Star Schema

Address

Dimension
Patient

Patient_visit

- visit_id (PK)
- Patient_id (FK to patient)
- Doctor_id (FK to doctor)
- Diagonis_id (FK to digres)
- treatment_id (FK to treatment)
- Date_id (FK to date)
- Cost_id (FK to billing)
- Length_stay
- total cost

Fact Table.

Specialty

Dimension
Doctor

Dimension
Diagnosis
/BED

**Disease
Category.**

Dimension
Treatment

Payment
Method

Dimension
Date.

Type

Queries.

1. count(*) on patient-visit group by doctor-id & date-id.

2. AVG (Length-Stay) grouped by diagnoses-id. & date-id.

3. Select (treat-id) Max (total_cost) AS H-c FROM (Patient-visit)

4. count (patient_id). group by doctor-id & date-id.

[Group by Treat-id order by H-C Desc limit 1]

Data Preprocessing / Cleaning Techniques.

Handling Missing.

Normalization / Scaling.

Encoding categorical Data (variable)

Feature Extraction.

Data Splitting.

Extract : From multiple resources.
Load : To use for analysis / insights (BI).
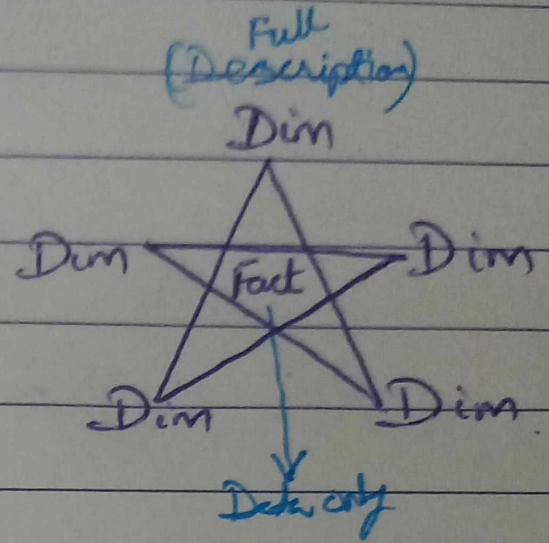Transform : Normalize, missing, (Clean).

09/02/2025.

Data Modeling

Establishing Relationships between tables.

Star Schema.

Type of DM technique used in data warehousing to represent data in a structured way.

Star Flake Schema.

Normalized version of star schema.