■ Project 2: HealthTrack – Patient Readmission Analysis

Introduction

This project analyzes a healthcare dataset (100k+ hospital records) to uncover patterns in patient readmissions. Using SQL, we answered 5 key business questions about patient age, diagnosis, hospital type, and length of stay. The findings provide insights for hospital administrators to improve patient care and reduce readmissions.

Dataset

Source: UCI Machine Learning Repository – Diabetes 130-US hospitals for years 1999–2008 File Size: ~300MB Format: CSV, cleaned to remove missing values and irrelevant columns. Key Columns: age, gender, admission_type, discharge_disposition, readmitted, time_in_hospital, diagnosis.

Q1: Which age group has the highest patient readmission rate?

Result: Middle-aged and senior groups (50–80) showed the highest readmission counts.

Q2: What are the top diagnoses leading to readmission?

Result: Diabetes-related and circulatory diseases were among the leading causes.

Q3: Which hospitals or admission types have the highest readmission rates?

Result: Emergency admissions and certain hospitals recorded higher return rates.

Q4: How does length of stay in hospital affect readmission likelihood?

Result: Patients staying 3–7 days were most likely to return, compared to very short or very long stays.

Q5: Does gender influence readmission rates?

Result: Male and female patients had nearly equal readmission rates; gender had no major effect.

Conclusion

The analysis provides actionable insights for healthcare providers. Hospitals should focus on middle-aged to elderly patients, prioritize better follow-up care for diabetes and circulatory issues, and refine protocols for emergency admissions. These steps could significantly reduce costly readmissions.

Prepared by: SK