DBMS

CRT- Technical Class

1. What is DBMS used for?

- DBMS, commonly known as Database Management System, is an application system whose main purpose revolves around the **data**. This is a system that allows its user to **store** the data, **define** it, **retrieve** it and **update** the information about the data inside the database.
- Two types of DBMS
    - RDBMS- eg. MySQL
    - NOSQL- MongoDB (JSON), Neo4j (Graph)

2. What is meant by a Database?

In simple terms, Database is a collection of data in some organized way to facilitate its user's to easily access, manage and upload the data

3. What are advantages of DBMS over traditional file based systems?

Database management systems were developed to handle the following difficulties of typical Fille-processing systems supported by conventional operating systems.

1. Data redundancy and inconsistency
2. Difficulty in accessing data
3. Atomicity of updates
4. Concurrent access by multiple users
5. Security problems

4. What is E-R model in the DBMS?

- E-R model is known as an **Entity-Relationship model** in the DBMS which is based on the concept of the Entities and the relationship that exists among these entities

5. Explain Entity, Entity Type, and Entity Set in DBMS?

- **Entity** is an object, place or thing which has its independent existence in the real world and about which data can be stored in a database. **For Example,** any person, book, etc.
- **Entity Type** is a collection of entities that have the same attributes. **For Example,** the STUDENT table contains rows in which each row is an entity holding the attributes like name, age, and id of the students, hence STUDENT is an Entity Type which holds the entities having the same attributes.
- **Entity Set** is a collection of entities of the same type. **For Example,** A collection of the employees of a firm.

6. What is RDBMS? What is the difference between DBMS and RDBMS?

- RDBMS is the Relational Database Management System which contains data in the form of the tables and data is accessed on the basis of the common fields among the tables.
- DBMS provides a systematic and organized way of storing, managing and retrieving from collection of logically related information. RDBMS also provides what DBMS provides but above that it provides relationship integrity

7. What are Instances and schemas?

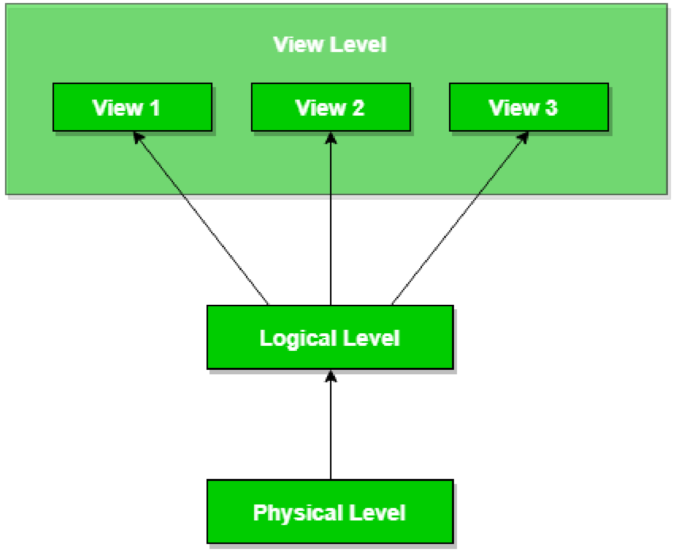
Instances: The data stored in database at a particular moment of time

- **Schema**
  - Logical representation of database
  - SQL is used to represent schema
  - The description of a data base is called the database schema , which is specified during database design and is not expected to change frequently .
  - Course:

Course_id	Course Name	Department

8. Describe the three levels of data abstraction?

- Physical level: The lowest level of abstraction describes how data are stored.(B+, INDEXING, HASHING) i.e. DATA STRUCTURE
- Logical level: The next higher level of abstraction, describes what data are stored in database and what relationship among those data.
- View level: The highest level of abstraction describes only part of entire database



9. What are database constraints for SQL?

- Constraints are used to limit the type of data that can go into a table.
- Constraints can be column level or table level

- `NOT NULL` - Ensures that a column cannot have a NULL value
- `UNIQUE` - Ensures that all values in a column are different
- `PRIMARY KEY` - A combination of a `NOT NULL` and `UNIQUE`. Uniquely identifies each row in a table
- `FOREIGN KEY` - Prevents actions that would destroy links between tables
- `CHECK` - Ensures that the values in a column satisfies a specific condition
- `DEFAULT` - Sets a default value for a column if no value is specified
- `CREATE INDEX` - Used to create and retrieve data from the database very quickly

10. What are the different keys? Explain each of them.

- **Foreign Key:** It’s a field or collection of fields which identifies a row of another table.
- **Super Key:** It’s a set of attributes of a relation schema. All attributes of the schema are functionally dependent.
- **Primary Key:** It’s one of the many candidate keys. The candidate key which is selected as the most important one eventually becomes the primary key.
- **Candidate Key:** It’s a column or set of columns which can identify any record in the database without referring to any data

11. What are the different type of relationships in the DBMS?

- **Relationships in DBMS depicts an association between the tables.**

- **Different types of relationships are:**

- **One-to-One:** This basically states that there should be a one-to-one relationship between the tables i.e. there should be one record in both the tables. **Example:** Among a married couple, both wife and husband can have only one spouse.

- **One-to-Many:** This states that there can be many relationships for one i.e. a primary key table hold only one record which can have many, one or none records in the related table. **Example:** A Mother can have many children.

- **Many-to-Many:** This states that both the tables can be related to many other tables. **Example:** One can have many siblings and so do they have.

12. What is an attribute? And its types

- It is a particular property, which describes the entity
- Single simple composite multivalued, stored, derived
- Name
- Fname lname

14. What is the difference between primary key and unique constraints?

- Primary key cannot have NULL value, the unique constraints can have NULL values. There is only one primary key in a table, but there can be multiple unique constrains

15. How do you communicate with an RDBMS?

- You communicate with an RDBMS using Structured Query Language (SQL)

16. What is SQL? What is the purpose of SQL?

- SQL is Structured Query Language designed for inserting and modifying in a relational database system.
- SQL stands for Structured Query Language whose main purpose is to interact with the relational databases in the form of inserting and updating/modifying the data in the database

17. What are the differences between DDL, DML and DCL in SQL?

- Following are some details of three.
  DDL stands for Data Definition Language. SQL queries like CREATE, ALTER, DROP and RENAME come under this.
  DML stands for Data Manipulation Language. SQL queries like SELECT, INSERT and UPDATE come under this.
  DCL stands for Data Control Language. SQL queries like GRANT and REVOKE come under this.
- DQL: data query language

18. What are the major functions of the database administrator?

- Managing database structure,
- controlling concurrent processing,
- managing processing rights and responsibilities, developing database security,
- providing for database recovery,
- managing the DBMS and maintaining the data repository

19. What is Enterprise Resource Planning (ERP),and what kind of a database is used in an ERP application?

- Enterprise Resource Planning (ERP) is an information system used in manufacturing companies and includes sales, inventory, production planning, purchasing and other business functions. An ERP system typically uses a multiuser database

20. What is the use of DROP command and what are the differences between DROP, TRUNCATE and DELETE commands?

- **DROP** command is a DDL command which is used to drop/delete the existing table, database, index or view from the database.
- **The major difference between DROP, TRUNCATE and DELETE commands are:**
- **DROP** and **TRUNCATE** commands are the **DDL** commands which are used to delete tables from the database and once the table gets deleted, all the privileges and indexes that are related to the table also get deleted. These 2 operations cannot be rolled back and so should be used only when necessary.
- **DELETE** command, on the other hand, is a **DML** Command which is also used to delete rows from the table and this can be rolled back.

DROP	**TRUNCATE**
The DROP command is used to remove table definition and its contents.	Whereas the TRUNCATE command is used to delete all the rows from the table.
In the DROP command, table space is freed from memory.	While the TRUNCATE command does not free the table space from memory.
In the DROP command, integrity constraints will be removed.	While in this command, integrity constraints will not be removed.

23. What is Relational Algebra?

- Relational Algebra is a Procedural Query Language consisting of a set of operations. It takes one or more relations (tables) as input and produces a new relation (table) as output.
- Some fundamental operations of Relation Algebra are:
  - Set difference
  - Select
  - Union
  - Rename
  - Project

24. State the difference between UNION and UNION ALL

- Both UNION and UNION ALL are used to join data from more than 2 tables.
- The difference is that UNION removes any duplicate rows in the tables while UNION ALL does not remove any duplicate data.

26. What is a checkpoint in DBMS?

- A checkpoint in DBMS is a point at which all the previous logs are deleted from the system and stored on the storage disk.
- These stored logs can be used in log recovery to recover the database engine

- **What is database normalization?**

- It is a process of analyzing the given relation schemas based on their functional dependencies and primary keys to achieve the following desirable properties:
  1) Minimizing Redundancy
  2) Minimizing the Insertion, Deletion, And Update Anomalies
  Relation schemas that do not meet the properties are decomposed into smaller relation schemas that could meet desirable properties

- **What is the purpose of normalization in DBMS?**

- Normalization is the process of analyzing the relational schemas which are based on their respective functional dependencies and the primary keys in order to fulfill certain properties.

- **What is the concept of sub-query in terms of SQL?**

- Sub-query is basically the query which is included inside some other query and can also be called as an inner query which is found inside the outer query

```
SELECT *
FROM EMPLOYEE
WHERE ID IN (SELECT ID
FROM EMPLOYEE
WHERE SALARY > 4500);
```

- **What is a functional dependency in the DBMS?**
- **What is 1NF in the DBMS?**
- **What is 2NF in the DBMS?**
- **What is 3NF in the DBMS?**
- **What is BCNF in the DBMS?**

Functional Dependency (FD) is a constraint that determines the relation of one attribute to another attribute in a Database Management System (DBMS).

$\alpha \to \beta$

R(ABCD)

NP → NP

- This is basically a constraint which is useful in describing the relationship among the different attributes in a relation.
- **Example:** If there is some relation 'R1' which has 2 attributes as Y and Z then the functional dependency among these 2 attributes can be shown as **Y->Z** which states that Z is functionally dependent on Y.

SK

AB → CK

TA → O

- 1NF is known as the **First Normal Form**.
- This is the easiest form of the normalization process which states that the **d**omain of an attribute should have only atomic values. The objective of this is to remove the duplicate columns that are present in the table.

AB → C

B → C

C → D

A → D

2NF

- **:** 2NF is the **Second Normal Form**.
- Any table is said to have in the 2NF if it satisfies the following 2 conditions:
- A table is in the 1NF.
- Each non-prime attribute of a table is said to be functionally dependent in totality on the primary key

NP	NP
SV	SM

- 3NF is the **Third Normal Form.**
- Any table is said to have in the 3NF if it satisfies the following 2 conditions:
- A table is in the 2NF.
- Each non-prime attribute of a table is said to be non-transitively dependent on every key of the table.
- BCNF is the **Boyce Codd Normal Form** which is stricter than the 3NF**.**
- Any table is said to have in the BCNF if it satisfies the following 2 conditions:
- A table is in the 3NF.
- For each of the functional dependency X->Y that exists, X is the super key of a table

- **What is a join in the SQL?**

- A Join is one of the SQL statements which is used to join the data or the rows from 2 or more tables on the basis of a common field/column among them

- **What are different types of joins in SQL?**

- **There are 4 types of SQL Joins:**

- **Inner Join:** This type of join is used to fetch the data among the tables which are common in both the tables.

- **Left Join:** This returns all the rows from the table which is on the left side of the join but only the matching rows from the table which is on the right side of the join.

- **Right Join:** This returns all the rows from the table which is on the right side of the join but only the matching rows from the table which is on the left side of the join.

- **Full Join:** This returns the rows from all the tables on which the join condition has put and the rows which do not match hold null values.

- **What are the Aggregate functions in SQL?**

- In database management an aggregate function is a function where the values of multiple rows are grouped together as input on certain criteria to form a single value of more significant meaning

1) Count()
2) Sum()
3) Avg()
4) Min()
5) Max()

- **What is order by clause?**

- ORDER BY clause helps to sort the data in either ascending order to descending

- **What is a transaction? What are ACID properties?**

- A transaction is a logical unit of database processing that includes one or more database access operations.
- A Database Transaction is a set of database operations that must be treated as whole, means either all operations are executed or none of them.
- An example can be bank transaction from one account to another account. Either both debit and credit operations must be executed or none of them.
- ACID (Atomicity, Consistency, Isolation, Durability) is a set of properties that guarantee that database transactions are processed reliably.

- **What are the disadvantages of using a dbms?**

- 1) High initial investments in h/w, s/w, and training.
  2) Generality that a DBMS provides for defining and processing data.
  3) Overhead for providing security, concurrency control, recovery, and integrity functions.

- **What is Identity?**

- Identity (or AutoNumber) is a column that automatically generates numeric values. A start and increment value can be set, but most DBA leave these at 1 Identity columns do not need to be indexed

- **What is a view in SQL? How to create one?**

A view is a virtual table based on the result-set of an SQL statement. We can create using create view syntax.

- `CREATE VIEW view_name AS`
- `SELECT column_name(s)`
- `FROM table_name`
- `WHERE condition`

Explain the Stored Procedure

- Stored Procedure is a group of SQL statements in the form of a function that has some unique name and is stored in relational database management systems(RDBMS) and can be accessed whenever required

- **What is SQL Deadlock?**

- Deadlock is a unique situation in a multi user system that causes two or more users to wait indefinitely for a locked resource

- **Explain the difference between an exclusive lock and a shared lock?**

- An exclusive lock prohibits other users from reading the locked resource; a shared lock allows other users to read the locked resource, but they cannot update it.

- Explain 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.

```
create trigger safety
on database
for
create_table,alter_table,drop_table
as
print'you can not create ,drop and alter table in this database’
rollback;
```

```
create trigger deep
on emp
for insert,update,delete
as
print'you can not insert,update and delete this table i’
rollback;
```

- **Name four applications for triggers.**

(1)Providing default values,

(2) enforcing data constraints,

(3) Updating views and
(4) enforcing referential integrity

- **What are serial, non serial schedule?**

- A schedule S is serial if, for every transaction T participating in the schedule, all the operations of T is executed consecutively in the schedule, otherwise, the schedule is called non-serial schedule.

- **What are conflict serializable schedules?**

- A schedule S of n transactions is serializable if it is equivalent to some serial schedule of the same n transactions.

- **What is the difference between having and where clause?**

- HAVING is used to specify a condition for a group or an aggregate function used in select statement. The WHERE clause selects before grouping. The HAVING clause selects rows after grouping. Unlike HAVING clause, the WHERE clause cannot contain aggregate functions

```
SELECT column1,
column2 FROM table1,
table2 WHERE [
conditions ] GROUP BY
column1, column2
HAVING [ conditions ]
ORDER BY column1,
column2
```

NOTE: The HAVING clause must follow the GROUP BY clause in a query and must also precede the ORDER BY clause if used

```
SQL > SELECT ID, NAME,
AGE, ADDRESS, SALARY
FROM CUSTOMERS
GROUP BY age
HAVING COUNT(age) >= 2;
```

- **What is the difference between Trigger and Stored Procedure?**

- Unlike Stored Procedures, Triggers cannot be called directly. They can only be associated with queries

- **What are indexes?**

- A database index is a data structure that improves the speed of data retrieval operations on a database table at the cost of additional writes and the use of more storage space to maintain the extra copy of data.

- Data can be stored only in one order on disk. To support faster access according to different values, faster search like binary search for different values is desired, For this purpose, indexes are created on tables. These indexes need extra space on disk, but they allow faster search according to different frequently searched values

- Foreign Key Syntax

- CREATE TABLE Orders (
-     OrderID int NOT NULL,
-     OrderNumber int NOT NULL,
-     PersonID int,
-     PRIMARY KEY (OrderID),
-     FOREIGN KEY (PersonID) REFERENCES Persons(PersonID)
-);
- Table1: Person
- Table2: Order