**Problem 1: Determining the correlation between the number of people vaccinated and the number of people dying each day from covid-19**

To solve this problem, I will query to get data comparing the trend of the number of vaccines administered and the number of new deaths each day.

By December 2020, the new covid-19 vaccine was licensed and started to be used in some countries. However, by querying, I found that 198 countries have been affected by the epidemic.

SELECT location, total\_cases

FROM PortfolioProject..CovidDeaths

WHERE date = '2020-12-30'

AND total\_cases != 0

AND location is not NULL

GROUP BY location, total\_cases

Therefore, getting data from all countries will not be accurate. So I'll do a query to identify the top 5 most vaccinated countries and compare.

First I will create a temporary table to import data from the two tables CovidDeaths and CovidVaccinations together.

WITH deavac AS (SELECT dea.date, dea.location,

CASE WHEN total\_cases is NULL THEN 0

ELSE total\_cases END AS total\_cases,

CASE WHEN new\_deaths is NULL THEN 0

ELSE new\_deaths END AS new\_deaths,

CASE WHEN total\_vaccinations is NULL THEN 0

ELSE total\_vaccinations END AS total\_vaccinations

FROM PortfolioProject..CovidDeaths AS dea

JOIN PortfolioProject..CovidVaccinations AS vac

ON dea.location = vac.location

AND dea.date = vac.date

WHERE dea.continent is not NULL)

Subsequently, a query will be executed to ascertain the top 5 countries leading in vaccination rates.

SELECT location, MAX(total\_vaccinations) AS total\_vaccinations

FROM deavac

GROUP BY location

ORDER BY total\_vaccinations DESC

**A screenshot of a computer

Description automatically generated**

The findings revealed that the top 5 countries in terms of vaccination are "China, United States, India, United Kingdom, Brazil." Following this, a query will be conducted to extract data encompassing the date, country, daily new death figures, and the cumulative total of daily vaccinated individuals.

SELECT date, location, new\_deaths, total\_vaccinations

FROM deavac

WHERE location IN ('China', 'United States', 'India', 'United Kingdom','Brazil')

ORDER BY location, date

Next I will import this data into Microsoft Excel, then Import into Power BI.

A graph showing the spread of vaccinations

Description automatically generated

A graph showing the spread of vaccinations

Description automatically generated

A graph of vaccinations and vaccinations

Description automatically generated

A graph of vaccinations

Description automatically generated

A graph of vaccinations and vaccinations

Description automatically generated

Examination of the resultant chart indicates a negative correlation in the United Kingdom and the United States. Conversely, in China, there appears to be no discernible correlation, while in Brazil and India, a positive correlation is evident. This observation prompted an exploration of historical data. The results showed that the main reason why the daily death toll in India increased was because many people stopped following the COVID-19 protocols, such as wearing masks, maintaining social distance, and avoiding large gatherings. The festival season, the election campaigns, and the religious events also contributed to the spread of the virus. The same reason also causes the death toll in Brazil to increase

In conclusion, there is a negative correlation between the number of vaccines administered and the number of new deaths per day. However, that is not the biggest impact factor in reducing the number of deaths