**Covid death**

**PPt notes**

**Source of data**

[Our World in Data](https://ourworldindata.org/explorers/coronavirus-data-explorer?zoomToSelection=true&facet=none&pickerSort=asc&pickerMetric=location&Metric=Confirmed+cases&Interval=7-day+rolling+average&Relative+to+Population=true&Color+by+test+positivity=false&country=USA~GBR~CAN~DEU~ITA~IND)

The data spans from the dates of 1.2.2020-28.11.22

**Objective**

**To showcase my skills in SQL And Tableau And parsing through data related to the corona virus**

**Tools used**

1. Microsoft sql server
2. Tableau Public

**Thing i learnt**

**SQL**

**a)Practical use case of a cte (Common Table Expression )**

Used it as a template as a to see the trend of the population vaccinated

**b)Using temp tables**

Used it for parsing through some Data that had not been created yet ,also using a drop table if statement to prevent running into errors

**c)Creating views**

**Which would used later on for tableau mp**

**d) Joins**

**e)Rolling lists**

I made a rolling list for the people that were vaccinated to help in showcasing a trend line for the vaccinated people ( I partitioned my data though location , but I also must order it by the location And date , for it to have the intended effect ,or else

**Tableau**

1. **Using map based info graphics**
2. **Setting up comparative trend lines**
3. **Making simple bar graphs**
4. **Making simple tabular based info graphics**
5. Tableau accepts nulls as strings ,so make them 0 in you excel sheet of SQL to prevent that

**PROCESS**

1. **Ask the right questions**

Set up a set of questions that you want answered from the Database(Example:- Which country has the highest death toll for covid ?)that you wish to ask

1. **Upload the Database into your Database Management System**
2. **Thin the herd**

Whittle down the data to the few categories that you need (Location,date,total\_cases ,new\_cases ,total\_deaths ,population)

1. **Find correlations**

Comparing attributes to get see the correlation between them

Example:- total cases vs. total deaths (mortality rate of those diagnosed with covid )

1. **Expand your search**

Join the Covid death table And the vaccination table to parse through the Data from there

1. **Store for later reference**

After making the appropriate queries store them as views

1. **Connect the data to your Data visualisation tool**

Upload the views for tableau for visualisation if you have the tableau paid version .

If not , use the tableau public And upload the views through excel

1. **Make the data make sense –Tell a story through the data**

Make visualisations through tableau

**Questions that shaped the query(What Data are we getting And why )**

a) How many new cases are added across the world per day ?

b)What is the COVID diagnosed mortality percentage in India ?

c)What is the percentage of the population that got covid in India ?

d)What are the top 5 countries in terms of infection rate ?

e)What are the bottom 5 countries in terms of infection rate ?

f)What are the countries with the highest death count per population

g) Which continent  has the highest death count ?

h)Global trend of diagnosis And deaths from the start?

I)How many deaths are happening  across the world per day ?

m)Total cases  ,total deaths And total death percentage as of now

n)What's the total vaccination percentage in relation to the population

o)Rolling count of the vaccinations all over the world

p)Rolling count of the vaccinations IN A SPECIFIC REGION

(What I cover)

* + Population Density
  + GDP per capita
  + Effect of diabetes on the disease
  + Human development index
  + Medan And mode of age of those who died in general And those who died in India And some other countries ,compared to the us, Canada and Uk

1. The way that the Data is structured for location is that sometimes it is NULL(showing the entire continent or grouping of nations .This helps in narrowing what kindof location we do need ,thus we can add filters accordingly

SELECT \*

FROM [Covid Project]..['Covid Deaths$']

order by 3,4

--SELECT \*

--FROM [Covid Project]..['Covid Vaccinations$']

--order by 3,4

--Filtered DATASET for covid mortality for the diagnosed

SELECT location,date ,total\_cases ,new\_cases ,total\_deaths ,population

FROM [Covid Project]..['Covid Deaths$']

ORDER by 1,2

--What is the COVID diagnosed mortality percentage in India ?

SELECT location,date ,total\_cases ,total\_deaths ,(total\_deaths/total\_cases)\*100 as Diagnosed\_Mortality

FROM [Covid Project]..['Covid Deaths$']

WHERE location='India'

ORDER by 1 ,2 DESC

--What is the percentage of the population that got covid in India ?

SELECT location,date ,total\_cases ,population ,(total\_cases/population)\*100 as Diagnosed\_Percentage

FROM [Covid Project]..['Covid Deaths$']

WHERE location='India'

ORDER by 1 ,2 DESC

--What are the top 5 countries in terms of infection rate

SELECT location,MAX(total\_cases) AS MAX\_CASES,population ,MAX(total\_cases/population)\*100 AS Diagnosed\_Percentage

FROM [Covid Project]..['Covid Deaths$']

GROUP BY location ,population

order by Diagnosed\_Percentage DESC

--What are the bottom 5 countries in terms of infection rate

SELECT location,MAX(total\_cases) AS Highest\_Infection\_Count,population ,MAX(total\_cases/population)\*100 AS Diagnosed\_Percentage

FROM [Covid Project]..['Covid Deaths$']

GROUP BY location ,population

order by Highest\_Infection\_Count

--What are the countries with the highest death count per poluation

SELECT location,MAX(cast(total\_deaths as int)) as Highest\_Death\_count,population,MAX(total\_deaths/population)\*100 as Death\_Percentage\_per\_population

FROM [Covid Project]..['Covid Deaths$']

where continent is not null

GROUP BY location,population

order by Highest\_Death\_count DESC

--CONTINENT BASED DATA

--Which coninent has the highest death count

SELECT LOCATION,MAX(cast(total\_deaths as int)) as Highest\_Death\_count

FROM [Covid Project]..['Covid Deaths$']

where location is not null

GROUP BY location

order by Highest\_Death\_count DESC

--Global trend of diagnosis And deaths from the start

--How many new cases are added across the world per day

SELECT date,SUM(new\_cases) AS TOTAL\_CASES\_TILL\_THAT\_DAY

FROM [Covid Project]..['Covid Deaths$']

where continent is not null

GROUP BY date

order by 1,2

--How many deaths are happenning acroos the world per day

SELECT date,SUM(new\_cases) TOTAL\_CASES\_Added\_THAT\_DAY,SUM(cast(new\_deaths as bigint)) AS TOTAL\_Deaths\_Added\_THAT\_DAY

FROM [Covid Project]..['Covid Deaths$']

where continent is not null

GROUP BY date

order by 3 desc

--Global Deaths added per day

SELECT date,SUM(new\_cases)as total\_cases,SUM(cast(new\_deaths as int))as total\_deaths

FROM [Covid Project]..['Covid Deaths$']

where continent is not null

GROUP BY date

order by 1,2

--Global Death percentage per day

SELECT date,SUM(new\_cases)as total\_cases,SUM(cast(new\_deaths as int))as total\_deaths,SUM(cast(new\_deaths as int))/SUM(new\_cases )\*100 as Death\_percenatge

FROM [Covid Project]..['Covid Deaths$']

where continent is not null

GROUP BY date

order by 1 desc

--Total cases ,total deaths And total death percentage as of now

SELECT SUM(new\_cases)as total\_cases,SUM(cast(new\_deaths as int))as total\_deaths,SUM(cast(new\_deaths as int))/SUM(new\_cases )\*100 as Death\_percenatge

FROM [Covid Project]..['Covid Deaths$']

where continent is not null

order by 1 desc

--Joining the vaccination And Death tables

SELECT\*

FROM [Covid Project]..['Covid Deaths$'] AS DEA

JOIN [Covid Project]..['Covid Vaccinations$']AS VACC

on DEA.location =vacc.location

And DEA.date=vacc.date

--What's the total vaccination percentage in relation to the population

SELECT dea.continent ,dea.location,VACC.new\_vaccinations, dea.population ,(VACC.new\_vaccinations/dea.population)\*100 as Vaccination\_percentage

FROM [Covid Project]..['Covid Deaths$'] AS DEA

JOIN [Covid Project]..['Covid Vaccinations$']AS VACC

on DEA.location =vacc.location

And DEA.date=vacc.date

WHERE DEA.continent IS NOT NULL

ORDER BY 3 DESC

--Rolling count of the vaccinations all over the world

SELECT dea.continent ,DEA.DATE,dea.location,VACC.new\_vaccinations, dea.population ,(VACC.new\_vaccinations/dea.population)\*100 as Vaccination\_percentage

,SUM(convert(bigint,vacc.new\_vaccinations)) OVER (PARTITION by dea.location ORDER BY dea.location,dea.date) as Rolling\_list\_4\_vaccinations

FROM [Covid Project]..['Covid Deaths$'] AS DEA

JOIN [Covid Project]..['Covid Vaccinations$']AS VACC

on DEA.location =vacc.location

And DEA.date=vacc.date

WHERE DEA.continent IS NOT NULL

ORDER BY 3 ,2

--Using a cte to clculate the rolling percentage of total vaccination vs. the popuoatio n of acountry

WITH VACCPERC(continent,date,location,New\_vacc,population,Vaccine\_percentage ,Rolling\_vaccines)

AS

(

SELECT dea.continent ,DEA.DATE,dea.location,VACC.new\_vaccinations, dea.population ,(VACC.new\_vaccinations/dea.population)\*100 as Vaccination\_percentage

,SUM(convert(bigint,vacc.new\_vaccinations)) OVER (PARTITION by dea.location ORDER BY dea.location,dea.date) as Rolling\_list\_4\_vaccinations

FROM [Covid Project]..['Covid Deaths$'] AS DEA

JOIN [Covid Project]..['Covid Vaccinations$']AS VACC

on DEA.location =vacc.location

And DEA.date=vacc.date

WHERE DEA.continent IS NOT NULL

)

--Rolling count of the vaccinations IN A SPECIFIC REGION

SELECT dea.continent ,DEA.DATE,dea.location,VACC.new\_vaccinations, dea.population ,(VACC.new\_vaccinations/dea.population)\*100 as Vaccination\_percentage

,SUM(convert(bigint,vacc.new\_vaccinations)) OVER (PARTITION by dea.location ORDER BY dea.location,dea.date) as Rolling\_list\_4\_vaccinations

FROM [Covid Project]..['Covid Deaths$'] AS DEA

JOIN [Covid Project]..['Covid Vaccinations$']AS VACC

on DEA.location =vacc.location

And DEA.date=vacc.date

WHERE DEA.continent IS NOT NULL

ORDER BY 3 ,2