Data Exploration and Visualization of Covid 19 Data.

In this Analysis the dataset is taken from “Our World in Data” and the raw data on confirmed cases and deaths for all countries is sourced from https://github.com/CSSEGISandData/COVID-19.

Select the data that we are going to be using in this analysis.

```
SELECT *
FROM Covid19Data..CovidDeaths
WHERE continent is not null
ORDER BY 3,4
```

Let's find the percentage of the total number of cases to the number of deaths in India as DeathPercentage.

```
SELECT Location, date, total_cases, total_deaths, (total_deaths/total_cases)*100 AS DeathPercentage
FROM Covid19Data..CovidDeaths
WHERE Location = 'India' AND continent is not null
ORDER BY 1,2
```

The percentage of infected individuals per total population, as calculated by the PopulationInfectedPercentage.

```
SELECT Location, date, population, total_cases, (total_cases/population)*100 AS PopulationInfectedPercentage
FROM Covid19Data..CovidDeaths
--WHERE Location = 'India'
ORDER BY 1,2
```

Percentage of Countries with the Highest Infection rate to their Total Population as PopulationInfectedPercentage

```
SELECT Location, population, MAX(total_cases) AS HighestInfectionCount, MAX(total_cases/population)*100 AS PopulationInfectedPercentag
FROM Covid19Data..CovidDeaths
--WHERE Location = 'India'
GROUP BY Location, Population
ORDER BY PopulationInfectedPercentage DESC
```

Count of Deaths to the respective population of a country as TotalDeathCount.

```
SELECT Location, MAX(CAST(total_deaths AS INT)) AS TotalDeathCount
FROM Covid19Data..CovidDeaths
--WHERE Location = 'India'
WHERE continent is not null
GROUP BY Location
ORDER BY TotalDeathCount DESC
```

Count of Deaths to the respective population by Continent as TotalDeathCount.

```
SELECT continent, MAX(CAST(total_deaths AS INT)) AS TotalDeathCount
FROM Covid19Data..CovidDeaths
--WHERE Location = 'India'
WHERE continent is not null
GROUP BY continent
ORDER BY TotalDeathCount DESC
```

Total population of a country to the total number of people vaccinated as RollingPeopleVaccinated.

```
SELECT D.continent, D.location, D.date, D.population, V.new_vaccinations, SUM(CONVERT(BIGINT,V.new_vaccinations)) OVER (PARTITION BY D
FROM Covid19Data..CovidDeaths D
JOIN Covid19Data..CovidVacs V
  ON D.location = V.location
  AND D.date = V.date
WHERE D.continent is not null
ORDER BY 2,3
SET ANSI_WARNINGS OFF;
GO
```

Queries used for Tabeau Visualization:

- Global Numbers such as the total number of cases, the total number of deaths, and the death rate.

```
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
FROM Covid19Data..CovidDeaths
--Where location = 'India'
WHERE continent is not null
--Group By date
ORDER BY 1,2
```

Output:

total_cases	total_deaths	DeathPercentage	
667557469	6704157	1.004281625	

- Exploration of Total Deaths per Continent.

```
SELECT location, SUM(CAST(new_deaths AS INT)) AS TotalDeathCount
FROM Covid19Data..CovidDeaths
--WHERE location = 'India'
WHERE continent IS NULL
AND location NOT IN ('World', 'European Union', 'International')
GROUP BY location
ORDER BY TotalDeathCount DESC
```

Output:

location	TotalDeathCount
Europe	2018890
North America	1560534
Asia	1522511

South America	1321492
Africa	257656
Oceania	23074

- Percent Population Infected per Country.

```
SELECT Location, Population, MAX(total_cases) AS HighestInfectionCount,  MAX((total_cases/population))*100 AS PercentPopulationInfecte
FROM Covid19Data..CovidDeaths
--Where location = 'India'
GROUP BY Location, Population
ORDER BY PercentPopulationInfected DESC
```

Output:

Location	Population	HighestInfectionCount	PercentPopulationInfected
Cyprus	896007	640729	71.50937437
San Marino	33690	23403	69.46571683
Faeroe Islands	53117	34658	65.24841388
Austria	8939617	5759515	64.42686527
Gibraltar	32677	20392	62.40474952
Slovenia	2119843	1320002	62.2688567
Brunei	449002	274217	61.07255647
Andorra	79843	47820	59.89253911
France	67813000	39532897	58.29692979

~ First 10 rows of 249 total rows.

- Percent Population Infected.

```
SELECT Location, Population,date, MAX(total_cases) AS HighestInfectionCount,  MAX((total_cases/population))*100 AS PercentPopulationIn
FROM Covid19Data..CovidDeaths
--Where location = 'India'
GROUP BY Location, Population, date
ORDER BY PercentPopulationInfected DESC
```

Output:

Location	Population	date	HighestInfectionCount	PercentPopulationInfected
Cyprus	896007	00:00.0	640729	71.50937437
Cyprus	896007	00:00.0	640729	71.50937437
Cyprus	896007	00:00.0	640729	71.50937437
Cyprus	896007	00:00.0	640729	71.50937437
Cyprus	896007	00:00.0	640729	71.50937437
Cyprus	896007	00:00.0	638062	71.21172044
Cyprus	896007	00:00.0	638062	71.21172044
Cyprus	896007	00:00.0	638062	71.21172044
Cyprus	896007	00:00.0	638062	71.21172044

~ First 10 rows of 252511 total rows.