DATA WAREHOUSE CONCEPTS – FAQ's

1. What is OLAP and OLTP

Online Analytical Processing (OLAP). It is a category of software technology that enables the users to gain insight into data through fast, consistent, interactive access to a wide variety of possible views of information.

Online Transaction Processing (OLTP). It is a process which is used for day to day transaction processing. Example: Operational systems, High volume data collection.

2. What is meant by DWH?

Data Warehouse integrates and aggregates data from various operational and external database maintained by different Business Units.

3. Give four layers of data warehouse architecture.

a) Operational : Functions as data storage
b) Informational : Stores business logic
c) Data access : Acts as a bridge between operational and informational layer
d) Meta data : Stores data dictionary

4. What are the components of data warehouse?

The various components of Data Warehouse are:

a) Data Source: The data is sourced from transaction processing systems like manufacturing, ERP, Sales.

b) Data staging area: Data staging area is the storage area as well as set of ETL process that extract data from source system. It is everything between source systems and Data warehouse.
c) Data Marts: Data Mart is a logical subset of a Data Warehouse that may make it simpler for users to access key corporate data. A Data Mart has a smaller data model, users only need a piece of data from the data warehouse.
d) End users.

5. What is meant by Star Schema?

A Star schema is a database design that contains a central table, called a fact table, which is in relationship with many tables called dimension tables. It is very efficient in the performance. It is best suited for MOLAP application tools. Despite the fact that the Star schema is the simplest Data warehouse architecture, it is most commonly used in the Data warehouse implementations about 90-95%, across the world today.

6. What is ETL Processes?

ETL is Extraction, Transformation, and Load

Extraction: - In extraction process data extraction done from various sources.

Transformation: - In transformation process data convert to format required by data warehouse.

Load: - In load process data load to data warehouse.

7. What are the steps involve in transformation process.

The steps involved in transformation process are:

(1) Selecting only certain columns to load.
(2) Cleansing the data to remove duplicates and enforce consistency.
(3) Translating coded values.
(4) Encoding free-form values.
(5) Deriving a new calculated value.

8. What is the use of OLAP?

OLAP Analysis is used for:

- Aggregation: (total sales, percent-to-total)
- Comparison: Budget versus Expenses
- Ranking: Top 10, quartile analysis
- Access to detailed and aggregate data
- Complex criteria specification

9. What is Data Mining?

Data mining is the way of analyzing data by exploring large databases. It helps in understanding the business by extracting necessary information from the databases. It allows you to understand the pattern and helps in predicting the behavior of it. Data mining helps in increasing the business and forecasting the chunks related to it at early stages. It includes finding patterns that are suitable for the organization. Data Mining is also known as Knowledge Discovery in Databases.

10. What are the Emerging Trends in Business Intelligence and Data Warehouse?

- Data Quality
- Enterprise Integration
- Metadata Management
- Packaged BI/DW Solutions
- Open Source BI/DW
- Data Warehouse Appliances