Data Mining and Business Intelligence

Assignment 2

Q1. What is the purpose of Multidimensional Data Model?

Q2. Differentiate Between the Snowflake and Star Schema.

Q3. Describe the fact tables and dimension table.

Q4. Define in detail OLAP.

Q5. Differentiate Between ROLAP and MOLAP.

Ans 1. A multidimensional model views data in the form of a data-cube. A data cube enables data to be modeled and viewed in multiple dimensions. It is defined by dimensions and facts.

The dimensions are the perspectives or entities concerning which an organization keeps records. For example, a shop may create a sales data warehouse to keep records of the store's sales for the dimension time, item, and location. These dimensions allow the save to keep track of things, for example, monthly sales of items and the locations at which the items were sold. Each dimension has a table related to it, called a dimensional table, which describes the dimension further. For example, a dimensional table for an item may contain the attributes item_name, brand, and type.

A multidimensional data model is organized around a central theme, for example, sales. This theme is represented by a fact table. Facts are numerical measures. The fact table contains the names of the facts or measures of the related dimensional tables.

Ans 2.

- **Star schema :** In the STAR Schema, the center of the star can have one fact table and a number of associated dimension tables. It is known as star schema as its structure resembles a star. The star schema is the simplest type of Data Warehouse schema. It is also known as Star Join Schema and is optimized for querying large data sets.

- **Snowflake schema :** SNOWFLAKE SCHEMA is a logical arrangement of tables in a multidimensional database such that the ER diagram resembles a snowflake shape. A Snowflake Schema is an extension of a Star Schema, and it adds additional

dimensions. The dimension tables are normalized which splits data into additional tables.

Ans 3.

- Fact table :
    1. A fact table is a primary table in a dimensional model.
    2. A Fact Table contains
    3. Measurements/facts
    4. Foreign key to dimension table

- Dimension tabvle :
    1. A dimension table contains dimensions of a fact.
    2. They are joined to fact table via a foreign key.
    3. Dimension tables are de-normalized tables.
    4. The Dimension Attributes are the various columns in a dimension table
    5. Dimensions offers descriptive characteristics of the facts with the help of their attributes
    6. No set limit set for given for number of dimensions
    7. The dimension can also contain one or more hierarchical relationships

Ans 4.

- **OLAP :** Online Analytical Processing Server (OLAP) is based on the multidimensional data model. It allows managers, and analysts to get an insight of the information through fast, consistent, and interactive access to information. This chapter cover the types of OLAP, operations on OLAP, difference between OLAP, and statistical databases and OLTP.

  Types of OLAP Servers :

  We have four types of OLAP servers –

    1. Relational OLAP (ROLAP)
    2. Multidimensional OLAP (MOLAP)
    3. Hybrid OLAP (HOLAP)
    4. Specialized SQL Servers

Ans 5.

- **Relational OLAP :** ROLAP servers are placed between relational back-end server and client front-end tools. To store and manage warehouse data, ROLAP uses relational or extended-relational DBMS.

  ROLAP includes the following −

    1. Implementation of aggregation navigation logic.
    2. Optimization for each DBMS back end.

3. Additional tools and services.

- **Multidimensional OLAP :** MOLAP uses array-based multidimensional storage engines for multidimensional views of data. With multidimensional data stores, the storage utilization may be low if the data set is sparse. Therefore, many MOLAP server use two levels of data storage representation to handle dense and sparse data sets.