

POSTGRE PIXELS

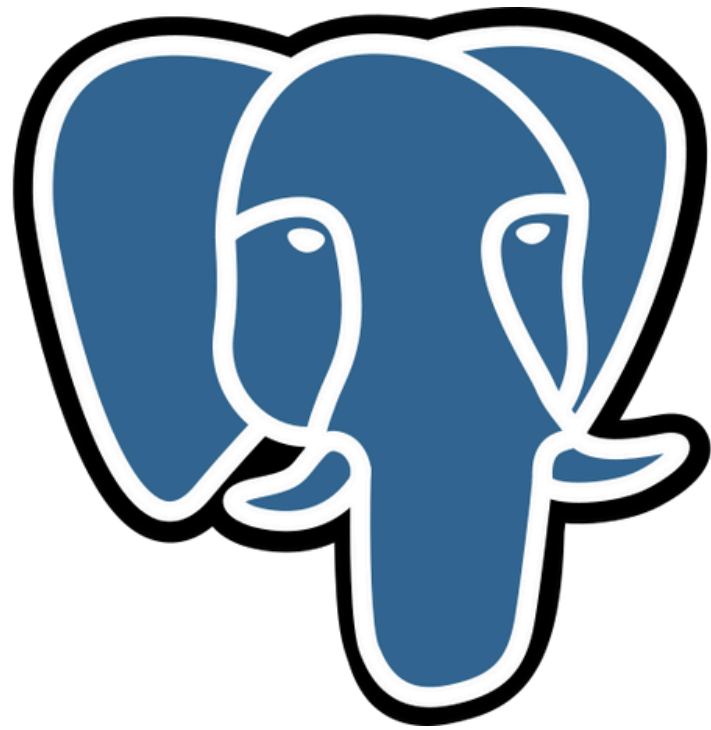
BY: DIMPLE SHARMA



WHAT YOU'LL LEARN?

In this project, my aim is to analyze hospital data using PostgreSQL for data management and Power BI for visualization, providing actionable insights into patient demographics, hospital operations, and financial performance. The dataset consists of patient records, including admission and discharge dates, diagnoses, bed occupancy details, medical tests, doctor assignments, follow-up dates, feedback scores, billing amounts, and insurance claims. Properly structuring and storing this data in PostgreSQL will allow efficient querying, analysis, and integration with Power BI for insightful dashboards.

TOOLS USED



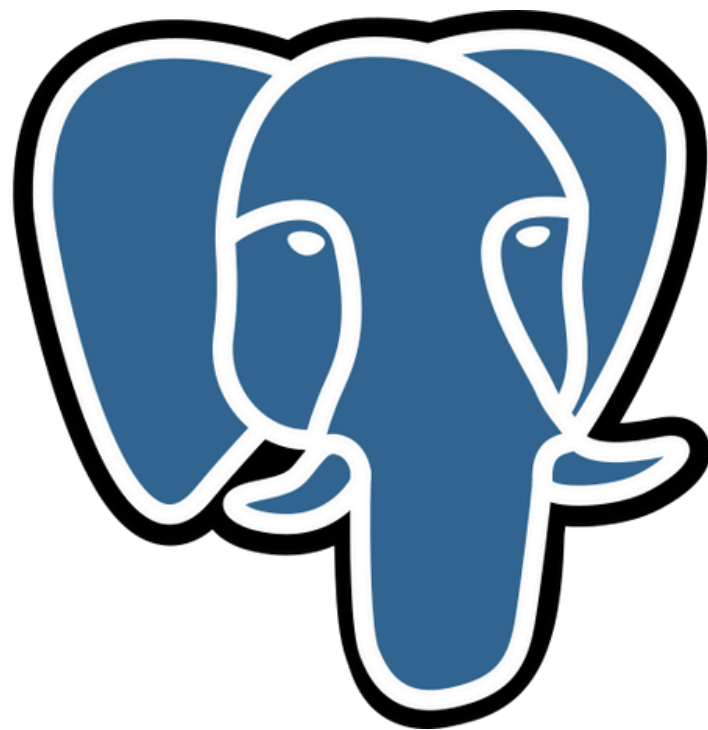
pgAdmin-4



Power BI

STEPS FOLLOWED

1



pgAdmin-4

Using PostgreSQL, I cleaned and preprocess the dataset, ensuring accurate data storage by handling missing values, standardizing date formats, and resolving inconsistencies. SQL queries will be leveraged to extract insights, such as the most common diagnoses, average patient stay duration, total revenue, and doctor efficiency metrics.

STEPS FOLLOWED

2



Power BI

For visualization, Power BI dashboards will provide an interactive representation of key hospital metrics. We will create reports that showcase daily admission trends, revenue breakdowns, patient feedback scores, and department-wise occupancy levels, enabling hospital administrators to make data-driven decisions.

PART-1 (SQL QUERIES)

Basic Queries

1. Retrieve all patient records
2. List distinct diagnoses in the hospital
3. Count the total number of patients admitted
4. Retrieve patients who were admitted to the ICU
5. List all patients along with their assigned doctor


QUERY-I

	patient_id [PK] integer	admit_date date	discharge_date date	diagnosis character varying (255)	bed_occupancy character varying (50)	test character varying (100)	doctor character varying (100)	followup_date date	feedback numeric (3,1)
1	23571	2022-12-31	2023-01-12	Viral Infection	General	MRI	Jay Sinha	2023-01-20	5.0
2	27781	2023-01-04	2023-01-12	Typhoid	ICU	MRI	Jaya Yaadav	2023-01-16	4.0
3	24413	2023-01-05	2023-01-12	Malaria	General	CT Scan	Jay Sinha	2023-01-15	5.0
4	27360	2023-01-05	2023-01-12	Flu	Private	X-Ray	Jaya Yaadav	2023-01-19	5.0
5	26097	2023-01-06	2023-01-12	Viral Infection	General	Blood Test	Jaya Yaadav	2023-02-06	4.9
6	28623	2023-01-06	2023-01-12	Flu	General	Blood Test	Naresh Goyenka	2023-01-20	5.0
7	23992	2023-01-07	2023-01-12	Malaria	General	X-Ray	Tejas Saxena	2023-01-16	5.0
8	25255	2023-01-07	2023-01-12	Pneumonia	Private	X-Ray	Mark Joy	2023-01-15	5.0
9	25676	2023-01-07	2023-01-12	Viral Infection	Private	Ultrasound	Jaya Yaadav	2023-01-19	5.0
10	26939	2023-01-07	2023-01-12	Malaria	ICU	Blood Test	Tejas Saxena	2023-02-07	5.0

Retrieve all patient records

```
SELECT * FROM patients;
```


QUERY-2

	diagnosis character varying (255) 
1	Fracture
2	Malaria
3	Flu
4	Typhoid
5	Viral Infection
6	Pneumonia

List distinct diagnoses in the hospital

```
SELECT DISTINCT Diagnosis FROM patients;
```

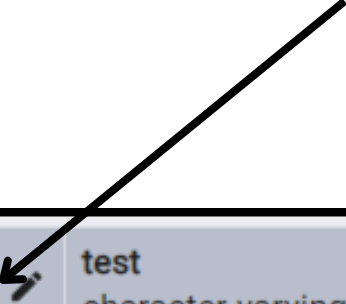
QUERY-3

	total_patients bigint 
1	7157

Count the total number of patients admitted

```
SELECT COUNT(*) AS Total_Patients FROM patients;
```


QUERY-4



	patient_id [PK] integer	admit_date date	discharge_date date	diagnosis character varying (255)	bed_occupancy character varying (50)	test character varying (100)	doctor character varying (100)	followup_date date	feedback numeric (3,1)
1	27781	2023-01-04	2023-01-12	Typhoid	ICU	MRI	Jaya Yaadav	2023-01-16	4.0
2	26939	2023-01-07	2023-01-12	Malaria	ICU	Blood Test	Tejas Saxena	2023-02-07	5.0
3	26098	2023-01-05	2023-01-13	Fracture	ICU	Ultrasound	Mark Joy	2023-01-20	5.0
4	23993	2023-01-08	2023-01-13	Flu	ICU	Blood Test	Jaya Yaadav	2023-01-20	5.0
5	28624	2023-01-10	2023-01-13	Malaria	ICU	MRI	Jay Sinha	2023-02-08	5.0
6	24835	2023-01-11	2023-01-13	Pneumonia	ICU	Ultrasound	Ravi D	2023-01-21	4.0
7	29887	2023-01-11	2023-01-13	Viral Infection	ICU	MRI	Tejas Saxena	2023-01-16	4.0
8	24415	2023-01-06	2023-01-14	Viral Infection	ICU	Ultrasound	Jay Sinha	2023-01-22	4.0
9	23994	2023-01-07	2023-01-14	Typhoid	ICU	CT Scan	Naresh Goyenka	2023-01-17	4.9
10	30681	2023-01-11	2023-01-14	Malaria	ICU	CT Scan	Niki Sharma	2023-01-18	5.0
11	29468	2023-01-06	2023-01-15	Flu	ICU	MRI	Ravi D	2023-01-18	5.0
12	27364	2023-01-08	2023-01-16	Malaria	ICU	MRI	Mark Joy	2023-01-19	3.5
13	23699	2023-01-09	2023-01-16	Viral Infection	ICU	Ultrasound	Ravi D	2023-01-23	3.5
14	30540	2023-01-14	2023-01-16	Viral Infection	ICU	Blood Test	Ravi D	2023-01-24	4.9
15	28628	2022-12-27	2023-01-17	Malaria	ICU	MRI	Jay Sinha	2023-01-20	5.0

Retrieve patients who were admitted to the ICU

```
SELECT * FROM patients WHERE Bed_Occupancy = 'ICU';
```

QUERY-5

	patient_id [PK] integer	doctor character varying (100)
21	24414	Ravi D
22	26940	Naresh Goyenka
23	23572	Jay Sinha
24	23993	Jaya Yaadav
25	25256	Tejas Saxena
26	27782	Jaya Yaadav
27	28203	Naresh Goyenka
28	26519	Mark Joy
29	29466	Ravi D
30	28624	Jay Sinha
31	24835	Ravi D
32	29045	Jay Sinha
33	29887	Tejas Saxena
34	27361	Naresh Goyenka
35	28625	Jay Sinha
36	30309	Jay Sinha
37	30660	Naresh Goyenka
38	23857	Ravi D
39	29467	Tejas Saxena
40	30538	Naresh Goyenka

	patient_id [PK] integer	doctor character varying (100)
21	24414	Ravi D
22	26940	Naresh Goyenka
23	23572	Jay Sinha
24	23993	Jaya Yaadav
25	25256	Tejas Saxena
26	27782	Jaya Yaadav
27	28203	Naresh Goyenka
28	26519	Mark Joy
29	29466	Ravi D
30	28624	Jay Sinha
31	24835	Ravi D
32	29045	Jay Sinha
33	29887	Tejas Saxena
34	27361	Naresh Goyenka
35	28625	Jay Sinha

	patient_id [PK] integer	doctor character varying (100)
1	23571	Jay Sinha
2	27781	Jaya Yaadav
3	24413	Jay Sinha
4	27360	Jaya Yaadav
5	26097	Jaya Yaadav
6	28623	Naresh Goyenka
7	23992	Tejas Saxena
8	25255	Mark Joy
9	25676	Jaya Yaadav
10	26939	Tejas Saxena
11	29465	Jaya Yaadav
12	29886	Tejas Saxena
13	28202	Tejas Saxena
14	24834	Tejas Saxena
15	30307	Mark Joy
16	26518	Jaya Yaadav
17	29044	Ravi D
18	30308	Naresh Goyenka
19	26098	Mark Joy
20	25677	Mark Joy

List all patients along with their assigned doctor

```
SELECT Patient_ID, Doctor FROM patients;
```

TO BE CONTINUE...

Intermediate Queries

1. Find the total revenue generated from patient billing
2. Calculate the average billing amount per diagnosis
3. Find the top 5 doctors with the highest number of patients
4. Retrieve all patients who had a follow-up after their discharge
5. List the most common medical test performed
6. Calculate the average length of stay for patients
7. Identify the doctor who has generated the highest revenue
8. Find the month with the highest number of admissions
9. Retrieve the top 3 diagnoses with the highest total billing

QUERY-1

	total_revenue numeric
1	190432682.00

QUERY-2

	diagnosis character varying (255)	avg_billing numeric
1	Fracture	26355.236933797909
2	Malaria	26559.078266946191
3	Flu	26719.580663948748
4	Typhoid	26508.939737991266
5	Viral Infection	26654.826846307385
6	Pneumonia	26555.251308900524

Find the total revenue generated from patient billing

```
SELECT SUM("Billing Amount") AS Total_Revenue  
FROM patients;
```

Calculate the average billing amount per diagnosis

```
SELECT Diagnosis, AVG("billing_amount") AS Avg_Billing  
FROM patients  
GROUP BY Diagnosis;
```

QUERY-3

	doctor character varying (100) 🔒	patient_count bigint 🔒
1	Jaya Yaadav	1023
2	Jay Sinha	1023
3	Mark Joy	1023
4	Tejas Saxena	1022
5	Niki Sharma	1022

Find the top 5 doctors with the highest number of patients

```
SELECT Doctor, COUNT(*) AS Patient_Count
FROM patients
GROUP BY Doctor
ORDER BY Patient_Count DESC
LIMIT 5;
```

QUERY-4

	test character varying (100) 🔒	test_count bigint 🔒
1	Blood Test	2236
2	MRI	1789
3	CT Scan	1342

List the 3-most common medical test performed

```
SELECT Test, COUNT(*) AS Test_Count
FROM patients
GROUP BY Test
ORDER BY Test_Count DESC
LIMIT 3;
```

QUERY-5

	patient_id [PK] integer	admit_date date	discharge_date date	diagnosis character varying (255)	bed_occupancy character varying (50)	test character varying (100)	doctor character varying (100)	followup_date date	feedback numeric (3,1)	
1	23571	2022-12-31	2023-01-12	Viral Infection	General	MRI	Jay Sinha	2023-01-20	5.0	
2	27781	2023-01-04	2023-01-12	Typhoid	ICU	MRI	Jaya Yaadav	2023-01-16	4.0	
3	24413	2023-01-05	2023-01-12	Malaria	General	CT Scan	Jay Sinha	2023-01-15	5.0	
4	27360	2023-01-05	2023-01-12	Flu	Private	X-Ray	Jaya Yaadav	2023-01-19	5.0	
5	26097	2023-01-06	2023-01-12	Viral Infection	General	Blood Test	Jaya Yaadav	2023-02-06	4.9	
6	28623	2023-01-06	2023-01-12	Flu	General	Blood Test	Naresh Goyenka	2023-01-20	5.0	
7	23992	2023-01-07	2023-01-12	Malaria	General	X-Ray	Tejas Saxena	2023-01-16	5.0	
8	25255	2023-01-07	2023-01-12	Pneumonia	Private	X-Ray	Mark Joy	2023-01-15	5.0	
9	25676	2023-01-07	2023-01-12	Viral Infection	Private	Ultrasound	Jaya Yaadav	2023-01-19	5.0	
10	26939	2023-01-07	2023-01-12	Malaria	ICU	Blood Test	Tejas Saxena	2023-02-07	5.0	

Retrieve all patients who had a follow-up after their discharge

```
SELECT * FROM patients
WHERE "followup_date" > "discharge_date";
```


QUERY-6

	month numeric	admissions bigint
1	1	1091
2	2	914

Find the top-2 months with the highest number of admissions

```
SELECT EXTRACT(MONTH FROM "admit_date") AS
Month, COUNT(*) AS Admissions
FROM patients
GROUP BY Month
ORDER BY Admissions DESC
LIMIT 2;
```

QUERY-7

	diagnosis character varying (255)	total_billing numeric
1	Viral Infection	53416273.00
2	Flu	45877520.00
3	Malaria	38006041.00

Retrieve the top 3 diagnoses with the highest total billing

```
SELECT Diagnosis, SUM("billing_amount") AS Total_Billing
FROM patients
GROUP BY Diagnosis
ORDER BY Total_Billing DESC
LIMIT 3;
```

QUERY-8

	avg_stay_duration numeric
1	8.2498253458152857

Calculate the average length of stay for patients

```
SELECT AVG("Discharge_Date" - "Admit_Date") AS  
Avg_Stay_Duration  
FROM patients;
```

QUERY-9

	doctor character varying (100)	total_revenue numeric
1	Ravi D	28629410.00
2	Mark Joy	27883041.00

Identify the top 2 doctors who has generated the highest revenue



```
SELECT Doctor, SUM("billing_amount") AS Total_Revenue  
FROM patients  
GROUP BY Doctor  
ORDER BY Total_Revenue DESC  
LIMIT 2;
```


TO BE CONTINUE...

Advanced Queries

1. Calculate the insurance coverage percentage for each patient
2. Rank doctors based on total revenue generated
3. Compare Insurance Coverage Across Different Diagnosis
4. Find the most profitable diagnosis for the hospital
5. Determine patient flow over time (rolling admissions count)
6. Identify seasonal trends in patient admissions
7. Find revenue growth over time
8. Identify Patients with Unusually High Billing Amounts (Outliers)

QUERY-I

	patient_id [PK] integer 	insurance_coverage_percentage numeric 
1	23571	90.000000000000000000000000
2	27781	90.000000000000000000000000
3	24413	90.000000000000000000000000
4	27360	90.000000000000000000000000
5	26097	90.000000000000000000000000
6	28623	90.000000000000000000000000
7	23992	90.000000000000000000000000
8	25255	90.000000000000000000000000
9	25676	90.000000000000000000000000
10	26939	90.000000000000000000000000

Calculate the insurance coverage percentage for each patient

```
SELECT Patient_ID,  
("health_insurance_amount" / "billing_amount") * 100 AS Insurance_Coverage_Percentage  
FROM patients;
```

QUERY-2

	doctor character varying (100) 🔒	total_revenue numeric 🔒	revenue_rank bigint 🔒
1	Ravi D	28629410.00	1
2	Mark Joy	27883041.00	2
3	Jaya Yaadav	27163686.00	3
4	Tejas Saxena	27005366.00	4
5	Niki Sharma	26897665.00	5
6	Naresh Goyenka	26713614.00	6
7	Jay Sinha	26139900.00	7

Rank Doctors Based on Total Revenue Generated

```
SELECT Doctor,  
SUM("billing_amount") AS Total_Revenue,  
RANK() OVER (ORDER BY SUM("billing_amount") DESC) AS Revenue_Rank  
FROM patients  
GROUP BY Doctor  
ORDER BY Revenue_Rank;
```

QUERY-3

	diagnosis character varying (255)	avg_coverage_percentage numeric
1	Fracture	90.000000000000000000000000
2	Malaria	90.000000000000000000000000
3	Flu	90.000000000000000000000000

Compare Insurance Coverage Across Different Diagnoses

```
SELECT Diagnosis,  
AVG("Health Insurance Amount" / NULLIF("Billing Amount", 0)  
* 100) AS Avg_Coverage_Percentage  
FROM patients  
GROUP BY Diagnosis  
ORDER BY Avg_Coverage_Percentage DESC;
```

QUERY-4

	diagnosis character varying (255)	total_revenue numeric
1	Viral Infection	53416273.00
2	Flu	45877520.00
3	Malaria	38006041.00

Find the 3-Most Profitable Diagnosis for the Hospital

```
SELECT Diagnosis,  
SUM("billing_amount") AS Total_Revenue  
FROM patients  
GROUP BY Diagnosis  
ORDER BY Total_Revenue DESC  
LIMIT 3;
```

QUERY-5

	admit_date date	rolling_weekly_admissions bigint
1	2022-12-05	1
2	2022-12-06	2
3	2022-12-07	3
4	2022-12-08	4
5	2022-12-09	5
6	2022-12-11	6
7	2022-12-11	7
8	2022-12-12	7
9	2022-12-13	7
10	2022-12-14	7

Determine Patient Flow Over Time (Rolling Admissions Count)

```
SELECT "admit_date",  
COUNT(*) OVER (ORDER BY "admit_date" ROWS BETWEEN 6 PRECEDING AND CURRENT ROW) AS  
Rolling_Weekly_Admissions  
FROM patients;
```





QUERY-6

	month numeric 🔒	total_admissions bigint 🔒
1	1	1091
2	2	914
3	3	648
4	12	587
5	8	513
6	5	509
7	7	505
8	10	501
9	4	493
10	11	472
11	6	466
12	9	458

Identify Seasonal Trends in Patient Admissions

```
SELECT EXTRACT(MONTH FROM "Admit_Date") AS Month, COUNT(*) AS Total_Admissions
FROM patients
GROUP BY MonthORDER BY Total_Admissions DESC;
```

QUERY-7

	month timestamp with time zone 	monthly_revenue numeric 	previous_month_revenue numeric 	revenue_growth numeric 
1	2022-12-01 00:00:00+05:30	1613162.00	[null]	[null]
2	2023-01-01 00:00:00+05:30	15712160.00	1613162.00	14098998.00
3	2023-02-01 00:00:00+05:30	13198088.00	15712160.00	-2514072.00
4	2023-03-01 00:00:00+05:30	16213144.00	13198088.00	3015056.00
5	2023-04-01 00:00:00+05:30	13325323.00	16213144.00	-2887821.00
6	2023-05-01 00:00:00+05:30	13447264.00	13325323.00	121941.00

Find Revenue Growth Over Time

```
SELECT DATE_TRUNC('month', "admit_date") AS Month,
       SUM("billing_amount") AS Monthly_Revenue,
       LAG(SUM("billing_amount")) OVER (ORDER BY DATE_TRUNC('month', "admit_date")) AS
       Previous_Month_Revenue,
       (SUM("billing_amount") - LAG(SUM("billing_amount")) OVER (ORDER BY DATE_TRUNC('month',
       "admit_date")))) AS Revenue_Growth
FROM patients
GROUP BY Month
ORDER BY Month;
```

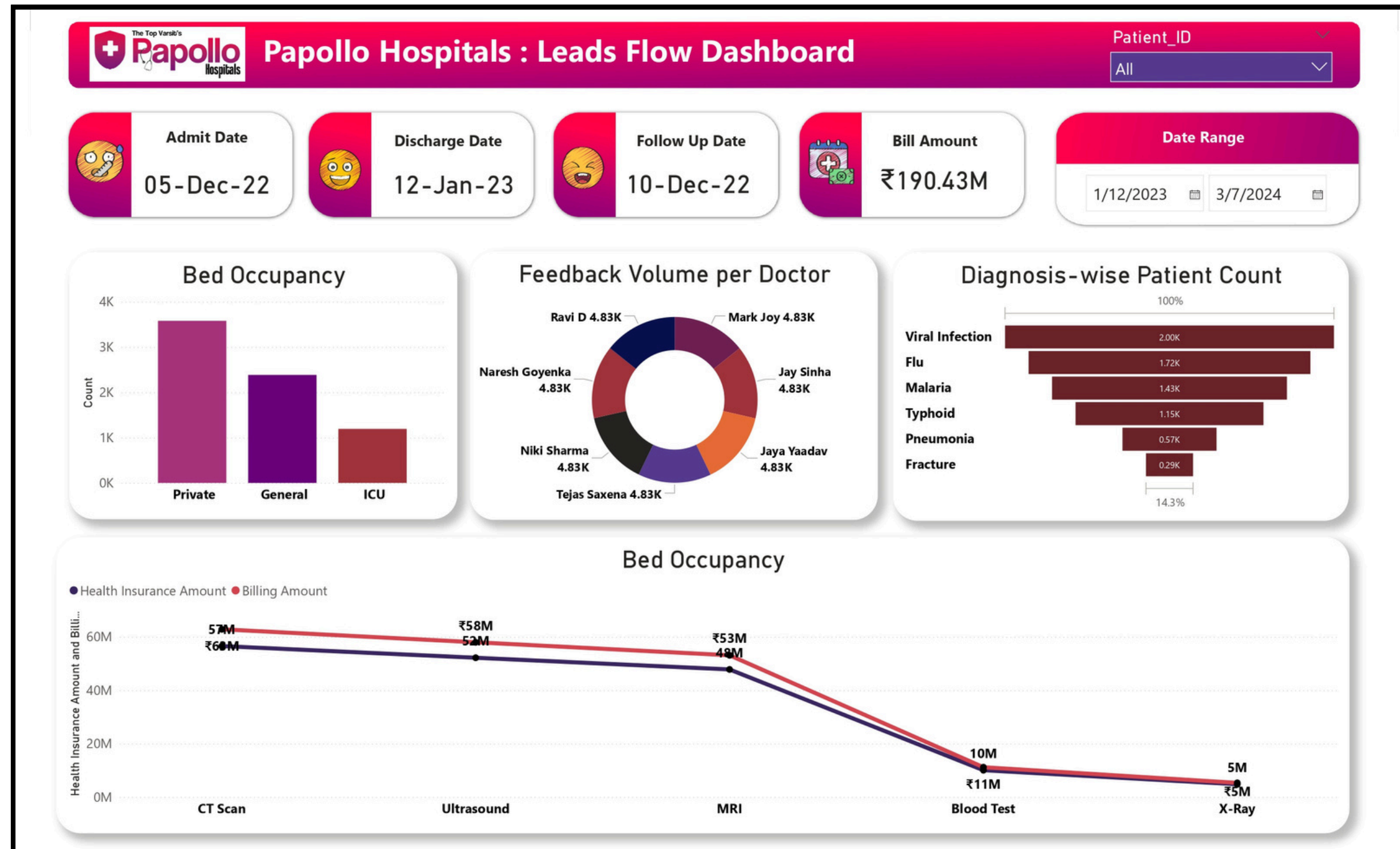
QUERY-8

	<small>_date</small>	<small>diagnosis</small> <small>character varying (255)</small>	<small>bed_occupancy</small> <small>character varying (50)</small>	<small>test</small> <small>character varying (100)</small>	<small>doctor</small> <small>character varying (100)</small>	<small>followup_date</small> <small>date</small>	<small>feedback</small> <small>numeric (3,1)</small>	<small>billing_amount</small> <small>numeric (10,2)</small>	<small>health_insurance_amount</small> <small>numeric (10,2)</small>
1	2	Typhoid	Private	Ultrasound	Mark Joy	2023-02-07	5.0	78960.00	71064.00
2	3	Flu	General	Ultrasound	Tejas Saxena	2023-02-07	4.5	85778.00	77200.20
3	3	Pneumonia	ICU	Ultrasound	Ravi D	2023-01-21	4.0	87895.00	79105.50
4	4	Viral Infection	Private	CT Scan	Jay Sinha	2023-01-18	4.9	87654.00	78888.60
5	4	Viral Infection	General	Ultrasound	Niki Sharma	2023-01-21	5.0	84764.00	76287.60
6	5	Pneumonia	General	CT Scan	Jay Sinha	2023-01-23	4.9	83754.00	75378.60
7	5	Flu	Private	Ultrasound	Jay Sinha	2023-01-18	5.0	87436.00	78692.40
8	7	Typhoid	ICU	Ultrasound	Naresh Goyenka	2023-01-20	5.0	86795.00	78115.50
9	7	Flu	ICU	CT Scan	Jaya Yaadav	2023-01-24	5.0	83754.00	75378.60
10	7	Flu	ICU	CT Scan	Mark Joy	2023-01-21	5.0	83754.00	75378.60

Identify Patients with Unusually High Billing Amounts (Outliers)

```
SELECT * FROM patients
WHERE "billing_amount" > (SELECT AVG("billing_amount") + 2 * STDDEV("billing_amount") FROM patients);
```


PART-2 (POWER BI)



INSIGHTS

Financial Overview

- **Total Bill Amount:** ₹190.43 million
- **Health Insurance Amount:** Significant contributions from various services:
 - CT Scan: ₹63 million
 - MRI: ₹53 million
 - Ultrasound: ₹5 million
 - Blood Test: ₹58 million
 - X-Ray: ₹11 million

Patient Flow and Occupancy

- **Bed Occupancy:**
 - The dashboard indicates a range from 0K to 4K beds occupied, suggesting varying levels of patient intake and discharge.
- **Admit and Discharge Dates:**
 - Tracking of patient admissions and discharges is crucial for understanding hospital capacity and patient turnover.

INSIGHTS

Patient Diagnosis Insights

- **Diagnosis-wise Patient Count:**
 - Viral Infection: 14.3% of total patients
 - Other common diagnoses include Flu, Malaria, Typhoid, Pneumonia, and Fractures, indicating prevalent health issues in the patient population.

Feedback and Performance Metrics

- **Feedback Volume per Doctor:**
 - All listed doctors (Mark Joy, Tejas Saxena, Ravi D, Jay Sinha, Jaya Yaadav, Niki Sharma, Naresh Goyenka) have a consistent feedback volume of 4.83K, suggesting a uniform level of patient engagement or satisfaction across these practitioners.

INSIGHTS

Trends and Future Considerations

- **Follow-Up Dates:**
 - Monitoring follow-up dates is essential for patient care continuity and can help in assessing the effectiveness of treatments.
- **Potential Areas for Improvement:**
 - Analyzing the feedback data can provide insights into areas where patient care can be enhanced, particularly focusing on doctors with lower feedback scores if applicable.

Data Range & Monitoring

Interpretation:

- Dashboard covers Jan 12, 2023, to March 7, 2024 → Provides over a year's worth of data for trends and decision-making.

CONCLUSION

- This project successfully integrates PostgreSQL and Power BI to analyze hospital operations, patient demographics, and financial insights at Papollo Hospitals. By leveraging structured queries and data visualization, we identified key trends in bed occupancy, diagnosis patterns, revenue generation, and doctor performance.
- The findings highlight that private wards have the highest occupancy, viral infections are the most common diagnosis, and billing amounts often exceed insurance coverage, indicating potential financial burdens on patients.
- Additionally, patient feedback appears uniformly distributed among doctors, suggesting a need for deeper satisfaction analysis. The project also emphasizes the importance of monitoring patient readmission rates and length of stay to optimize hospital efficiency.

THANK YOU!