




COVID-19 DATA ANALYTICS USING SQL



1. Introduction:

COVID-19, caused by the novel coronavirus, has drastically impacted global health and economics. Analyzing its spread and impact using data analytics can provide valuable insights for managing future outbreaks. This project aims to perform an end-to-end SQL-based analytical study of COVID-19 data from multiple countries across multiple days.

2. Abstract:

This project involves building a SQL-based data model to analyze key trends in COVID-19 cases. The dataset includes daily figures of confirmed, active, recovered, and death cases across countries. Using SQL queries, we extract summaries, trends, and rate calculations including rolling averages and growth percentages. The analysis helps identify country-wise performances, peak periods, and daily fluctuations

3. Tools Used:

- **SQL (MySQL / SQL Workbench)**: Main tool for data creation, manipulation, and querying.
- **Schema Design**: For structuring the COVID_DATA table.
- **Views and CTEs**: Used for creating reusable, complex aggregations.
- **Window Functions**: For advanced analytics like rolling averages and rankings.

4. Steps Involved:

- ❖ **Schema & Table Creation:** A schema covid_analysis was created with a main table covid_data having columns for date,
- ❖ country, confirmed, recovered, deaths, and active cases.
- Data Insertion:** Sample records for India, USA, and Brazil over three days were inserted to simulate real-world scenarios.
- ❖ **Analysis Performed:**
 - Country-wise totals for confirmed, recovered, and death cases.
 - Country-wise totals for confirmed, recovered, and death cases.
 - 3-day rolling average of daily new cases using window functions.
 - Calculation of recovery and death rates.
 - Country growth rate analysis.
 - Views for daily and country-wise summaries.
 - Views for daily and country-wise summaries.
 - Identification of top country per day by cases using ranking.

5. Conclusion:

This project demonstrates the effective use of SQL for pandemic data analytics. With simple yet powerful SQL operations, we derived valuable insights from raw data. The logical separation of queries and the creation of views ensured reusability and better data management. These insights can aid in decision-making and contribute to proactive health policy planning.