COVID-19 Data Exploration

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

```
--
SELECT *
From `sigma-celerity-358814.PortfolioProject.CovidDeaths`
WHERE continent IS NOT NULL
order by 3,4

SELECT *
From `sigma-celerity-358814.PortfolioProject.CovidVaccinations`
WHERE continent IS NOT NULL
order by 3,4


-- Select Data that we are going to be using

SELECT Location, date, total_cases,new_cases, total_deaths, population
From `sigma-celerity-358814.PortfolioProject.CovidDeaths`
WHERE continent IS NOT NULL
order by 1,2

-- Looking at Total Cases vs Total Deaths
-- Shows likelihood of dying if you contract Covid in your country

SELECT Location, date, total_cases, total_deaths, (total_deaths/total_cases)*100 AS
DeathPercentage
From `sigma-celerity-358814.PortfolioProject.CovidDeaths`
WHERE continent IS NOT NULL
AND location like '%States%'
order by 1,2

-- Looking at Total Cases vs Population
-- Shows what percentage of US population contracted Covid
```

```sql
SELECT Location, date, population, total_cases, (total_cases/population)*100 AS
DeathPercentage
From `sigma-celerity-358814.PortfolioProject.CovidDeaths`
WHERE location like '%States%' and continent IS NOT NULL
order by 1,2


-- Looking at countries with Highest Infection Rate compared to Population

SELECT Location, population, MAX(total_cases) AS HighestInfectionCount,
MAX((total_cases/population))*100
AS PercentPopulationInfected
From `sigma-celerity-358814.PortfolioProject.CovidDeaths`
WHERE continent IS NOT NULL
GROUP BY Location, population
order by PercentPopulationInfected DESC

-- Showing Countries with Highest Death Count per Population


SELECT location, MAX(cast(total_deaths AS int)) AS TotalDeathCount
From `sigma-celerity-358814.PortfolioProject.CovidDeaths`
WHERE continent IS NOT NULL
GROUP BY Location
order by TotalDeathCount DESC

-- LET'S BREAK THINGS DOWN BY CONTINENT

SELECT continent, MAX(cast(total_deaths AS int)) AS TotalDeathCount
From `sigma-celerity-358814.PortfolioProject.CovidDeaths`
WHERE continent IS NOT NULL
GROUP BY continent
order by TotalDeathCount DESC


-- Showing the continents with the Highest Death Count
```

```sql
SELECT continent, MAX(cast(total_deaths AS int)) AS TotalDeathCount
From `sigma-celerity-358814.PortfolioProject.CovidDeaths`
WHERE continent IS NOT NULL
GROUP BY continent
order by TotalDeathCount DESC


-- GLOBAL NUMBERS

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 DeathPercentage
From `sigma-celerity-358814.PortfolioProject.CovidDeaths`
WHERE continent IS NOT NULL
GROUP BY date
order by 1,2

-- Looking at Total Population Vs Vaccinations

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 RollingPeopleVaccinated
From `sigma-celerity-358814.PortfolioProject.CovidDeaths` dea
Join `sigma-celerity-358814.PortfolioProject.CovidVaccinations` vac
 ON dea.location = vac.location
 AND dea.date = vac.date
WHERE dea.continent IS NOT NULL
order by 2,3

-- USE CTE

WITH Pop_Vac 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 RollingPeopleVaccinated
```

```sql
FROM `sigma-celerity-358814.PortfolioProject.CovidDeaths` dea
JOIN `sigma-celerity-358814.PortfolioProject.CovidVaccinations` vac
 ON dea.location = vac.location
 AND dea.date = vac.date
WHERE dea.continent IS NOT NULL
 )
SELECT *, (RollingPeopleVaccinated/Population)*100
From Pop_Vac

-- TEMP TABLE

CREATE TABLE sigma-celerity-358814.PortfolioProject.PercentPopulationVaccinated
  (continent STRING,
 location STRING,
 Date datetime,
 Population numeric,
 new_vaccinations numeric,
 RollingPeopleVaccinated numeric
 )

INSERT INTO sigma-celerity-358814.PortfolioProject.PercentPopulationVaccinated
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 RollingPeopleVaccinated
FROM `sigma-celerity-358814.PortfolioProject.CovidDeaths` dea
JOIN `sigma-celerity-358814.PortfolioProject.CovidVaccinations` vac
 ON dea.location = vac.location
 AND dea.date = vac.date
-- WHERE dea.continent IS NOT NULL

SELECT *, (RollingPeopleVaccinated/Population)*100
From sigma-celerity-358814.PortfolioProject.PercentPopulationVaccinated


-- Creating view to store data for later visualization
```

```sql
CREATE VIEW `sigma-celerity-358814.PortfolioProject.PercentPopulationVaccinated`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 RollingPeopleVaccinated
FROM `sigma-celerity-358814.PortfolioProject.CovidDeaths` dea
JOIN `sigma-celerity-358814.PortfolioProject.CovidVaccinations` vac
 ON dea.location = vac.location
 AND dea.date = vac.date
WHERE dea.continent IS NOT NULL
```

Queries used in Tableau

--

1.

```
-- GLOBAL NUMBERS

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 DeathPercentage
From `sigma-celerity-358814.PortfolioProject.CovidDeaths`
WHERE continent IS NOT NULL
GROUP BY date
order by 1,2
```

2.

```
SELECT location, SUM(cast(new_deaths AS int)) AS TotalDeathCount
From `sigma-celerity-358814.PortfolioProject.CovidDeaths`
WHERE continent IS NOT NULL
And location not in _'World', 'European Union', 'International')
GROUP BY Location
order by TotalDeathCount DESC
```

3.

```
SELECT Location, population, MAX(total_cases) AS HighestInfectionCount,
MAX((total_cases/population))*100
AS PercentPopulationInfected
From `sigma-celerity-358814.PortfolioProject.CovidDeaths`
WHERE continent IS NOT NULL
```

```
GROUP BY Location, population
order by PercentPopulationInfected DESC
```

4.

```
SELECT Location, population, date, MAX(total_cases) AS HighestInfectionCount,
MAX((total_cases/population))*100
AS PercentPopulationInfected
From `sigma-celerity-358814.PortfolioProject.CovidDeaths`
WHERE continent IS NOT NULL
GROUP BY Location, population, date
order by PercentPopulationInfected DESC
```