Proyecto analisis de datos covid.

Se pasaron los archivos excell a csv con el objetivo de que fuera factible utilizarlos en big query. Este es el paso de exploración de datos.

--Looking at total deaths vs total cases

SELECT location, date, total\_cases, total\_deaths, (total\_deaths / total\_cases)\*100 AS DeathPercentage

FROM `analisisdedatos-367714.proyectou2.CovidDeaths`

order by 1,2

--Looking at total cases vs population

-- Shows what percentage of the population got Covid in Mexico

SELECT location, date, total\_cases, population, (total\_cases / population)\*100 AS contagious\_percentage

FROM `analisisdedatos-367714.proyectou2.CovidDeaths`

WHERE location ="Mexico"

order by 1,2

--Looking at Countries with Highest Infection rate compared to population

SELECT location, population, MAX(total\_cases) as Highest\_infection\_count, MAX((total\_cases / population))\*100 AS PercentPopulationInfected

FROM `analisisdedatos-367714.proyectou2.CovidDeaths`

GROUP BY location, population

order by PercentPopulationInfected DESC

Andorra had the highest infection count with 17%

Showing countries with highest infection rate count per population

SELECT location, MAX(total\_deaths) as TotalDeathCount

FROM `analisisdedatos-367714.proyectou2.CovidDeaths`

GROUP BY location

order by TotalDeathCount DESC

Para quitar de los resultados los continents enteros.

SELECT location, MAX(total\_deaths) as TotalDeathCount

FROM `analisisdedatos-367714.proyectou2.CovidDeaths`

WHERE continent is not null

GROUP BY location

order by TotalDeathCount DESC

para buscar casos en los que solo hay continents

SELECT continent, MAX(total\_deaths) as TotalDeathCount

FROM `analisisdedatos-367714.proyectou2.CovidDeaths`

WHERE continent is not null

GROUP BY continent

order by TotalDeathCount DESC

Numeros globales

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 `analisisdedatos-367714.proyectou2.CovidDeaths`

WHERE continent is not null

GROUP BY DATE

order by 1, 2

porcentaje de Muerte global:

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 DeathPercentage

FROM `analisisdedatos-367714.proyectou2.CovidDeaths`

WHERE continent is not null

order by 1, 2

looking at total population vs Vaccinations

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 RollingPeopleVaccinated

FROM `analisisdedatos-367714.proyectou2.CovidDeaths` dea

JOIN `analisisdedatos-367714.proyectou2.CovidVaccination` vac

ON dea.location = vac.location

AND dea.date = vac.date

WHERE dea.continent is not null

order by 2, 3

Usando una temp table

CREATE TABLE `analisisdedatos-367714.proyectou2.PercentPopulationVaccinated`

(

Continent STRING(255),

Location STRING(255),

Date DATETIME,

Population NUMERIC,

new\_vaccinations NUMERIC,

RollingPeopleVaccinated NUMERIC,)

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 RollingPeopleVaccinated

FROM `analisisdedatos-367714.proyectou2.CovidDeaths` dea

JOIN `analisisdedatos-367714.proyectou2.CovidVaccination` vac

ON dea.location = vac.location

AND dea.date = vac.date

WHERE dea.continent is not null

order by 2, 3

SELECT \*,(RollingPeopleVaccinated/Population)\*100

FROM `analisisdedatos-367714.proyectou2.PercentPopulationVaccinated`

Creating View to store data for later visualizations

CREATE VIEW `analisisdedatos-367714.proyectou2.PercentPopulationVaccinateds` AS

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 RollingPeopleVaccinated

FROM `analisisdedatos-367714.proyectou2.CovidDeaths` dea

JOIN `analisisdedatos-367714.proyectou2.CovidVaccination` vac

ON dea.location = vac.location

AND dea.date = vac.date

WHERE dea.continent is not null