Report On Covid-19 Analysis Project

The Project

The aim of the project is the analysis of a Covid-19 Dataset to draw insights and answer questions on the effect of Covid-19. Some of the questions I sought to answer are:

1. How many recorded cases have there been globally from Covid19?
2. How many recorded deaths have there been globally?
3. Which countries have had the highest deaths and cases?
4. Which continents have had the highest deaths and cases?
5. Which countries have had the highest deaths and cases per population? Are they the same with the countries with the highest deaths and cases?
6. How many people (what percentage of the world's population) have had at least one vaccination?
7. How many people (what percentage of the world's population) have been fully vaccinated?

I also wanted to be able to drill Questions 1, 2, 6 and 7 down to a country level

The Dataset

The dataset was obtained from ourworldindata.org and is the Covid figures as at the date of the download (18th August, 2021). It contains columns like location, continent, date, total_cases, new_cases, vaccinations e.t.c.

The data was downloaded as an excel file and then divided into deaths/cases and vaccinations.

The Analysis

The data was analysed using SQL. I wrote different queries based on the questions I sought to answer. For e.g. to answer the question of the Global Deaths, I wrote the below query:

```
SELECT SUM(Pop) AS GlobalPop, SUM(deaths) AS Globaldeaths, (SUM(deaths)/SUM(Pop)*100) AS GlobalDeathPercent
FROM(SELECT location, continent, MAX(population) AS Pop, MAX((CAST(total_deaths AS INT))) AS deaths, MAX(total_deaths/population)*100 AS DeathPerPop
FROM [Covid deaths]
WHERE continent is NOT NULL
GROUP BY location, continent)t1
```

To answer the question on Global Vaccination figures:

```
SELECT SUM(Pop) AS GlobalPop, SUM(people_vacc) AS GlobalPeopleVacc, SUM(people_fullvacc) AS GlobalPeopleFullVacc, (SUM(people_vacc)/SUM(Pop))*100 AS GlobalVaccPerc,
(SUM(people_fullvacc)/SUM(Pop))*100 AS GlobalFullVaccPerc
FROM(SELECT v.location, v.continent, MAX(d.population) AS Pop, MAX(CAST(v.total_vaccinations AS BIGINT)) AS vaccinations, MAX(CAST(v.people_vaccinated AS BIGINT)) AS people_vacc,
MAX(CAST(v.people_fully_vaccinated AS BIGINT)) AS people_fullvacc,
MAX(CAST(v.people_vaccinated AS BIGINT))/MAX(d.population)*100 AS VaccPerPop,
MAX(CAST(v.people_fully_vaccinated AS BIGINT))/MAX(d.population)*100 AS FullyVaccPerPop
FROM [Covid vaccinations] v
JOIN [Covid deaths] d
ON v.location = d.location
```

WHERE v.continent IS NOT NULL
GROUP BY v.location, v.continent)T1

The full details of all the queries I wrote are on my Github: github/hikwuneme/Covid-19-Data-Analysis

The Visualisation

The resulting tables from the queries were imported into PowerBI which was used for the visualisation. The visualisation was done in four sections: Cases, Deaths, Vaccinations and Deaths/Cases per Population. I included a slicer which I synchronised across all pages of the visuals to toggle between 2020 and 2021.

The Results

I believe the analysis and visuals did a good job of answering the questions I set out to answer. We can clearly see that as at the date this data was gotten (18th August, 2021), there have been 208.5Million recorded cases of Covid-19, leading to approximately 4.4Million deaths. The United States of America has had the highest recorded cases followed by India and Brazil. The United States has also had the highest recorded cases followed again by Brazil and India in that other.

Surprisingly though, Andorra has had the highest Cases Per Population, followed by Seychelles and Montenegro and Peru has had the highest Deaths Per Population followed by Hungary and Bosnia and Herzegovina. (This is most likely due to the small populations of these countries)

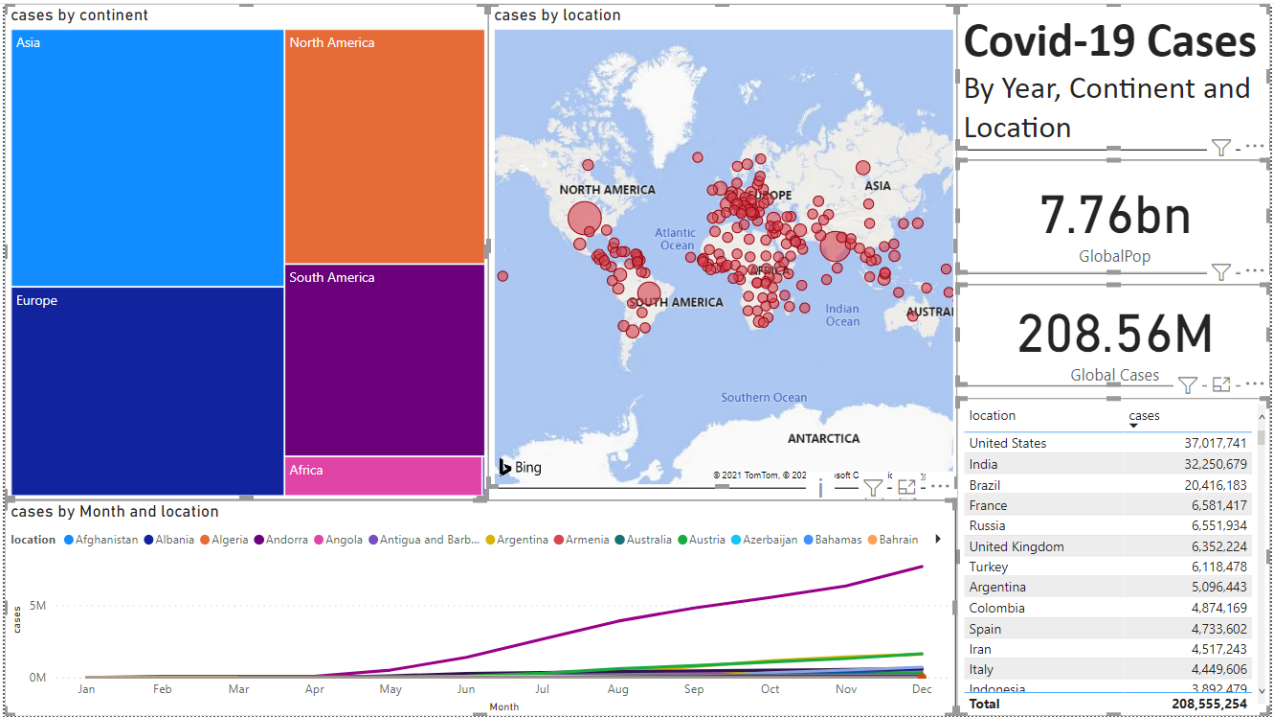

Fig. 1: Visualisation of Analysis on Covid-19 Cases

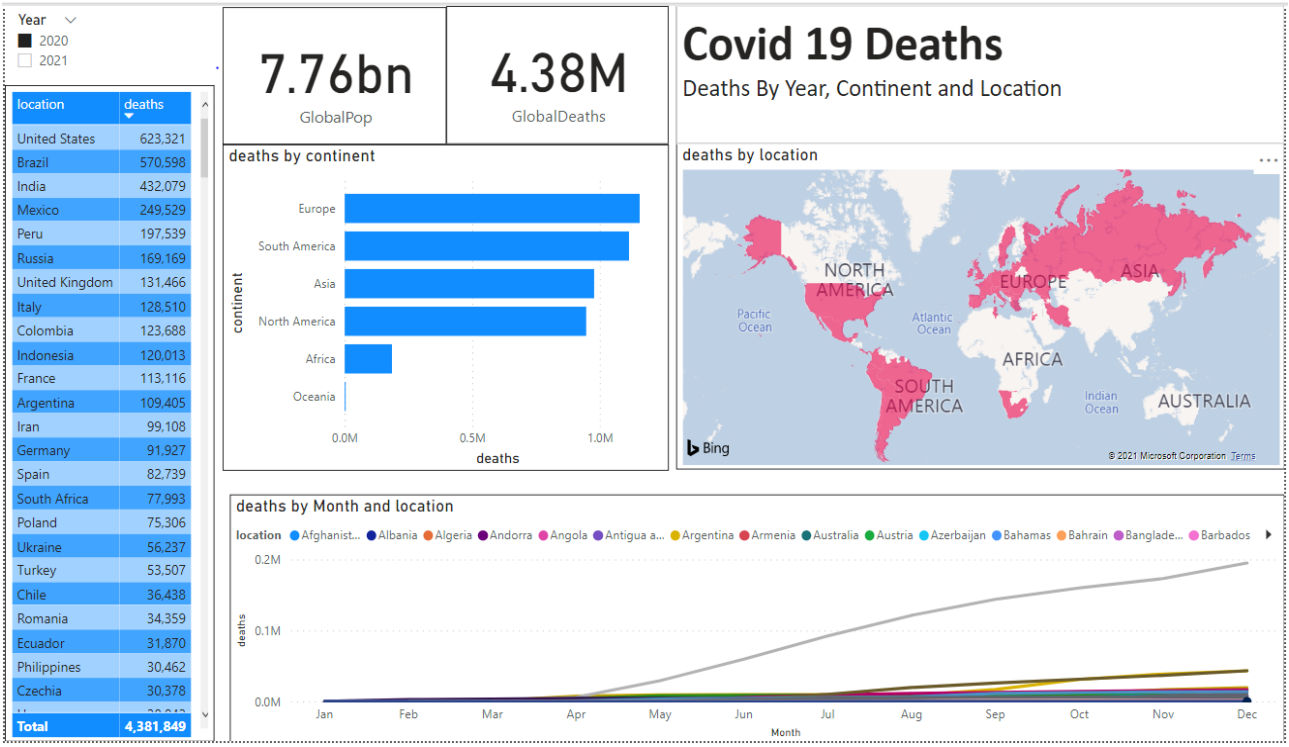

Fig. 2: Visualisation of Analysis on Covid-19 Deaths

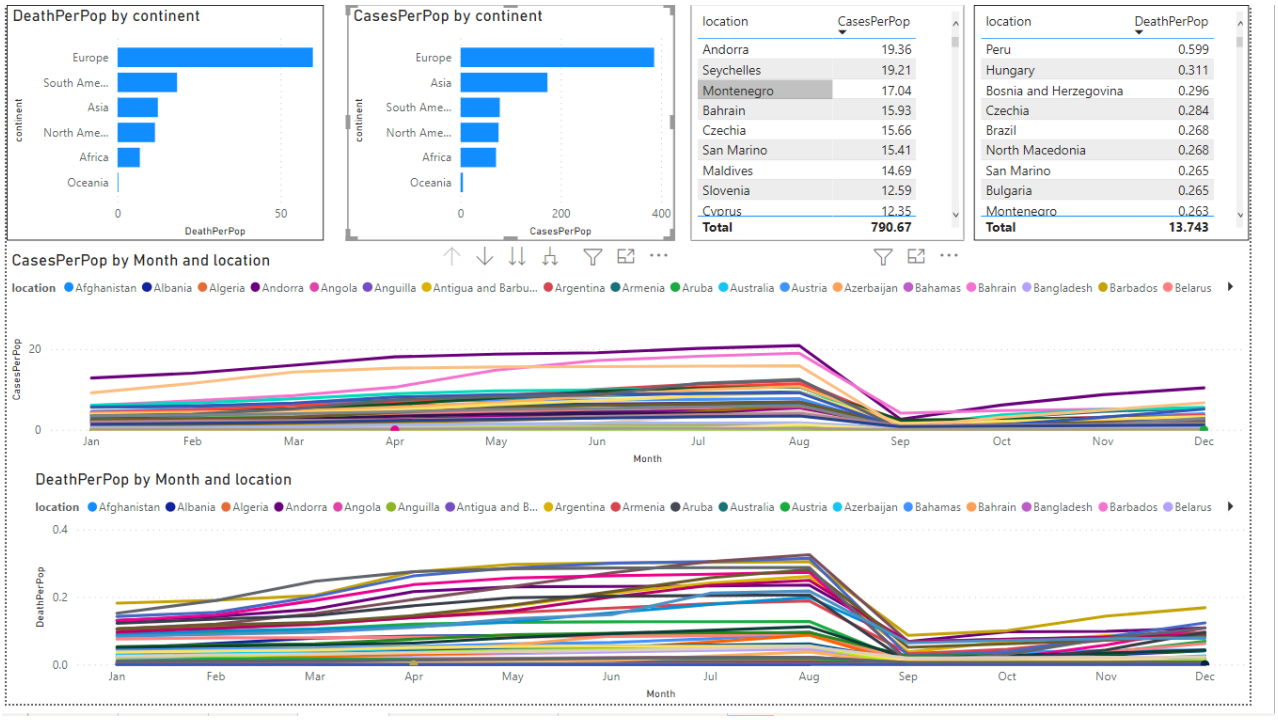

Fig. 3 Visualisation of Analysis on Covid-19 Cases and Deaths Per Population

The result of the analysis on vaccinations was also quite surprising to me. Less than half of the world's population has had at least one Covid-19 vaccinnation and less than 25% have been fully vaccinated

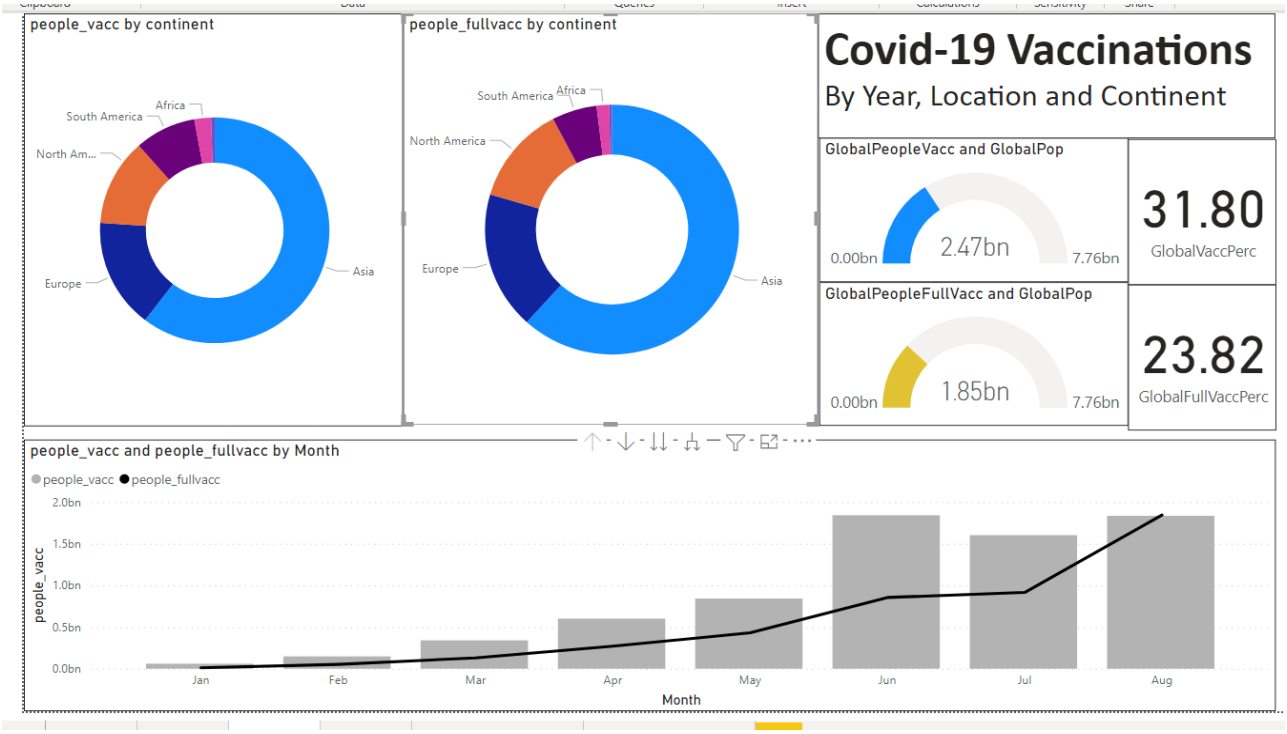

Fig. 4 Visualisation on Covid-19 Vaccinations

Conclusion and Acknowledgment

I had a lot of fun working on this project. I wanted to show my skills in analysing data and presenting insights from that data.

I would like to thank Alex Freberg (Alex The Analyst) whose tutorial was the initial inspiration behind this project and who pointed me to the data source.