

Hospital Database Management System

Project Overview

This project focuses on building a Hospital Database Management System using SQL.

The system is designed with 7 interconnected tables, each having proper relationships to store and manage hospital-related information such as patients, doctors, appointments, treatments, billing, and more.

The dataset was AI-generated in CSV format, which was then imported into SQL. This provided a realistic and structured dataset for database design, query execution, and testing.

In total, the project includes 40 SQL queries developed for data retrieval, analysis, and insights. However, only a selected set of queries have been showcased here

Key Database Tables and Primary Keys

- Departments Table → Primary Key: **Department_ID**
- Doctors Table → Primary Key: **Doctor_ID**
- Patients Table → Primary Key: **Patient_ID**
- Appointments Table → Primary Key: **Appointment_ID**
- Treatments Table → Primary Key: **Treatment_ID**
- Billing Table → Primary Key: **Billing_ID**
- Visits/Records Table → Primary Key: **Visit_ID**

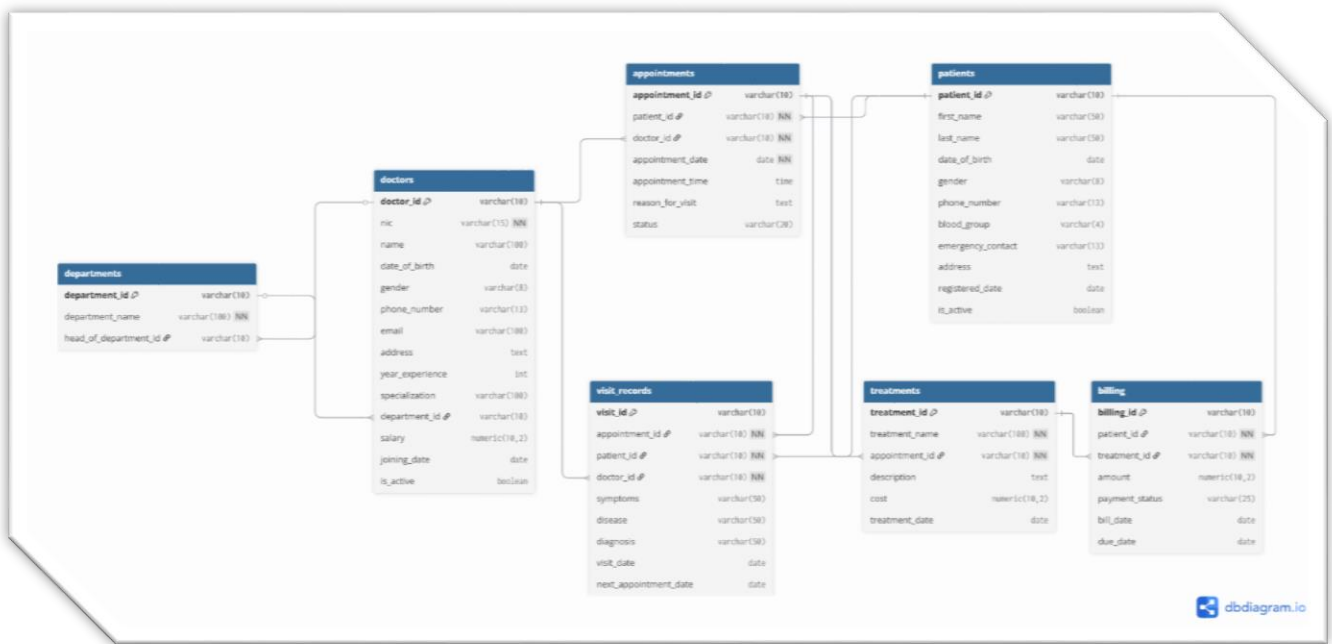
Tools & Technologies

- SQL – Database creation, table design, relationships, and queries
- Excel – Data review, cleaning, and verification
- AI Tools – For synthetic dataset generation

Project Highlights

- Designed a relational database schema with 7 tables
- Established primary and foreign key relationships
- Imported and managed CSV data into SQL for analysis
- Performed queries for hospital operations such as patient records, doctor details, appointments, and billing

Entity Relationship Diagram (ERD)



Top 10 SQL Queries for Analysis

Revenue Analysis

1. Doctor who generated the highest total revenue.

Query

```

1  --1. Doctor who generated the highest total revenue.
2  SELECT
3      d.doctor_id, d.name, SUM(b.amount) AS total_revenue
4  FROM
5      billing b
6      JOIN treatments t ON b.treatment_id = t.treatment_id
7      JOIN appointments a ON a.appointment_id = t.appointment_id
8      JOIN doctors d ON a.doctor_id = d.doctor_id
9  GROUP BY
10     d.doctor_id, d.name
11  ORDER BY
12     SUM(b.amount) DESC
13  LIMIT 1;

```

Output

	doctor_id [PK] character varying (10)	name character varying (100)	total_revenue numeric
1	DOC065	Dr. Asim Hussain	65634.00

2. Total revenue generated from treatments in each department

Query

```

15  --2. Total revenue generated from treatments in each department
16  SELECT
17      dep.department_name, SUM(t.cost) AS total_revenue
18  FROM
19      treatments t
20      JOIN appointments a ON a.appointment_id = t.appointment_id
21      JOIN doctors d ON a.doctor_id = d.doctor_id
22      JOIN departments dep ON dep.department_id = d.department_id
23  GROUP BY
24      dep.department_name;

```

Output

	department_name character varying (100)	total_revenue numeric
1	Cardiology	172576.00
2	ENT	205789.00
3	Surgery	204886.00
4	Oncology	139034.00
5	Orthopedics	179579.00
6	Gynecology	154456.00
7	Ophthalmology	173152.00
8	Pulmonology	177887.00
9	Neurology	154069.00
10	Anesthesiology	161373.00

3. Patient who spent the maximum amount on treatments

Query

```
--3. Patient who spent the maximum amount on treatments
26
27 SELECT
28     p.patient_id, CONCAT(p.first_name, ' ', p.last_name) AS patient_name,
29     p.gender, SUM(t.cost) AS total_amount
30 FROM
31     treatments t
32 JOIN appointments a ON a.appointment_id = t.appointment_id
33 JOIN patients p ON a.patient_id = p.patient_id
34 GROUP BY
35     p.patient_id, patient_name, p.gender
36 ORDER BY
37     SUM(t.cost) DESC
38 LIMIT 1;
```

Output

	patient_id [PK] character varying (10)	patient_name text	gender character varying (8)	total_amount numeric
1	PAT638	Farhan Tariq	Male	21369.00

Doctor Performance Analysis

4. Top 5 doctors who handled the most appointments

Query

```
--4. Top 5 doctors who handled the most appointments
40
41 SELECT
42     a.doctor_id, d.name, COUNT(a.appointment_id) AS Number_of_handle
43 FROM
44     appointments a
45 JOIN doctors d ON a.doctor_id = d.doctor_id
46 GROUP BY
47     a.doctor_id, d.name
48 ORDER BY
49     COUNT(a.appointment_id) DESC
50 LIMIT 5;
```

Output

	doctor_id character varying (10)	name character varying (100)	number_of_handle bigint
1	DOC117	Dr. Azhar Butt	33
2	DOC049	Dr. Yasir Qureshi	32
3	DOC078	Dr. Tasneem Fatima	31
4	DOC067	Dr. Kashif Mahmood	29
5	DOC062	Dr. Asma Yousuf	29

5. Top 5 doctors with highest completed appointments

Query

```
52 --5. Top 5 doctors with highest completed appointments
53 SELECT
54     a.doctor_id, d.name, COUNT(*) FILTER(WHERE a.status = 'Completed') AS completed
55 FROM
56     appointments a
57 JOIN doctors d ON d.doctor_id = a.doctor_id
58 GROUP BY
59     a.doctor_id, d.name
60 ORDER BY
61     completed DESC
62 LIMIT 5;
```

Output

	doctor_id character varying (10)	name character varying (100)	completed bigint
1	DOC104	Dr. Sana Khan	15
2	DOC117	Dr. Azhar Butt	14
3	DOC062	Dr. Asma Yousuf	12
4	DOC101	Dr. Shahid Raza	12
5	DOC056	Dr. Shaista Tariq	11

6. Doctor with the highest number of unique patients treated

Query

```
64 --6. Doctor with the highest number of unique patients treated
65 SELECT
66     a.doctor_id, d.name, COUNT(DISTINCT a.patient_id) AS number_of_treat
67 FROM
68     appointments a
69 JOIN doctors d ON d.doctor_id = a.doctor_id
70 WHERE
71     a.status = 'Completed'
72 GROUP BY
73     a.doctor_id, d.name
74 HAVING
75     COUNT(DISTINCT a.patient_id) > 0
76 ORDER BY
77     number_of_treat DESC LIMIT 1;
```

Output

	doctor_id character varying (10)	name character varying (100)	number_of_treat bigint
1	DOC104	Dr. Sana Khan	15

Patient Behavior Analysis

7. Patients who spent more than 10,000 in total treatment costs

Query

```
79 --7. Patients who spent more than 10,000 in total treatment costs
80 SELECT
81     p.patient_id, p.first_name, p.last_name, p.gender, SUM(t.cost)
82 FROM
83     patients p
84 JOIN appointments a ON p.patient_id = a.patient_id
85 JOIN treatments t ON a.appointment_id = t.appointment_id
86 GROUP BY
87     p.patient_id, p.first_name, p.last_name, p.gender
88 HAVING
89     SUM(t.cost) > 10000;
```

Output

	patient_id [PK] character varying (10)	first_name character varying (50)	last_name character varying (50)	gender character varying (8)	sum numeric
1	PAT941	Sana	Hussain	Male	12206.00
2	PAT1272	Omar	Shah	Male	10842.00
3	PAT1088	Nadia	Khan	Female	10196.00
4	PAT1150	Sara	Rehman	Female	10285.00
5	PAT1085	Zahid	Hussain	Female	10506.00
6	PAT638	Farhan	Tariq	Male	21369.00
7	PAT824	Usman	Siddiqui	Female	10599.00
8	PAT1125	Omar	Mahmood	Female	15524.00
9	PAT1122	Ayesha	Rehman	Female	10125.00

8. Patients with more than 3 visits in the last 6 months

Query

```
--
--8. Patients with more than 3 visits in the last 6 months
91 SELECT
92     p.patient_id, p.first_name, p.last_name, p.gender, COUNT(v.visit_id) AS no_of_visit
93 FROM
94     patients p
95 JOIN visit_records v ON p.patient_id = v.patient_id
96 WHERE
97     visit_date BETWEEN CURRENT_DATE - INTERVAL '6 months' AND CURRENT_DATE
98 GROUP BY
99     p.patient_id, p.first_name, p.last_name, p.gender
100 HAVING
101     COUNT(v.visit_id) > 3
102 ORDER BY
103     no_of_visit DESC limit 10;
```

Output

	patient_id [PK] character varying (10)	first_name character varying (50)	last_name character varying (50)	gender character varying (8)	no_of_visit bigint
1	PAT203	Sana	Mahmood	Female	7
2	PAT1285	Bilal	Tariq	Female	7
3	PAT1327	Farhan	Iqbal	Male	7
4	PAT741	Aisha	Tariq	Male	7
5	PAT1065	Imran	Tariq	Male	6
6	PAT569	Ayesha	Akhtar	Male	6
7	PAT1250	Imran	Siddiqui	Female	6
8	PAT240	Hassan	Siddiqui	Female	6
9	PAT560	Sadia	Mahmood	Female	6
10	PAT131	Mehwish	Khan	Female	6

Department Efficiency Analysis

9. Ratio of 'No-show' vs 'Completed' appointments per department

Query

```
106 --9. Ratio of 'No-show' vs 'Completed' appointments per department
107 SELECT
108     dep.department_name,
109     COUNT(CASE WHEN a.status = 'Completed' THEN 1 END) AS completed,
110     COUNT(CASE WHEN a.status = 'No-show' THEN 1 END) AS no_show,
111     SUM(COUNT(*) OVER()) AS total_appointment,
112     CONCAT(ROUND((COUNT(CASE WHEN a.status = 'Completed' THEN 1 END) / SUM(COUNT(*)
113 OVER()) * 100, 2), '%') AS completed_ratio,
114     CONCAT(ROUND((COUNT(CASE WHEN a.status = 'No-show' THEN 1 END) / SUM(COUNT(*) OVER())
115 * 100, 2), '%') AS no_show_ratio
116 FROM
117     appointments a
118 JOIN doctors d ON d.doctor_id = a.doctor_id
119 JOIN departments dep ON d.department_id = dep.department_id
120 GROUP BY
121     dep.department_name;
```

Output

	department_name character varying (100)	completed bigint	no_show bigint	total_appointment numeric	completed_ratio text	no_show_ratio text
1	Cardiology	45	1	3000	1.50%	0.03%
2	ENT	47	4	3000	1.57%	0.13%
3	Surgery	50	2	3000	1.67%	0.07%
4	Oncology	39	6	3000	1.30%	0.20%
5	Orthopedics	52	5	3000	1.73%	0.17%
6	Gynecology	46	3	3000	1.53%	0.10%
7	Ophthalmology	55	5	3000	1.83%	0.17%
8	Pulmonology	54	5	3000	1.80%	0.17%
9	Neurology	38	7	3000	1.27%	0.23%
10	Pediatrics	41	8	3000	1.37%	0.27%

10. Department with the highest number of appointments

Query

```
121 --10. Department with the highest number of appointments
122 SELECT
123     dep.department_name, COUNT(a.appointment_id) AS num_of_app
124 FROM
125     appointments a
126     JOIN doctors d ON d.doctor_id = a.doctor_id
127     JOIN departments dep ON dep.department_id = d.department_id
128 GROUP BY
129     dep.department_name
130 ORDER BY
131     num_of_app DESC
132 LIMIT 1;
```

Output

	department_name character varying (100)	num_of_app bigint
1	Ophthalmology	183

Conclusion

Our comprehensive analysis of the hospital management system reveals that Dr. Asim Hussain stands out as the highest revenue generator, while the ENT, Surgery, and Radiology departments collectively drive the majority of institutional revenue, with Mr. Farhan Tariq being the highest-spending patient. In terms of operational efficiency, Dr. Azhar Butt and Dr. Yasir Qureshi handled the most appointments, whereas Dr. Sana Khan excelled in both appointment completion rates and treating the highest number of unique patients. Patient engagement was notably high for Sana Mahmood and Bilal Tariq, each with 7 visits in the last six months. Departmentally, Ophthalmology achieved the highest appointment completion rate (1.83%) and the most completed appointments, while Anesthesiology reported the highest no-show ratio, indicating a need for improved scheduling or patient communication strategies. These insights collectively underscore a well-performing institution with clear strengths in specialized departments and opportunities for enhancing patient retention and operational efficiency in underperforming areas.

Full Project on GitHub

This repository folder contains the complete Hospital Database Management System project, including the database schema, AI-generated dataset (CSV files), and SQL queries (40 in total). If you would like to view and explore the full project, please [click here](#).