-- COVID-19 DATA EXPLORATION

-- This project will explore global Covid deaths and vaccination data between 01/01/2020 and 30/04/2021

-- Skills used: Joins, CTE's, Temp Tables, Windows Functions, Aggregate Functions, Creating Views, Converting Data Types

USE Covid19;

-- Here, I am making sure both data sets have imported correctly.

SELECT \*

FROM dbo.CovidDeaths;

SELECT \*

FROM dbo.CovidVaccinations;

-- Testing out random bits of data to see if data set has imported correctly.

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

FROM dbo.CovidDeaths

ORDER BY 1,2;

-- TOTAL CASES VS TOTAL DEATHS.

-- Shows the likelihood of dying from Covid, by date, in the UK in descending order. (See DeathPercentage rate)

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

ROUND((total\_deaths/total\_cases)\*100,2) as DeathPercentage

FROM dbo.CovidDeaths

WHERE Location like '%Kingdom%'

AND continent IS NOT NULL

ORDER BY 5 DESC

-- TOTAL CASES VS TOTAL POPULATION.

-- Shows the percentage of the population infected with Covid in the UK.

SELECT location, date, population, total\_cases,

ROUND((total\_cases/population)\*100,4) as PopulationInfected

FROM dbo.CovidDeaths

WHERE Location like '%Kingdom%'

AND continent IS NOT NULL

ORDER BY 5 DESC

-- COUNTRIES WITH HIGHEST INFECTION RATE COMPARED TO POPULATION

-- Shows the percetage of the population affected by Covid

SELECT location, population,

MAX(total\_cases) as HighestInfectionCount,

ROUND(MAX((total\_cases/population)\*100),3) as PercetagePopulationInfected

FROM dbo.CovidDeaths

WHERE Continent IS NOT NULL

AND total\_cases IS NOT NULL

GROUP BY Location, population

ORDER BY PercetagePopulationInfected DESC;

-- BREAKING THINGS DOWN BY CONTINENT

-- Showing contintents with the highest death count per population

SELECT continent,

SUM(CAST(new\_deaths as INT)) as TotalDeathCount

FROM dbo.CovidDeaths

WHERE continent is NOT NULL

GROUP BY continent

ORDER BY TotalDeathCount DESC;

-- GLOBAL NUMBERS

-- Showing the percentage rate of new cases dying globally per day each day

SELECT date,

SUM(new\_cases) as TotalCases,

SUM(cast(total\_deaths as INT)) as TotalDeaths,

SUM(cast(total\_deaths as INT))/SUM(new\_cases)\*100 as DeathPercentage

FROM dbo.CovidDeaths

WHERE Continent IS NOT NULL

GROUP BY date

Order by 1,2

-- LOOKING AT TOTAL POPULATION vs TOTAL VACCINATIONS

-- Looking at the total number of people in the world who are vaccinated.

-- When partitioned by location the Rolling\_People\_Vaccinated shows the number of new people vaccinated each day as a running total.

SELECT dea.continent, dea.location, dea.date, dea.population, vac.new\_vaccinations,

SUM (CAST(vac.new\_vaccinations as INT)) OVER (PARTITION BY dea.location

ORDER BY dea.location, dea.date) AS Rolling\_People\_Vaccinated

FROM dbo.CovidDeaths dea

JOIN dbo.CovidVaccinations vac

ON dea.location = vac.location

AND dea.date = vac.date

WHERE dea.continent IS NOT NULL

ORDER BY 2,3;

-- USING CTE

WITH PopvsVac (Continent, Location, Date, Population, New\_Vaccinations, RollingPeopleVaccinated)

as

(

SELECT dea.continent, dea.location, dea.date, dea.population, vac.new\_vaccinations,

SUM(CONVERT(int,vac.new\_vaccinations)) OVER (Partition by dea.Location

ORDER BY dea.location, dea.Date) as RollingPeopleVaccinated

FROM dbo.CovidDeaths dea

JOIN dbo.CovidVaccinations vac

ON dea.location = vac.location

AND dea.date = vac.date

WHERE dea.continent is not null

)

SELECT \*,

CONCAT(ROUND(RollingPeopleVaccinated / Population \* 100,2),'%') AS Vaccination\_Percentage

-- I like seeing percentage results with the % sign so this is why I used Concat and Round.

FROM PopvsVac;

-- TEMP TABLE

DROP TABLE if exists #PercentPopulationVaccinated

CREATE TABLE #PercentPopulationVaccinated

(

Continent nvarchar(255),

Location nvarchar(255),

Date datetime,

Population numeric (18,0),

new\_vaccination numeric (10,0),

RollingPeopleVaccinated numeric

)

INSERT INTO #PercentPopulationVaccinated

SELECT dea.continent, dea.location, dea.date, dea.population, vac.new\_vaccinations,

SUM(CONVERT(int,vac.new\_vaccinations)) OVER (Partition by dea.Location

ORDER BY dea.location, dea.Date) as RollingPeopleVaccinated

FROM dbo.CovidDeaths dea

JOIN dbo.CovidVaccinations vac

ON dea.location = vac.location

AND dea.date = vac.date

WHERE dea.continent is not null

SELECT \*,

ROUND(RollingPeopleVaccinated / Population \* 100,2) AS Vaccination\_Percentage

FROM #PercentPopulationVaccinated;

-- CREATING A VIEW

CREATE VIEW PercentPopulationVaccinated as

SELECT dea.continent, dea.location, dea.date, dea.population, vac.new\_vaccinations,

SUM(CONVERT(int,vac.new\_vaccinations)) OVER (Partition by dea.Location

ORDER BY dea.location, dea.Date) as RollingPeopleVaccinated

FROM dbo.CovidDeaths dea

JOIN dbo.CovidVaccinations vac

ON dea.location = vac.location

AND dea.date = vac.date

WHERE dea.continent is not null

SELECT \* FROM PercentPopulationVaccinated