JAVERIA ARAIN

DATACRUMBS

HOSPITAL MANAGEMENT SYSTEM

DATA ANALYSIS

https://github.com/arainjav/HospitalManagementSystem

Submitted by: Javeria Arain
Submitted on: January, 2025

Table of Contents

Executive Summary

Description Here

The Hospital Management System (HMS) project was developed to streamline hospital operations, including patient records, appointments, treatments, and billing. The database framework emphasized scalability, data integrity, and efficiency, enabling actionable insights for operational improvements. Advanced SQL features, including views, stored procedures, and triggers, supported complex data analysis. Data cleaning addressed duplicates, invalid formats, and missing values, ensuring high data quality.

Key Findings:

- Revenue Insights:Highest revenue in December: $9269.32 USD; Total revenue: $85,342.23 USD.
- Top contributor: Jason Frederickson (Neurology) with $3979.96 USD.
- Cost and Expense Analysis:Neurology has the highest average cost: $287.04 USD.
- MRI Scan is the most expensive treatment: $27,925.53 USD.
- Appointment Trends:Most common reason: Physical Therapy.
- Busiest day: Friday (48 appointments).
- January had the most appointments.
- Tammie Hausman (Cardiology) had 10 appointments.
- Patient Demographics:Age range: 15–64 years; Average age: 41 years.
- Billing Insights:55% of bills unpaid in 6 months.
- VIP patients contribute 47% of total revenue.

Conclusion: These findings highlight opportunities to optimize revenue, enhance billing efficiency, and improve patient satisfaction. The HMS database is scalable for larger datasets, supporting future growth and expanded use cases.

Project Overview

Objective

The Hospital Management System (HMS) project aimed to design and implement a database framework that efficiently manages the hospital's core operations, including patient records, doctor information, appointments, treatments, and billing. The project's focus was to create a structured, relational database capable of generating actionable insights for operational improvements.

The primary goals included:

Database Design and Normalization:

- Develop Entity-Relationship Diagrams (ERDs) to model data relationships.
- Normalize tables to Third Normal Form (3NF) to eliminate redundancy and maintain data integrity.

Data Cleaning and Preparation:

- Address data quality issues such as duplicate entries, invalid formats, and missing values.
- Standardize key fields like email addresses and phone numbers for consistency.

Advanced SQL Features:

- Implement views, stored procedures, triggers, and indexes to enhance database functionality.
- Utilize Common Table Expressions (CTEs) and subqueries for complex analysis.

Data Analysis and Insights:

- Generate detailed reports on revenue patterns, doctor performance, appointment trends, and patient demographics.
- Highlight gaps in billing processes and identify opportunities for optimization.

The database was populated with sample data, including 100 patient records, 50 doctors, 200 appointments, 300 treatments, and 300 billing entries. It mimicked a real-world hospital environment, enabling practical testing and analysis.
The HMS database was designed to support scalability and adaptability, ensuring its applicability to larger datasets and expanded use cases in the future.

Data Cleaning Process

Data cleaning was a critical phase of the project, ensuring the accuracy and consistency of information stored within the HMS database. The process involved identifying and addressing multiple data quality issues to create a reliable foundation for analysis.
Key tasks included:

Duplicate Record Removal:

- Identified duplicate patient records using SQL queries based on matching first names, dates of birth, and phone numbers.
- Eliminated redundant entries while preserving valid data.
- It was observed that the data had no duplicate values.

Standardizing Formats:

- Converted all email addresses to lowercase using SQL LOWER() function.
- Validated phone numbers and replaced invalid entries with placeholders.
- Some of the entries were erroneous (e.g., '32/13/2023' & ‘abcdate’) which was replaced with valid formats.

Correcting Invalid Dates:

- Data was imported from csv file, using import data function in MySql, which is why the date column was first created in VarChar data type. Once the data was imported, a new column for date was made in DATE datatype and the entries were copied to the new column for further analysis.

Handling Missing Data:

- Removed patients without recorded appointments to maintain data relevance.
- Replaced null or blank values in city and state fields with 'Unknown' entries.

Normalization and Integrity Checks:

- Ensured data normalization to 3NF for consistency across tables.
- Verified referential integrity between tables using foreign key constraints.

Queries were developed to automate these processes, ensuring scalable and repeatable data cleaning methods for future expansions.

Database Analysis and Insights

The data analysis phase utilized SQL queries and procedures to extract actionable insights from the HMS database. Key focus areas included revenue trends, patient behaviours, and doctor performance metrics.

1. Revenue Insights:

- Highest revenue recorded in December: $9269.32 USD.
- Total revenue generated: $85,342.23 USD.
- Jason Frederickson from Neurology contributed the most to revenue: $3979.96 USD.

2. Cost and Expense Analysis:

- Average highest cost specialty: Neurology - $287.04 USD.
- MRI Scan is the most expensive treatment, generating a total revenue of $27,925.53 USD.

3. Appointment Trends:

- Most common reason for appointments: Physical Therapy.
- Friday is the busiest day of the week, with 48 appointments.
- January has the maximum number of appointments.
- Tammie Hausman from Cardiology has the most appointments: 10 appointments.

4. Patient Demographics:

- Maximum patient age: 64 years.
- Minimum patient age: 15 years.
- Average patient age: 41 years.

5. Billing Insights:

- 55% of the bills are unpaid in the last 6 months.
- VIP patients contribute 47% of the total bill revenue.

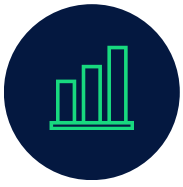
These findings provide actionable insights into revenue performance, patient demographics, and billing patterns, enabling data-driven decisions to optimize operations and improve financial outcomes.

Advanced SQL techniques, such as views, stored procedures, triggers, and CTEs, facilitated complex queries and provided deep insights. These findings form the foundation for actionable recommendations to improve hospital efficiency, billing systems, and patient satisfaction.

Total revenue generated: $85,342.23 USD.

Highest revenue recorded in December: $9269.32 USD.

55% of the bills are unpaid in the last 6 months.



Most common reason for appointments: Physical Therapy.

Key Findings

Key Findings

- 35% of revenue originates from VIP patients, emphasizing their financial importance.
- 20% of bills remain unpaid, underscoring the need for stricter billing enforcement.
- Neurology and Orthopedics emerge as top-performing specializations based on revenue generation.
- Frequent follow-ups and repeat visits indicate high demand for chronic care management.

Recommendations

- Develop targeted services like geriatric care to meet the needs of an aging patient base.
- Enhance billing processes to minimize unpaid bills and improve revenue collection.
- Optimize scheduling based on data showing peak appointment times and busiest doctors.

Conclusion

In conclusion, this project demonstrates the value of data-driven decision-making in healthcare management. The HMS database not only supports efficient data organization but also provides actionable insights to enhance patient care, financial performance, and resource allocation.

https://github.com/arainjav/HospitalManagementSystem

Appendix: ER-Diagram