--SELECT \*

--FROM CovidDeaths$

--Order by 3,4

--SELECT \*

--FROM covidvaccinations$

--Order by 3,4

-- Select the data we are going to be using

SELECT location,date,total\_cases,new\_cases,total\_deaths,population

FROM CovidDeaths$

Order by 1,2

--Looking at Total Cases vs. Total Deaths

SELECT location,date,total\_cases,total\_deaths, (total\_deaths/total\_cases)

FROM CovidDeaths$

Order by 1,2

--Looking at Total Cases vs. Population

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

FROM CovidDeaths$

WHERE location like '%states%'

ORDER BY 1,2

Quantify of what success looks like

Chapgpt- Data Analysis

Organizations/Cheer

Volunteer

**Project Quick Summary**

* I'll be acting as a junior data analyst at Cyclistic, a fictional company.
* My goal is to address critical business questions by following the data analysis process: ask, prepare, process, analyze, share, and act.
* Tools used: Excel, SSMS 19, Tableau.
* Skills used: Joins, CTE's, Temp Tables, Windows Functions, Aggregate Functions, Creating Views, Converting Data Types

[**Project Introduction**](https://github.com/GaneshHonrao/Google-Data-Analytics-Capstone-Project-on-Cyclistic-Bike-Share-With-SQL-Tableau#project-introduction)

**Business Needs**

* 1. What percentage of the population got Covid?
  2. What is the likelihood of dying if you contract covid in the United States.
  3. What percentage of the population got Covid in the United States.
  4. Display Countries with Highest Infection Rate compared to its Population.
  5. Display Countries with the Highest Death Count per Population
  6. Show Continents with the Highest Death Counts

**Data Gathering**

[**Data Gathering**](https://github.com/GaneshHonrao/Google-Data-Analytics-Capstone-Project-on-Cyclistic-Bike-Share-With-SQL-Tableau#data-gathering)

I plan to utilize Covid 19 data to conduct an analysis. This data can be obtained through Alex the Anaylst and <https://ourworldindata.org/covid-deaths>.   
Note - This is public data that you can use to explore how Covid 19 has affected different countries and their population.

[**Concerned Source Data**](https://github.com/GaneshHonrao/Google-Data-Analytics-Capstone-Project-on-Cyclistic-Bike-Share-With-SQL-Tableau#concerned-source-data)

1. **Data Timeframe:** September 15th - November 17th
2. **Source Data Description and Structure:**
   1. The two datasets are organized into individual .xsls files.
   2. These files share common fields relevant to our analysis.
3. After my initial examination, we can confidently assert that the source data is Reliable, Original, Comprehensive, Current and Cited.

[**Tool Selection for Analysis**](https://github.com/GaneshHonrao/Google-Data-Analytics-Capstone-Project-on-Cyclistic-Bike-Share-With-SQL-Tableau#tool-selection-for-analysis)

1. **Tool for Data Cleaning: Excel** 
   1. I chose Excel for this task because the dataset will be split into two tables, Covid Deaths and Covid Vaccinations.
2. **Tool for Data Exploration: SSMS 19**
   1. exceeded 5.8 million rows, making it impractical to manage with Excel (which is typically suitable for datasets of less than 1 million rows). SQL's capacity to handle large volumes of data makes it the preferred choice for data wrangling.
3. **Tool for Analysis and Visualization: Tableau**
   1. While there are several tools available for data visualization, including Excel and Power BI, our selection of Tableau is deliberate. Tableau offers dynamic capabilities and a diverse range of visualization options, making it the ideal choice for our analytical and visualization needs.

[**Data Loading**](https://github.com/GaneshHonrao/Google-Data-Analytics-Capstone-Project-on-Cyclistic-Bike-Share-With-SQL-Tableau#data-loading)

1. **File Naming Convention:** Files are named in the format of CovidDeaths.csv and CovidVaccinations.csv .

**Data Observation**

**Columns and Their Data Types**

Covid Deaths Table

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated

Covid Vaccinations Table

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated