



**SHRI VILEPARLE KELAVANI MANDAL'S
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING**
(Autonomous College Affiliated to the University of Mumbai)
NAAC ACCREDITED with "A" GRADE (CGPA : 3.18)
DEPARTMENT OF INFORMATION TECHNOLOGY

**COURSE CODE: DJS22ITL503****DATE:20-08-24****COURSE NAME: Data Warehousing and Mining****CLASS: T Y B.TECH****NAME: ANISH SHARMA****Roll no.: I011****LAB EXPERIMENT NO.02****CO/LO:** Design a data warehouse model for a given problem statement**AIM / OBJECTIVE:**Design a real life problem statement and draw star scheme using erdplus tool**DESCRIPTION OF EXPERIMENT:**

In this experiment, the objective is to design a data warehouse model for a hospital management system. The model will consolidate data from various sources, including patient records, treatments, medications, and hospital departments. The star schema design approach will be employed to create a central fact table for patient visits and dimension tables for patients, doctors, treatments, and departments. The schema will be visualized using the ERDPlus tool to depict the relationships between the fact and dimension tables.

INPUT DATA / DATASET:**PROCEDURE / ALGORITHM:****TECHNOLOGY STACK USED:**ERDPlus tool**SOURCE CODE (OPTIONAL):****OBSERVATIONS / DISCUSSION OF RESULT:**

Identify real-life problem statement for building a data warehouse. Include following details in the document.

1. Introduction

The hospital management system needs a robust data warehouse to integrate patient records, treatments, medications, and departmental data. This system will enable efficient data retrieval, analysis, and decision-making, improving patient care and operational efficiency through a well-structured star schema design.



**SHRI VILEPARLE KELAVANI MANDAL'S
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING**
(Autonomous College Affiliated to the University of Mumbai)
NAAC ACCREDITED with "A" GRADE (CGPA : 3.18)



2. Need for building a data warehouse for a given problem statement.

Building a data warehouse for hospital management centralizes data from various sources, such as electronic health records and billing systems, into a single repository. This consolidation improves data quality and consistency, enabling advanced reporting and in-depth analysis. Key benefits include enhanced patient care through a comprehensive view of health data, improved operational efficiency by optimizing resource use and workflow, and better strategic decision-making with accurate performance metrics. Additionally, it supports regulatory compliance, data security, and enables predictive analytics for future planning. Overall, a data warehouse facilitates efficient data management, supports better decision-making, and enhances both patient outcomes and hospital operations.

3. Features of data warehouse for a given problem statement.

1. **Data Integration:**

- **Consolidate Data Sources:** Integrate data from various sources such as hospital management systems, billing systems, and inventory systems.
- **ETL Process:** Extract, Transform, and Load (ETL) processes to clean and integrate data from multiple operational systems into the data warehouse.

2. **Historical Data Storage:**

- **Temporal Data:** Store historical data to analyze trends over time, such as revenue growth, patient volume, and operational costs.

3. **Advanced Analytics:**

- **Aggregation and Summarization:** Provide aggregated views of data, such as total revenue by month or the number of beds occupied by quarter.
- **Trend Analysis:** Enable trend analysis on patient demographics, room utilization, and financial metrics.

4. **Data Accessibility:**

- **Reporting and Dashboards:** Develop dashboards and reports for hospital management to monitor key performance indicators (KPIs) and operational metrics.
- **Ad-hoc Querying:** Allow users to perform ad-hoc queries to explore data from different angles.

5. **Data Quality and Consistency:**

- **Data Validation:** Implement data quality checks to ensure accuracy and consistency.



**SHRI VILEPARLE KELAVANI MANDAL'S
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING**
(Autonomous College Affiliated to the University of Mumbai)
NAAC ACCREDITED with "A" GRADE (CGPA : 3.18)



- **Standardization:** Standardize data formats and definitions across the data warehouse.
 - 6. **Scalability and Performance:**
 - **Scalability:** Ensure the data warehouse can handle growing data volumes and increasing user queries.
 - **Performance Optimization:** Optimize query performance through indexing, partitioning, and efficient data storage techniques.
 - 7. **Security and Privacy:**
 - **Access Control:** Implement access control mechanisms to protect sensitive patient and financial data.
 - **Data Encryption:** Encrypt sensitive data both at rest and in transit to ensure confidentiality.
 - 8. **Data Warehouse Architecture:**
 - **Data Storage:** Use appropriate data storage solutions such as relational databases, cloud storage, or a hybrid approach.
 - **Data Modeling:** Employ proper data modeling techniques to ensure data integrity and support complex queries.
4. List the dimensions and facts and give their significance for a given problem statement.
1. **Time Dimension**
 - **Significance:**
 - **Temporal Analysis:** Allows for analysis of data over various time periods such as days, weeks, months, quarters, and years. This helps in identifying trends and patterns in hospital operations.
 - **Performance Tracking:** Enables the tracking of performance metrics over time, such as revenue growth, patient admission rates, and operational costs.
 - **Seasonality and Trends:** Facilitates the study of seasonal variations and long-term trends in hospital metrics.
 2. **Patients Dimension**
 - **Significance:**
 - **Patient Segmentation:** Provides detailed information about patients, which can be used to segment and analyze patient demographics, diseases, and medical histories.
 - **Treatment Analysis:** Helps in analyzing treatment outcomes, disease prevalence, and patient demographics.
 - **Personalized Services:** Supports personalized treatment plans and targeted healthcare services based on patient profiles.
 3. **Employees Dimension**



**SHRI VILEPARLE KELAVANI MANDAL'S
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING**
(Autonomous College Affiliated to the University of Mumbai)
NAAC ACCREDITED with "A" GRADE (CGPA : 3.18)



- **Significance:**
 - **Staff Performance:** Allows for analysis of staff performance, qualifications, and salary, aiding in human resource management and planning.
 - **Role Analysis:** Facilitates understanding of the distribution of staff across various roles and departments, which can help in optimizing workforce allocation.
 - **Cost Management:** Supports analysis of payroll costs and budget planning.
- 4. **Rooms Dimension**
 - **Significance:**
 - **Room Utilization:** Provides information on room availability, size, and facilities, which is essential for managing room occupancy and planning.
 - **Cost Analysis:** Helps in analyzing the cost associated with different types of rooms and optimizing room allocation.
 - **Facility Management:** Assists in evaluating the quality and facilities of different rooms, which can impact patient satisfaction.
- 5. **Medicines Dimension**
 - **Significance:**
 - **Inventory Management:** Facilitates tracking of medicine inventory, including expiry dates and supply levels, to ensure timely procurement and avoid stockouts.
 - **Cost Analysis:** Helps in analyzing the cost of medicines and understanding expenditure on pharmaceuticals.
 - **Usage Patterns:** Enables analysis of medicine usage patterns, which can help in forecasting demand and managing supplies.
- 6. **Payment Dimension**
 - **Significance:**
 - **Revenue Tracking:** Tracks payment transactions and modes, providing insights into revenue sources and financial health.
 - **Financial Analysis:** Supports analysis of payment methods and their impact on hospital cash flow.
 - **Transaction Monitoring:** Helps in monitoring and auditing financial transactions for accuracy and compliance.

Facts

1. **Revenue**
 - **Significance:**
 - **Financial Performance:** Measures the total revenue generated by the hospital, which is a key indicator of financial health.



**SHRI VILEPARLE KELAVANI MANDAL'S
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING**
(Autonomous College Affiliated to the University of Mumbai)
NAAC ACCREDITED with "A" GRADE (CGPA : 3.18)



- **Revenue Trends:** Helps in tracking revenue trends over time, assessing the impact of various factors on financial performance.
- **Profitability Analysis:** Supports analysis of profitability and financial sustainability.

2. Number of Beds Occupied

○ **Significance:**

- **Capacity Utilization:** Indicates how effectively the hospital's bed capacity is utilized, which can impact operational efficiency and patient care.
- **Resource Planning:** Assists in planning for future capacity needs and optimizing resource allocation.
- **Operational Metrics:** Provides insights into patient flow and hospital occupancy rates.

3. Operational Cost

○ **Significance:**

- **Cost Management:** Measures the total operational costs of running the hospital, including administrative and clinical expenses.
- **Budget Analysis:** Supports budget planning and cost control measures by providing detailed insights into various cost components.
- **Efficiency Improvement:** Helps in identifying areas where operational costs can be reduced or optimized.

4. Number of Patients

○ **Significance:**

- **Patient Volume Analysis:** Tracks the total number of patients served by the hospital, which is essential for evaluating demand and capacity.
- **Service Utilization:** Provides insights into the utilization of hospital services and patient load trends.
- **Resource Allocation:** Aids in planning for staffing, bed allocation, and other resources based on patient volume.

5. Draw the information package.



**SHRI VILEPARLE KELAVANI MANDAL'S
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING**
(Autonomous College Affiliated to the University of Mumbai)
NAAC ACCREDITED with "A" GRADE (CGPA : 3.18)



Subject: Hospital Data Warehouse						
Facts: Revenue, No of Beds Occupied, Operational Cost, No of patients						
Dimensions	Attributes					
Time	Year	Quarter	Month	Week	Day	
Patients	Name	Age	Gender	Disease	Medical Records	Patient_ID
Employees	Name	Age	Designation	Qualifications	Salary	Employee_ID
Rooms	Availability	Size	Cost	Beds	Facilities	Room_No
Payment	Mode of Payment	Date	Transaction_No			
Medicines	Name	Date of Expiry	Cost	Type	Supply	

6. Create Dimension Tables and Fact Table.

Payment Dimension

Transaction no.
Mode of Payment
Date

Rooms Dimension

Number
Availability
Size
Cost
Number of beds
Facilities

Time Dimension

Year
Month
Week
Day

Patient Dimension

Patient id
Name
Age
Gender
Disease
Medical Records
Valid Insurance



**SHRI VILEPARLE KELAVANI MANDAL'S
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING**
(Autonomous College Affiliated to the University of Mumbai)
NAAC ACCREDITED with "A" GRADE (CGPA : 3.18)



Employee Dimension

Employee id
Name
Age
Designation
Qualification
Salary

Medicine Dimension

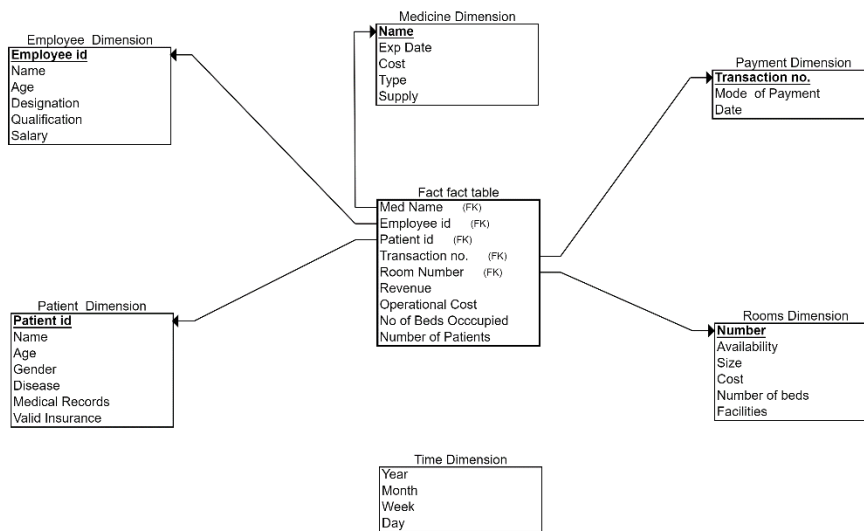
Name
Exp Date
Cost
Type
Supply

Fact fact table

Med Name (FK)
Employee id (FK)
Patient id (FK)
Transaction no. (FK)
Room Number (FK)
Revenue
Operational Cost
No of Beds Occupied
Number of Patients

7. Draw Star Schema and Snowflake schema.

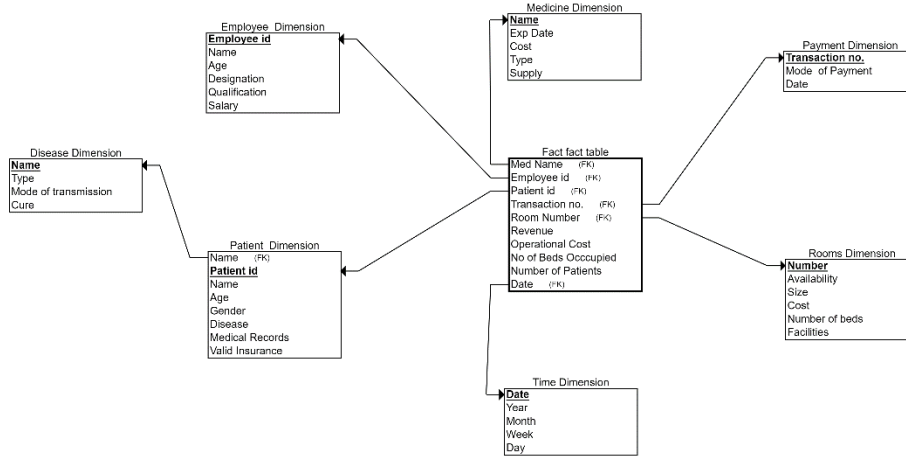
Star Schema:



Snowflake schema:



**SHRI VILEPARLE KELAVANI MANDAL'S
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING**
(Autonomous College Affiliated to the University of Mumbai)
NAAC ACCREDITED with "A" GRADE (CGPA : 3.18)



8. Example queries to extract data from datawarehouse.

1. "Can you tell me how many patient visits occurred in each department?"
2. "What is the average cost of treatments provided by each doctor?"
3. "How many unique patients received each type of treatment?"
4. "Show me the number of patient admissions for each month in the year 2023."
5. "Which are the top five most common diagnoses among patients?"
6. "How much has each patient spent on medications in total?"
7. "Give me the breakdown of patients by age group and gender within each department."
8. "What is the readmission rate for patients with chronic conditions, broken down by diagnosis?"



**SHRI VILEPARLE KELAVANI MANDAL'S
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING**
(Autonomous College Affiliated to the University of Mumbai)
NAAC ACCREDITED with "A" GRADE (CGPA : 3.18)



CONCLUSION:

The hospital data warehouse, structured using a star schema, provides efficient data retrieval and analysis capabilities. It supports critical decision-making by enabling quick access to comprehensive patient, treatment, and operational data, ultimately enhancing patient care, optimizing resource allocation, and improving overall hospital management.

REFERENCES:

- [1] Ponniah P., “Data Warehousing: Fundamentals for IT Professionals”, 2nd Edition, Wiley India, 2013.
- [2] Ageed, Z. S., Zeebaree, S. R., Sadeeq, M. M., Kak, S. F., Yahia, H. S., Mahmood, M. R., & Ibrahim, I. M. (2021), “Comprehensive survey of big data mining approaches in cloud systems”, Qubahan Academic Journal, 1(2), 29-38.

Website References:

Author's Last Name, First Initial. Middle Initial. (Date of Publication or Update). Title of work. Site name. Retrieved Month Day, Year, from URL from Homepage

- [3] U.S. Census Bureau. U.S. and world population clock. U.S. Department of Commerce. Retrieved July 3, 2019, from <https://www.census.gov/popclock>.