**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?

* **A relational database is a collection of information organizing data which is the subject of interest by mapping out the data points of the entire dataset bringing out the relationships between them through storing them in one or more tables (“or relations”) of columns and rows. The software used to store, manage, query and retrieve data stored in relational database is called as Relational database management system (RDBMS).**

**Here are few advantages of Database management system over a file system: 1. risk of data loss is low.**

**2. Multiple access/working possible.**

**3. Data manipulation is smooth.**

**4. Data integrity is maintained.**

**5. changes to data can be recorded(versioning).**

**6. Diverse connectivity.**

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

* **ACID stands for “Atomicity”, “Consistency”, “Isolation”, “Durability”.**

**Atomicity : From a Database perspective means “Either All or None” from a user perspective one could understand that before a set of operations gets committed if there is a failure at any operation then the commit (DB update) doesn’t happen and the process is rolled back to step one (back end) rather than resumption from the interrupted point unlike downloading operation. E.g., online transaction.**

**Consistency: From a Database perspective means “Before and after commit count remains the same” from a user perspective could be understood by an example of transaction from one account to another E.g. Acc “A” has 3000 INR and has to make a 1000 INR transaction to Acc “B” which has 5000 INR, before transaction(commit) the record (count) of both Acc “A” and Acc “B” is 8000 INR, after transaction(commit) the record(count) ) of both Acc “A” and Acc “B” will remain the same i.e. 8000 INR (A = 2000 INR and B = 6000 INR).**

**Isolation: From a Database perspective means “parallel/concurrent operations to serial operations” from a user perspective could be understood by relating the “Database” to a “Bank” which holds the asset/cash of all it’s customers and relating the “customer operations” (deposit/withdrawal) to “commits” even though there are multiple operations performed by customers on the bank changing the asset/cash it holds remains isolated/unaffected to the individual customer asset/cash after a specific operation performed by the customer. Giving all the customers the illusion that the bank just works for Them specifically.**

**Durability: From a Database perspective means “permanency” from a user perspective could be understood any changes (commits) done to their bank account (data base) is permanent even if there is a system failure/server crash the changes done remains permanent.**

1. Explain the concept of normalization.

* **Normalization in DBMS is the concept of reducing data dependencies or data replications with respect to a specific table. This is important in database design, as an industry standard the DB design should at least satisfy “3NF” There are several Normalization Factors out of which the most important ones are:**

**1NF (first normal form); all records (rows) must be unique and each cell in the table must have only one value.**

**2NF (second normal form); 1NF to be satisfied and A single column primary key shouldn’t functionally depend on subset of any candidate key in other words single column primary key shouldn’t be a composite key.**

**3NF (third normal form); 2NF to be satisfied and shouldn’t have any transitive functional dependency (non-key columns to be independent of each other).**

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

* **DQL (Data query language); performs queries on the data within schema objects. E.g., “SELECT”**
* **DML (Data manipulation language); used for managing data within schema objects. E.g., “INSERT”, “UPDATE” , “DELETE”**
* **DCL (Data control language); used to provide/deny access to the data stored in the databases. E.g., “GRANT” , “REVOKE”**
* **DDL (Data definition language); used to define the database structure or schema. E.g., “CREATE TABLE” , “ALTER TABLE” , “DROP TABLE”**
* **TCL (Transaction Control Language); used to manage transactions in the database. These are used to manage the changes made by DML-statements. E.g., “COMMIT” , “ROLLBACK” , “SAVEPOINT”**

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

* **Main key is a single column primary key whereas composite key is a multiple column primary key.**

**Primary key can be used on a single column making it the Main key E.g., “employeeID” can be made the main key if the table consists of columns “employeeID”, “job-role”, “salary” whereas employeeID and shift column can be made composite key for a schema like “employeeID” , “shift” , “age” since the same employee can have multiple shifts (Day / Night) associated eliminating the possibility of making a main key (single column primary key).**