SQL Covid Portfolio

Starting my project I took a look to make sure my datasets had transferred over correctly:

```
SELECT location, date, total_cases, new_cases, total_deaths, population
FROM `neural-cairn-393519.Portfolio.CovidDeaths`
ORDER BY location

--Selecting the data that I will be using
```

From there I wanted to look at Total Cases vs Total Deaths.

To figure this out I had taken total_deaths/ total_cases and then multiplied it by 100. I have this figure the alias of 'DeathPercentage'

```

--Looking at Total Cases vs Total Deaths

SELECT location, date, total_cases, total_deaths,(total_deaths/total_cases)*100 as DeathPercentage
FROM `neural-cairn-393519.Portfolio.CovidDeaths`
ORDER BY 1,2
```

This made it so when you viewed the dataset there was now a DeathPercentage field.To take it one step further I wanted to isolate the data so that I could view Canada specifically.

```

--Looking at Canada's DeathPercentage

SELECT location, date, total_cases, total_deaths,(total_deaths/total_cases)*100 as DeathPercentage
FROM `neural-cairn-393519.Portfolio.CovidDeaths`
WHERE location = 'Canada'
ORDER BY 1,2
```

Then going to the end of the data which was on 4/30/2021 we can see there were 1,228,367 total cases. From that amount 24,220 had died. Which puts our DeathPercentage at 1.97%.

This shows the likelihood of dying if you contract covid in Canada.

Row	location	date	total_cases	total_deaths	DeathPercentage
453	Canada	2021-04-22	1164108	23810	2.04534287196...
454	Canada	2021-04-23	1172697	23870	2.03547890034...
455	Canada	2021-04-24	1180065	23912	2.02632905814...
456	Canada	2021-04-25	1186991	23950	2.01770695818...
457	Canada	2021-04-26	1195827	24011	2.00789913591...
458	Canada	2021-04-27	1202672	24048	1.99954767384...
459	Canada	2021-04-28	1210918	24110	1.99105141718...
460	Canada	2021-04-29	1220108	24165	1.98056237644...
461	Canada	2021-04-30	1228367	24220	1.97172343444...

Next I wanted to look at Canada's Total Cases vs their Population

```sql

--Total Cases vs Population
--Shows what % of the population has gotten covid
SELECT location, date, population,total_cases,(total_cases/population)*100 as PercentageInfected
FROM `neural-cairn-393519.Portfolio.CovidDeaths`
WHERE location = 'Canada'
ORDER BY 1,2
```

From this query we can see 3.25% of the population had confirmed cases of covid

Row	location	date	population	total_cases	PercentofPopulation
453	Canada	2021-04-22	37742157	1164108	3.08437061506...
454	Canada	2021-04-23	37742157	1172697	3.10712766098...
455	Canada	2021-04-24	37742157	1180065	3.12664959768...
456	Canada	2021-04-25	37742157	1186991	3.14500043015...
457	Canada	2021-04-26	37742157	1195827	3.16841191667...
458	Canada	2021-04-27	37742157	1202672	3.18654813502...
459	Canada	2021-04-28	37742157	1210918	3.20839638285...
460	Canada	2021-04-29	37742157	1220108	3.23274581259...
461	Canada	2021-04-30	37742157	1228367	3.25462850467...

Next I wanted to take a look at countries with the highest infection rate compared to population.

```sql
--Looking at countries with Highest Infection Rate compared to poulation

SELECT Location, Population, MAX(total_cases) as HighestInfectionCount, MAX((total_cases/population))*100 as PercentofPopulationInfected
FROM `neural-cairn-393519.Portfolio.CovidDeaths`
GROUP BY Location, population
ORDER BY PercentofPopulationInfected DESC
```

For this I made sure to use the GROUP BY function to collapse the data into distinct values. I also changed the order so that I could see the data which countries had the highest percent of the population infected.

Row	Location	Population	HighestInfectionCou	PercentofPopulation
1	Andorra	77265	13232	17.1254772536...
2	Montenegro	628062	97389	15.5062716738...
3	Czechia	10708982	1630758	15.2279460363...
4	San Marino	33938	5066	14.9272202251...
5	Slovenia	2078932	240292	11.5584348117...
6	Luxembourg	625976	67205	10.7360346083...
7	Bahrain	1701583	176934	10.3981997939...
8	Serbia	6804596	689557	10.1336949320...
9	United States	331002647	32346971	9.77242064169...