

NAME

HAMMAD PIR

ID

6961

SUBJECT

DATA WARE HOUSE

DATE

29-06 2019

Q:1: Differentiate between OLTP vs OLAP?

Ans: **(ONLINE TRANSACTION PROCESS) OLTP:**

Online transaction processing provides transaction-oriented applications in a 3-tier architecture. OLTP administers day to day transaction of an organization.

Examples – Uses of OLTP are as follows:

- ATM center is an OLTP application.
- OLTP handles the ACID properties during data transaction via the application.
- OLTP handles the ACID properties during data transaction via the application.
- It's also used for Online banking, Online airline ticket booking, sending a text message, add a book to the shopping cart.

Online Analytical Processing (OLAP):

Online Analytical Processing consists of a type of software tools that are used for data analysis for business decisions. OLAP provides an environment to get insights from the database retrieved from multiple database systems at one time.

EXAMPLE: Any type of Data warehouse system is an OLAP system.

Uses of OLAP are as follows:

- Spottily analyzed songs by users to come up with the personalized homepage of their songs and playlist.
- Netflix movie recommendation system.

Q2: Differentiate between Expert system and DSS?

Ans: Expert System (ES) :

- ES is based on simple rule-based logic.
- Problem is completely defines'
- There is clear way for the solution method.
- ES represent precisely what is needed, the extraction of the expertise from those who know and making that knowledge available to those who don't know, with very positive additional connotations of top-down technology transfer within organizations.
- **Example:**

Locating critical areas for non-point leakage of nitrogen and phosphorus. The principle is to use GIS and expert systems to integrate landscape concept which consider hydrological and hydrochemical processes into account.

Decision Support System (DSS):

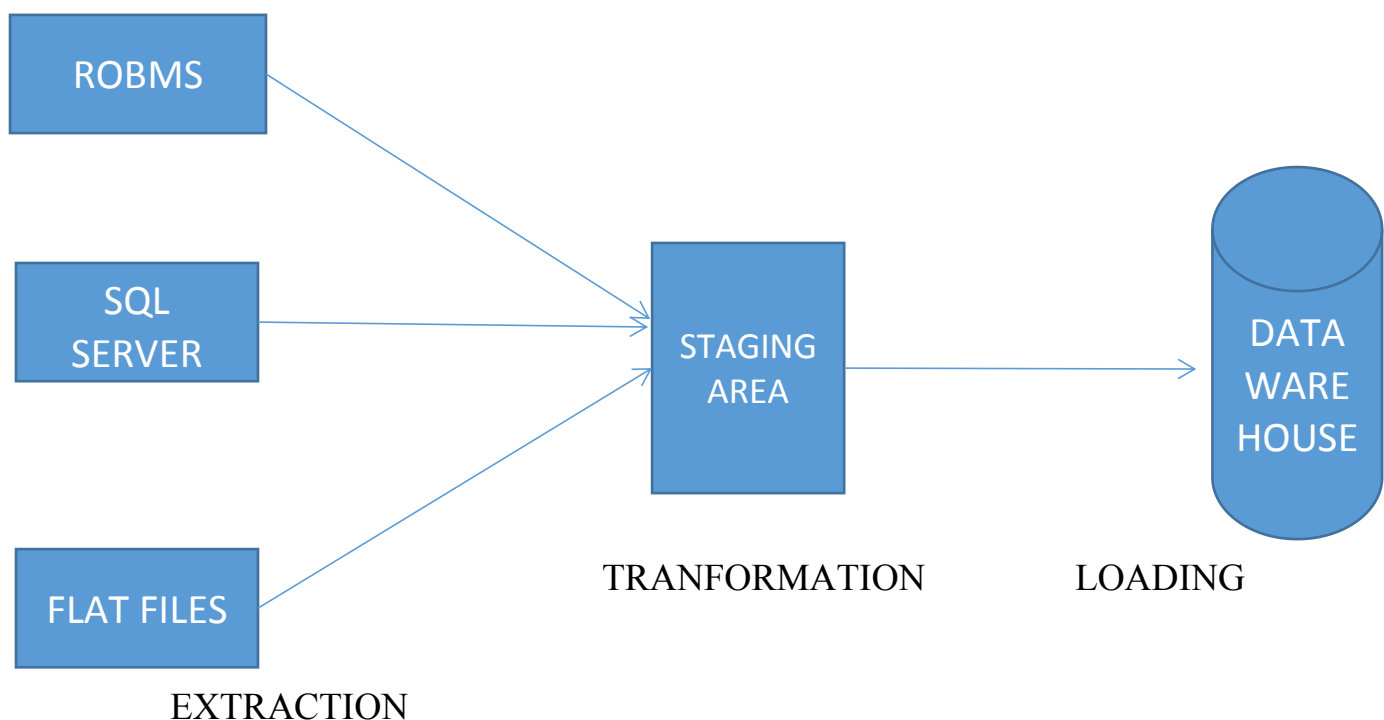
- The problem is open-ended
- The evaluation required to solve it is also incompletely defined/ill-defined problems. Its characteristics: the solution involving a mixture of methods and dependent on the perspective of the user
- One way of method: Multi Criteria Evaluation in IDRISI Andes software

- DSS has flexibility in the form of choices of data, procedures, and displays
- Example: Use MCE in ADRISI Andes to determine the best site location for ecovillage

Q3: Explain ETL process?

ANS:

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.



- When used with an enterprise data warehouse (data at rest), ETL provides deep historical context for the business.
- By providing a consolidated view, ETL makes it easier for business users to analyze and report on data relevant to their initiatives.
- ETL can improve data professionals' productivity because it codifies and reuses processes that move data without requiring technical skills to write code or scripts.

- ETL has evolved over time to support emerging integration requirements for things like streaming data.

Organizations need both ETL and ELT to bring data together, maintain accuracy and provide the auditing typically required for data warehousing, reporting .

Q4: What is the relation between datamining and data warehousing?

Ans: While closely related, both concepts have their own specific roles. Data mining is the automated process of analyzing large data sets to find these patterns, relationships and trends and ultimately to generate business insights – which will be used to solve challenges and identify new opportunities, so organizations can use past patterns to predict future behaviors and results. A data warehouse – where the data from the various sources is combined and stored – allows data mining to be used throughout the organization, from sales and marketing applications to research, product development and finance. Data warehousing and data mining are the cornerstones of modern business decisions: How Your Data Warehouse Can Make Data Mining Easier and More Efficient? Data mining techniques can be carried with any traditional database, but because a data warehouse contains quality data that has already been sanitized and tested, it makes sense to have data mining over a data warehouse system. A data warehouse is a description for specific server and storage capacities, mostly used to store big and/or unstructured data. The idea is that data is stored in a easy to find and easy to extract way - like goods in the shelf's of a warehouse. Data mining is a collection of techniques to find meaningful relationships in large data sets. These relationships, such as correlations, can be used for diverse means. So, you'd first store data in a data warehouse (or somebody does that for you). Now

this data can be downloaded and used for data mining.
Starting by loading data into data warehouse, (you can build your data warehouse in any DB engine, only you should respect some norms to build your star or snowflake, or constellation schema the schema (schema = Fact tables linked into Dimension tables).
Then you can extract the meaning of this data, or the information behind it by using Data mining. Data warehouse is not mandatory to do data mining; you can extract information using DM only by using Excel file, or any other data format. The link between "Data Mining" and "Data Warehousing" is the same as the link between metal mining and gathering metal bearing ore in a place and format conducive to easy processing.
Data mining requires *data preparation*. A data warehouse (DW) supports and performs the *data preparation* processes to support the one or more data mining projects.