**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?

Answer:

Data Sharing: File system does not allow sharing of data. 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.

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

Answer:

Atomicity: The term atomicity defines that the data remains atomic. It means if any operation is performed on the data, either it should be performed or executed completely or should not be executed at all. It further means that the operation should not break in between or execute partially. In the case of executing operations on the transaction, the operation should be completely executed and not partially.

Consistency: The word consistency means that the value should remain preserved always. In DBMS, the integrity of the data should be maintained, which means if a change in the database is made, it should remain preserved always. In the case of transactions, the integrity of the data is very essential so that the database remains consistent before and after the transaction. The data should always be correct.

Isolation: The term 'isolation' means separation. In DBMS, Isolation is the property of a database where no data should affect the other one and may occur concurrently. In short, the operation on one database should begin when the operation on the first database gets complete. It means if two operations are being performed on two different databases, they may not affect the value of one another. In the case of transactions, when two or more transactions occur simultaneously, the consistency should remain maintained. Any changes that occur in any particular transaction will not be seen by other transactions until the change is not committed in the memory.

Durability: Durability ensures the permanency of something. In DBMS, the term durability ensures that the data after the successful execution of the operation becomes permanent in the database. The durability of the data should be so perfect that even if the system fails or leads to a crash, the database still survives. However, if gets lost, it becomes the responsibility of the recovery manager for ensuring the durability of the database. For committing the values, the COMMIT command must be used every time we make changes.

1. Explain the concept of normalization.

Answer:

Normalization: Normalization is the process of organizing the data in the database. Normalization is used to minimize the redundancy from a relation or set of relations. It is also used to eliminate the undesirable characteristics like Insertion, Update and Deletion Anomalies. Normalization divides the larger table into the smaller table and links them using relationship. The normal form is used to reduce redundancy from the database table.

First Normal Form (1NF): A relation will be 1NF if it contains an atomic value. It states that an attribute of a table cannot hold multiple values. It must hold only single-valued attribute. First normal form disallows the multi-valued attribute, composite attribute, and their combinations.

Second Normal Form (2NF): In the 2NF, relational must be in 1NF.In the second normal form, all non-key attributes are fully functional dependent on the primary key.

Third Normal Form (3NF): A relation will be in 3NF if it is in 2NF and not contain any transitive partial dependency.3NF is used to reduce the data duplication. It is also used to achieve the data integrity. If there is no transitive dependency for non-prime attributes, then the relation must be in third normal form.

Boyce Codd normal form (BCNF): BCNF is the advance version of 3NF. It is stricter than 3NF.A table is in BCNF if every functional dependency X → Y, X is the super key of the table. For BCNF, the table should be in 3NF.

Fourth normal form (4NF): A relation will be in 4NF if it is in Boyce Codd normal form and has no multi-valued dependency. For a dependency A → B, if for a single value of A, multiple values of B exists, then the relation will be a multi-valued dependency.

Fifth normal form (5NF): A relation is in 5NF if it is in 4NF and not contains any join dependency and joining should be lossless.5NF is satisfied when all the tables are broken into as many tables as possible in order to avoid redundancy.5NF is also known as Project-join normal form (PJ/NF).

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

Answer:

i. DQL: DQL stands for Data Query Language. DQL statements are used for performing queries on the data within schema objects.

* Select: It is used to retrieve data from the database.

ii. DML: DML stands for Data Manipulate Language. DML is a set of SQL commands used to (Insert, Update and Delete Data). DML deals with data manipulation.

* Insert: It is used to insert a new record inside table.
* Update: It is used to update a record inside table.
* Delete: It is used to delete a record from table.

iii. DCL: DCL stands for Data Control Language. DCL is a set of SQL commands (Grant and Revoke). DCL is used to give rights, permissions and other controls of the database system.

* Grant: allow users to access privileges to the database.
* Revoke: withdraw user access privileges given by using the GRANT command.

iv. DDL: DDL stands for Data Definition Language. DDL is a set of SQL commands used to (create, alter, drop and rename table). DDL deals with database schemas and descriptions of how the data should reside in the database.

* Create: It is used to create a database and its objects like (table, index, views, store procedure, function, and triggers).
* Alter: It is used to alter the structure of the existing database.
* Drop: It is used to delete objects from the database.
* Rename: It is used rename an object.

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

Answer:

PRIMARY KEY: A PRIMARY KEY constraint uniquely identifies each record in a table. A Primary keys column must contain unique values and cannot have null values. A table can have only one primary key, which may consist of single or multiple columns.

COMPOSITE KEY: A COMPOSITE KEY is a combination of two or more columns in a table that can be used to uniquely identify each row in the table when the columns are combined uniqueness is guaranteed, but when it taken individually it does not guarantee uniqueness.

EXAMPLE:

Creating Primary Key

CREATE TABLE STUDENT (

ID INT,

FIRSTNAME VARCHAR (255) NOT NULL,

LASTNAME VARCHAR (255) NOT NULL,

AGE INT,

PRIMARY KEY (ID)

);

A primary key is specified on column ID, it means that ID column cannot have NULL values and duplicate values.

Creating Composite Key

CREATE TABLE STUDENT (

ID INT,

FIRSTNAME VARCHAR (255) NOT NULL,

LASTNAME VARCHAR (255) NOT NULL,

AGE INT,

CONSTRAINT IDAGE PRIMARY KEY (ID, AGE)

);

A primary key is specified on column ID and AGE, it means that ID and AGE column cannot have NULL values and duplicate values.

1. Create a table with a primary key, a column default value, and a column unique constraint in SQL.

Answer:

CREATE TABLE STUDENT (

ID INT,

FIRSTNAME VARCHAR (255) NOT NULL,

LASTNAME VARCHAR (255) NOT NULL,

AGE INT,

DEPARTMENT VARCHAR (255),

CITY VARCHAR (255) DEFAULT 'ABC',

PRIMARY KEY (ID),

UNIQUE (DEPARTMENT)

);