**EDA** using SQL

Covid dataset

1. Death percentage query:

Select location,date,total\_cases,total\_deaths,(total\_deaths/total\_cases)\*100 as death\_percentage

from coviddeaths

order by 1,2;

A screenshot of a computer

Description automatically generated

1. Death percentage in India

select location,date,total\_cases,total\_deaths,(total\_deaths/total\_cases)\*100 as death\_percentage

from coviddeaths

where location='India'

order by 1,2;

A screenshot of a computer

Description automatically generated

1. Percentage of people infected with covid

select location,date,total\_cases,population,(total\_cases/population)\*100 as infected\_percentage

from coviddeaths

where location='India'

order by 1,2;

A screenshot of a computer

Description automatically generated

1. Countries with highest infection rate based on population

select location,population,max(total\_cases) as highest\_infected\_count, max((total\_cases/population))\*100 as infected\_percentage

from coviddeaths

group by location,population

order by infected\_percentage desc

A screenshot of a computer

Description automatically generated

1. Countries with highest death count per population

select location,max(cast(total\_deaths as signed)) as total\_death\_count

from coviddeaths

where continent != ""

group by location

order by total\_death\_count desc

A screenshot of a computer

Description automatically generated

1. Continents with highest deaths per population

select location,max(cast(total\_deaths as signed)) as total\_death\_count

from coviddeaths

where continent = ""

group by location

order by total\_death\_count desc;

A screenshot of a computer

Description automatically generated

1. Total number of cases and deaths globally per day

select date,sum(new\_cases) as total\_cases,sum(new\_deaths) as total\_deaths

from coviddeaths

where continent!=""

group by date

order by date;

A screenshot of a computer

Description automatically generated

1. Percentage of deaths globally

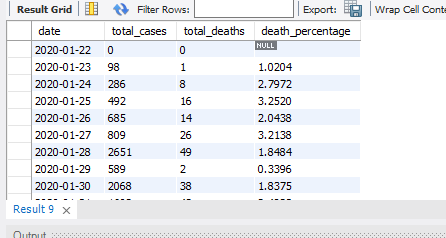
select date,sum(new\_cases) as total\_cases,sum(new\_deaths) as total\_deaths, (sum(cast(new\_deaths as signed))/sum(new\_cases))\*100 as death\_percentage

from coviddeaths

where continent!=""

group by date

order by date;



1. Total number of cases, deaths and death percentage globally

select sum(new\_cases) as total\_cases,sum(new\_deaths) as total\_deaths, (sum(cast(new\_deaths as signed))/sum(new\_cases))\*100 as death\_percentage

from coviddeaths

where continent!=""

order by date;

A screenshot of a computer

Description automatically generated

1. Rolling count of new vaccination

select dea.continent,dea.location,dea.date,dea.population,vac.new\_vaccinations,sum(vac.new\_vaccinations) over (partition by dea.location order by dea.location,dea.date) as rolling\_vaccination\_count

from coviddeaths dea join covidvaccinations vac

on dea.date = vac.date and dea.location = vac.location

where dea.continent !=""

order by 2,3;

A screenshot of a computer

Description automatically generated

1. CTE for percentage of vaccinated people based on population

with pop\_vs\_vac(Continent,Location,Date,Population,New\_vaccinations,Rolling\_vaccination\_count)

as

(

select dea.continent,dea.location,dea.date,dea.population,vac.new\_vaccinations,sum(vac.new\_vaccinations) over (partition by dea.location order by dea.location,dea.date) as rolling\_vaccination\_count

from coviddeaths dea join covidvaccinations vac

on dea.date = vac.date and dea.location = vac.location

where dea.continent !=""

)

select \*,(Rolling\_vaccination\_count/population)\*100 as percentage\_vaccinated from pop\_vs\_vac;

A screenshot of a computer

Description automatically generated

1. Temp table for percentage of vaccinated people based on population

drop table if exists percent\_pop\_vac\_temp\_table;

create table percent\_pop\_vac\_temp\_table(

Continent char(255),

Location char(255),

Date datetime,

Population numeric,

New\_vaccinations numeric null,

Rolling\_vaccination\_count numeric

);

Insert into percent\_pop\_vac\_temp\_table

select dea.continent,dea.location,dea.date,dea.population,vac.new\_vaccinations,sum(vac.new\_vaccinations) over (partition by dea.location order by dea.location,dea.date) as rolling\_vaccination\_count

from coviddeaths dea join covidvaccinations vac

on dea.date = vac.date and dea.location = vac.location

where dea.continent !="";

select \*,(Rolling\_vaccination\_count/population)\*100 as percentage\_vaccinated from percent\_pop\_vac\_temp\_table;

A screenshot of a computer

Description automatically generated

1. Creating view

create view percentage\_infected\_view as

select location,population,max(total\_cases) as highest\_infected\_count, max((total\_cases/population))\*100 as infected\_percentage

from coviddeaths

group by location

order by infected\_percentage desc;