Axon Healthcare Analysis: Presentation

Presented By Group No. 1

Team Members:

1. **Bairla Yaswanth**
2. **Vikrant Hirave**
3. **Vadyarapu Prashanth Kumar**
4. **Banashankari Narayan Bandurgi**
5. **Uma Babaleshwar**
6. **Soumya Keshri**
7. **Anil Kumar Reddy**

Healthcare Project Summary

This project delivered an end-to-end healthcare data analysis solution, encompassing data cleaning, processing, and integration into a **SQL database**. Interactive dashboards were then developed using **Excel**, **Power BI**, and **Tableau** to visualize key healthcare trends, performance metrics, and actionable insights.

The dashboards provide detailed insights across critical healthcare parameters, including patient and diagnosis counts, patient demographics, treatment costs, doctor workload, and patient follow-up rates. Financial factors such as hospital revenue, visit costs, and comprehensive treatment cost analysis are also included.

These insights empower healthcare providers to minimize resource and medical equipment wastage by optimizing allocation, analyze patient demographics and health factors, and ultimately maximize hospital efficiency.

Healthcare data Analysis KPI's

The following are some of the key KPIs included in the dashboard

- **Total Patients**
- **Total Doctors**
- **Total Visits**
- **Average Age of Patients**
- **Top 5 Diagnosed Conditions**
- **Follow-Up Rate**
- **Treatment Cost Per Visit (Avg.)**
- **Total Lab Tests Conducted**
- **Percentage of Abnormal Lab Results**
- **Doctor Workload (Average Patients Per Doctor)**
- **Total Revenue - Sum(Treatment Cost) + Sum(Visit Charges)**

Tools used for Data Storage, Visualization & analysis

MS Excel

Power BI

Tableau

MySQL

Excel Dashboard

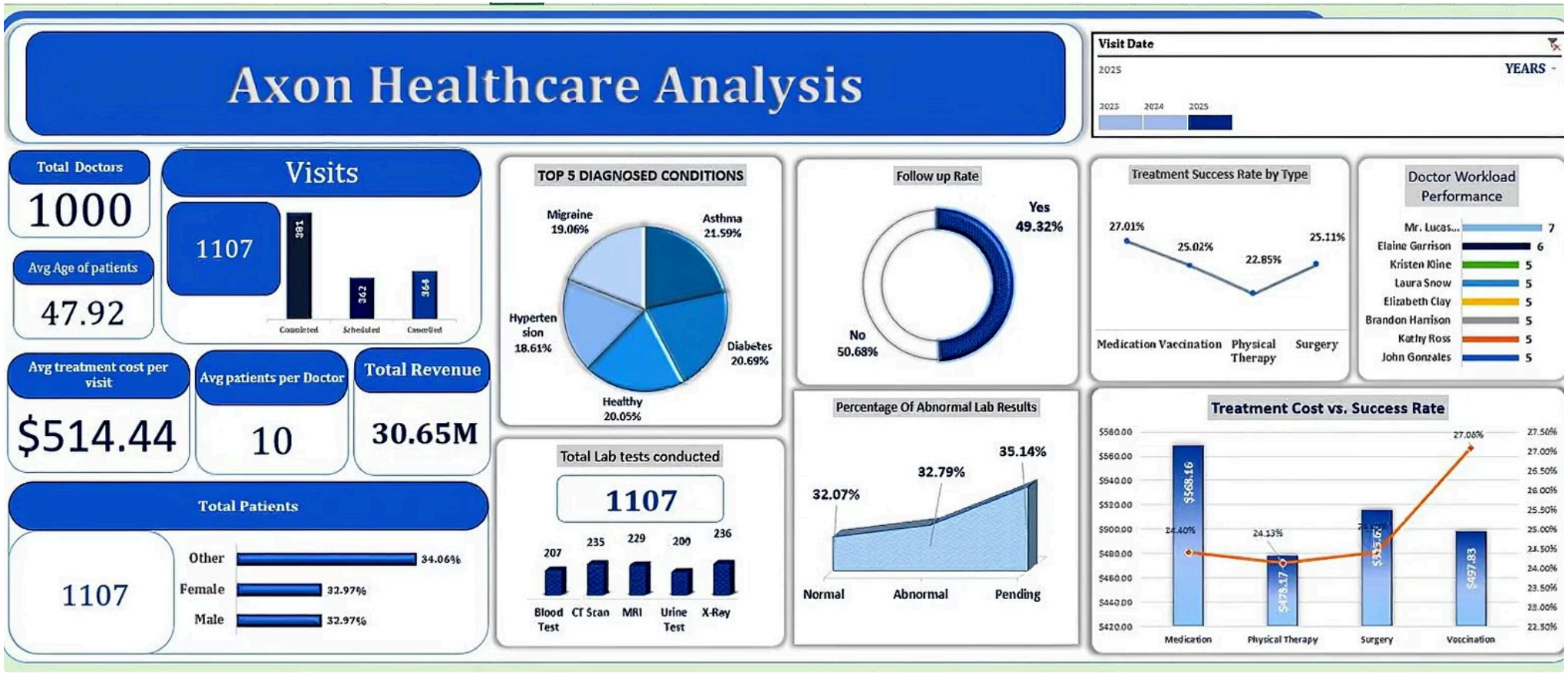
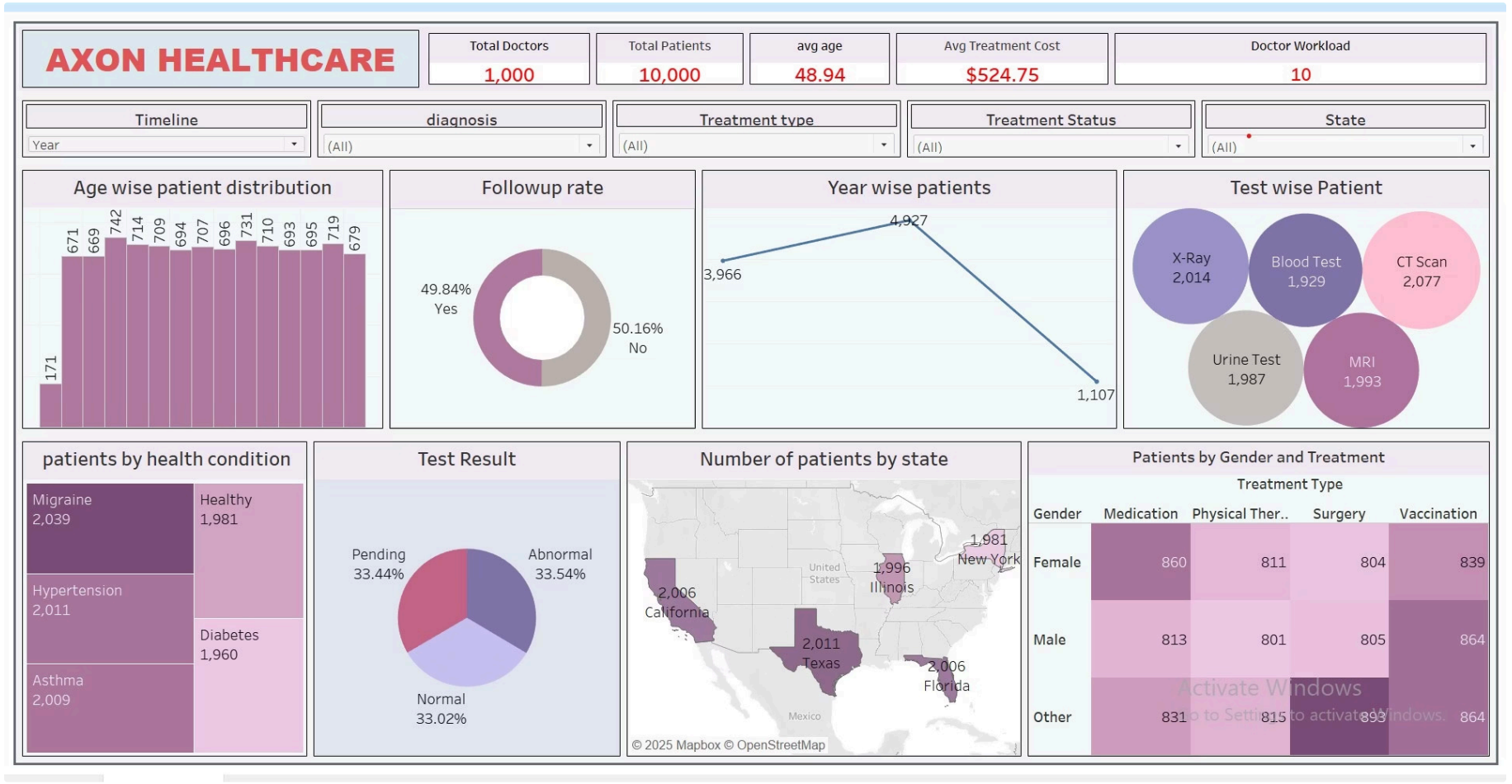
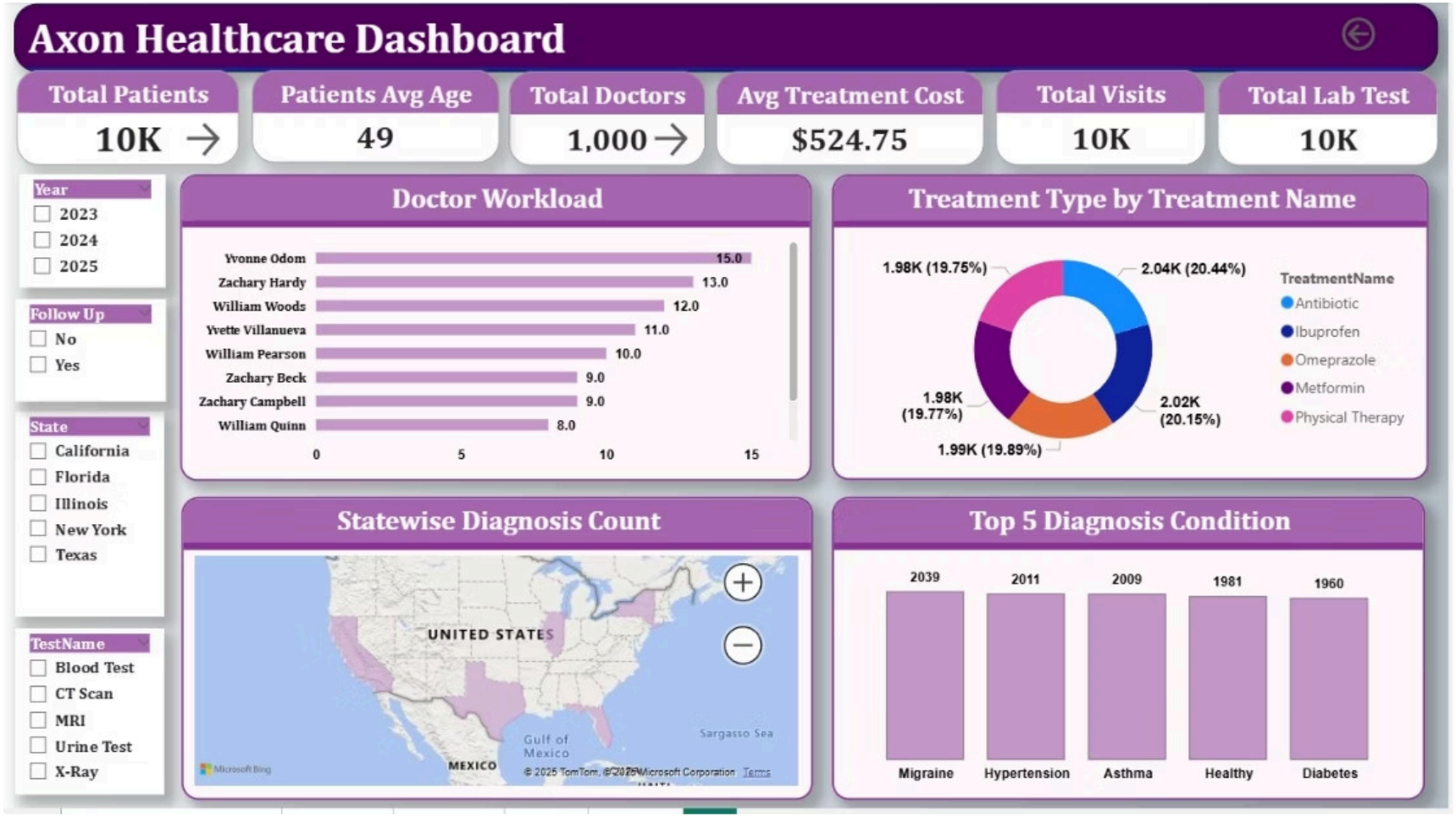


Tableau Dashboard



Power BI Dashboard



MySQL Healthcare Model

The following query was used to build the complete healthcare database model encompassing all the provided data

```
# -------------- HEALTHCARE DATA MODEL -------------------------

SELECT *
FROM patients AS p
JOIN visits AS v ON p.PatientID = v.PatientID
JOIN doctor AS d ON v.DoctorID = d.DoctorID
LEFT JOIN treatments AS t ON v.VisitID = t.VisitID
LEFT JOIN labresults AS l on v.VisitID = l.VisitID
ORDER BY v.VisitDate DESC;
```

MySQL Queries

SQL Queries and Output (KPI 1, 2)

```sql
-- KPI -1  Total Patients Count
SELECT COUNT(*) AS TotalPatients
FROM patients;
--- KPI -2 Average Patient Age
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

TotalPatients
9999

```sql
--- KPI -2 Average Patient Age
SELECT AVG(Age) AS AveragePatientAge
FROM patients;
--- KPI -3 -Gender Distribution
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

AveragePatientAge
48.9367

MySQL Queries

SQL Queries and Output (KPI 3,4)

```
25    --- KPI -3 -Gender Distribution
26    SELECT Gender, COUNT(*) AS PatientCount
27    FROM patients
28    GROUP BY Gender;
29    DESC patients;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Gender	PatientCount
Male	3282
Female	3314
Other	3403

```
31    -- KPI - 4 Most Common Blood Type
32    SELECT `Blood Type`, COUNT(*) AS Count
33    FROM patients
34    GROUP BY `Blood Type`
35    ORDER BY Count DESC
36    LIMIT 1;
37     -- KPI 5 Insurance Provider Coverage
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Blood Type	Count
O+	1311

MySQL Queries

SQL Queries and Output (5, 6)

```
37    -- KPI 5 Insurance Provider Coverage
38    SELECT `Insurance Provider`, COUNT(*) AS PatientCount
39    FROM patients
40    GROUP BY `Insurance Provider`
41    ORDER BY PatientCount DESC;
42    -- KPI 6--Patients per State
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

Insurance Provider	PatientCount
Smith Ltd	18
Smith LLC	18
Smith and Sons	17
Smith PLC	14
Davis and Sons	12
Johnson and Sons	12
Smith Inc	12
Brown and Sons	12
Smith Group	9

```
42    -- KPI 6--Patients per State
43    SELECT State, COUNT(*) AS PatientCount
44    FROM patients
45    GROUP BY State
46    ORDER BY PatientCount DESC;
47    -- KPI 7 Total Visits per Month
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

State	PatientCount
Texas	2011
California	2006
Florida	2006
Illinois	1996
New York	1980

MySQL Queries

SQL Queries and Output (7,8)

```sql
-- KPI 7 Total Visits per Month
SELECT DATE_FORMAT(STR_TO_DATE(`Visit Date`, '%Y-%m-%d'), '%Y-%m') AS Month,
       COUNT(*) AS VisitCount
FROM visit
GROUP BY Month
ORDER BY Month;
-- KPI 8 Patients with Most Visits
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Month	VisitCount
2023-03	136
2023-04	403
2023-05	404
2023-06	413
2023-07	455
2023-08	464
2023-09	415
2023-10	436
2023-11	404
2023-12	436

Result 16

```sql
-- KPI 8 Patients with Most Visits
SELECT `Patient ID`, COUNT(*) AS VisitCount
FROM visit
GROUP BY `Patient ID`
ORDER BY VisitCount DESC
LIMIT 5;
--  KPI 9 -Most Common Visit Reasons
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Patient ID	VisitCount
236037	1
299316	1
466120	1
123920	1
860284	1

MySQL Queries

SQL Queries and Output (9, 10)

```sql
-- KPI 9 -Most Common Visit Reasons
SELECT `Reason for Visit`, COUNT(*) AS Count
FROM visit
GROUP BY `Reason for Visit`
ORDER BY Count DESC
LIMIT 5;
-- KPI 10 Treatment-Based KPIs
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Reason for Visit	Count
Flu Symptoms	2097
Emergency	2016
Follow-up	1991
Chronic Pain	1965
Routine Checkup	1931

```sql
-- KPI 10 Treatment-Based KPIs
SELECT `Treatment Name`, COUNT(*) AS Count
FROM `treatment (1)`
GROUP BY `Treatment Name`
ORDER BY Count DESC
LIMIT 5;
-- Average Treatment Cost
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Treatment Name	Count
Antibiotic	2044
Ibuprofen	2015
Omeprazole	1989
Metformin	1977
Physical Therapy	1975

MySQL Quality Analysis Queries

The following three SQL queries were used for qualitative analysis, including data count validation, completeness checks, and consistency checks. All QA parameters were successfully validated

```sql
# 1. Data Count Validation

Select count(*) from doctor;
Select count(*) from patients;
Select count(*) from visits;
Select count(*) from labresults;
Select count(*) from treatments;

# 2. Data Completeness Check
# Identify MISSING or NULL values in key columns
Select * from doctor where DoctorID IS NULL OR DoctorName IS NULL;
Select * from patients where PatientID IS NULL;
Select * from visits where VisitID IS NULL OR VisitDate IS NULL;
Select * from labresults where LabResultID IS NULL or TestName IS NULL;
Select * from treatments where TreatmentID IS NULL OR TreatmentName IS NULL;
```

```sql
# 3. Data Consitstency Check. (Ensuring data consistency across all the tables)

SELECT v.VisitID, v.PatientID, p.PatientID
FROM visits AS v
LEFT JOIN patients AS p ON v.PatientID = p.PatientID
WHERE p.PatientID IS NULL;

SELECT t.TreatmentID, t.VisitID, v.VisitID
FROM Treatments t
LEFT JOIN Visits v ON t.VisitID = v.VisitID
WHERE v.VisitID IS NULL;
```

MySQL Quality Analysis Queries

The following SQL queries were used for quality analysis, covering duplicate record checks, dashboard aggregation validation, and performance testing

```sql
# 4. Duplicate Records Check (Identifying duplicate entry)

Select PatientID, count(*)
From patients
Group by PatientID
Having count(*) > 1;

Select VisitID, count(*)
From visits
group by VisitID
Having count(*) >1;

Select TreatmentID, Count(*)
From treatments
Group by TreatmentID
Having count(*);
```

```sql
# 5. Dashboard Aggregation Check
# Comparing sum or average values between SQL and Power BI.

Select ROUND(SUM(TreatmentCost), 2) AS Total_Treatment_Cost from treatments;

Select AVG(age) AS Average_age from patients;

# 6. Performance Testing (Query Execution Time)

EXPLAIN ANALYZE
SELECT * FROM Visits WHERE VisitDate BETWEEN '2023-01-01' AND '2023-12-31';
```

Key Takeaways

1. **Patient Demographics & Volume**: Patient base is around **10,000+**, with an average age close to **49 years**, giving insights into the healthcare needs of middle-aged populations.
2. **Doctor Workload**: Average doctor workload is around **10 patients per doctor**, but variations exist, suggesting the need for better **resource allocation** among doctors.
3. **Treatment Costs & Revenue**: The **average treatment cost is ~$514–$525**, with total revenue exceeding **30M**, highlighting revenue potential but also areas to optimize cost-effectiveness.
4. **Top Diagnosed Conditions**: The most common conditions are **Asthma, Diabetes, Hypertension, and Migraine**, indicating where healthcare resources and preventive measures should be prioritized
5. **Lab Test Utilization**: High usage of lab tests (Blood, CT, MRI, Urine, X-Ray) reflects strong diagnostic dependence; however, **~33% abnormal results** raise red flags for patient health monitoring.

Key Takeaways

6. **Treatment Success Rate**: Some treatments like **Medication and Vaccination show higher success rates**, while **Physical Therapy lags**, suggesting areas for clinical improvement.

7. **Geographical Insights (Power BI & Tableau)**: State-wise diagnosis and patient distribution show hotspots (e.g., **Texas, Illinois, New York**) where additional healthcare support may be needed.

Thank You

Presented by Group 1