UNIT I

1.Define data warehouse?

A data warehouse is a repository of multiple heterogeneous data sources organized under a unified schema at a single site to facilitate management decision making .(or)A data warehouse is a subject-oriented, time-variant and nonvolatile collection of data in support of management's decision-making process.

2.What are operational databases?

Organizations maintain large database that are updated by daily transactions are called operational databases.

3.Define OLTP?

If an on-line operational database systems is used for efficient retrieval, efficient storage and management of large amounts of data, then the system is said to be on-line transaction processing.

4.Define OLAP?

Data warehouse systems serves users (or) knowledge workers in the role of data analysis and decision-making. Such systems can organize and present data in various formats. These systems are known as on-line analytical processing systems.

5.How a database design is represented in OLTP systems?

Entity-relation model

6. How a database design is represented in OLAP systems?

- Star schema
- Snowflake schema
- Fact constellation schema

7.Write short notes on multidimensional data model?

Data warehouses and OLTP tools are based on a multidimensional data model.This model is used for the design of corporate data warehouses and department data marts. This model contains a Star schema, Snowflake schema and Fact constellation schemas. The core of the multidimensional model is the data cube.

8.Define data cube?

It consists of a large set of facts (or) measures and a number of dimensions.

9.What are facts?

Facts are numerical measures. Facts can also be considered as quantities by which can analyze the relationship between dimensions.

10.What are dimensions?

Dimensions are the entities (or) perspectives with respect to an organization for keeping records and are hierarchical in nature.

11.Define dimension table?

A dimension table is used for describing the dimension. (e.g.) A dimension table for item may contain the attributes item_ name, brand and type.

12.Define fact table?

Fact table contains the name of facts (or) measures as well as keys to each of the related dimensional tables.

13.What are lattice of cuboids?

In data warehousing research literature, a cube can also be called as cuboids. For different (or) set of dimensions, we can construct a lattice of cuboids, each showing the data at different level. The lattice of cuboids is also referred to as data cube.

14.What is apex cuboid?

The 0-D cuboid which holds the highest level of summarization is called the apex cuboid. The apex cuboid is typically denoted by all.

15.List out the components of star schema?

- A large central table (fact table) containing the bulk of data with no redundancy.
- A set of smaller attendant tables (dimension tables), one for each dimension.

16.What is snowflake schema?

The snowflake schema is a variant of the star schema model, where some dimension tables are normalized thereby further splitting the tables in to additional tables.

17.List out the components of fact constellation schema?

This requires multiple fact tables to share dimension tables. This kind of schema can be viewed as a collection of stars and hence it is known as galaxy schema (or) fact constellation schema.

18.Point out the major difference between the star schema and the snowflake schema?

The dimension table of the snowflake schema model may be kept in normalized form to reduce redundancies. Such a table is easy to maintain and saves storage space.

19.Which is popular in the data warehouse design, star schema model (or) snowflake schema model?

Star schema model, because the snowflake structure can reduce the effectiveness and more joins will be needed to execute a query.

20.Define concept hierarchy?

A concept hierarchy defines a sequence of mappings from a set of low-level concepts to higher-level concepts.