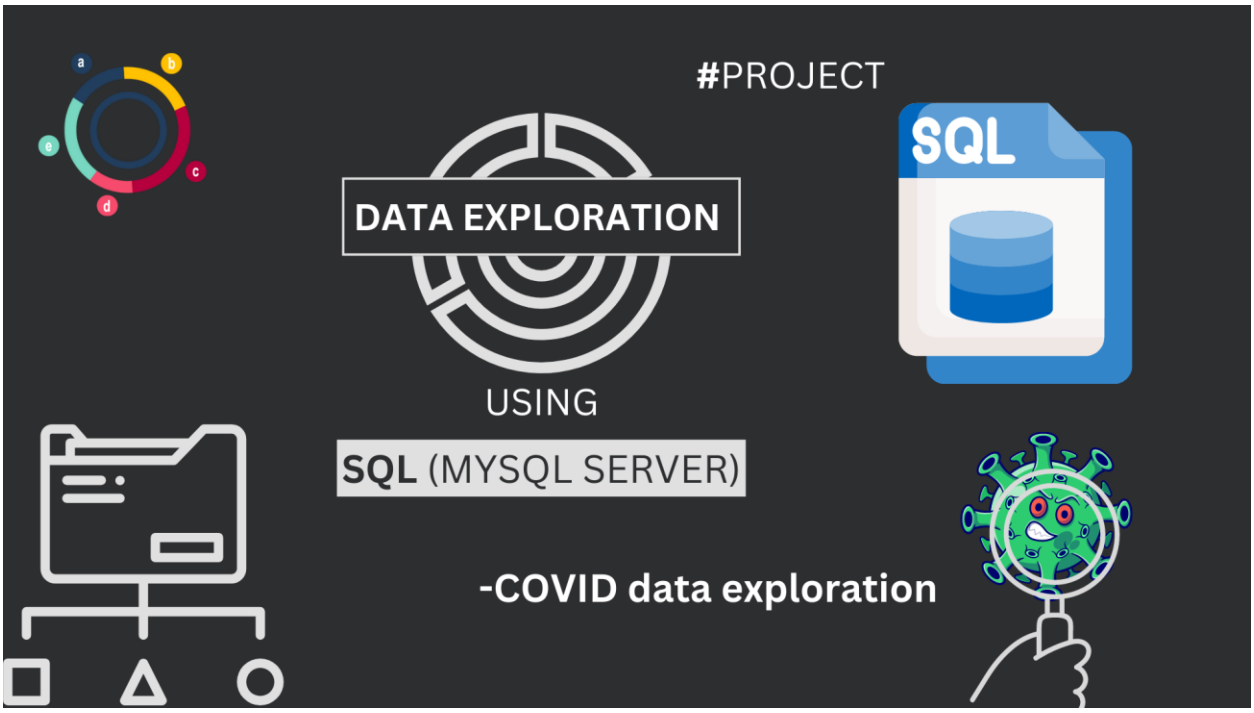
COVID Data Exploration: SQL in MySQL Analysis and Insights



Description:

In this project, I used SQL in MySQL to analyze COVID-19 data. I explored overall trends, calculated cases, deaths, and recoveries, and examined geographical distribution. I also analyzed case progression over time, identified growth rates, and investigated factors affecting outcomes. This project improved my SQL skills and provided valuable insights into the COVID-19 pandemic.

Skills: Data Exploration, SQL.

🎯Objective:

🌟The objective of this project was to explore and analyze COVID-19 data using SQL in SQL server. The project aimed to gain insights into various aspects of the COVID-19 pandemic, such as cases, deaths, vaccinations, and population percentages.

🎯Data gathering & cleaning:

🌟The COVID-19 data was downloaded as a CSV file from the website "https://ourworldindata.org/covid-deaths." The dataset was imported into the SQL server and divided into two tables: CovidDeaths and CovidVaccinations. The data cleaning process involved ensuring data accuracy, resolving any inconsistencies or errors, and preparing the dataset for analysis.

🎯Data Processing:

🌟SQL queries were utilized to process and analyze the COVID-19 data. The queries performed calculations, aggregations, and comparisons to derive meaningful insights from the dataset. Temporary tables and views were created to store intermediate results and facilitate further analysis.

🎯Questions asked:

1. Total Cases Vs Total Deaths: Examined the relationship between total COVID-19 cases and total deaths.
2. Likelihood of dying if you contract COVID in your Country: Calculated the percentage likelihood of death if a person contracts COVID-19 in a specific country.
3. Total Cases Vs Populations: Explored the percentage of the population affected by COVID-19 in different locations.
4. What percentage of the population got COVID?: Determined the percentage of the population that contracted COVID-19.
5. Countries with the highest infection rate compared to the population: Identified countries with the highest infection rates relative to their population.
6. Countries with the highest death count per population: Identified countries with the highest death count per population.
7. Break things down by Continent: Analyzed COVID-19 data by continent.
8. Continents with the highest death count per population: Identified continents with the highest death count per population.
9. Total death percentage: Calculated the total death percentage based on COVID-19 data.
10. Total population Vs Vaccinations: Explored the relationship between the total population and the number of COVID-19 vaccinations administered.

🎯Key Highlights:

- Imported the COVID-19 data into the SQL server and divided it into separate tables for deaths and vaccinations.

- Utilized SQL queries to explore and analyze the data, addressing questions such as total cases vs total deaths, the likelihood of dying from COVID-19 in different countries, total cases vs populations, percentage of the population affected by COVID-19, countries with the highest infection rates, and death counts, breakdown by continent, and global statistics.
- Created temporary tables and views to store and retrieve data for visualization and further analysis.
- Overcame challenges and errors encountered during the project, demonstrating problem-solving skills and a willingness to learn.

🎯Conclusion:

👉Through the utilization of SQL in the SQL server, this project successfully explored and analyzed COVID-19 data. The objectives were met by gathering and cleaning the data, processing it using SQL queries, and answering key questions related to COVID-19 cases, deaths, vaccinations, and population percentages. The project highlighted the significance of data exploration and demonstrated the effectiveness of SQL for extracting valuable insights from large datasets.