

DAY: _____

DATE: _____

Name :- Abdullah Tahir

ID :- 13789

Subject :- Dataware housing :-

Question 1

OLAP :-

Online Analytical processing, a category of software tool which provide analysis of data for business decisions. OLAP system allow user to analyze database information from multiple database system at one time. The primary objective is data analysing and not data processing.

The benigits of using OLAP service are as follows

- * OLAP creates a single platform for all type of business analytical needs which includes planning, budgeting, forecasting and analysis
- * The main benifit of OLAP is the consistency of information and calculation.

• 1) Drawbacks :-

- 1) Implementation and maintance are dependant on IT professional because the traditionnal OLAP tools require a complicated modeling procedure
- 2) OLAP tools need cooperation between people of various department to be effective which might always be not possible

Any datawarehouse system is an OLAP system

* OLTP :-

Online transaction processing shortly known as OLTP supports transaction oriented application in a 3 tier architecture. OLTP administers day to day transaction of an organization. The primary objective is data processing and not data analysing.

* Benefits :-

- 1) It ~~administers~~ administers daily transaction of an organization
- 2) OLTP widens the customer base of an organization by ~~eg~~ simplifying individual process.

* Drawbacks :-

- 1) If OLTP system faces hardware failure, then online transaction get severely affected
- 2) OLTP system allow multiple users to access and change the same data at the same time which many time created unprecedented situation.

DAY: _____

DATE: _____

Question 2:-

Expert System vs DSS

* Expert System :-

- 1) Expert system is based on simple rule based logic
 - 2) Problem is completely defines.
 - 3) Expert system 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 with organization.
- The objective is to replicate humans.

* Decision System Support (DSS) :-

- 1) The problem is open-ended.
- 2) One way of method : multi Criteria Evaluation in IDRTSP
- 3) DSS has flexibility in the form of choices of data, procedures, and displays

DAY: _____

DATE: _____

* The evaluation required to solve it is also incompletely defined / ill-defined problems. Its characteristics, the solution involving a mixture of method and dependent on perspective of the user.

* The Objective is to assist human.

* The query in is Human queries.

* The DSS is complex and wide.

Question 3

• Relation between data mining and data warehousing :-

• The link between "Data mining" & "Data warehousing" is the same as the link between mental mining and gathering mental bearing are in a place and format conducive to easy processing.

• Data mining requires data preparation. A data warehouse supports and performs the data preparation processes to support the one or more data mining projects. 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.

DAY: _____

DATE: _____

Question 4 :-

ETL :-

1) Extraction :-

2) Transformation :-

3) Loading :-

* Extraction :-

The first step of the ETL process is extraction. In the step, data from various source system is extracted which can be in various formats like relational database, NoSQL, XML and flat files 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 roll back will be much more difficult. So it is important step of ETL process.

2) Transformation :-

The second step of the ETL process is transformation. In this step, a set of rule or functions are applied on the extracted

data to convert it into a single standard format. It may involve following processes.

- * **Filtering**: Loading only certain attributes into the data warehouse.
- * **Cleaning**: filling up the NULL values with some default values, mapping USA, United states of America into USA etc.
- * **Joining** :- joining multiple attributes into one.
- * **Splitting**: Splitting a single attribute.
- * **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. Sometime the data is updated by loading into the data warehouse very frequently and sometimes it is done after longer but regular intervals. The rate & period of loading solely

DAY: _____

DATE: _____

depends on the requirements and varies from system to system

RDBMS

