1 DATA PERPARATION

I downloaded COVID-19 deaths from "ourworldindata.org/covid-deaths" as one CSV file. Afterthat, I split it into two CSV files: one for COVID-19 vaccinations and the other for COVID-19 deaths. Since I use pgAdmin 4 for data management, I needed to create tables with the column names from the CSV files. To simplify the process, I wrote a Python script to automatically extract the column names, saving time and effort.

```python
import pandas as pd
df = pd.read_csv("CovidDeaths.csv")

lst = list(df.columns)

for col in lst:
    print(col +",")
```

```
iso_code,
continent,
location,
date,
population,
total_cases,
new_cases,
new_cases_smoothed,
total_deaths,
new_deaths,
new_deaths_smoothed,
total_cases_per_million,
new_cases_per_million,
new_cases_smoothed_per_million,
total_deaths_per_million,
new_deaths_per_million,
new_deaths_smoothed_per_million,
reproduction_rate,
icu_patients,
icu_patients_per_million,
hosp_patients,
hosp_patients_per_million,
weekly_icu_admissions,
weekly_icu_admissions_per_million,
```

```python
df = pd.read_csv("CovidVaccinations.csv")


lst = list(df.columns)

for col in lst:
    print(col +",")
```

```
iso_code,
continent,
location,
date,
population,
total_tests,
new_tests,
total_tests_per_thousand,
new_tests_per_thousand,
new_tests_smoothed,
new_tests_smoothed_per_thousand,
positive_rate,
tests_per_case,
tests_units,
total_vaccinations,
people_vaccinated,
people_fully_vaccinated,
total_boosters,
new_vaccinations,
new_vaccinations_smoothed,
total_vaccinations_per_hundred,
people_vaccinated_per_hundred
```

This allowed me to quickly set up the necessary tables in pgAdmin 4 for further analysis of the COVID-19 data. Subsequently, I utilized the extracted column names and incorporated them into my SQL query. Further processes is explained in the SQL code.