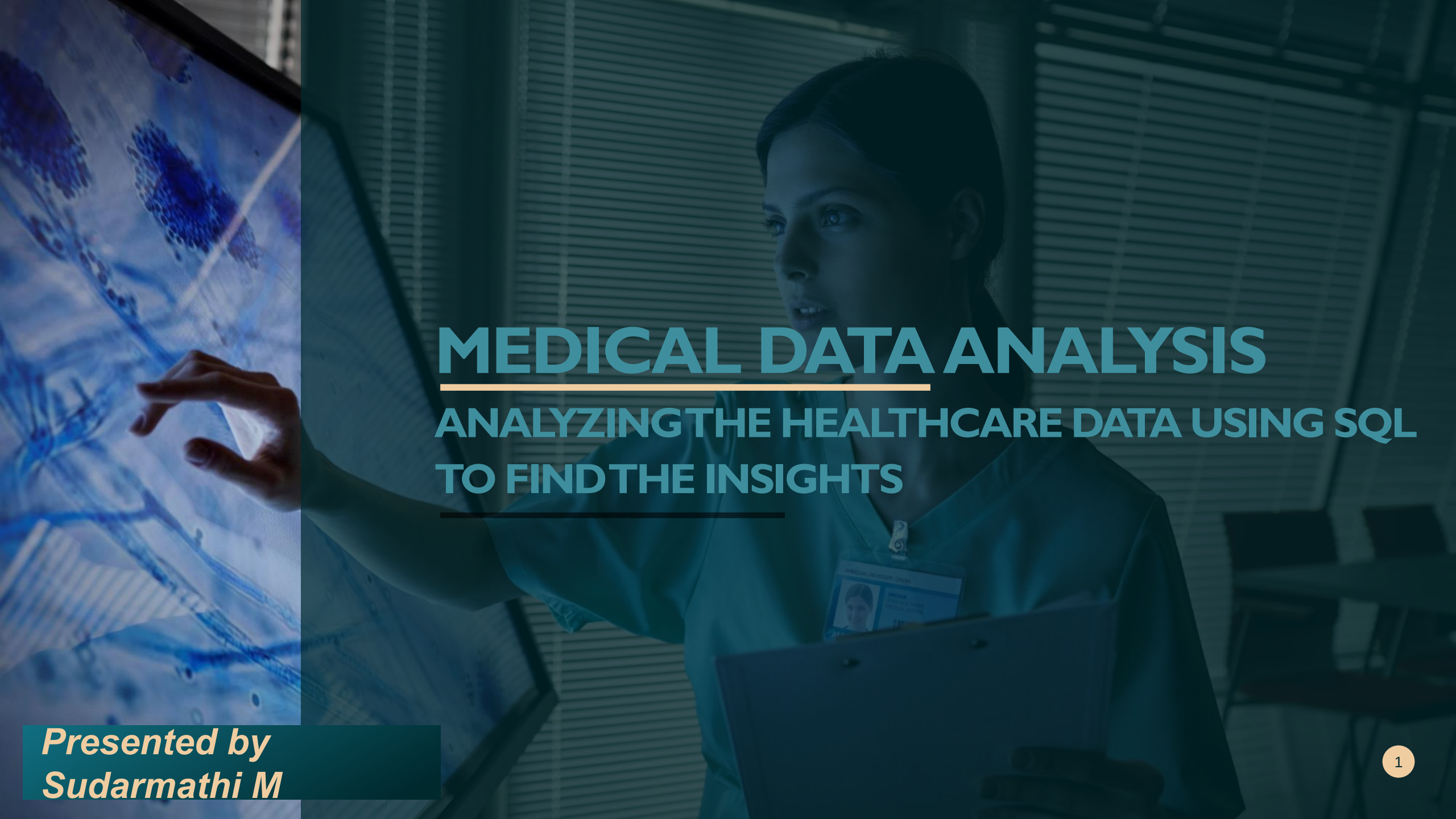


MEDICAL DATA ANALYSIS

ANALYZING THE HEALTHCARE DATA USING SQL TO FIND THE INSIGHTS

Presented by
Sudarmathi M

CONTENT

SUMMARY

This SQL project focuses on analysing Medical Data history to extract meaningful insights using various SQL techniques. By querying the patients details, attending doctors, admissions we aim to understand common allergies, frequent diagnosis, admission patterns, geographical concentration of patients also Doctor- patient relationship . This project involves filtering, calculating and searching data to answer targeted questions.

ABOUT THE DATASET

Over 10,000 records inside 4 entities

Admissions

- Patient ID
- Admission Date
- Discharge Date
- Diagnosis
- Attending Doctor id

Doctors

- Doctor id
- First name
- Last name
- Speciality

Patients

- Patient id
- First name
- Last name
- Gender
- Birth date
- City
- Province id
- Allergies
- Height
- Weight

Province Name

- Province Id
- Province Name

* According to a Survey

TOOLS AND TECHNOLOGIES

01 MySQL Workbench : Primary query engine for complex data analysis

02 SQL : Core language for all Data transformations and insights

03 Excel : Basic preview and basic analysis of dataset. It helps to understand

* According to a Surve

DATA CLEANING

Handling Missing Values:

Using conditional formatting we can easily spot missing values.
These values can be imputed with techniques like mean, median, mode.

Removing Duplicate :

Duplicate records can be identified and removed to ensure data accuracy and prevent skewed analysis.

Check Data Type :

Converting data types from text to Numeric for regression analysis ensures consistency and allows for accurate computations and analysis.

*Based on 1st year projections

DATA ANALYSIS

1 Extract Dementia Patients List:

```
SELECT
   pat.patient_id, pat.first_name,
pat.last_name,
ad.diagnosis
FROM patients pat join admissions
ad ON pat.patient_id =
ad.patient_id
WHERE ad.diagnosis= 'Dementia';
```

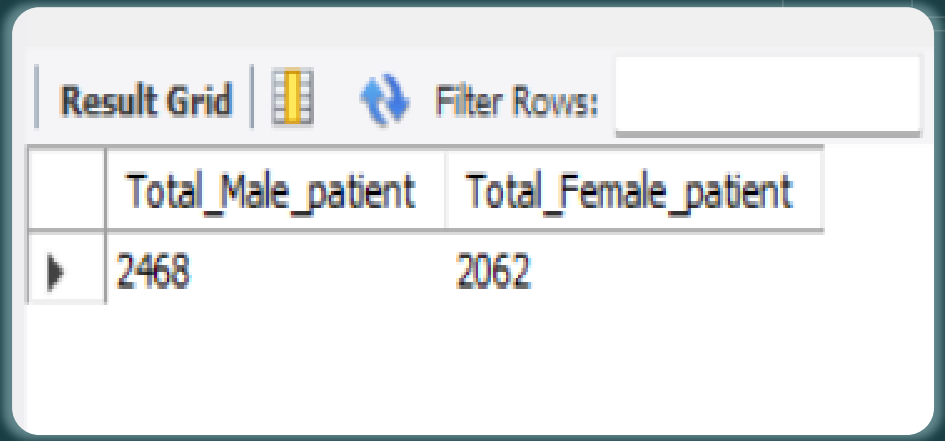
patient_id	first_name	last_name	diagnosis
160	Miranda	Delacour	Dementia
178	David	Bustamonte	Dementia
207	Matt	Celine	Dementia
613	Jaki	Granger	Dementia
836	Montana	Vimes	Dementia
924	Simon	Spellman	Dementia
1201	Irene	Murphy	Dementia
1264	Jillian	Valentine	Dementia
1402	Kathryn	Hallow	Dementia
1491	Doris	McGrew	Dementia
1585	Alex	Cantropus	Dementia
1749	Alejandro	Mellie	Dementia
1798	Sister	Trenton	Dementia

* According to a Survey

2

Show the total amount of male patients and the total amount of female patients in the patient's table. Display the two results in the same row :

```
SELECT
  COUNT(CASE WHEN gender = 'M' THEN
1 END ) AS Total_Male_patient,
  COUNT(CASE WHEN gender = 'F' THEN
1 END ) AS Total_Female_patient
FROM patients;
```



* According to a Survey

3 Show first name, last name and role of every person that is either patient or doctor. The roles are either "Patient" or "Doctor":

```
SELECT first_name, last_name,
'Doctors' AS role from Doctors
UNION ALL
SELECT first_name, last_name,
'Patients' AS Role
FROM patients;
```

first_name	last_name	role
Simon	Santiago	Doctors
Heather	Beck	Doctors
Jenny	Pulaski	Doctors
Jeanette	Sites	Doctors
Larry	Miller	Doctors
Donna	Greenwood	Doctors
Donald	Waterfield	Patients
Mickey	Baasha	Patients
Jiji	Sharma	Patients
Blair	Diaz	Patients
Charles	Wolfe	Patients
Sue	Falcon	Patients
Thomas	ONeill	Patients

* According to a Survey

4

We want to display each patient's full name in a single column. Their last name in all upper letters must appear first, then first name in all lowercase letters. Separate the last name and first name with a comma. :

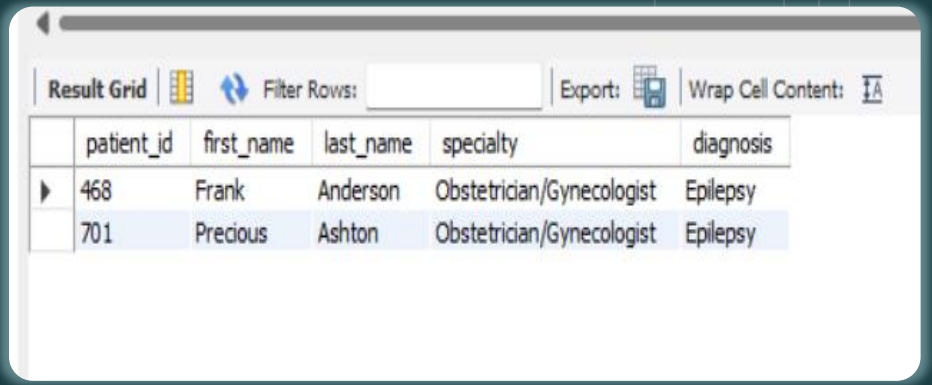
```
SELECT
 CANCAT
(UPPER(last_name),',',LOWER(first_n
ame)) AS "Patient Name"
FROM patients
ORDER BY LOWER(first_name) DESC;
```

Patient Name
MILLER,zoe
CORBIE,ziva
KOBAYAKAWA,zenigata
OVERSTREET,zenigata
BENNETT,zen
MEPHESTO,zelda
MORRIS,zelda
THOMAS,zefram
FLUTE,zefram
MARONEY,zatanna
TYRELL,zatanna
CHE,zane
WONG,zane
DREW,zack
BAKSHI,zachary
PRESTON,yusuke
CHURCHILL,yuri
LAZARUS,yuri
PROVENZA,yuri
RANDALL,yuko

* According to a Survey

5 Show patient id, first name, last name, and attending doctor's specialty. Show only the patients who has a diagnosis as 'Epilepsy' and the doctor's first name is 'Lisa'.

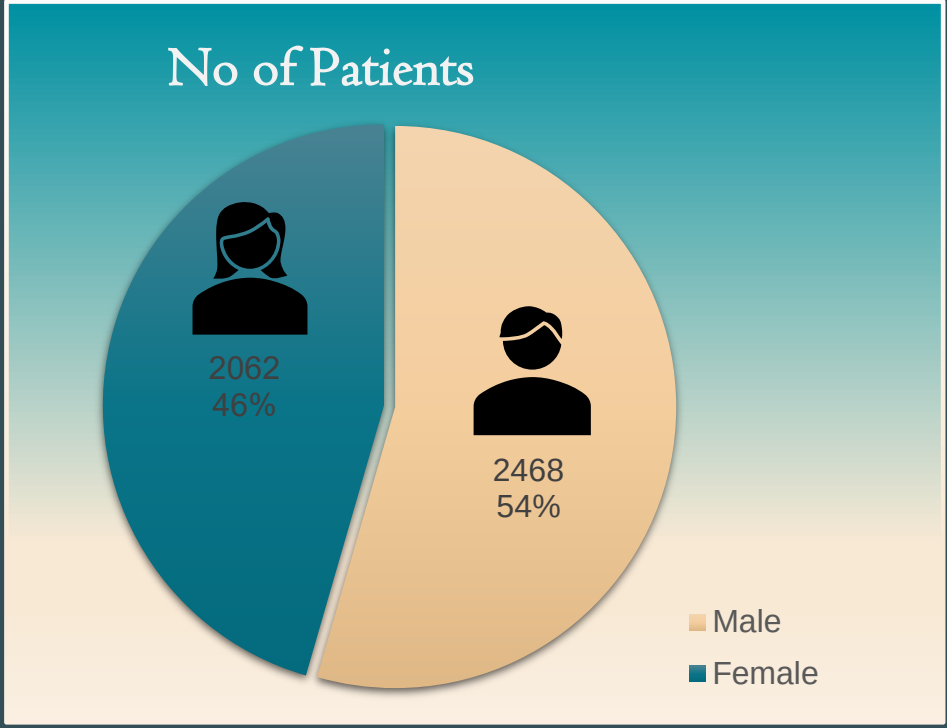
```
SELECT pat.patient_id, pat.first_name,
pat.last_name,
doc.specialty,ad.diagnosis
FROM patients pat
JOIN admissions ad ON ad.patient_id =
pat.patient_id
JOIN doctors doc ON
ad.attending_doctor_id = doc.doctor_id
WHERE ad.diagnosis = 'Epilepsy' AND
doc.first_name = 'Lisa';
```



patient_id	first_name	last_name	specialty	diagnosis
468	Frank	Anderson	Obstetrician/Gynecologist	Epilepsy
701	Precious	Ashton	Obstetrician/Gynecologist	Epilepsy

* According to a Survey

ANALYSIS INSIGHTS



⚠ Patients without any allergies:

2,059 A significant portion (45%) of patients reported no known allergies, which may streamline certain treatment protocols.

Weight Analysis :

Patients weighing between 100–120 kg:
952 Over 20% of patients fall into this weight range, indicating a need to focus on weight-related healthcare strategies and potential risks associated with obesity.

* According to a Survey

ANALYSIS INSIGHTS

👶 **Pediatric & Rare Cases:**

Patients born in 2010: 55Young patients form a small subset requiring pediatric care protocols.

📏 **Exceptionally tall patient:**

Name: Sam HarukoHeight: 220 cm A rare outlier in height data, which may require custom medical attention.

🕰️ **Historical Range of Records:**

Patient birth years range from 1918 to 2018The data spans a full century, offering broad generational insights into health patterns.

* According to a Survey

13

ANALYSIS INSIGHTS

Multiple Admissions for Same Diagnosis:

♻ 11 patients repeatedly admitted for issues like:

- Pregnancy
- Congestive Heart Failure
- Appendicitis
- Severed Spine
- Shoulder Surgery

* According to a Survey

14

CONCLUSION

The analysis of our Medical Data History reveals critical patterns in patient demographics, health conditions, and treatment trends.

Key takeaways include:

- A nearly balanced gender distribution.
- A large group of patients with no known allergies, which may simplify certain medical procedures.
- A considerable proportion of patients falling in the 100–120 kg weight range, emphasizing the importance of obesity prevention programs.
- Identification of repeat admissions for serious medical conditions, indicating the need for improved chronic care management.

These insights can guide more data-driven decisions in resource planning, patient care, and targeted healthcare interventions. By leveraging this analysis, we can move toward a more efficient, personalized, and proactive healthcare delivery system

* According to a Survey

15