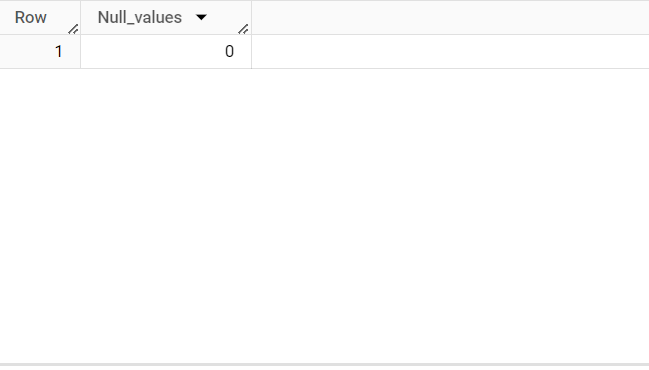
* Q1 checking the null values

SELECT count(\*) as Null\_values FROM `corona\_virus.corona`

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;



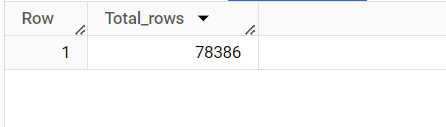
Insight – There are no null values present in the data

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

Insight- Because there a no null values so we didn’t update the value

--Q3. check the number of rows Present in the dataset.

select count(\*) as Total\_rows FROM `corona\_virus.corona`;

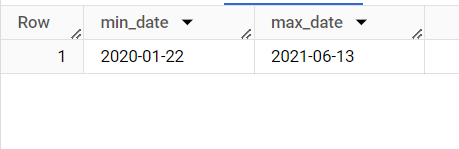


Insight- There are total 78386 rows are present in the data set

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

SELECT MIN(Date) as min\_date, MAX(Date) as max\_date FROM

`corona\_virus.corona` ;

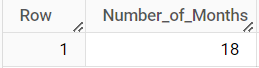


Insight- The data is from 22-01-2020 till 13-06-2021

-- Q5. Number of month present in dataset

SELECT COUNT(DISTINCT CONCAT(EXTRACT(YEAR FROM Date), '-', EXTRACT(MONTH FROM Date))) as Months

FROM `corona\_virus.corona`;



Insight- There are total 18 Months are present in the data

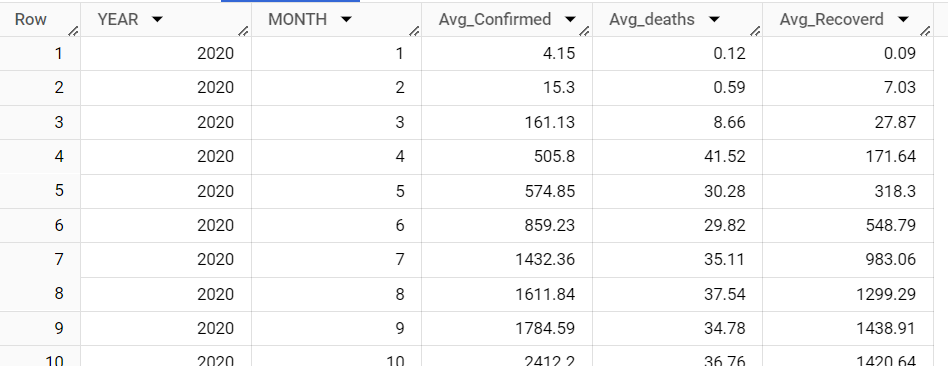
--Q6. Find monthly average for confirmed, deaths, recovered

SELECT EXTRACT(Year FROM DATE) AS YEAR,EXTRACT(MONTH FROM DATE) AS MONTH,ROUND(AVG(Confirmed),2) as Avg\_Confirmed, ROUND(AVG(deaths),2) as Avg\_deaths, ROUND(AVG(Recovered),2) as Avg\_Recoverd

from `corona\_virus.corona`

group by 1 ,2

ORDER BY 1 ,2 ASC;



Insight - This is the monthly average data for confirmed, deaths, recovered

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

SELECT \*

FROM (

    SELECT \*, RANK() OVER (ORDER BY Year) AS rank

    FROM (

        SELECT EXTRACT(YEAR FROM date) AS Year, EXTRACT(MONTH FROM date) AS month, COUNT(confirmed) AS confirmed, COUNT(deaths) AS deaths, COUNT(recovered) AS recovered

        FROM `corona\_virus.corona`

        GROUP BY 1, 2

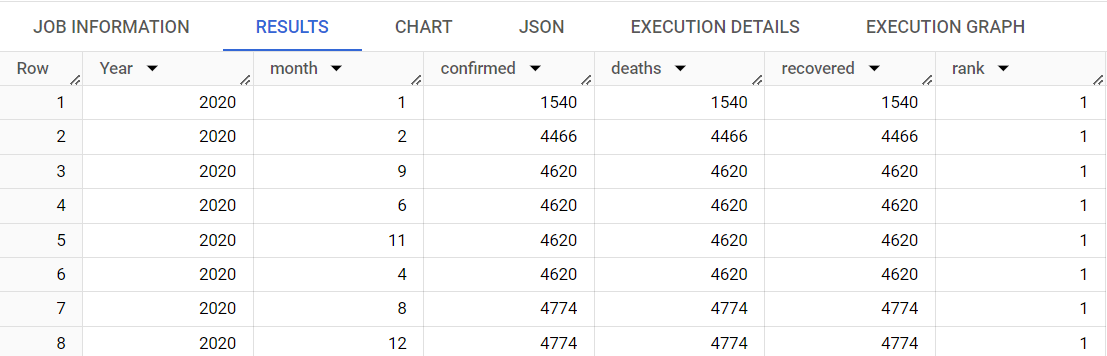
        ORDER BY 1, 2

    ) tbl

) ranked\_data

WHERE rank = 1

order by confirmed;

  
  
Insight – As you can see 9,6,11,4 months in year 2020 have most frequent value for confirmed, deaths, recovered each month

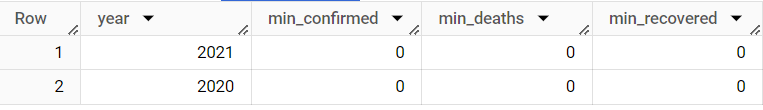
-- Q8. Find minimum values for confirmed, deaths, recovered per year

select extract(year from date) as year, min(confirmed) as min\_confirmed, min(deaths) as min\_deaths, min(recovered) as min\_recovered

from `corona\_virus.corona`

group by 1

order by 1 desc;



Insight – The minimum value for confirmed, deaths, recovered per year

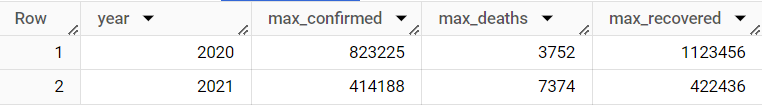
-- Q9. Find maximum values of confirmed, deaths, recovered per year

select extract(year from date) as year, max(confirmed) as max\_confirmed, max(deaths) as max\_deaths, max(recovered) as max\_recovered

from `corona\_virus.corona`

group by 1

order by 1 ;



Insight – The Maximum value for confirmed, deaths, recovered per year

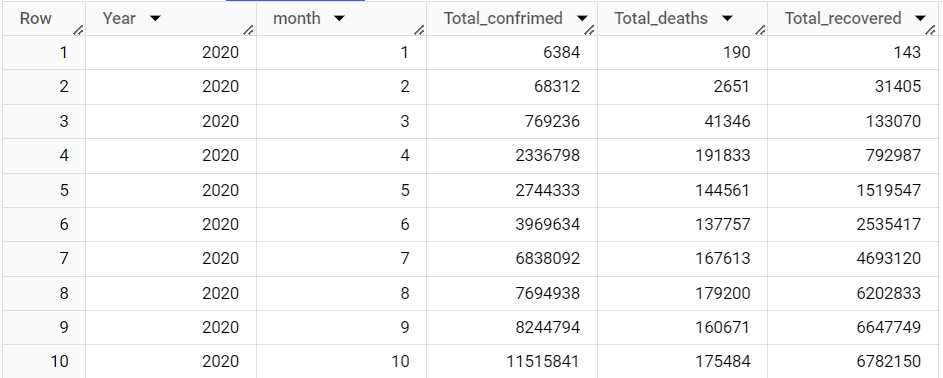
-- Q10. The total number of case of confirmed, deaths, recovered each month

select extract (year from Date) as Year,extract (month from date) as month, sum(confirmed) as Total\_confrimed,sum(deaths) as Total\_deaths, sum(recovered) as Total\_recovered

from `corona\_virus.corona`

group by 1,2

order by 1,2 ;



Insight – These are the total number of case of confirmed, deaths, recovered each month

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

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

select

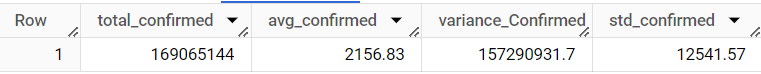
sum(Confirmed) as total\_confirmed,

round(avg(Confirmed),2) as avg\_confirmed,

round(variance(Confirmed),2)as variance\_Confirmed,

round(stddev(Confirmed),2) as std\_confirmed

from `corona\_virus.corona`;



Insight – as per the given question the total confirmed ,average confirmed, variance confirmed and standard deviation is displayed in above image

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

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

select

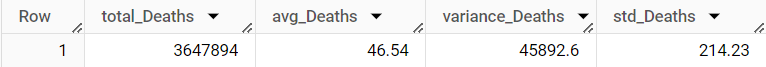
sum(Deaths) as total\_Deaths,

round(avg(Deaths),2) as avg\_Deaths,

round(variance(Deaths),2)as variance\_Deaths,

round(stddev(Deaths),2) as std\_Deaths

from `corona\_virus.corona`;



Insight – as per the given question the total deaths, average deaths, variance deaths and standard deviation is displayed in above image

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

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

select

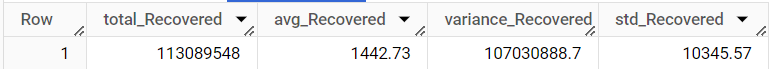
sum(Recovered) as total\_Recovered,

round(avg(Recovered),2) as avg\_Recovered,

round(variance(Recovered),2)as variance\_Recovered,

round(stddev(Recovered),2) as std\_Recovered

from `corona\_virus.corona`;



Insight – as per the given question the total recovered, average recovered, variance recovered and standard deviation is displayed in above image

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

select Country\_Region,sum(confirmed) as max\_confirmed

from `corona\_virus.corona`

group by 1

order by 2 desc



Insight – USA have the maximum confirmed cases followed by India, Brazil ,France etc

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

select Country\_Region,sum(Deaths) as min\_Deaths

from `corona\_virus.corona`

group by 1

order by 2

;



Insight – These are country with minimum deaths

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

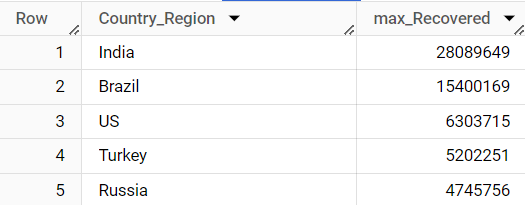
select Country\_Region,sum(Recovered) as max\_Recovered

from `corona\_virus.corona`

group by 1

order by 2 desc

limit 5;



Insight – These are top 5 country with maximum recovered cases