**Project: Healthcare Data Analysis Using SQL**

**Ramakrishnan CL**

Introduction:

The aim of this project is to extract valuable insights and trends from a sample healthcare dataset using SQL. By leveraging SQL queries and analytical skills, we explore various aspects of patient records, medical conditions, treatments, and billing information. At the end of this project, we seek to identify meaningful insights that improve healthcare decision-making and patient outcomes.

Importance of the project:

The project holds significant importance as it allows us to gain valuable insights from healthcare data, which can ultimately contribute to improving patient care and healthcare outcomes. By analysing patient records, medical conditions, treatments, and billing information, we can identify trends, patterns, and correlations that may help healthcare providers make informed decisions. These insights can aid in optimizing resource allocation, enhancing operational efficiency, and tailoring treatment plans to individual patient needs.

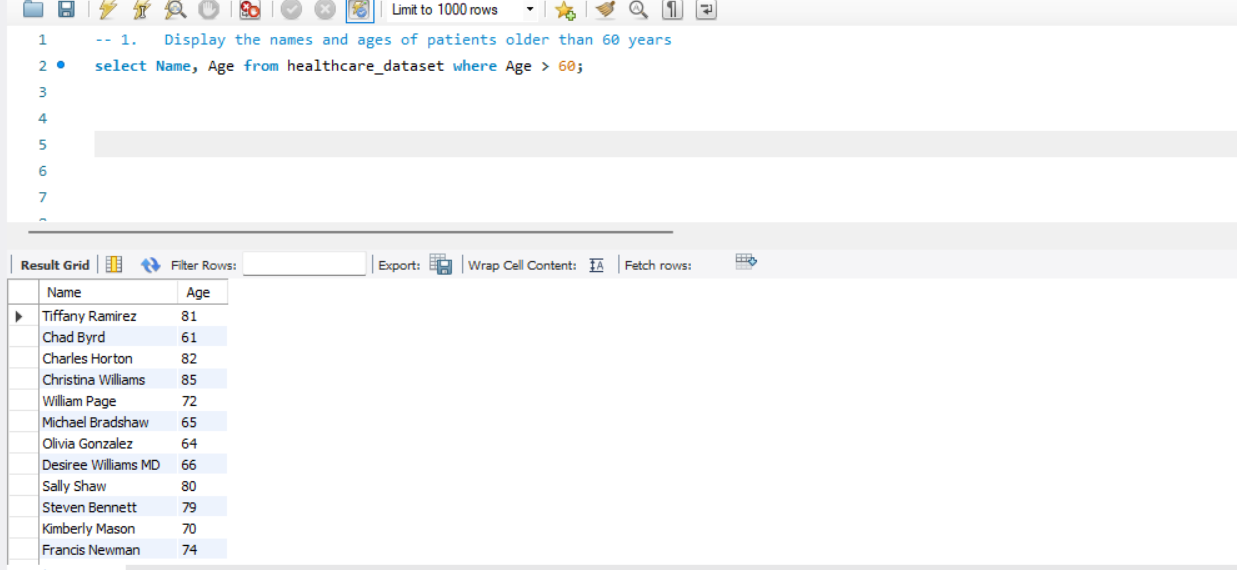
Steps involved:

* The healthcare dataset was obtained from Kaggle and underwent a thorough cleaning process to ensure data quality and consistency.
* The dataset consists of around 10,000 rows with important columns such as name, age and gender of patients along with medical condition, billing amount, test results, etc.
* The cleaned and organized dataset was then loaded into SQL Workbench.
* Executed a series of SQL queries to extract relevant information and insights from the healthcare data.
* Obtained key findings and insights aimed at improving patient care and enhancing the operational efficiency of healthcare organizations.

SQL Queries and Results:

1. **Display the names and ages of patients older than 60 years**

**Query:**

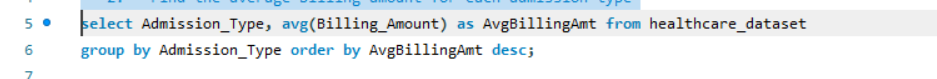
****

**Result:**

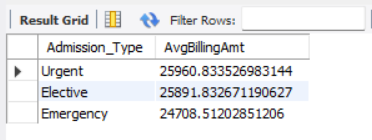
****

**2. Find the average billing amount for each admission type**

**Query:**

****

**Result:**

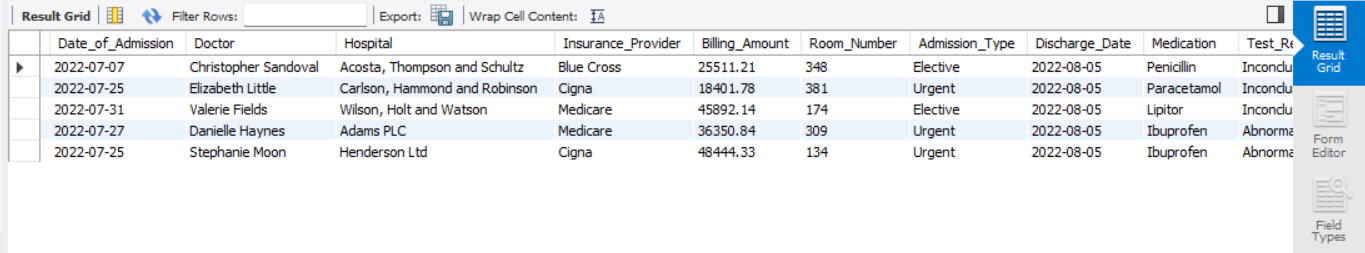
****

**3. Display the details of patients discharged on a particular date, say 05-08-2022**

**Query:**

****

**Result:**

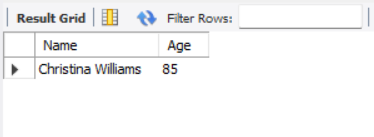
****

**4. Display the name and age of the oldest patient**

**Query:**

****

**Result:**

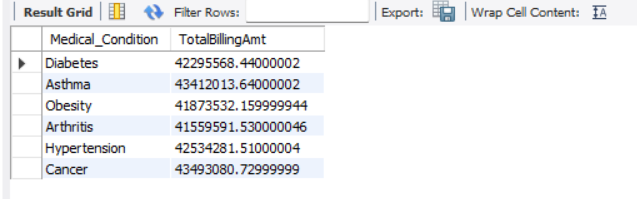
****

**5.Retrieve the total billing amount for each medical condition**

**Query:**

****

**Result:**

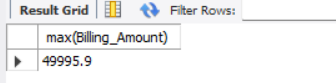
****

**6. Display the highest billing amount.**

**Query:**

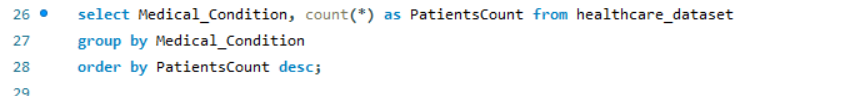
****

**Result:**

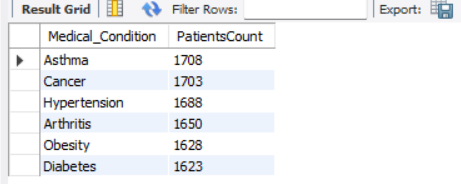
****

**7. Find the number of patients for each medical condition**

**Query:**

****

**Result:**

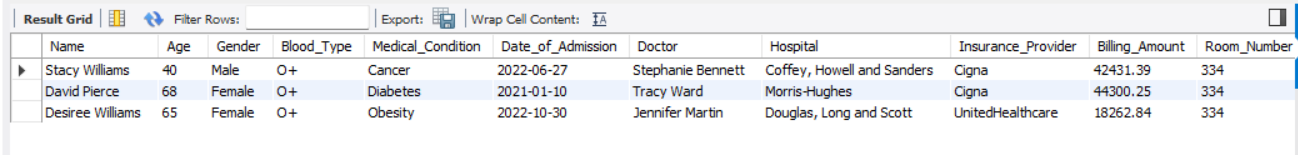
****

**8. Display the details of patients in room number 334 with blood type O+**

**Query:**

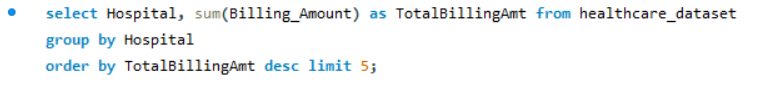
****

**Result:**

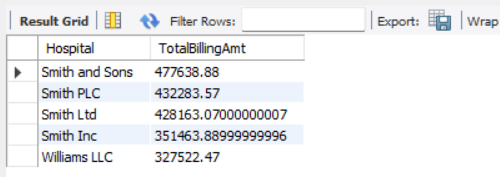
****

**9. Display the top 5 hospitals with highest total billing amount.**

**Query:**

****

**Result:**

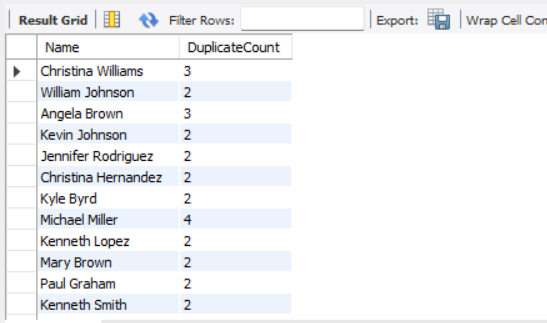
****

**10. Find patients with duplicate names**

**Query:**

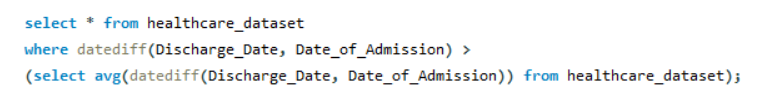
****

**Result:**

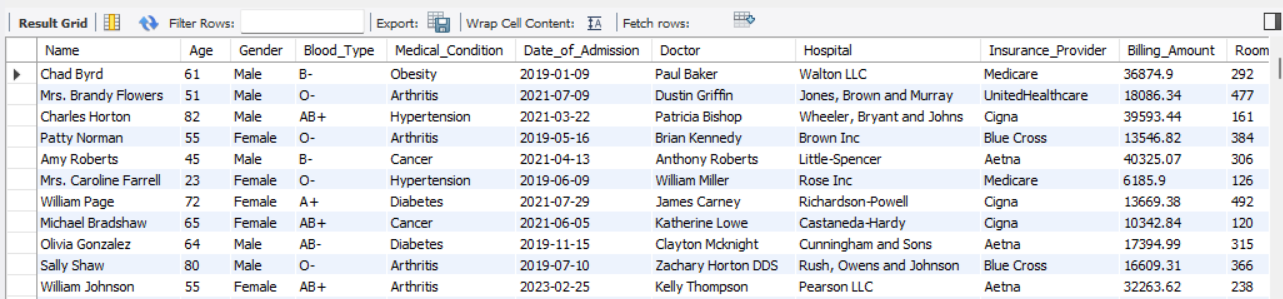
****

**11. Find patients who stayed in hospitals longer than average.**

**Query:**

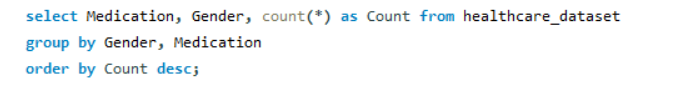
****

**Result:**

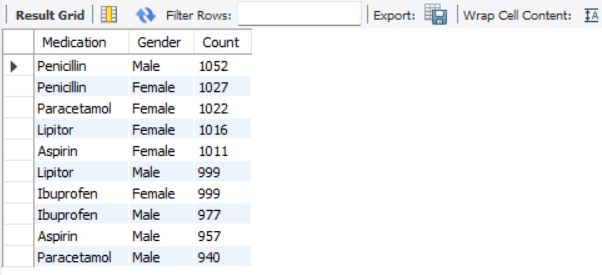
****

**12. Analyse medication usage based on gender.**

**Query:**

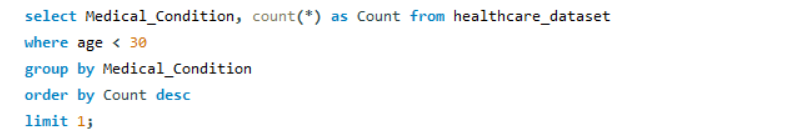
****

**Result:**

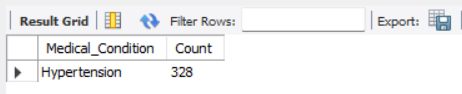
****

**13. Find the most common medical condition among patients below 30 years old.**

**Query:**

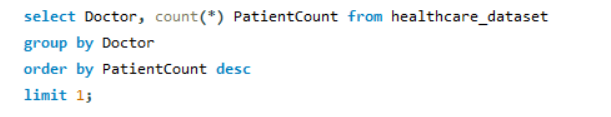
****

**Result:**

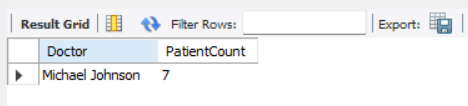
****

**14. Find the doctor having the most number of patients.**

**Query:**

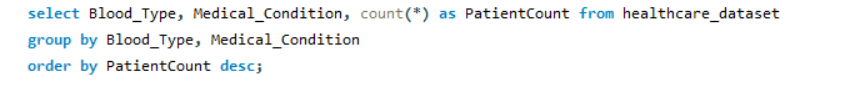
****

**Result:**

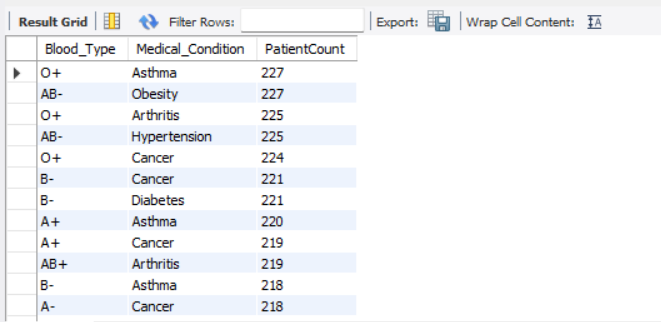
****

**15. Identify patients with same blood type and medical condition.**

**Query:**

****

**Result:**

****

Key findings:

* There were 5025 female patients and 4975 male patients out of which 1000 patients were above 60 years old.
* The oldest patient is 85 years old and the youngest patient is 18 years old.
* With 1708 patients, Asthma is the most common medical condition followed by cancer with 1703 patients.
* Diabetes is the least common medical condition with 1623 patients.
* Hypertension is the most common medical condition of patients with less than 30 years old.
* Average billing amount is highest for Diabetes and lowest for Arthritis.
* Highest billing amount recorded was $ 49995.9.
* Smith and Sons hospital generated the highest revenue with $ 477638.88.
* 1000 patients stayed in the hospitals for more days than average.
* There were 522 patients with duplicate names.
* Patients with AB- blood type were the highest in number, totalling 1275 individuals and A- the lowest with 1238 patients.

Conclusion:

Through comprehensive data analysis, valuable insights were obtained from the healthcare dataset. These findings provide valuable information for healthcare providers and stakeholders to improve patient care, optimize resource allocation, and enhance healthcare delivery.

By leveraging data-driven approaches, healthcare organizations can make informed decisions, improve customer service, and address emerging healthcare challenges effectively. This project underscores the importance of data analysis in healthcare decision-making and highlights the potential for data-driven strategies to drive positive outcomes in patient care and healthcare management.