**COVID-19 Data Analysis Project**

**Objective**

The aim of this project is to perform an end-to-end analysis of COVID-19 data using SQL for querying and Tableau for visualization. The data was sourced from “Our World in Data,” covering over 85,000 records with 59 columns.

**Tools & Technologies Used**

* **MySQL Server** and **Microsoft SQL Server** for database management
* **Azure Data Studio** (via Docker Desktop on macOS) for SQL Server connectivity and querying
* **Excel** (Microsoft 365) for intermediate data storage
* **Tableau** for data visualization and dashboarding

**Data Preparation**

1. **Dataset Source**: [Our World in Data](https://ourworldindata.org/)
2. **Records**: 85,172 rows | 59 columns
3. **Table Creation**:
   * Created two separate tables: Covid\_Deaths and Covid\_Vaccinations.
   * Data was cleaned and formatted using Excel before importing to SQL Server.
   * Handled data type constraints and null values during import.

**SQL Analysis**

1. **Death Rate Analysis**

Select Location, date,total\_cases,total\_deaths, (cast(total\_deaths as float)/total\_cases)\*100 as DeathPercantage

from Covid\_Deaths

Where location like '%states%'

order by date

1. **Infection Rate by Population**

Select Location, date,total\_cases,population, (cast(total\_cases as float)/population)\*100 as PercentPopulationInfected

from Covid\_Deaths

order by date

1. **Countries with Highest Infection Rate**

Select Location,population,max(total\_cases) as HighestInfectedCount, Max((cast(total\_cases as float)/population)

)\*100 as PercentPopulationInfected

from Covid\_Deaths

Group by Location, population

order by PercentPopulationInfected desc

1. **Continent-Wise Death Analysis**

Select continent, Max(Total\_deaths) as TotalDeathCount

from Covid\_Deaths

where continent is not null

Group by continent

order by TotalDeathCount desc

1. **Total Global Cases and Death Rate**

Select Location, Max(Total\_deaths) as TotalDeathCount

from Covid\_Deaths

where continent is not null

Group by Location

order by TotalDeathCount desc

Select SUM(new\_cases) as total\_cases, SUM(cast(new\_deaths as int)) as total\_deaths, SUM(cast(new\_deaths as float))/sum(new\_cases)\*100 as DeathPercantage

From Covid\_Deaths

Where continent is not null

1. **Population vs Vaccination Tracking**

Select D.continent,D.location,D.date,d.population,V.new\_VACCINATIONS, SUM(V.new\_Vaccinations) Over (Partition by D.location Order by D.Location, D.Date) as RollingPeopleVaccinated

from Covid\_Vaccinations V join Covid\_Deaths D on V.location =D.location and V.date=D.date

Where D.continent is not null

order by D.location,D.date

1. **CTE for Vaccination Percentage**

With PopvsVac (Continent,Location,Date,Population,New\_VACCINATIONS,ROLLINGPEOPLEVACCINATED)

as

(

Select D.continent,D.location,D.date,d.population,V.new\_VACCINATIONS, SUM(V.new\_Vaccinations) Over (Partition by D.location Order by D.Location, D.Date) as RollingPeopleVaccinated

from Covid\_Vaccinations V join Covid\_Deaths D on V.location =D.location and V.date=D.date

Where D.continent is not null

)

Select \*,(RollingPeopleVaccinated/cast(Population as float))\*100

From PopvsVac

1. **SQL View Creation**

Create View PercentPopulationVaccinated as

Select D.continent,D.location,D.date,d.population,V.new\_VACCINATIONS, SUM(V.new\_Vaccinations) Over (Partition by D.location Order by D.Location, D.Date) as RollingPeopleVaccinated

from Covid\_Vaccinations V join Covid\_Deaths D on V.location =D.location and V.date=D.date

Where D.continent is not null

Select \* FRom PercentPopulationVaccinated

**Tableau Visualization**

**Steps Taken:**

1. Executed SQL queries in Azure Data Studio
2. Exported result sets to Excel using Microsoft 365
3. Imported 4 Excel sheets into Tableau
4. Designed and built:
   * Infection Rate Charts
   * Death Percentage Analysis
   * Vaccination Rollout Over Time
   * Population vs Vaccination Percentage Comparison

**Dashboard Creation:**

All 4 sheets were combined into a single interactive dashboard providing a consolidated view of global COVID trends.

A screenshot of a computer

AI-generated content may be incorrect.

**Conclusion**

This project demonstrates the power of SQL for data extraction and analysis, and Tableau for storytelling through visualization. By integrating real-world data with database tools and BI platforms, we were able to uncover key insights into the global COVID-19 pandemic — from infection rates to vaccine distribution.