CAPSTONE PROJECT-1

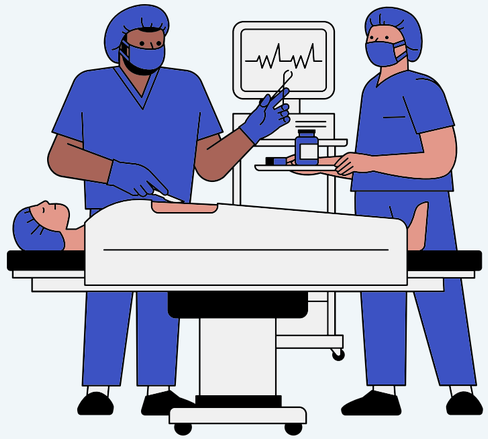
HOSPITAL MANAGEMENT SYSTEM

LEVEL :INTERMEDIATE

Designed and Analyzed a hospital database using SQL. Extracted insights to support efficient healthcare management.

VENKATA SAHITH KOTARU
SAHITHKOTARU@GMAIL.COM

OVERVIEW



This SQL Capstone Project analyzes hospital data across patients, doctors, treatments, and billing to understand operations. Using joins, subqueries, and aggregate functions, key insights on patient flow, doctor performance, and costs were derived.

The project highlights how SQL analysis improves hospital efficiency and data-driven decision-making.

VENKATA SAHITH KOTARU
SAHITHKOTARU@GMAIL.COM

Dataset Overview

Table Name	Description
patients	Stores patient details like name, contact, and insurance.
Doctors	Contains doctor information such as name, specialization, and experience.
Appointments	Records when and why patients visit doctors.
Treatments	Includes details of treatments, their type, date, and cost.
Billing	Tracks payments, amounts, and payment status.

VENKATA SAHITH KOTARU
SAHITHKOTARU@GMAIL.COM

ER-Diagram and relationships

billing

billing_id	int
patient_id	int
treatment_id	int
bill_date	date
amount	decimal
payment_method	varchar
payment_status	varchar

treatments

treatment_id	int
appointment_id	int
treatment_type	varchar
description	text
cost	decimal...
treatment_date	date

appointments

appointment_id	varchar
patient_id	int
doctor_id	int
appointment_date	date
appointment_time	time
reason_for_visit	varchar
status	varchar

Patients

patient_id	int
first_name	varchar
last_name	varchar
gender	char
date_of_birth	date
contact_number	int
address	varchar
registration_date	date
insurance_provider	varchar
insurance_number	varchar
email	varchar

doctors

doctor_id	int
first_name	varchar
last_name	varchar
specialization	varchar
phone_number	int
years_of_experience	int
hospital_branch	varchar
email	varchar

VENKATA SAHITH KOTARU
SAHITHKOTARU@GMAIL.COM

Queries and insights

1.Patients registered each month

```
select monthname(registration_date) as month,
count(registration_date) as registered_patients
from patients group by month;
```

month	registered_patients
June	6
January	4
February	1
March	3
September	10
October	4
December	4
May	5
August	3
April	5
July	5

Insights:

- In the month September there are more number of patient registrations.
- In the months Jan,Oct,Dec,May,Jul, Apr has same number of new registrations

2.Gender distribution

```
select gender,count(*) as distribution from patients
group by gender;
```

gender	distribution
F	19
M	31

Insights:

- Male gender shows a clear dominance in patient registrations.
- It means male gender is most suffering than female as their registrationsare more.

VENKATA SAHITH KOTARU
SAHITHKOTARU@GMAIL.COM

3.Age group of patients

```
select case
when age<18 then "Adolscents(<18)"
when age<30 then "Youth(<30)"
when age<59 then "working citizen(<50)"
when age<100 then "senior citizen(<100)"
end as age_group,count(*) as num_of_patients
from age_of_patients group by age_group;
```

age_group	num_of_patients
senior citizen(<100)	13
working citizen(<50)	28
Youth(<30)	9

Insights:

- The Working citizen age group has seen more number of registration in hospitals
- The Adolscents age group has no registrations in any hospital.

4.Insurance Providers

```
select insurance_provider,count(*) as num_of_patients
from patients group by insurance_provider
order by insurance_provider;
```

insurance_provider	num_of_patients
HealthIndia	6
MedCare Plus	18
PulseSecure	10
WellnessCorp	16

Insights:

- WellnessCorp provided most insurances to the patients
- Healthindia provided less insurances to the patients

VENKATA SAHITH KOTARU
SAHITHKOTARU@GMAIL.COM

5.patients from same address registered

```
SELECT address, COUNT(*) AS num_of_patients
FROM patients
GROUP BY address;
```

address	num_of_patients
789 Pine Rd	15
321 Maple Dr	19
123 Elm St	9
456 Oak Ave	7

Insights:

- From address 321 Maple Dr more number of patients are registered in the hospital.
- There are less registration from 456 Oak Ave
- So publicity should be done at 456 Oak Ave,123 Elm St places to boost registrations.

6.Doctors with most Appointments

```
select
concat(d.first_name,' ',d.last_name) as doctor_name,
count(a.appointment_id)as appointments
from doctors d join appointments a
on d.doctor_id=a.doctor_id
group by doctor_name
order by appointments desc limit 5;
```

doctor_name	appointments
Sarah Taylor	29
David Taylor	25
Alex Davis	24
Jane Smith	22
Jane Davis	21

Insights:

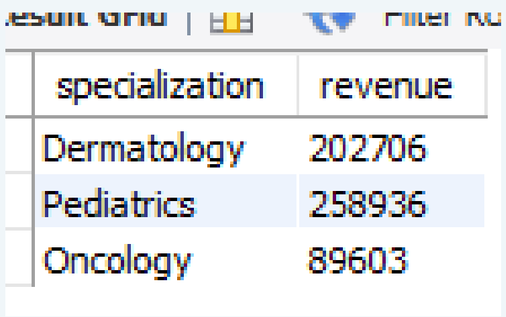
- Most Patients are choosing Sarah Taylor as their primary appointment doctor than any one.

VENKATA SAHITH KOTARU
SAHITHKOTARU@GMAIL.COM

7.Specilization's Revenue

```
select d.specialization,sum(b.amount) as revenue
from doctors d join appointments a on d.doctor_id=a.doctor_id
join treatments t on a.appointment_id=t.appointment_id
join billing b on t.treatment_id=b.treatment_id
group by d.specialization;
```

Insights:



specialization	revenue
Dermatology	202706
Pediatrics	258936
Oncology	89603

- The specialization Pediatrics bought most Revenue than anyone,so the hospital has to hire more pediatricians to boost revenue.
- Oncology bought least revenue, so the management have to take care on oncology improvement.

8.Patients per Doctor

```
select concat(d.first_name,' ',d.last_name) as doctor_name,
count(a.patient_id) as total_patients
from doctors d join appointments a on d.doctor_id=a.doctor_id
group by doctor_name;
```

Insights:

- Sarah Taylor has treated more number of patients.
- Robert Davis has treated less number of patients.
- patients are prefering sarah taylor,david taylor,alex davis,jane smith,jane davis as their treatment doctors most.

doctor_name	total_patients
David Taylor	25
Jane Davis	21
Jane Smith	22
David Jones	14
Sarah Taylor	29
Alex Davis	24
Robert Davis	13
Linda Brown	16
Sarah Smith	17
Linda Wilson	19

VENKATA SAHITH KOTARU
SAHITHKOTARU@GMAIL.COM

9.Most active hospital branch

```
select d.hospital_branch,count(a.appointment_id) as
 num_of_appointments,sum(b.amount) as revenue from
 doctors d join appointments a on d.doctor_id=a.doctor_id
join billing b on a.patient_id=b.patient_id
group by d.hospital_branch
order by d.hospital_branch asc;
```

hospital_branch	num_of_appointments	revenue
Central Hospital	420	1139030
Eastside Clinic	309	849448
Westside Clinic	273	759781

Insights:

- **Central hospital is most Active branch in terms of appointments and revenue.**
- **Westside clinic has got least active branch.**
- **All two branches ae performing well in terms of Revenue and appointments.**

10.Doctor’s experience VS patients

```
select d.years_of_experience,count(a.patient_id)
as num_of_patients from
doctors d join appointments a
on d.doctor_id=a.doctor_id
group by d.years_of_experience;
```

years_of_experience	num_of_patients
17	25
24	21
19	22
28	14
26	59
23	24
5	16
21	19

Insights:

- **Patients are prefering the doctors with an experience >20 years to get treatment.**
- **A doctor with exp 26 years is choosen by the patients as he/she treated a total of 59 patients.**

VENKATA SAHITH KOTARU
SAHITHKOTARU@GMAIL.COM

11.Appointments vs Status

```
select status,count(*) as
num_of_appointments
from appointments
group by status;
```

status	num_of_appointments
Scheduled	51
No-show	52
Cancelled	51
Completed	46

Insights:

- Most appointments are under cancelled status and sheduled status.
- Canceled appointments showing dominance ,it this increases then no one will book an appointment in this hospital.

12.Peak days for appointments

```
select dayname(appointment_date) as day,
count(appointment_id) as total_appointments
from appointments group by day
order by total_appointments desc;
```

day	total_appointments
Wednesday	37
Tuesday	37
Thursday	28
Monday	26
Sunday	26
Friday	23
Saturday	23

Insights:

- Wednesday,Tuesday are considered as peak days for most appointments.
- Patients are prefering week starting days than weekends for appointments.

VENKATA SAHITH KOTARU
SAHITHKOTARU@GMAIL.COM

13.Top5 doctors with cancelled appointments

```
select concat(d.first_name,' ',d.last_name) as name,
count(a.appointment_id) as cancelled_appointments
from doctors d join appointments a
on d.doctor_id=a.doctor_id
where a.status ="cancelled" group by name
order by cancelled_appointments desc limit 5;
```

name	cancelled_appointments
Jane Davis	8
David Taylor	7
Sarah Taylor	6
Alex Davis	6
Robert Davis	5

Insights:

- **Jane davis got more cancelled appointments than all doctors,the management has to take action to reduce this cancellations.**

14.Most common reason visit reasons

```
select reason_for_visit,count(*) as total_reasons
from appointments
group by reason_for_visit
order by total_reasons desc;
```

reason_for_visit	total_reasons
Checkup	45
Consultation	43
Therapy	42
Follow-up	41
Emergency	29

Insights:

- **patients are visiting hospitals mostly for checkup,so managemnt can increase cost for checkup.**
- **Emergency cases are less,so management has to decrease the cost and do more publicity on this cases to boost visits.**

VENKATA SAHITH KOTARU
SAHITHKOTARU@GMAIL.COM

15.Total appointments took by patient

```
select concat(p.first_name,' ',p.last_name)
as patient_name, count(a.appointment_id) as
total_appointments from patients p join appointments a
 on  p.patient_id=a.patient_id
group by patient_name
order by total_appointments desc;
```

patient_name	total_appointm
Michael Taylor	16
David Wilson	15
Michael Wilson	14
Laura Davis	14
David Smith	9
Robert Williams	9
Alex Moore	8
David Moore	7
John Taylor	6
Linda Johnson	6
Laura Wilson	6
Emily Miller	5
Robert Wilson	5
Jane Moore	5
Linda Miller	5
Jane Wilson	5
Sarah Miller	5
David Williams	4
John Brown	4

Insights:

- Michael Taylor has the most appointments (16)
- The top 4 patients all have 14 or more appointments, while the rest have 9 or fewer

16.Most common treatment

```
select treatment_type,count(treatment_id)
as num_of_treatments from treatments
group by treatment_type
order by num_of_treatments desc;
```

treatment_type	num_of_treatments
Chemotherapy	49
X-Ray	41
ECG	38
MRI	36
Physiotherapy	36

Insights:

- Chemotherapy is the most commontreatment type in this hospital and its branches.
- X-ray is maintaining a middile role in terms of treatments.

VENKATA SAHITH KOTARU
SAHITHKOTARU@GMAIL.COM

17.Avg cost per treatment_type

```
select treatment_type,round(avg(cost),0)
as avg_cost from treatments
group by treatment_type;
```

treatment_type	avg_cost
Chemotherapy	2630
MRI	3225
ECG	2532
Physiotherapy	2762
X-Ray	2699

Insights:

- MRI treatment has got highest avg cost.
- The treatments MRI,Xray,ECG,Chemotherapy are maintaining the higher level ofavg cost(>2500).

18.Revenue by treatment

```
select t.treatment_type,sum(b.amount) as revenue
from treatments t join billing b
on t.treatment_id=b.treatment_id
group by t.treatment_type;
```

treatment_type	revenue
Chemotherapy	128854
MRI	116098
ECG	96223
Physiotherapy	99418
X-Ray	110652

Insights:

- Chemotherapy has bought highest revenue of 1,28,854 than alltreatments.
- MRI,ECG treatments are maintaining same level (>1,00,000).
- ECG's cost should be decreased to boost appointments and revenue too.

VENKATA SAHITH KOTARU
SAHITHKOTARU@GMAIL.COM

19.Cost variation across branches

```
select d.hospital_branch,t.treatment_type,max(t.cost) as
 max_cost,min(t.cost) as min_cost from
doctors d join appointments a on d.doctor_id=a.doctor_id
join treatments t on a.appointment_id=t.appointment_id
group by d.hospital_branch,t.treatment_type;
```

hospital_branch	treatment_type	max_cost	min_cost
Westside Clinic	X-Ray	4652	935
Westside Clinic	Physiotherapy	4331	1287
Westside Clinic	ECG	4961	968
Westside Clinic	Chemotherapy	4965	534
Westside Clinic	MRI	4672	1864
Eastside Clinic	MRI	4716	663
Eastside Clinic	Physiotherapy	4541	1556
Eastside Clinic	ECG	4672	582
Eastside Clinic	Chemotherapy	4945	616
Eastside Clinic	X-Ray	4833	807
Central Hospital	MRI	4966	894
Central Hospital	Physiotherapy	4846	956
Central Hospital	Chemotherapy	4479	695
Central Hospital	X-Ray	4974	864
Central Hospital	ECG	4550	606

Insights:

- MRI is the most expensive treatment in all branches.
- Physiotherapy has the biggest difference between high and low costs.

20.Treatment trend over time

```
select monthname(treatment_date)as month,
count(treatment_id) as total_treatments
from treatments group by month;
```

Insights:

- April month has seen most number of treatments than all months.
- Most Treatments are happened in summer season' months.

month	total_treatments
August	15
June	18
September	11
July	16
April	25
May	19
March	19
January	20
November	17
February	14
December	12
October	14

VENKATA SAHITH KOTARU
SAHITHKOTARU@GMAIL.COM

21.Total revenue per month

```
select monthname(bill_date) as month_name,
sum(amount) as revenue from billing
group by month_name
order by revenue desc;
```

month_name	revenue
April	64273
January	58698
June	56888
November	52476
May	48791
March	47303
October	43314
August	41958
July	39879
February	36670
September	33426
December	27569

Insights:

- April month bought highest revenue than all months as tthe number of treatments are more.
- December bought least revenue so the month december is not prefered by patients more.

22.Revenue by payment method

```
select payment_method,sum(amount) as revenue
from billing
group by payment_method;
```

payment_method	revenue
Insurance	182156
Credit Card	201381
Cash	167708

Insights:

- Patients are mostly prefering Credit card to pay the amount.
- Only less people are prefering to pay through cash as it bought least revenue.

VENKATA SAHITH KOTARU
SAHITHKOTARU@GMAIL.COM

23.pending vs paid vs failed bills

```
select payment_status,count(*) as
num_of_bills,sum(amount) as total_amount
from billing
group by payment_status;
```

payment_status	num_of_bills	total_amount
Pending	69	184610
Paid	64	173423
Failed	67	193212

Insights:

- Most bills are pending, showing delays in payment processing.
- Failed payments have the highest total amount, which may need quick attention to recover revenue.

24.Top billing patients

```
select p.patient_id,concat(p.first_name,' ',p.last_name)
 as patient_name,sum(b.amount) as total_bill from
 patients p join billing b on p.patient_id=b.patient_id
group by p.patient_id,patient_name
order by total_bill desc limit 10;
```

Insights:

- Two patients are named "David Wilson" and two are named "Michael Taylor".
- The highest total is 30,052 (Laura Davis), which is more than double the lowest total of 14,850 (Jane Jones).

patient_id	patient_name	total_bill
P012	Laura Davis	30052
P049	David Moore	23555
P016	Michael Taylor	22967
P036	Michael Wilson	21583
P025	Robert Wilson	19512
P005	David Wilson	18610
P035	David Wilson	18408
P048	Emily Miller	17083
P010	Michael Taylor	15929
P017	Jane Jones	14850

VENKATA SAHITH KOTARU
SAHITHKOTARU@GMAIL.COM

25.Doctors contributing most revenue

```
select d.doctor_id,concat(d.first_name,' ',d.last_name)
as doctor_name,sum(b.amount) as revenue from doctors d
join appointments a on d.doctor_id=a.doctor_id
join treatments t on a.appointment_id=t.appointment_id
join billing b on t.treatment_id=b.treatment_id
group by d.doctor_id,doctor_name
order by revenue desc limit 10;
```

doctor_id	doctor_name	revenue
D005	Sarah Taylor	82694
D006	Alex Davis	69585
D001	David Taylor	66586
D002	Jane Davis	59802
D008	Linda Brown	53426
D003	Jane Smith	52793
D010	Linda Wilson	49436
D007	Robert Davis	40167
D004	David Jones	39317
D009	Sarah Smith	37439

Insights:

- The top four doctors earned over $59,000 each, while the rest earned less than $54,000.
- Three doctors share the last name "Davis," making it the most common name in the list.

VENKATA SAHITH KOTARU
SAHITHKOTARU@GMAIL.COM

Conclusion

This project helped me to understand how the hospital operates through data.

Using SQL, I analyzed patient records, doctor performance, treatments, and billing details.

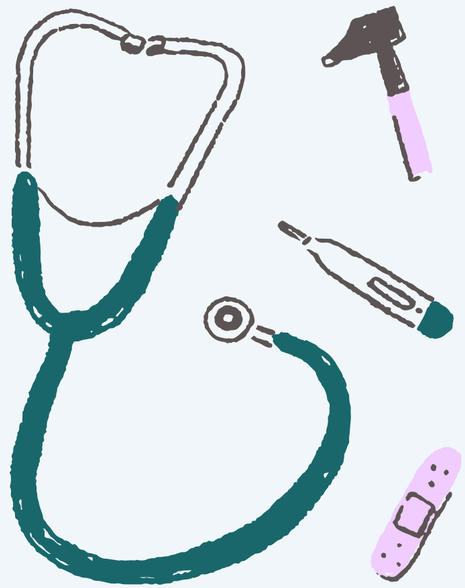
The insights show which doctors and treatments perform best, how patients use services, and where improvements can be made.

Overall, this analysis helps the hospital plan better, reduce waiting times, and improve service quality.

VENKATA SAHITH KOTARU
SAHITHKOTARU@GMAIL.COM

Tools Used

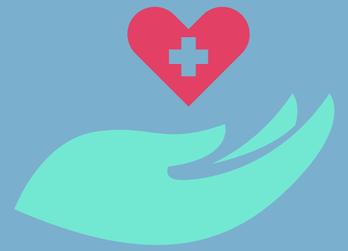
- Database: MySQL
- Techniques: Joins, Aggregate Functions, Subqueries, Case Statements
- Data Preparation: Cleaning & Validation using SQL
- Analytics: KPI Calculation & Trend Analysis



Future Scope

- Integrate with Power BI or Tableau to visualize key performance indicators (KPIs).
- Automate SQL queries for real-time hospital performance reports.
- Implement predictive models to forecast patient inflow and treatment demand.
- Expand dataset with staff details, patient feedback, and regional analysis for deeper insights.

VENKATA SAHITH KOTARU
SAHITHKOTARU@GMAIL.COM

THANK YOU 

For your attention