

Assignment 1

Q .1 What Is Rdbms?

The software used to store, manage, query, and retrieve data stored in a relational database is called a relational database management system (RDBMS). The RDBMS provides an interface between users and applications and the database, as well as administrative functions for managing data storage, access, and performance.

Q.2 What Is 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.

Q.3 What Are Different Normalization Forms?

Normalization works through a series of stages called Normal forms. The normal forms apply to individual relations. The relation is said to be in particular normal form if it satisfies constraints.

1. First Normal Form –

If a relation contain composite or multi-valued attribute, it violates first normal form or a relation is in first normal form if it does not contain any composite or multi-valued attribute. A relation is in first normal form if every attribute in that relation is singled valued attribute.

1NF (First Normal Form) Rules :

- Each table cell should contain a single value.
- Each record needs to be unique.

2. Second Normal Form –

To be in second normal form, a relation must be in first normal form and relation must not contain any partial dependency. A relation is in 2NF if it has No Partial Dependency, i.e., no non-prime attribute (attributes which are not part of any candidate key) is dependent on any proper subset of any candidate key of the table.

Partial Dependency – If the proper subset of candidate key determines non-prime attribute, it is called partial dependency.

2NF (Second Normal Form) Rules :

- Rule 1- Be in 1NF
- Rule 2- Single Column Primary Key that does not functionally dependant on any subset of candidate key relation / There shouldn't be any partial dependencies

3. Third Normal Form –

The third Normal Form is an upgrade to the Second Normal Form. When a table is in the Second Normal Form and has no transitive dependency, then it is in the Third Normal Form.

For a table to be in the third normal form,

It should be in the Second Normal form.

And it should not have Transitive Dependency.

3NF (Third Normal Form) Rules

- Rule 1- Be in 2NF
- Rule 2- Has no transitive functional dependencies

Q.4 What Is Trigger?

A trigger is a stored procedure in database which automatically invokes whenever a special event in the database occurs. For example, a trigger can be invoked when a row is inserted into a specified table or when certain table columns are being updated.

Syntax:

```
create trigger [trigger_name]
[before | after]
{insert | update | delete}
on [table_name]
[for each row]
[trigger_body]
```

BEFORE and AFTER of Trigger:

BEFORE triggers run the trigger action before the triggering statement is run.

AFTER triggers run the trigger action after the triggering statement is run.

Q.5 What Is Database?

A database is an organized collection of structured information, or data, typically stored electronically in a computer system. A database is usually controlled by a database

management system (DBMS). Together, the data and the DBMS, along with the applications that are associated with them, are referred to as a database system, often shortened to just database.

Data within the most common types of databases in operation today is typically modeled in rows and columns in a series of tables to make processing and data querying efficient. The data can then be easily accessed, managed, modified, updated, controlled, and organized. Most databases use structured query language (SQL) for writing and querying data.

Q.6 Disadvantage In File Processing System?

Data duplication: Same data is used by multiple resources for processing, thus creating multiple copies of the same data wasting the space.

Security: Less security is provided to the data.

High maintenance: Accessing control and verifying data consistency needs a high maintenance cost.

Accessing Anomalies: Accessing anomalies means that it is not easy to access data in a desired or efficient way. It makes supervision of the department very difficult. If a user wants information in a specific manner then he requires creating a program for it.

Data Isolation: Data is isolated in the File Processing System and data is stored in different files. These files can be in different formats. If you want to extract data from two files then you are required to which part of the file is needed and how they are related to each other.

Q.7 What Is System R? What Are Its Two Major Subsystems?

System R was designed and developed over the period of 1974-79 at IBM San Jose Research Center. It is a prototype and its purpose was to demonstrate that it is possible to build a Relational System that can be used in a real-life environment to solve real life problems, with a performance at least comparable to that of the existing system.

Its two subsystems are

- Research Storage
- System Relational Data System.

Q.8 What are the different features of an RDBMS?

- Provides data to be stored in tables

- Persists data in the form of rows and columns.
- Provides facility primary key, to uniquely identify the rows.
- Creates indexes for quicker data retrieval.
- Provides a virtual table creation in which sensitive data can be stored and simplified query can be applied.(views).
- Sharing a common column in two or more tables(primary key and foreign key).
- Provides multi user accessibility that can be controlled by individual users
- SQL DDL (Data Definition Language) Support
- SQL DML (Data Manipulation Language) Enhancements

Q.9 What are different Codd's 12 rules for Relational Database?

Rule zero

This rule states that for a system to qualify as an RDBMS, it must be able to manage the database entirely through the relational capabilities.

Rule 1: Information rule

All information(including metadata) is to be represented as stored data in cells of tables. The rows and columns have to be strictly unordered.

Rule 2:Guaranteed Access

Each unique piece of data(atomic value) should be accessible by : Table Name + Primary Key(Row) + Attribute(column).

Rule 3: Systematic treatment of NULL

Null has several meanings, it can mean missing data, not applicable or no value. It should be handled consistently. Also, Primary key must not be null, ever. Expression on NULL must give null.

Rule 4: Active Online Catalog

Database dictionary(catalog) is the structure description of the complete Database and it must be stored online. The Catalog must be governed by same rules as rest of the database. The same query language should be used on catalog as used to query database.

Rule 5: Powerful and Well-Structured Language

One well structured language must be there to provide all manners of access to the data stored in the database. Example: SQL, etc. If the database allows access to the data without the use of this language, then that is a violation.

Rule 6:View Updation Rule

All the views that are theoretically updatable should be updatable by the system as well.

Rule 7: Relational Level Operation

There must be Insert, Delete, Update operations at each level of relations. Set operation like Union, Intersection and minus should also be supported.

Rule 8: Physical Data Independence

The physical storage of data should not matter to the system. If, say, some file supporting table is renamed or moved from one disk to another, it should not affect the application.

Rule 9: Logical Data Independence

If there is change in the logical structure(table structures) of the database the user view of data should not change. Say, if a table is split into two tables, a new view should give result as the join of the two tables. This rule is most difficult to satisfy.

Rule 10: Integrity Independence

The database should be able to enforce its own integrity rather than using other programs. Key and Check constraints, trigger etc, should be stored in Data Dictionary. This also make RDBMS independent of front-end.

Rule 11: Distribution Independence

A database should work properly regardless of its distribution across a network. Even if a database is geographically distributed, with data stored in pieces, the end user should get an impression that it is stored at the same place. This lays the foundation of distributed database.

Rule 12: Non-subversion Rule

If low level access is allowed to a system it should not be able to subvert or bypass integrity rules to change the data. This can be achieved by some sort of locking or encryption.

Q.10 What are the roles of databases and what is the need of databases in a project?

Role

- Creating and maintaining database standards and policies
- Supporting database design, creation, and testing activities
- Managing the database availability and performance, including incident and problem management
- Administering database objects to achieve optimum utilization
- Defining and implementing event triggers that will alert on potential database performance or integrity issues
- Performing database housekeeping, such as tuning, indexing, etc.
- Monitoring usage, transaction volumes, response times, concurrency levels, etc.
- Identifying reporting, and managing database security issues, audit trails, and forensics
- Designing database backup, archiving, and storage strategy

Need of database :

- Databases are useful in many different scenarios for storing data. It is typical to use a database when different sets of data needs to be linked together, such as:
 - Pupils in a school and their grades
 - Customer records and sales information
 - Patients' and doctors' records
 - Transactions between different bank accounts
 - Taxpayers and income tax payments
- The three main advantages that databases have over other, simpler data storage systems (such as text files and spreadsheets) are access, integrity, and security.

Q.11 List out different open source database software used for data storing purpose.

A database is backend storage for an application — like, for instance, a web app (which is the primary app type we'll reference throughout this piece). The database itself will sit on your server, alongside the other backend elements such as the core files of your site, any media you use, and the server configuration files.

1. MySQL
2. PostgreSQL
3. Redis
4. SQLite
5. MongoDB

