows**

SELECT COUNT(\*) AS Total\_No\_Rows

FROM corona.dbo.Covid19Data

Ans:Total No. Of Rows are 78386

**Q4. Check what is start\_date and end\_date**

SELECT MIN(Date) AS start\_date,

MAX(Date) AS End\_Date

FROM corona.dbo.Covid19Data

Ans--Start\_date - 2020-01-22 and End\_Date is 2021-06-13

**Q5. Number of month present in dataset**

SELECT

COUNT(DISTINCT CONCAT(YEAR(Date), '-', MONTH(Date))) AS Number\_of\_Months

FROM

corona.dbo.Covid19Data;

Ans:No. Of Months are 18

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

SELECT

YEAR(Date) AS Year,

DATENAME(MONTH, Date) AS Month,

AVG(Confirmed) AS Avg\_Confirmed,

AVG(Deaths) AS Avg\_Deaths,

AVG(Recovered) AS Avg\_Monthly\_Recovered

FROM

Covid19Data

WHERE

YEAR(Date) IN (2020, 2021)

GROUP BY

YEAR(Date),

DATENAME(MONTH, Date)

ORDER BY

YEAR(Date),

Avg\_Confirmed, Avg\_Deaths, Avg\_Monthly\_Recovered, Month;

**Q7. Find most frequent value for confirmed, deaths, recovered each month**

SELECT Month(Date) As Months, Confirmed, Deaths, Recovered, COUNT(\*) AS Frequency

FROM Covid19Data

GROUP BY Month(Date), Confirmed, Deaths, Recovered

ORDER BY Month(Date), Count(\*) DESC;

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

SELECT MONTH(Date) As Months,

MIN(Confirmed) AS Min\_Confirmed,

Min(Deaths) AS Min\_Deaths,

MIN(Recovered) AS Min\_Recovered

FROM Covid19Data

GROUP BY MONTH(Date)

ORDER BY MONTH(Date);

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

SELECT MONTH(Date) As Months,

MAX(Confirmed) AS Max\_Confirmed,

MAX(Deaths) AS Max\_Deaths,

MAX(Recovered) AS Max\_Recovered

FROM Covid19Data

GROUP BY MONTH(Date)

ORDER BY MONTH(Date);

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

SELECT Month(Date) As Month,

SUM(Confirmed) AS Total\_Confirmed,

SUM(Deaths) AS Total\_Deaths,

SUM(Recovered) AS Toatl\_Recovered

from Covid19Data

GROUP BY MONTH(Date)

ORDER BY MONTH(Date);

SELECT

SUM(Confirmed) AS Total\_Confirmed,

SUM(Deaths) AS Total\_Deaths,

SUM(Recovered) AS Total\_Recovered

from Covid19Data

**Q11. Check how corona virus spread out with respect to confirmed case**

**(Eg.: total confirmed cases, their average, variance & STDEV )**

SELECT SUM(Confirmed) AS Total\_ConfirmedCases,

AVG(Confirmed) AS AVG\_ConfirmedCases,

VAR(Confirmed) AS Variance\_ConfirmedCases,

STDEV(Confirmed) AS STDev\_ConfirmedCases

FROM Covid19Data

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

**(Eg.: total confirmed cases, their average, variance & STDEV )**

SELECT MONTH(Date) AS Month,

SUM(Deaths) AS Total\_Deaths,

AVG(Deaths) AS AVG\_Deaths,

VAR(Deaths) AS Variance\_Deaths,

STDEV(Deaths) AS STDev\_Deaths

FROM Covid19Data

GROUP BY MONTH(Date)

ORDER BY MONTH(Date);

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

**(Eg.: total confirmed cases, their average, variance & STDEV )**

SELECT

SUM(Recovered) AS Total\_Recovered,

AVG(Recovered) AS AVG\_Recovered,

VAR(Recovered) AS Variance\_Recovered,

STDEV(Recovered) AS STDev\_Recovered

FROM Covid19Data

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

SELECT TOP 1 Country\_Region,

SUM(Confirmed) AS Highest\_ConfirmedCases

FROM Covid19Data

GROUP BY Country\_Region

ORDER BY Highest\_ConfirmedCases DESC;

Ans--USA with 33.46millions Confirmed cases

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

SELECT TOP 1 Country\_Region,

SUM(Deaths) AS Lowest\_DeathCases

FROM Covid19Data

GROUP BY Country\_Region

ORDER BY Lowest\_DeathCases ASC

Ans:Marshall Islands had 0 Death Cases

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

SELECT TOP 5 Country\_Region,

SUM(Recovered) AS Highest\_RecoveredCases

FROM Covid19Data

GROUP BY Country\_Region

ORDER BY Highest\_RecoveredCases DESC;

Ans:India with 28.09 million, Brazil with 15.40 million, US with 6.30 million, Turkey with 5.20 million and Russia with 4.75 million, highest recovered cases