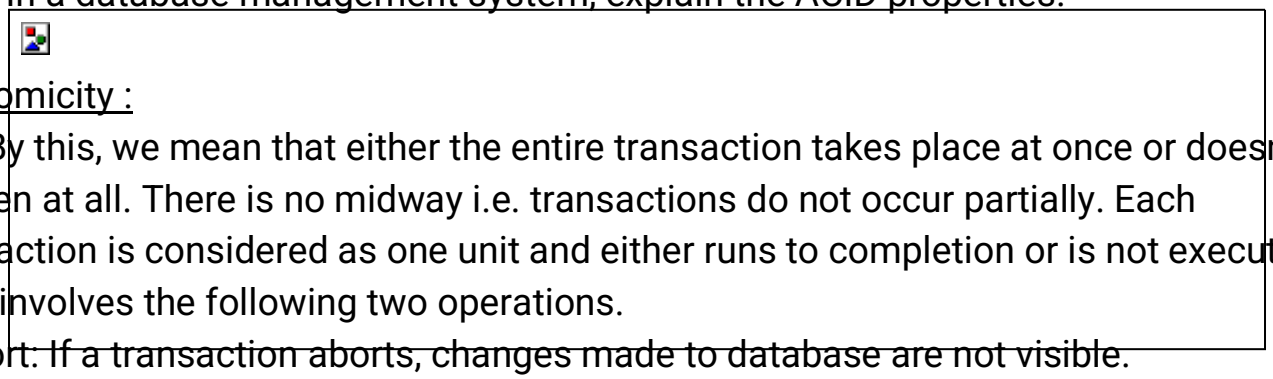
SQL Assignment 1

1. What is a relational database management system (RDBMS)? What are the advantages of a database management system over a file system?

   RDBMS expanded as Relational Database Management System is a type of database that stores data in a well-organized composite arrangement, comprising of rows and columns, which makes it effortless for the processing of data such as fetching, viewing, updating, deleting and manipulating the contents of the tables, as all the data in each and every table will be holding a relationship with the contents of other tables in the system and all the tables in the database system will be associated to the tables with similar properties.

2. In a database management system, explain the ACID properties.



Atomicity :

By this, we mean that either the entire transaction takes place at once or doesn't happen at all. There is no midway i.e. transactions do not occur partially. Each transaction is considered as one unit and either runs to completion or is not executed at all. It involves the following two operations.

—Abort: If a transaction aborts, changes made to database are not visible.

—Commit: If a transaction commits, changes made are visible.

Atomicity is also known as the 'All or nothing rule'.

Consistency:

This means that integrity constraints must be maintained so that the database is consistent before and after the transaction. It refers to the correctness of a database. Referring to the example above,

The total amount before and after the transaction must be maintained.

Isolation :

This property ensures that multiple transactions can occur concurrently without leading to the inconsistency of database state. Transactions occur independently without interference. Changes occurring in a particular transaction will not be visible to any other transaction until that particular change in that transaction is written to memory or has been committed. This property ensures that the execution of transactions concurrently will result in a state that is equivalent to a state achieved these were executed serially in some order.

Durability:
This property ensures that once the transaction has completed execution, the updates and modifications to the database are stored in and written to disk and they persist even if a system failure occurs. These updates now become permanent and are stored in non-volatile memory. The effects of the transaction, thus, are never lost.

3. Explain the concept of normalization.

Normalization is a database design technique that reduces data redundancy and eliminates undesirable characteristics like Insertion, Update and Deletion Anomalies. Normalization rules divides larger tables into smaller tables and links them using relationships. The purpose of Normalisation in SQL is to eliminate redundant (repetitive) data and ensure data is stored logically.
Database Normal Forms
Here is a list of Normal Forms in SQL:

1NF (First Normal Form)
2NF (Second Normal Form)
3NF (Third Normal Form)
BCNF (Boyce-Codd Normal Form)
4NF (Fourth Normal Form)
5NF (Fifth Normal Form)
6NF (Sixth Normal Form)

4. Explain the many types of query languages used in relational databases. DQL, DML, DCL, and DDL are some examples.

These SQL commands are mainly categorized into four categories as:

1. DDL – Data Definition Language
2. DQl – Data Query Language
3. DML – Data Manipulation Language
4. DCL – Data Control Language

Though many resources claim there to be another category of SQL clauses **TCL – Transaction Control Language**. So we will see in detail about TCL as well.

DDL (Data Definition Language):

DDL or Data Definition Language actually consists of the SQL commands that can be used to define the database schema. It simply deals with descriptions of the database schema and is used to create and modify the structure of database objects in the database.DDL is a set of SQL commands used to create, modify, and delete database structures but not data. These commands are normally not used by a general user, who should be accessing the database via an application.

DQL (Data Query Language):

DQL statements are used for performing queries on the data within schema objects. The purpose of the DQL Command is to get some schema relation based on the query passed to it. We can define DQL as follows it is a component of SQL statement that allows getting data from the database and imposing order upon it. It includes the SELECT statement. This command allows getting the data out of the database to perform operations with it. When a SELECT is fired against a table or tables the result is compiled into a further temporary table, which is displayed or perhaps received by the program i.e. a front-end

DML(Data Manipulation Language):

The SQL commands that deals with the manipulation of data present in the database belong to DML or Data Manipulation Language and this includes most of the SQL statements. It is the component of the SQL statement that controls access to data and to the database. Basically, DCL statements are grouped with DML statements.

DCL (Data Control Language):

DCL includes commands such as GRANT and REVOKE which mainly deal with the rights, permissions, and other controls of the database system.

Though many resources claim there to be another category of SQL clauses TCL – Transaction Control Language. So we will see in detail about TCL as well. TCL commands deal with the transaction within the database.

5. What is the difference between the main key and a composite key? Give instances of how primary key and composite are used.

Primary Key

First, a primary key uniquely identifies each record in a database table. Any individual key that does this can be called a candidate key, but only one can be chosen by database engineers as a primary key.

Composite Key
Next, there's the composite key, which is composed of two or more attributes that collectively uniquely identify each record.

An example would be a list of homes on a real estate market. In a well-ordered database, there should be a primary key that uniquely identifies each record.

How this works may have to do with the sophistication of the database.

In some cases, the homes may only be uniquely identified by a mortgage number — all other data (towns, streets, house numbers) is not unique to each record. The mortgage number would be the primary key. Suppose, however, that an MLS realtor's listing technology assigns its own unique numbers to the records in the table.

Then, there will be two keys that developers might identify as "candidate keys":

The mortgage number.
The MLS number.
One of them will qualify as the "primary key" in what some would consider an arbitrary way.

A composite key, then, would be the combination of two keys.

For example: the combination of house number and street might qualify as a composite key, given that the market listings are local. If so, then when someone searches using both the house number and the street, they should only get one single record returned.