CDBM 280 Learning Outcome 1 Review

# Questions

1. Give 5 examples of DBMS software.
   1. Access, Oracle, SQLite, MySql, SQL
2. What is SQL and what are 3 components of SQL?
   1. Structured Query language, 4th gen language, most common language to interact with relational databases
3. What are some of the disadvantages of the early file-based systems for handling data?
   1. Seperaion and isolation of data, duplication of data, data dependence, fixed queries, incompatible formats
4. What are the advantages of a relational database management system? What are the disadvantages?
   1. Control data redundancy, Improved consistency, improved security, improved sharing
   2. Complexity, Cost of software/hardwarte, high impact of a failure, scalibility
5. Briefly describe the hierarchical and network DBMS structures. What were the disadvantages to these approaches?
   1. Hierarchical Model is a family tree, each node can only have 1 parent node(balanced binary tree)
   2. Network is a family tree but each node can have more than 1 parent
6. Describe three levels of a DBMS architecture. What is the objective of this architecture model?
   1. External, Conceptual, Internal. Views are provided to the users, and these views hide how the data is stored so that users need not worry about the data.
7. What are the main functions of a DBMS?
   1. It manages and controls access to the database, translates user data requests to the physical storage, set of programs used to defind,administrate, and process database data
8. Describe the characteristics of a data warehouse.
   1. Data taken from a multitude of varying inputs and stored in a single uniform format for ease of use
9. Describe the ETL process.
   1. Extract the data from varying sources, Transform the data into a uniform format, Load the data into the warehouse
10. Discuss how a data mart is different from a data warehouse.
    1. Data mart is data that is useable for the specific user, HR person wants HR releated data, not anything to do with sales, or part specifications
    2. Data warehouse IS ALL THE DATA

# Terminology

Define the following terms:

|  |
| --- |
| 4GL – 4th Generation Language – Non-procedural language concerned with the “What” not the “How”. SQL is 4th Gen |
| Candidate key – A possible Primary Key, has minimum set of attributes to unique ID a row in a table. |
| Column – An attribute/field. Describes the type of data |
| Composite key – Any other key that has 2 parts. Ex the PK is both the first and last name fields |
| Constraint – Rules placed on a database |
| Data – Data is information |
| Database – Something or somewhere data is stored in tables |
| Data dictionary - metadata repository, where the meta data is stored |
| Data mart – Subsets of the DW, small DW focused on a specific area of interest. |
| Strategic Mart – A subset of DW, with data that will be used for determining business direction and setting strategic goals. |
| Data mining – Analyzing collected data (often stored in a data warehouse) for patterns. |
| Data warehouse – Used to provide aggregate data in a suitable format for decision makers. |
| DBA – DataBase Administrator |
| DBMS – DataBase Management System – used by organization to define who can use what and query it |
| DCL – Data Control Language – Controls permissions in the DB (GRANT, REVOKE) |
| DDL – Data Definition Language – Commands to specify the Schema (CREATE TABLE) etc |
| DML – Data manipulation Language – Commands used to work with the data, read/update the db. (SELECT, INSERT, UPDATE, DELETE) |
| Domain – All the valid choices you can make for the data, datatype size of data type, etc |
| ETL – Extraction Transformation Loading (get the data, make it useful, save it to the warehouse) |
| Foreign key – Primary Key of another table, put into this table. PK from table 1 in table 2 is the FK of table 2 |
| Information |
| Metadata – Data about data |
| OLAP- Equivalent of a crosstab table or pivot table |
| Object oriented – database - Makes the tables of the Db into classes for coding/programming |
| Relational database- based upon primary keys and foreign keys that are used instead of pointer to define relationships |
| Primary key – Unique value used to be able to uniquely identify (differentiate) rows/records in a DB |
| Relation or Table – has columns(attributes, fields) and rows(records, tuples), but not the data |
| RDBMS – Relational DataBase Management system – Organizes the data in tables(relations) |
| Row, tuple, record – and individual piece of data in a DB, stores info about a specific item or instance |
| Schema – another word for the entire structure of the DB, how the Db is structured |
| SQL – Structured Query Language. Most common language for interacting with a relational database, 3 components are DDL, DML, DCL |
| Star schema – the Dimensional approach |
| Surrogate key – used when you don’t have any good candidate keys, so you come up with one, equivalent of a AutoNumber |
| Transaction – A unit of work that either succeeds or fails in whole, it ALL works it none of it does. It will rollback the successful ones if some fail. |
| Unique key – Something that will be managed to be unique, a way to make values in a column unique |