Project Report

COVID-19 Data Analytics

Abstract

The COVID-19 pandemic has significantly impacted the world, leading to the need for data-driven analysis to monitor and predict trends. This project focuses on analyzing COVID-19 statistics using SQL. A database was created to store country-wise daily records of confirmed, deaths, recovered, and active cases. Analytical SQL queries were applied to derive insights such as top affected countries, daily and weekly trends, and recovery/death rates.

Introduction

The project "COVID-19 Data Analytics" aims to analyze COVID-19 data using SQL queries. The dataset was imported into a relational database system and cleaned for null values and inconsistencies. By applying SQL features like aggregation, grouping, and window functions, meaningful insights were extracted.

Tools Used

- **Database**: MySQL
- **Interface**: MySQL Workbench
- **SQL Features**: Aggregation, GROUP BY, Window Functions, Views.

Steps Involved

1. **Dataset Import** – Kaggle dataset loaded into the database.
2. **Database Creation** – covid19_analytics created.
3. **Table Design** – covid_data for raw data, country_summary for summarized data.
4. **Data Cleaning** – Handled null values and formatted date.
5. **Data Insertion** – Inserted sample COVID-19 records.
6. **Queries Implemented**:
    - Top 10 affected countries
    - Country-wise recovery and death rates
    - Daily and weekly COVID-19 trends

    - New confirmed cases using window functions
7. **Views Creation** – Created weekly_summary for weekly analysis.
8. **Summary Updates** – Used ON DUPLICATE KEY UPDATE to refresh country summary automatically.

Conclusion

The project demonstrates how SQL can be effectively used for real-world data analytics. By leveraging queries and views, the system provides insights into the global COVID-19 trends, country-wise comparisons, and recovery rates. This project highlights the importance of database systems in health data monitoring and decision-making.