UNIT 1

1. What is Data and Information?

Data refers to raw facts and figures without any context. Information is data that has been processed and organized to be meaningful and useful.

2. What is Database?

A Database is a structured collection of data stored electronically, designed for easy access, management, and updating.

3. What is DBMS?

A Database Management System (DBMS) is software that manages and controls access to databases, ensuring data is organized and easily retrievable.

4. Define Entity Set

An Entity Set is a collection of similar types of entities that share the same attributes.

5. Define Entity

An Entity is a real-world object or thing that can be identified distinctly and has attributes.

6. Define Attribute

An Attribute is a property or characteristic that describes an entity.

7. List the types of Relationships

The types of relationships are: - One-to-One - One-to-Many - Many- to-One - Many-to-Many

8. What is Data Model? Explain types of Data Models

A Data Model is a conceptual framework for organizing and structuring data.

Types include:

- Hierarchical Model
- Network Model
- Relational Model
- Entity-Relationship Model
- Object-Oriented Model

9. What is Inheritance?

Inheritance is a feature where an entity can inherit attributes and relationships of another entity, typically used in database design and object-oriented modeling.

10. Write the difference between File System and Database System

File System vs Database System:

- File System: Data is stored in separate files, redundancy is high, no security/integrity constraints, difficult to retrieve related data.
- Database System: Data is centralized, redundancy is minimized, provides security/integrity constraints, easy and efficient data retrieval.

UNIT-2

1. **Define Primary Key** A Primary Key is a field or a combination of fields in a table that uniquely identifies each record. It cannot have NULL values.

2. What is Referential Integrity?

Referential Integrity ensures that a foreign key value always points to an existing, valid record in another table.

3. Define Selection operation in Relational Algebra

Selection is an operation that retrieves rows (tuples) from a relation that satisfy a given condition.

4. Define Projection operation in Relational Algebra

Projection is an operation that retrieves specific columns (attributes) from a relation, removing duplicates.

5. What is Join operation in Relational Algebra?

Join is a relational algebra operation that combines related tuples from two relations based on a common attribute.

6. Compare Unique and Not Null Constraints

Unique Constraint vs Not Null Constraint:

- Unique Constraint: Ensures all values in a column are different; allows one NULL value.
- Not Null Constraint: Ensures that a column cannot have NULL values.

7. Define Relational Calculus

Relational Calculus is a non-procedural query language that uses mathematical predicates to describe the desired information without specifying how to retrieve it.

8. List the DML operations

The DML operations are:

- INSERT
- UPDATE
- DELETE
- SELECT

9. What is constraint? Explain types of Constraints

A Constraint is a rule enforced on data in a database to ensure validity, consistency, and integrity.

Types of Constraints: - Primary Key - Foreign Key - Unique - Not Null - Check

10. List the DDL operations?

DDL stands for **Data Definition Language**, and it refers to SQL commands used to define or modify the structure of database objects (like tables, indexes, schemas, etc.).

Here's a list of the main **DDL operations**:

1. **CREATE**
    - Used to create new database objects (tables, views, indexes, databases, etc.).
2. **ALTER**

    Used to modify existing database objects (add or remove columns, change column types, rename objects, etc.).
3. **DROP**
    - Used to delete database objects permanently.
4. **TRUNCATE**
    - Used to delete all rows from a table **quickly** without logging individual row deletions (structure remains).
5. **RENAME**

    Used to rename database objects.

UNIT-3

1. Define View?

A view in SQL is a virtual table that is based on the result-set of an SQL statement. It contains rows and columns, just like a real table. The fields in a view are fields from one or more real tables in the database. Views are used to simplify complex queries, enhance security by restricting access to specific data, and present data in a specific format.

2. List the SQL Functions?

SQL functions are built-in functions used to perform operations on data. These include: - String functions (e.g., CONCAT, LENGTH, SUBSTRING) - Numeric functions (e.g., ROUND, CEIL, FLOOR) - Date functions (e.g., NOW, CURDATE, DATEADD) - Aggregate functions (e.g., COUNT, SUM, AVG) - Conversion functions (e.g., CAST, CONVERT)

3. What is Nested query?

A nested query, also known as a subquery, is a query within another SQL query. It is used to perform operations that require multiple steps of data retrieval. The result of the inner query is used by the outer query.

Nested queries can appear in SELECT, INSERT, UPDATE, or DELETE statements.

4. What is join? Explain its types?

A join in SQL is used to combine rows from two or more tables based on a related column.

Types of joins include:

- INNER JOIN: Returns records with matching values in both tables.
- LEFT JOIN: Returns all records from the left table and matched records from the right.
- RIGHT JOIN: Returns all records from the right table and matched records from the left.
- FULL OUTER JOIN: Returns records with a match in either table.
- CROSS JOIN: Returns the Cartesian product of both tables.

5. List the Aggregate functions in SQL?

Aggregate functions perform a calculation on a set of values and return a single value. They are

commonly used with the GROUP BY clause.

Examples include:

- COUNT(): Counts the number of rows
  Syntax: Select count(coulumn_name) from table_name;

- SUM(): Calculates the total sum
  Syntax: Select sum(coulumn_name) from table_name;

- AVG(): Calculates the average value

  Syntax: Select avg(coulumn_name) from table_name;

- MAX(): Returns the maximum value

  Syntax: Select max(coulumn_name) from table_name;

- MIN(): Returns the minimum value

  Syntax: Select min(coulumn_name) from table_name;

6. Explain Relational set operator and its types?

Relational set operators are used to combine the results of two or more

queries. The main set operators in SQL are:

- UNION: Combines the result of two queries and removes duplicates

- UNION ALL: Combines the result of two queries including duplicates

- INTERSECT: Returns only the rows that are common to both queries -

EXCEPT (or MINUS): Returns rows from the first query that are not in the second

7. Explain about basic SQL querying "select" and "project"?

In SQL, SELECT and PROJECT refer to data retrieval concepts:

- SELECT: Retrieves rows that satisfy a given condition. It is used with WHERE clause.

- PROJECT: Refers to the selection of specific columns from a table.

In SQL, projection is done by listing the column names after the SELECT keyword.

8. Write the Difference between views and joins?

Views and joins are both used to retrieve data from multiple tables, but they differ

: - Views are virtual tables created with a SELECT statement and stored in the database.

- Joins are used in real-time to fetch data from multiple tables based on related keys.

Views can encapsulate complex queries and enhance security, while joins are more dynamic.

9. Implement of integrity constraints with syntax?

Integrity constraints ensure data accuracy and consistency. Types include:

- PRIMARY KEY: Uniquely identifies each record.

    Syntax: Create table table_name(attribute datatype primary key);

- FOREIGN KEY: Ensures referential integrity

    Syntax: Create table table_name(attribute datatype references Parent(attribute));

UNIQUE: Ensures all values in a column are different.

- CHECK: Ensures that values meet a specific

condition. Syntax Example:

CREATE TABLE Students (

ID INT PRIMARY KEY,

Name VARCHAR(50) NOT NULL, Age

INT CHECK (Age >= 18));

10.Define sub query with example?

A subquery is a query within another query. It is used to return data to be used in the main query.

Example: SELECT name FROM Employees WHERE salary > (SELECT AVG(salary) FROM Employees);

This query returns the names of employees whose salary

UNIT-4

1. Define Functional dependency?

Functional dependency is a relationship that exists when one attribute uniquely determines another attribute. In a database, if attribute A functionally determines attribute B (written as A -> B), it means that for every value of A, there is exactly one corresponding value of B. This concept is crucial for database normalization.

IF T1.x=T2.x then T1.Y=T2.Y

2. What is Transactive dependency?

Transitive dependency occurs when a non-prime attribute is dependent on another non-prime attribute, which is further dependent on the primary key. For example, if A -> B and B -> C, then A -> C is a

transitive dependency. It is removed in the third normal form (3NF) of database normalization.

3. Define Partial functional dependency?

Partial functional dependency exists when a non-prime attribute is functionally dependent on part of a candidate key and not on the whole key. This occurs in tables where there is a composite key and the dependency is only on one part of it. It is eliminated in the second normal form (2NF).

4. What is Join dependency?

Join dependency is a constraint that ensures that a table can be reconstructed by joining multiple projections without losing any data. A table is said to have a join dependency if it can be recreated by joining its decomposed tables. It is handled in the fifth normal form (5NF).

5. What is Normalization?

Normalization is the process of organizing the fields and tables of a relational database to minimize redundancy and dependency. It involves decomposing a table into smaller tables without losing information and ensures data integrity and efficient access.

6. Define BCNF Boyce-Codd Normal Form?

a higher version of the third normal form. A table is in BCNF if every determinant is a candidate key. It deals with anomalies that 3NF does not cover and ensures a stricter form of database

normalization.

7. Define Candidate key?

A candidate key is a minimal set of attributes that can uniquely identify a tuple in a table. There can

be multiple candidate keys in a relation, but one of them is chosen as the primary key. Each candidate key must be unique and minimal.

8. Difference between lossy join and lossless join?

A lossless join ensures that when a relation is decomposed and then joined back, no data is lost and the original relation is recovered. A lossy join, on the other hand, results in loss of data or generation of spurious tuples when relations are joined. Lossless joins are essential for reliable database design.

9. What is Denormalization?

Denormalization is the process of combining normalized tables to reduce the complexity of queries and improve performance. It is often used for optimization where speed is more important than eliminating redundancy. However, it can lead to data anomalies if not handled properly.

10. What is Fully Functional Dependency?

A fully functional dependency occurs when a non-prime attribute is functionally dependent on the whole candidate key and not on any part of it. It is an essential requirement for a relation to be in the second normal form (2NF).

UNIT-5

1.Define Transaction

A transaction is a logical unit of work that contains one or more SQL statements. It is executed as a single, indivisible operation that must either be completed entirely or not at all. Transactions ensure data integrity and consistency, especially in multi-user environments. They follow the ACID properties-Atomicity, Consistency, Isolation, and Durability.

2.List the States of Transaction

The states of a transaction include:

- Active: The transaction is currently being executed.
- Partially Committed: The final statement of the transaction has been executed.
- Committed: The transaction has been successfully completed.
- Failed: The transaction cannot proceed due to some error.
- Aborted: The transaction has been rolled back due to failure.

3.What is Lock Granularity

Lock granularity refers to the size of the data item that a lock can be placed on. It can vary from a single row or record, a page, a table, or even an entire database. Finer granularity (e.g., row-level locks) allows for more concurrency but requires more system overhead, while coarser granularity (e.g., table-level locks) reduces overhead but limits concurrency.

4.What is Two Phase Locking

Two-Phase Locking (2PL) is a concurrency control protocol that ensures serializability. It involves two phases:

- Growing Phase: The transaction can acquire locks but cannot release any.
- Shrinking Phase: The transaction can release locks but cannot acquire any new locks.

This protocol guarantees that once a lock is released, no new locks can be obtained, preventing non-serializable schedules.

5.Define Atomicity

Atomicity is one of the ACID properties of a transaction. It ensures that a transaction is treated as a single unit of operation. Either all the operations within the transaction are completed successfully, or none of them are. If any part of the transaction fails, the entire transaction is rolled back to maintain database consistency.

6.What is Time Stamping

Time stamping is a concurrency control method used to manage the execution of concurrent transactions. Each transaction is assigned a unique timestamp when it begins. The system uses these timestamps to determine the serial order of transaction execution, ensuring consistency and avoiding conflicts.

7.What are the Concurrency Protocols?

Concurrency protocols are methods used to manage simultaneous operations without conflicting. Common protocols include:

- Two-Phase Locking (2PL)
- Timestamp Ordering
- Validation-Based Protocols

These protocols help maintain data integrity and prevent problems like lost updates and dirty reads.

8.What is Deadlock?

Deadlock is a situation where two or more transactions are waiting for each other to release locks, and none can proceed. It typically occurs in systems with multiple resources and concurrent transactions. Deadlock must be detected and resolved, often by aborting one or more transactions.

9.What is Database Recovery

Database recovery is the process of restoring a database to a correct state in case of failure. Failures may be due to

hardware malfunctions, software bugs, or system crashes. Recovery techniques include rollback, rollforward, and the use of backup copies and logs to bring the database back to a consistent state.

10.List the Operations on B+ Trees

Operations on B+ trees include:

- Search: Efficiently locating a key in a sorted structure.
- Insert: Adding a new key while maintaining the balanced tree structure.
- Delete: Removing a key and reorganizing the tree if necessary.
- Traversal: Scanning all keys in sorted order, often used in range queries.

B+ trees are widely used in databases for indexing due to their balanced and ordered nature.