**Assignment 1 SQL**

1Ans: A database management system (or DBMS) is essentially nothing more than a computerized data-keeping system. Users of the system are given facilities to perform several kinds of operations on such a system for either manipulation of the data in the database or the management of the database structure itself. Database Management Systems (DBMSs) are categorized according to their data structures or types. RDBMS is basically a program that allows us to create, delete, and update a relational database. Relational Database is a database system that stores and retrieves data in a tabular format organized in the form of rows and columns. It is a smaller subset of DBMS which was designed by E.F Codd in the 1970s. The major DBMS like SQL, My-SQL, ORACLE are all based on the principles of relational DBMS.

Relational DBMS owes its foundation to the fact that the values of each table are related to others. It has the capability to handle larger magnitudes of data and simulate queries easily.

Relational Database Management Systems maintains data integrity by simulating the following features: 

* **Entity Integrity:** No two records of the database table can be completely duplicate.
* **Referential Integrity:** Only the rows of those tables can be deleted which are not used by other tables. Otherwise, it may lead to data inconsistency.
* **User-defined Integrity:** Rules defined by the users based on confidentiality and access.
* **Domain integrity:**The columns of the database tables are enclosed within some structured limits, based on default values, type of data or ranges

The following are the advantages of DBMS over file system

**Data redundancy and inconsistency** –

Redundancy is the concept of repetition of data i.e. each data may have more than a single copy. The file system cannot control redundancy of data as each user defines and maintains the needed files for a specific application to run. There may be a possibility that two users are maintaining same files data for different applications. Hence changes made by one user does not reflect in files used by second users, which leads to inconsistency of data. Whereas DBMS controls redundancy by maintaining a single repository of data that is defined once and is accessed by many users. As there is no or less redundancy, data remains consistent.

**Data sharing** –

File system does not allow sharing of data or sharing is too complex. Whereas in DBMS, data can be shared easily due to centralized system.

**Data concurrency** –

Concurrent access to data means more than one user is accessing the same data at the same time. Anomalies occur when changes made by one user gets lost because of changes made by other user. File system does not provide any procedure to stop anomalies. Whereas DBMS provides a locking system to stop anomalies to occur.

**Data searching** –

For every search operation performed on file system, a different application program has to be written. While DBMS provides inbuilt searching operations. User only have to write a small query to retrieve data from database.

**Data integrity** –

There may be cases when some constraints need to be applied on the data before inserting it in database. The file system does not provide any procedure to check these constraints automatically. Whereas DBMS maintains data integrity by enforcing user defined constraints on data by itself.

**System crashing** –

In some cases,systems might have crashes due to various reasons. It is a bane in case of file systems because once the system crashes, there will be no recovery of the data that’s been lost. A DBMS will have the recovery manager which retrieves the data making it another advantage over file systems.

**Data security** –

A file system provides a password mechanism to protect the database but how longer can the password be protected?No one can guarantee that. This doesn’t happen in the case of DBMS. DBMS has specialized features that help provide shielding to its data.

2Ans: In order to maintain consistency in a database, before and after the transaction, certain properties are followed. These are called **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.

**Consistency**   
This means that integrity constraints must be maintained so that the database is consistent before and after the transaction

**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.

The **ACID** properties, in totality, provide a mechanism to ensure correctness and consistency of a database in a way such that each transaction is a group of operations that acts a single unit, produces consistent results, acts in isolation from other operations and updates that it makes are durably stored.

3Ans: Normalization is the process of organizing data in a database. This includes creating tables and establishing relationships between those tables according to rules designed both to protect the data and to make the database more flexible by eliminating redundancy and inconsistent dependency.

Redundant data wastes disk space and creates maintenance problems. If data that exists in more than one place must be changed, the data must be changed in exactly the same way in all locations. A customer address change is much easier to implement if that data is stored only in the Customers table and nowhere else in the database.

What is an "inconsistent dependency"? While it is intuitive for a user to look in the Customers table for the address of a particular customer, it may not make sense to look there for the salary of the employee who calls on that customer. The employee's salary is related to, or dependent on, the employee and thus should be moved to the Employees table. Inconsistent dependencies can make data difficult to access because the path to find the data may be missing or broken.

There are a few rules for database normalization. Each rule is called a "normal form." If the first rule is observed, the database is said to be in "first normal form." If the first three rules are observed, the database is considered to be in "third normal form." Although other levels of normalization are possible, third normal form is considered the highest level necessary for most applications.

As with many formal rules and specifications, real world scenarios do not always allow for perfect compliance. In general, normalization requires additional tables and some customers find this cumbersome. If you decide to violate one of the first three rules of normalization, make sure that your application anticipates any problems that could occur, such as redundant data and inconsistent dependencies.

4Ans: These SQL commands are mainly categorized into four categories as:

DDL – Data Definition Language

DQl – Data Query Language

DML – Data Manipulation Language

DCL – Data Control Language

**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.

List of DDL commands:

CREATE: This command is used to create the database or its objects (like table, index, function, views, store procedure, and triggers).

DROP: This command is used to delete objects from the database.

ALTER: This is used to alter the structure of the database.

TRUNCATE: This is used to remove all records from a table, including all spaces allocated for the records are removed.

COMMENT: This is used to add comments to the data dictionary.

RENAME: This is used to rename an object existing in the database.

**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.

List of DQL:

SELECT: It is used to retrieve data from the database.

**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.

List of DML commands:

INSERT : It is used to insert data into a table.

UPDATE: It is used to update existing data within a table.

DELETE : It is used to delete records from a database table.

LOCK: Table control concurrency.

CALL: Call a PL/SQL or JAVA subprogram.

EXPLAIN PLAN: It describes the access path to data.

**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.

List of DCL commands:

GRANT: This command gives users access privileges to the database.

REVOKE: This command withdraws the user’s access privileges given by using the GRANT command.

5Ans: **Primary Key**

Primary key is a set of one or more fields/columns of a table that uniquely identify a record in a database table. It can not accept null, duplicate values. Only one Candidate Key can be Primary Key.

**Composite/Compound Key**

Composite Key is a combination of more than one fields/columns of a table. It can be a Candidate key, Primary key.