



ID: 14105

Name: Hamid

Subject: Date Warehouse

Program: BS(cs)

Date: 29/06/2020.

Q1: Differentiate between OLTP vs OLAP

ANS:

OLTP(Online transaction processing)	OLAP(Online analytical processing)
Online transaction processing (OLTP) captures, stores, and processes data from transactions in real time.	Online analytical processing (OLAP) uses complex queries to analyze aggregated historical data from OLTP systems.
OLTP is an online database modifying system.	OLAP is an online database query answering system.
OLTP is a class of software programs capable of supporting transaction-oriented applications on the Internet.	OLAP often is used in data mining. The chief component of OLAP is the OLAP server, which sits between a client and a database management systems (DBMS).
OLTP database systems are commonly used for order entry, financial transactions, customer relationship management and retail sales via the Internet. Almost any business that has a large number of users who conduct short online transactions needs an OLTP system	OLAP a category of software tools that provides analysis of data stored in a database. OLAP tools enable users to analyze different dimensions of multidimensional data. For example, it provides time series and trend analysis views. OLAP often is used in data mining.

Q2: Differentiate between Expert system and DSS. (10)

ANS:

Expert system	DSS(Decision Support System)
An expert system is a problem-solving computer program that achieves good performance in a specialized problem domain that is considered difficult and requires specialized knowledge and skill.	A DSS is an interactive system that helps decision-makers utilize data and models to solve unstructured or semi-structured problems.
A knowledge-based information system (KBIS) that uses its knowledge about a specific, complex application to act as an expert consultant to end users	A Decision Support System (DSS) is an interactive computer-based system or subsystem intended to help decision makers use communications technologies, data, documents, knowledge and/or models to

	identify and solve problems, complete decision process tasks, and make decisions.
An Expert System (ES) that does such exploration and also makes a recommendation would be a Decision Making System.	Decision support systems are helpful for middle level executives to make decisions. However, Decision making tools are for every one in organization to make decision.
Expert systems, also called knowledge-based systems, neural networks used to simulate the expert knowledge and use it effectively to solve a particular problem.	The DSS is one of the most emblematic of the Business Intelligence tools because, among other properties, allow to solve many of the limitations of management programs.

Q3: What is the relation between datamining and data warehousing?

ANS:

data warehousing and data mining is that data warehousing is the process of compiling and organizing data into one common database, whereas data mining is the process of extracting meaningful data from that database.

Data mining is considered as a process of extracting data from large data sets, whereas a Data warehouse is the process of pooling all the relevant data together. Data mining is the process of analyzing unknown patterns of data, whereas a Data warehouse is a technique for collecting and managing data.

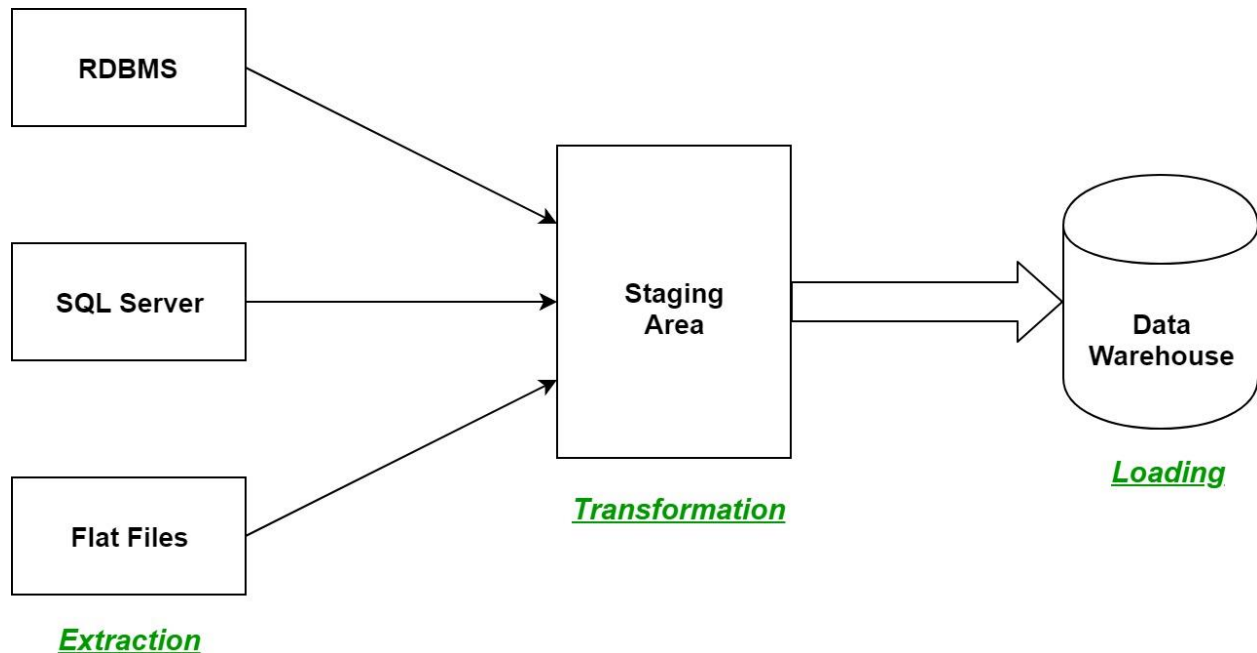
Data Warehouse is a relational database that is designed for query and analysis rather than for transaction processing. It usually contains historical data derived from transaction data.

Q4: Explain ETL process.

ANS:

ETL (Extract, Transform and Load):

ETL is a process in Data Warehousing and it stands for Extract, Transform and Load. It is a process in which an ETL tool extracts the data from various data source systems, transforms it in the staging area and then finally, loads it into the Data Warehouse system.



- Extraction:

The first step of the ETL process is extraction. In this step, data from various source systems is extracted which can be in various formats like relational databases, No SQL, XML and flat files into the staging area. It is important to extract the data from various source systems and store it into the staging area first and not directly into the data warehouse because the extracted data is in various formats and can be corrupted also. Hence loading it directly into the data warehouse may damage it and rollback will be much more difficult. Therefore, this is one of the most important steps of ETL process.

- Transformation:

The second step of the ETL process is transformation. In this step, a set of rules or functions are applied on the extracted data to convert it into a single standard format. It may involve following processes/tasks:

- ❖ Filtering – loading only certain attributes into the data warehouse.
- ❖ Cleaning – filling up the NULL values with some default values, mapping U.S.A ,United States and America into USA, etc.
- ❖ Joining – joining multiple attributes into one.
- ❖ Splitting – splitting a single attribute into multiple attributes.
- ❖ Sorting – sorting tuples on the basis of some attribute (generally key-attribute).

- Loading:

The third and final step of the ETL process is loading. In this step, the transformed data is finally loaded into the data warehouse. Sometimes the data is updated by loading into the data warehouse very frequently and sometimes it is done after longer but regular intervals. The rate and period of loading solely depends on the requirements and varies from system to system.