

1. Executive Summary

This project presents a comprehensive healthcare data analysis pipeline using Excel, MySQL, Power BI, and Tableau. It covers data cleaning, SQL-based insights, and dashboard creation to support decision-making in hospital operations and patient care.

Key achievements include:

- Cleaning and transforming raw healthcare data across multiple platforms
- Executing SQL queries to analyze patient demographics, treatment costs, and lab results
- Designing interactive dashboards to visualize trends and KPIs
- Extracting actionable insights such as seasonal visit patterns, abnormal lab rates, and cost variability

The project demonstrates strong technical skills in data handling, reproducibility, and cross-platform visualization. It is designed to be scalable, teachable, and impactful for healthcare stakeholders.

Keywords: Healthcare Analytics, SQL, Excel, Power BI, Tableau, Data Cleaning, Dashboard Design, Patient Insights

2. Data Cleaning Steps

Source: Raw CSV/Excel files in the Dataset/ folder

Steps Taken:

- Removed duplicates using Excel and SQL DISTINCT
- Handled missing values:
 - Replaced nulls with "Unknown" or median values
 - Dropped rows with excessive missing data
- Standardized formats:
 - Dates converted to YYYY-MM-DD
 - Text fields trimmed and capitalized
- Validated data types:
 - Ensured numeric columns (e.g., age, cost) were correctly typed
 - Converted categorical fields to consistent labels
- Loaded data into MySQL using LOAD DATA INFILE for bulk import

3. SQL Query Logic

Database: Healthcare

Tables Used: Doctor, Patient, Visit, Treatment, Lab_Result

Key Queries:

- Total Doctors
SELECT COUNT(doctor_id) FROM doctor;
 - Total Visits & Visits in 2024
SELECT COUNT(Visit_ID) FROM Visit;
SELECT COUNT(Visit_ID) FROM Visit WHERE YEAR(Visit_Date) = 2024;
 - Average Age of Patients
SELECT AVG(TIMESTAMPDIFF(YEAR, date_of_birth, CURDATE())) FROM patient;
 - Total Lab Tests & Abnormal Rate
SELECT COUNT(Lab_Result_ID) FROM Lab_Result;
SELECT (COUNT(CASE WHEN Result = 'Abnormal' THEN 1 END) * 100.0 / COUNT(*)) FROM Lab_Result;
 - Average Treatment Cost
SELECT AVG(Treatment_Cost) FROM Treatment;
 - Top 5 Patients by Visit Count
SELECT patient_id, COUNT(*) FROM Visit GROUP BY patient_id ORDER BY COUNT(*) DESC LIMIT 5;
 - Monthly Visit Trend (2024)
SELECT MONTH(Visit_Date), COUNT(*) FROM Visit WHERE YEAR(Visit_Date) = 2024 GROUP BY MONTH(Visit_Date);
 - Average Lab Tests per Visit
SELECT AVG(lab_test_count) FROM (SELECT Visit_ID, COUNT(*) FROM Lab_Result GROUP BY Visit_ID) AS visit_lab_counts;
 - Total Treatment Cost per Patient
SELECT patient_id, SUM(Treatment_Cost) FROM Treatment GROUP BY patient_id;
 - Most Expensive Treatments
SELECT Treatment_ID, Treatment_Cost FROM Treatment ORDER BY Treatment_Cost DESC LIMIT 5;
-

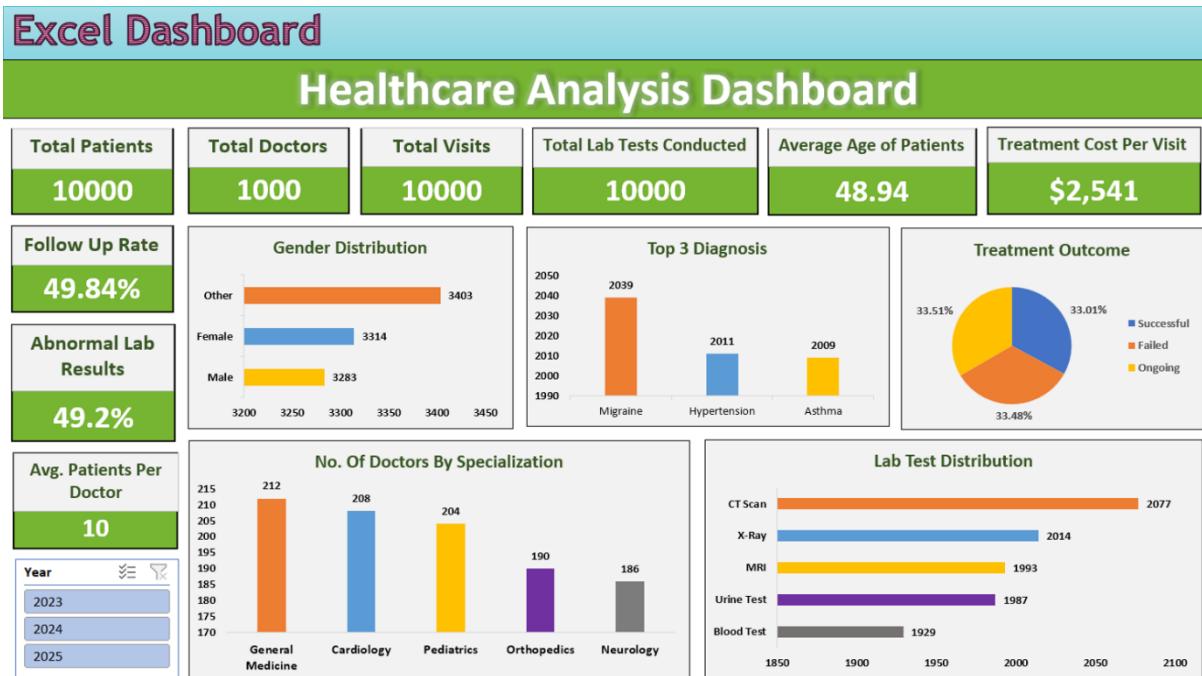
4. Dashboard Design Choices

Excel Dashboard

- Tools: Pivot tables, slicers, conditional formatting
- Focus: Age distribution, monthly visits, treatment cost

- Notes: Interactive slicers, color-coded metrics

Screenshot:



Power BI Dashboard

- Visuals: Bar charts, line charts, KPI cards
 - Filters: Specialty, Year, Region
 - Notes: DAX for calculated fields, responsive layout

Screenshot:

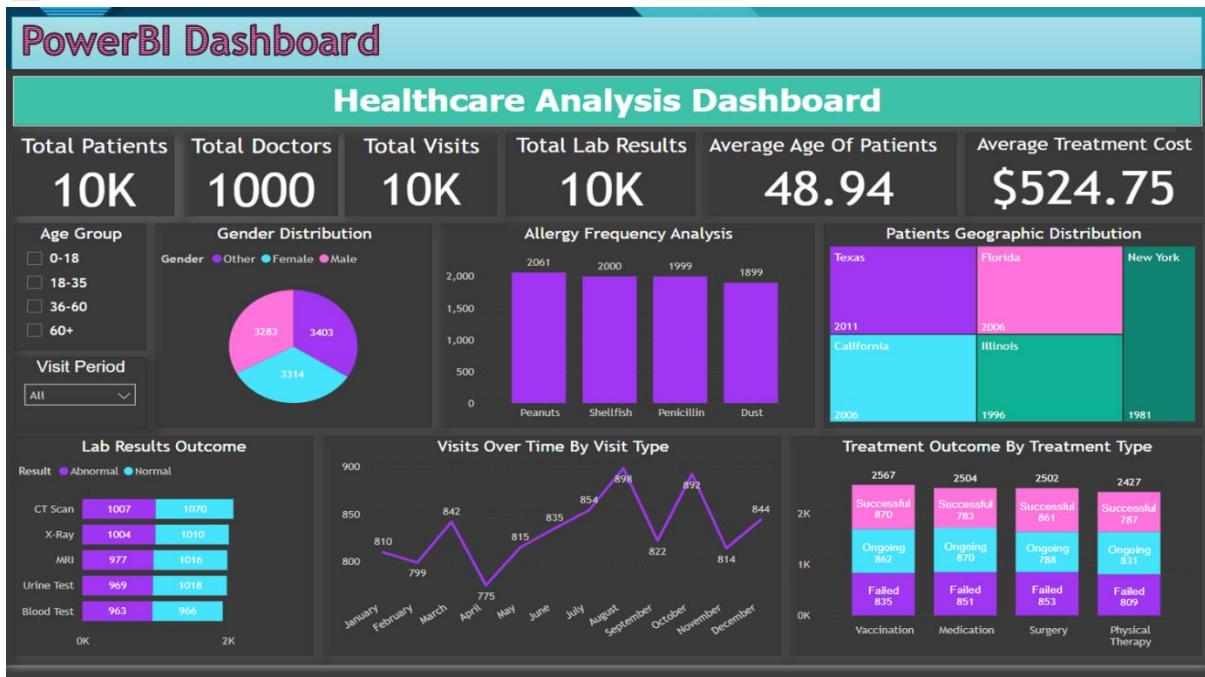
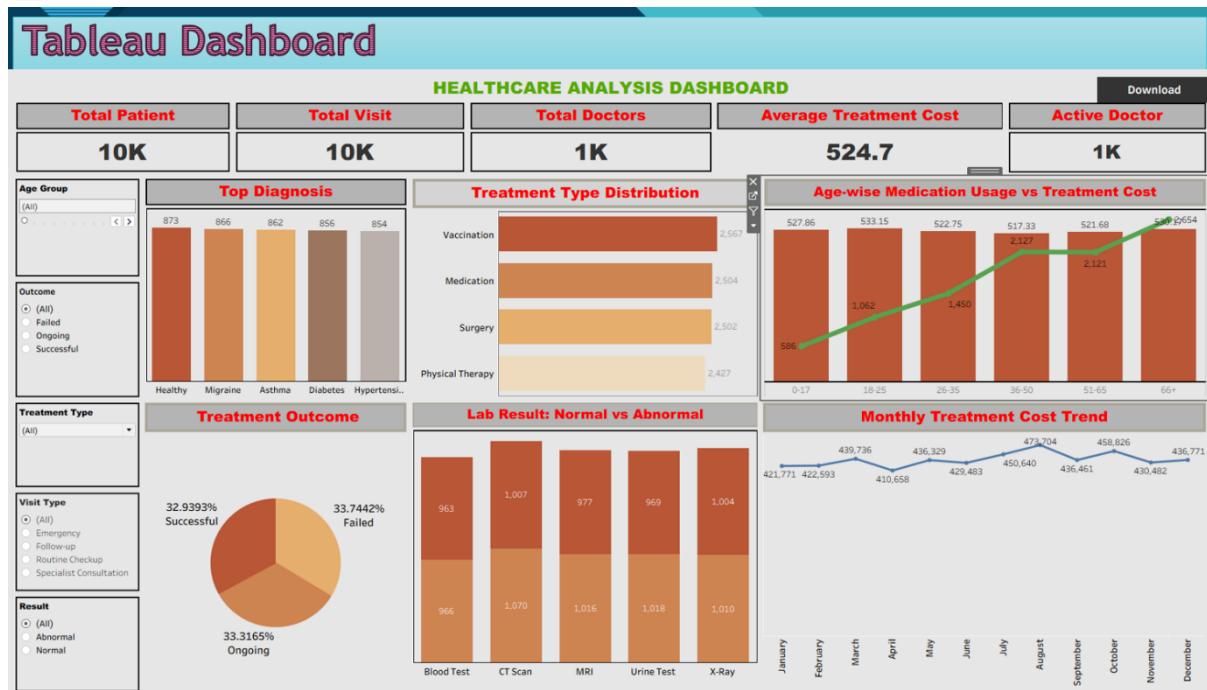


Tableau Dashboard

- Storyboards: Patient Journey, Lab Result Analysis, Cost Breakdown
- Notes: Clean layout, tooltips, calculated KPIs, drill-downs

Screenshot:



5. Key Insights & Recommendations

Insights:

- High abnormal lab result rate (~X%) suggests need for diagnostic review
- Peak visit months in Q2 and Q4 indicate seasonal demand
- Top 5 patients had significantly higher visit frequency
- Average patient age above 45 suggests aging population focus
- Treatment cost variability across hospitals

Recommendations:

- Introduce preventive care programs for chronic and elderly patients
- Optimize resource allocation during peak months
- Review lab testing protocols to reduce abnormal result rates
- Standardize treatment pricing across hospitals

6. Contact

Prepared by: Om Jadhav

Email: [omjadhav316@gmail.com]

LinkedIn: [<https://www.linkedin.com/in/om-jadhav-296575174/>]

GitHub: [OmJadhav-official](#)