INTRODUCTION:-

source code is a fundamental skill for programmers and software developers. It involves not only writing code but also reading, analyzing, debugging, and optimizing it to create efficient and reliable software systems.

Q1.Write a code to check NULL values ?

Syntax:-

SELECT \*FROM `Corona Virus Dataset`WHERE Confirmed IS NULL;

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

Syntax:-

UPDATE `Corona Virus Dataset` SET Confirmed = COALESCE (Confirmed, 0), Deaths = COALESCE(Deaths, 0), Recovered = COALESCE(Recovered, 0);

SELECT \*FROM `Corona Virus Dataset`;

Q3. check total number of rows ?

Syntax:-

SELECT COUNT(\*)FROM "Corona virus Dataset" ;

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

Syntax:-

SELECT MIN('2020-01-22') AS start\_date, MAX('2021-06-13') AS end\_date FROM "Corona Virus Dataset";

Q5 .Number of month present in dataset ?

Syntax:-

SELECT COUNT(DISTINCT SUBSTR(Date, 5)) AS Number \_Month FROM `Corona Virus Dataset`;

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

Syntax:-

SELECT SUBSTR(Date,5) AS Month,

AVG(Confirmed) AS Avg\_Confirmed,

AVG(Deaths) AS Avg\_Deaths,

AVG(Recovered) AS Avg\_Recovered

FROM `Corona Virus Dataset`

GROUP BY SUBSTR(Date, 5);

Q7.Find Most Frequent Value For Confirmed, Deaths, Recovered Each Month ?

Syntax:-

SELECT SUBSTR(Date, 5) AS Month,

MAX(Confirmed) AS Most\_Frequent\_Confirmed,

MAX(Deaths) AS Most\_Frequent\_Deaths,

MAX(Recovered) AS Most\_Frequent\_Recovered

FROM `Corona Virus Dataset`GROUP BY SUBSTR(Date,5);

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

Syntax:-

SELECT

SUBSTR(Date, 7, 5) AS Year,

Min(Confirmed) AS Min\_Confirmed,

Min(Deaths) AS Min\_Deaths,

Min(Recovered) AS Min\_Recovered

FROM `Corona Virus Dataset`

GROUP BY SUBSTR(Date, 7, 5);

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

Syntax:-

SELECT

SUBSTR(Date, 7, 5) AS Year,

Max(Confirmed) AS Min\_Confirmed,

Max(Deaths) AS Min\_Deaths,

Max(Recovered) AS Min\_Recovered

FROM `Corona Virus Dataset`

GROUP BY SUBSTR(Date, 7, 5);

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

Syntax:-

SELECT

SUBSTR(Date ,4 ) AS Month,

SUM(Confirmed) AS Total\_Confirmed,

SUM(Deaths) AS Total\_Deaths,

SUM(Recovered) AS Total\_Recovered

FROM `Corona Virus Dataset`

GROUP BY SUBSTR(Date , 4);

11 Check how corona virus spread out with respect to confirmed case ?

11a- Calculate total confirmed cases ?

Syntax:-

SELECT

SUM(Confirmed) AS Total\_Confirmed\_Cases

FROM

`Corona Virus Dataset`;

11B:- Calculate total confirmed cases ?

Syntax:-

SELECT avg(Confirmed) AS Total\_Confirmed\_Cases

FROM `Corona Virus Dataset`;

11c-- Calculate the variance ?

Syntax:-

SELECT SUM((Confirmed - Mean\_Confirmed) \* (Confirmed - Mean\_Confirmed)) / COUNT(\*) AS Variance\_Confirmed\_Cases

FROM

`Corona Virus Dataset`,

(SELECT AVG(Confirmed) AS Mean\_Confirmed FROM `Corona Virus Dataset`);

VARIANCE FUNCTION:-

It seems that the SQL engine you're using doesn't support the VARIANCE function. In such cases, you can calculate the variance manually using other SQL functions. One common approach is to calculate the variance using the

FORMULA:

Variance=1𝑁∑𝑖=1(𝑥𝑖−𝜇)2Variance=*N*1​∑*i*=1*N*​(*xi*​−*μ*)2

Where:

* 𝑁*N* is the number of observations.
* 𝑥𝑖*xi*​ is each observation.
* 𝜇*μ* is the mean of the observations.
* This formula calculates the standard deviation manually using the formula:

1.Standard Deviation=VarianceStandard Deviation=Variance​

2.Where the variance is calculated as previously explained.

3.Adjust the table name and column name (**Date** and **Deaths**) if they differ in your dataset.

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

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

Syntax:-

12A- Calculate total death cases per month

SELECT

SUBSTR(Date,2) AS Month,

SUM(Deaths) AS Total\_Death\_Cases

FROM

`Corona Virus Dataset`

GROUP BY

SUBSTR(Date,4);

12b-- -- Calculate avg death cases per month

Syntax:-

SELECT

SUBSTR(Date,2) AS Month,

avg(Deaths) AS Total\_Death\_Cases

FROM

`Corona Virus Dataset`

GROUP BY

SUBSTR(Date,4);

Q13A. Calculate variance of death cases per month

Syntax:-

SELECT SUBSTR(Date, 2) AS Month,

(sum(Confirmed \* Confirmed) / COUNT(\*)) - (SUM(Confirmed) \* SUM(Confirmed) / COUNT(\*) / COUNT(\*)) AS Variance\_Recovered\_Cases

FROM

`Corona Virus Dataset`

GROUP BY

SUBSTR(Date, 4);

13B- Calculate Total of Recovered cases per month

Syntax:-

SELECT

SUBSTR(Date, 2) AS Month,

SUM(Recovered) AS Total\_Recoverd\_Cases

FROM

`Corona Virus Dataset`

GROUP BY

SUBSTR(Date,4);

13C- Calculate Average of Recovered cases per month

Syntax:-

SELECT

SUBSTR(Date, 2) AS Month,

Avg(Recovered) AS Total\_Recoverd\_Cases

FROM

`Corona Virus Dataset`

GROUP BY

SUBSTR(Date,4);

13D- Calculate variance of death cases per month

Syntax:-

SELECT

SUBSTR(Date, 2) AS Month,

(SUM(Recovered \* Recovered) / COUNT(\*)) - (SUM(Recovered) \* SUM(Recovered) / COUNT(\*) / COUNT(\*)) AS Variance\_Recovered\_Cases

FROM

`Corona Virus Dataset`

GROUP BY

SUBSTR(Date, 4);

14-Find top 5 countries having highest recovered case ?

Syntax:-

SELECT "Country/Region" AS Country, MAX(Confirmed) AS Highest\_Confirmed\_Cases FROM `Corona Virus Dataset`;

15- Find Country having lowest number of the death case

Syntax:-

SELECT "Country/Region" AS Country, Min(Deaths) AS lowest\_Deaths\_Cases FROM `Corona Virus Dataset`;

16-Find top 5 countries having highest recovered case

Syntax:-

SELECT "Country/Region" AS Country,SUM(Recovered) AS Total\_Recovered\_Cases FROM `Corona Virus Dataset`

GROUP BY "Country/Region" ORDER BY Total\_Recovered\_Cases DESC LIMIT 5;