COVID-19 Analytics Project Report

Introduction

The COVID-19 pandemic presented one of the most significant global challenges of recent times, affecting millions of lives and economies. Data analytics plays a crucial role in understanding the spread, impact, and management of the virus. This project focuses on analyzing COVID-19 data using SQL to extract meaningful insights such as total cases, deaths, recovery rates, and country-wise comparisons.

Abstract

This project analyzes COVID-19 data using SQL queries to summarize and visualize key metrics. It processes real-world datasets containing country-wise daily records of COVID-19 cases, deaths, recoveries, and population statistics. By executing aggregation, filtering, window functions, and formatting techniques, the project delivers crucial analytical insights like total cases, death rates, and country-specific comparisons, which are essential for public health assessment and response strategies.

Tools Used

Database: SQLite / MySQL

Language: SQL

IDE: DB Browser for SQLite / MySQL Workbench

Visualization (Optional): Power BI / Tableau / Python Matplotlib

Dataset: Public COVID-19 dataset (daily reported cases, deaths, recoveries)

Steps Involved in Building the Project

Dataset Preparation:

Imported COVID-19 dataset containing fields like country, cases, deaths, recovered, population, day, and time.

Database Creation:

Created a table named covid_19 with necessary columns and populated it with dataset entries.

Data Cleaning:

- Checked for NULL values using SQL queries.
- Replaced missing data or removed corrupted rows.
- Ensured consistency in country names and date formats.

Analytics Using SQL Queries:
- Total and country-wise case counts.
- Top countries by deaths and recovery rates.
- Death rate and recovery rate calculations.
- Cases per 100,000 population.
- Latest reporting dates per country.
- Daily global summaries.
- Data validation using NULL checks.

Insights Presentation:
Results were structured as summary tables using SQL's GROUP BY, ORDER BY, and window functions.

Optional Visualization:
Analytical results were visualized using external tools like Power BI/Tableau to enhance interpretation.

Conclusion

The COVID-19 Analytics Project effectively demonstrates the power of SQL in real-world data analysis. Using fundamental and advanced SQL techniques, critical pandemic insights were derived directly from raw data. This project underscores the significance of data management and analysis in supporting public health decisions, highlighting SQL as a valuable tool for handling and interpreting large datasets.