Project Report – Healthcare Analytics with SQL

# Executive Summary

This project involved analyzing healthcare data using advanced SQL techniques to derive insights into patient behavior, doctor performance, appointment trends, and medication patterns. The primary goal was to support healthcare decision-making through efficient data querying and analysis.

# Data Methodology

The project utilized a structured healthcare dataset comprising multiple tables: Patients, Doctors, Appointments, Diagnoses, and Medications. Relationships between tables were established using primary and foreign keys, supported by an ER diagram.  
  
Key SQL methodologies applied:  
- Joins: Inner, Left, Right, and Full Outer Joins to combine data across tables.  
- Window Functions: RANK and DENSE\_RANK for ranking doctor and patient activities.  
- Aggregate Functions: COUNT, SUM for patient and doctor performance metrics.  
- Subqueries: For complex filtering, such as medication-specific analysis.  
- Conditional Logic: CASE statements to segment patients by age group.  
- Date Functions: DATEDIFF to calculate medication durations.

# Key Insights and Findings

## 1. Patient Analysis

- Identified patients who never had appointments using LEFT JOINs with NULL handling.  
- Segmented patients into age groups (18–30, 31–50, 51+) using CASE logic.  
- Retrieved patients with contact numbers ending in ‘1234’, applying string functions.

## 2. Doctor Performance

- Evaluated doctors based on number of diagnoses, including those with zero diagnoses using RIGHT JOIN + COUNT.  
- Ranked doctors based on the number of unique patients attended using COUNT(DISTINCT).  
- Used window functions to rank patients for each doctor by appointment frequency.

## 3. Appointment Trends

- Retrieved completed appointments with patient, doctor, and specialization details using INNER JOINs.  
- Detected mismatches between appointments and diagnoses using FULL JOIN to ensure data consistency.

## 4. Treatment & Medication Insights

- Identified top-prescribed medications by diagnosis.  
- Used subqueries to isolate patients only prescribed ‘Insulin’.  
- Calculated average medication durations with DATEDIFF to evaluate treatment length.

# Key Performance Analysis for Stakeholders

|  |  |
| --- | --- |
| Area | KPI |
| Patient Management | Count of inactive patients, Age-group segmentation |
| Doctor Performance | Total diagnoses per doctor, Top doctors by patient volume |
| Appointment Efficiency | Completion rates, Scheduling gaps |
| Treatment Analysis | Most prescribed medication, Avg. treatment duration |

# Recommendations

- Patient Engagement: Focus outreach on patients with no appointment history.  
- Doctor Load Balancing: Distribute patients more evenly among high-performing and underutilized doctors.  
- Medication Tracking: Monitor high-frequency prescriptions (e.g., Insulin) for better inventory and diagnosis mapping.  
- Data Quality Checks: Regularly reconcile appointments and diagnoses to eliminate mismatches.