211720243002

BIG DATA ANALYTICS

Data Warehousing

Homework 7

Q1. What is a data warehouse? List the types of Data warehouse architectures.

A **data warehouse** is a type of data management system that is designed to enable and support business intelligence (BI) activities, especially analytics. Data warehouses are solely intended to perform queries and analysis and often contain large amounts of historical data.

There are three approaches to constructing a data warehouse:

Single-tier architecture, which aims to deduplicate data to minimize the amount of stored data

Two-tier architecture, which separates physical data sources from the data warehouse, making it incapable of expansion or supporting many end users.

Three-tier architecture:

- The bottom tier, the database of the data warehouse servers
- The middle tier, an online analytical processing (OLAP) server providing an abstracted view of the database for the end-user
- The top tier, a front-end client layer consisting of the tools and APis used to extract data

Q2. What does OLAP stand for?

OLAP stands for online analytical processing.

Q3. What does OLTP stand for?

OLTP stands for Online Transaction Processing.

Q4. What is a star schema?

A star schema is a database organizational structure optimized for use in a data warehouse or business intelligence that uses a single large fact table to store transactional or measured data, and one or more smaller dimensional tables that store attributes about the data.

Q5. What is a snow flake schema?

A snowflake schema is a multi-dimensional data model that is an extension of a star schema, where dimension tables are broken down into sub dimensions**. Snowflake schemas are commonly used for business intelligence and reporting in OLAP data warehouses, data marts, and relational databases.**

Q6. Define fact-less fact.

Fact less facts are those fact tables that have no measures associated with the transaction**.** Fact less facts are a simple collection of dimensional keys which define the transactions or describing condition for the time period of the fact.

Q7. What do you understand by dimensional modeling?

Dimensional Modeling (DM) is a data structure technique optimized for data storage in a Data warehouse**.** The purpose of dimensional modeling is to optimize the database for faster retrieval of data.

Q8. What is a data mart?

A data mart is a simple form of data warehouse focused on a single subject or line of business**.** With a data mart, teams can access data and gain insights faster, because they don't have to spend time searching within a more complex data warehouse or manually aggregating data from different sources.