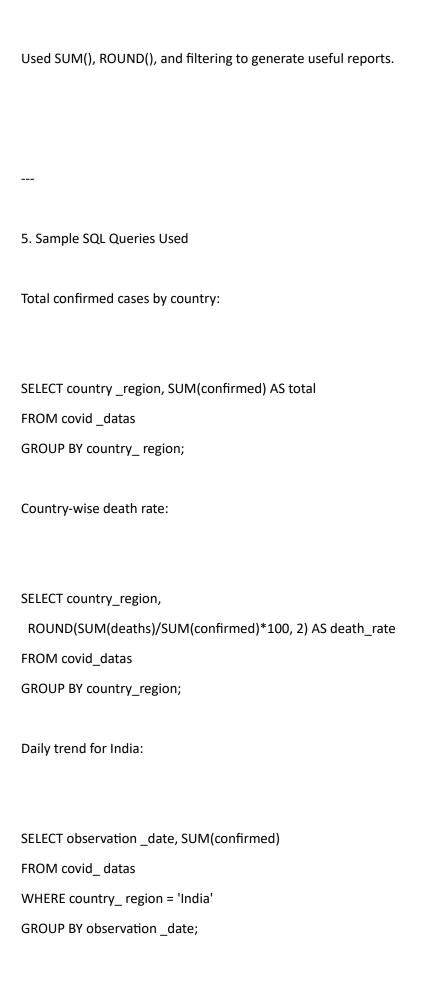# 📝 COVID-19 Data Analytics Using MySQL

Internship Project Phase – SQL Internship

---

1. Introduction

The COVID-19 pandemic has impacted the world significantly, leading to the generation of vast amounts of data on a daily basis. This project aims to leverage SQL to analyze such real-world data and draw meaningful insights. Using MySQL as the backend, we created a COVID-19 database, inserted data, and performed analytics to understand country-wise trends, death rates, and recovery patterns.

---

2. Abstract

The project revolves around importing, cleaning, and analyzing COVID-19 data using MySQL. We used a simplified dataset representing country-wise case reports over time. After preparing the dataset, we created a relational table named covid_datas and used SQL queries to extract key metrics like total confirmed cases, death rates, recovery numbers, and time-based trends.

This project showcases the importance of structured databases in organizing global-scale data and the power of SQL in performing quick and effective analysis.

---

3. Tools Used

MySQL Workbench 8.0 – For writing and executing SQL queries.

MySQL Server – As the RDBMS to store and manage data.

Microsoft Excel – To format and inspect the raw dataset.

CSV Dataset – Cleaned sample file from Kaggle based on global COVID-19 statistics.

---

4. Steps Involved in Building the Project

Created a new database: covid_analysis.

Defined a table: covid_datas with fields such as observation_date, country_region, confirmed, deaths, and recovered.

Inserted more than 20 sample data records manually using INSERT INTO statements.

Wrote SQL queries to:

Calculate total confirmed cases per country.

Determine country-wise death rate percentages.

Analyze recovery trends.

Extract India's daily COVID trend using GROUP BY and ORDER BY.

Used SUM(), ROUND(), and filtering to generate useful reports.

---

5. Sample SQL Queries Used

Total confirmed cases by country:

```
SELECT country _region, SUM(confirmed) AS total
FROM covid _datas
GROUP BY country_ region;
```

Country-wise death rate:

```
SELECT country_region,
   ROUND(SUM(deaths)/SUM(confirmed)*100, 2) AS death_rate
FROM covid_datas
GROUP BY country_region;
```

Daily trend for India:

```
SELECT observation _date, SUM(confirmed)
FROM covid_ datas
WHERE country_ region = 'India'
GROUP BY observation _date;
```

---

6. Conclusion

Through this project, I developed a solid understanding of how to build and query relational databases using MySQL. It helped me strengthen concepts like data aggregation, filtering, sorting, and group-based analysis. More importantly, it simulated a real-world scenario where SQL is used to analyze large-scale health data — a valuable experience for any aspiring data analyst or backend developer.

Prepared by: Anamika Saini

Course: B.Sc. IT – Final Year

Internship Period: July 2025

Project Title: COVID-19 SQL Analytics