Assignment 1

Q1 Difference between OLTP System and Data Warehouse?

	OLTP	DW
1)	Transactional data (current)	Data Analysis (historical)
2)	Stores detailed data	Stores summarized data
3)	Data is dynamic (insert, update)	Data is largely static (no updates)
4)	Transaction are repetitive	Ad hoc reporting
5)	Application-oriented design	Subject-oriented design.
6	It is product for real-time business operations	~~It can be optimized~~ It is producted for analysis of business measures by categories and attributes.
7	It provides thousands of concurrent users.	It provides some concurrent users relative to OLTP.

Q2 Explain the characteristics of Data warehouse?

⇒ Characteristics of DW :-

1. Subject-Oriented :-
   Data warehouse are design around major subjects or areas of interest, such as sales, finance, inventory, and customer information, rather than the day-to-day operation.

2. Integrated :-
   Data from various sources is cleaned,

transformed, and integrated to provide a unified view.
This integration ensures that the data is consistent in terms of naming conventions, formats, and structures, which is crucial for accurate analysis.

2. Time - Variant :-
Data warehouses store historical data, allowing for analysis of trends over time. This is different from operational database, which typically store current data.

4. Non- Volatile :-
Once data is entered into the data ware house, it is not updated or deleted. This ensures a stable, consistent environment for historical analysis.

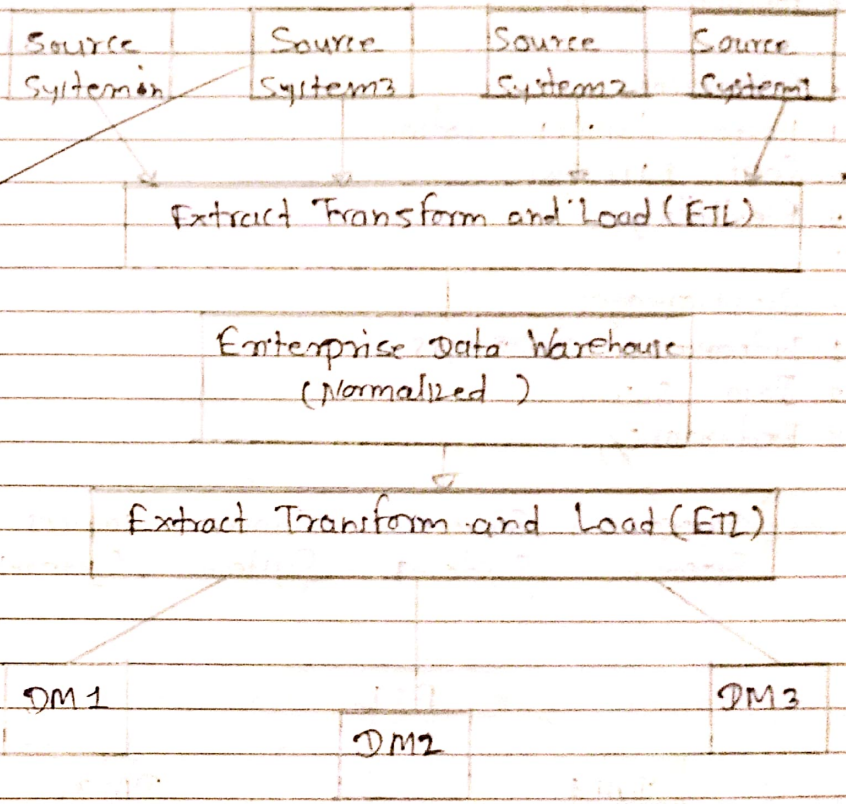
Q3. What is top down and bottom - up development methodology in data warehouse?

→ • Top-Down Development Methodology :-
Top-Down methodology, also known as the enterprise data warehouse (EDW) approach, involves creating a comprehensive, unified data warehouse before developing individual data marts.

• Key features :
1 Centralized Data Warehouse
2 Subject - oriented
3 ETL Process
4 Data Marts :

* Advantages :
1 Unified View
2 Data Consistency
3 Scalability

* Disadvantage :
1 Time - Consuming
2 Costly
3. Complexity



Top-down Design Approach.

• Bottom-Up Development Methodology

Bottom-up methodology, also known as the data mart approach, involves building individual data mart for specific business units or functions and then integrating them into a comprehensive data warehouse.

- Key features:

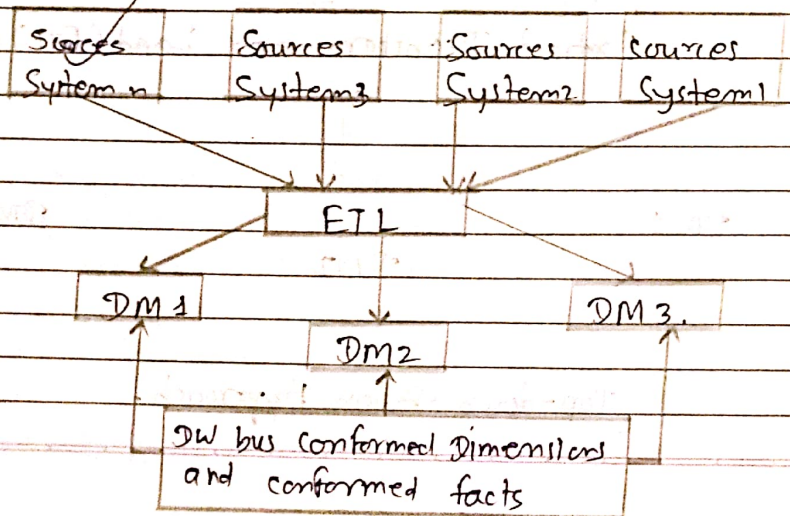
1 Independent Data Marts
2 Incremental Development
3. Focused Scope

• Advantages

1 Faster Implementation
2. Cost-Effective
3. Flexibility

• Disadvantages

1 Integration challenges
2. Data Silos
3 Redundancy.



Q4. Which tools are used for data warehouse?

→ 1 Amazon Redshift :-

- Fully managed, scalable data warehouse service in the cloude
- Optimized for high-performance querying and analysis.

2. Google BigQuery :-

- Severless, highly scalable data warehouse service by Google Cloud.
- Support real-time analytics and extensive machine learning capabilities.

3. Snowflake : -

- Cloud-based data warehousing solution that separates compute and storage.

4. Microsoft Azure synapse Analytics :-

- Integrated analytics service that combines big data warehousing.

5. Oracle Exadata :-

:: High-performance data warehouse appliance by oracle.

Q5 What is project planning and management in data warehouse?

1 Project planning :-

- Define Objectives and scope

Clearly define the goals of the data warehouse project, such as improving decision-making,

integrating disparate data sources, or enhancing data analytics capabilities.

- stakeholder Identification and Requirements Gathering:

- Develop a project plan :-
  Break down the project into manageable takes and define milestones to track progress.

- Risk management :-
  List potential risk that could impact the project, such as data quality issues, integration challenges, or scope creep.

3. Project Management:
- Project Execution :-
  Ensure effective communication and coordination among team members.

- Data Integration and ETL Process :-
  Design and implement the ETL processes to gather data from various source, transform it into the desired and load it into the data warehouse.

- Testing and validation
- Deployment and maintenance.

30/7/24

Q1 Explain
warcho
→
Data
ope
sys
op
S
A
Se
1.
2.