NAME: AKSHAT MEHTA

TOPIC: CORONA VIRUS ANALYSIS

Batch Name: MIP-DA-07

-- Q1. Write a code to check NULL values

SELECT \*

FROM coronadata

WHERE

Province IS NULL

OR Country 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

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

UPDATE coronadata

SET Province = COALESCE (Province, 0),

Country = COALESCE (Country, 0),

Latitude = COALESCE (Latitude, 0),

Longitude = COALESCE (Longitude, 0),

Date = COALESCE (Date, 0),

Confirmed = COALESCE (Confirmed, 0),

Deaths = COALESCE (Deaths, 0),

Recovered = COALESCE (Recovered, 0);

-- Q3. check total number of rows

SELECT COUNT(\*)

FROM coronadata

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

SELECT min(Date) as start\_date,

max(Date) as end\_date

FROM coronadata

-- Q5. Number of month present in dataset

SELECT

COUNT(distinct substr(date, 6, 2)) months

FROM coronadata;

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

SELECT

substr(date, 1, 4) as year,

substr(date, 6, 2) as month,

avg(Confirmed) as confirmed\_avg,

avg(Deaths) as death\_avg,

avg(Recovered) as recorded\_avg

FROM coronadata

GROUP BY year, month

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

SELECT

substr(date, 1, 4) as year,

substr(date, 6, 2) as month,

max(confirmed) as confirmed\_max,

max(deaths) as death\_max,

max(recovered) as recovered\_max

FROM coronadata

GROUP BY year, month

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

SELECT

substr(date, 1, 4) as year,

substr(date, 6, 2) as month,

min(confirmed) as min\_confirmed,

min(deaths) as min\_deaths,

min(recovered) as min\_recovered

FROM coronadata

GROUP BY year

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

SELECT

substr(date, 1, 4) as year,

substr(date, 6, 2) as month,

max(confirmed) as max\_confirmed,

max(deaths) as max\_deaths,

max(recovered) as max\_recovered

FROM coronadata

GROUP BY year

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

SELECT

substr(date, 1, 4) as year,

substr(date, 6 ,2) as month,

count(confirmed),

count(deaths),

count(recovered)

FROM coronadata

GROUP BY year, month

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

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

SELECT

SUM(Confirmed) AS TotalConfirmedCases,

AVG(Confirmed) AS AverageConfirmedCases,

VARIANCE(Confirmed) AS VarianceConfirmedCases,

STDDEV(Confirmed) AS StdDevConfirmedCases

FROM coronadata

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

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

SELECT

SUM(Deaths) as total\_deaths,

AVG(Deaths) as avg\_deaths,

VARIANCE(Deaths) as var\_deaths,

STDDEV(Deaths) as stddev\_deaths

from coronadata

-- 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,

VARIANCE(Recovered) as var\_recovered,

STDDEV(Recovered) as stddev\_recovered

from coronadata

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

SELECT Country,

MAX(Confirmed) AS Highest\_Confirm

from coronadata

GROUP BY Country

ORDER BY Highest\_Confirm DESC

LIMIT 1

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

SELECT Country,

MIN(Deaths) AS Lowest\_Death

FROM coronadata

GROUP BY Country

ORDER BY Lowest\_Death DESC

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

SELECT Country,

MAX(Recovered) AS Highest\_Recovered

FROM coronadata

GROUP BY Country

ORDER BY Highest\_Recovered DESC

LIMIT 5