Advantages of DBMS over file system and relational model:

- Data redundancy and consistency
- Data sharing
- Easy access to data
- Data concurrency
- Data searching
- Privacy & data security
- Easy access to data
- Data integrity

Types of database users:	**DBMS Schemas:**
● System administrators ● Database administrators ● Database designers	● Internal level ● Conceptual or logical level ● External or view level

ER model

- ER model stands for an Entity-Relationship model. It is a high-level data model. This model is used to define the data elements and relationship for a specified system.
- It develops a conceptual design for the database. It also develops a very simple and easy to design view of data.

Advantages:

- Allows us to draw database design
- It is an easy to use graphical tool for modelling data
- Widely used in database design
- Helps us to identify entities which exist in the system

Entity: A real world thing either living or non living that is easily recognizable or non - recognizable. Entities have attributes, which can be considered as properties describing it.

STRONG ENTITY	WEAK ENTITY
● An entity that contains sufficient attributes to uniquely identify all its entities ● It contains primary key	● An entity that does not contain sufficient attributes for unique identification ● It contains partial discriminator key

The relationship between string and weak entity is called an **identifying relationship.**

Constraints

Participation: total partial	Cardinality: 1 to 1, 1 to many, many to 1, many to many

EER: extended entity relationship

- generalisation: is a bottom-up approach in which multiple lower-level entities are combined to form a single higher-level entity. Subclasses are combined to form an upper class.
- specialisation: is a top-down approach in which a higher-level entity is divided into multiple specialized lower-level entities.
- aggregation: refers to when relation between 2 entities is considered as a single entity.

Relational model: represents data in the form of relations or tables.

Relational Algebra: is a procedural query language, which takes instances of relations as input and yields instances of relations as output. It uses operators to perform queries. An operator can be either unary or binary.

Two categories

BASIC	FUNDAMENTAL
Select	Natural join
Project	Theta join
Union	Equi join
Set difference	Left , right, full outer join
Cartesian product	Intersection
rename	division

SQL:

- Structured query language
- Used to communicate with relational dbms
- Like creating able, querying the database for information, modifying data

Advantages:

1. Faster Query Processing –
2. No Coding Skills –
3. Standardised Language –
4. Portable –

5. Interactive Language –
6. Multiple data views -

SQL is divided as:

- **DDL:**DDL changes the structure of the table like creating a table, deleting a table, altering a table, etc.
  -create
  -drop
  -alter
  -truncate
- **DML:**DML commands are used to modify the database. It is responsible for all forms of changes in the database.
  -insert
  -update
  -delete
  -merge
- **DCL:**DCL commands are used to grant and take back authority from any database user.
  -grant
  -revoke
- **TCL:** These are used to manage transactions in a database.
  -commit: used to permanently save the transaction
  -savepoint: used to temporarily save a transaction
  -rollback: restores the database to the last committed point

SET OPERATIONS	**ARITHMETIC OPERATIONS**	**COMPARISON CONDITIONS**	**LOGICAL CONDITIONS**
Union	Add	Between … and..	And
Intersect	Subtract	In	Or
minus	Multiply	Like	not
	divide	Is null	

VIEWS: a view is a virtual table based on the result-set of an SQL statement
Why use it?

- To restrict data access
- To make complex queries easy
- To provide data independency
- To provide diff views of the same data

JOINS:

ORACLE	SQL
Equi join(inner join) Outer join	Cross join, Natural join, Left outer join, Right outer join, Full outer join

TRIGGERS: Triggers are stored programs, which are automatically executed or fired when some events occur.

Triggers can be written for the following purposes −

- Generating some derived column values automatically
- Enforcing referential integrity
- Event logging and storing information on table access
- Auditing
- Synchronous replication of tables
- Imposing security authorizations
- Preventing invalid transactions

Normalisation: it is a set of rules that structures and tables need to follow to be well structured.

Characteristics:

- Scalar values in each fields
- Absence of redundancy
- Minimal use of null values
- Minimal loss of information

Functional Dependency: is a constraint that determines the relation of one attribute to another attribute in a Database Management System. Functional Dependency helps to maintain the quality of data in the database. It plays a vital role to find the difference between good and bad database design

Levels of Normalisation:

1NF: disallows-composite attributes,Multivalued attributes & Nested relations

2NF:must be-1NF & No partial dependencies

3NF:must be-2NF & No transitive dependency

BCNF(boyce-codd normal form): Must be-3NF & For every functional dependency, LHS attribute must be a superkey

4NF:must be-BCNF & Should not have any multi-valued dependency

TRANSACTION:

It is a unit of program execution that accesses and possibly updates various data items. It is a set of logically related operations. It contains a group of tasks.

Operations: Retrieve, insert, delete, update, commit, rollback, savepoint

Transaction states:

- Active: transaction stays in this state while execution
- Partially committed: after the final statement has been executed
- Failed: when normal execution can no longer proceed
- Aborted: after transaction has been rolled back and restored to its initial state
- Committed: after successful transaction

Transaction properties:ACID

- **ATOMICITY:** either all operations are properly executed or none are
- **CONSISTENCY:** execution of a transaction performed in isolation preserves the consistency in database
- **ISOLATION:** although multiple transactions may execute concurrently, each transaction must be unaware of each other concurrently executing transactions.
- **DURABILITY:** After a transaction completes successfully, the changes it has made must persist even if there are system failures.

SCHEDULES: Sequences that indicate the chronological order on which instructions of concurrent transactions are executed.

1. It must consist of all instructions of those transactions
2. Must preserve the order in which the instruction will appear

SERIALIZABILITY: Schedule is serialisable if it is equivalent to a serial schedule

CONFLICT SERIALIZABILITY	**VIEW SERIALIZABILITY**
If all the conflicting operations in 2 schedules get executed in the same manner then they are conflict serializable.	If the order of initial read , final write and update operations in 2 schedules are the same then they are view serializable.

CONCURRENCY CONTROL:is the process of managing simultaneous execution of transactions without conflicting with each other.

This allows many transactions to access the same database at the same time without interfering with each other.

Problems:

- Lost updates
- Dirty read(uncommitted data)
- Unrepeatable read(inconsistent retrievals)

LOCK-BASED PROTOCOL	TIMESTAMP-BASED PROTOCOL
A lock is a mechanism to control concurrent access to a data item A data item can be locked in 2 ways: ● Exclusive:data can be both read as well as written ● Shared: data can only be read	It manages concurrent execution such that the timestamp determines the serialisability order.protocol maintain 2 timestamp values: ● W-timestamp: that executed write successfully ● R-timestamp: that executed read successfully

Starvation: Starvation is the situation when a transaction needs to wait for an indefinite period to acquire a lock.

Deadlock: Deadlock refers to a specific situation where two or more processes are waiting for each other to release a resource or more than two processes are waiting for the resource in a circular chain.
Necessary conditions:

- Mutual exclusion
- Hold and wait
- Circular wair
- No pre emption

Deadlock avoidance: used to detect any deadlock situation in advance by using some techniques.The key of Deadlock avoidance approach is when the request is made for resources then the request must only be approved in the case if the resulting state is also a safe state.

Wait for graph: a graph is created based on transactions and their locks.
While creating a graph if a cycle is encountered then there is a deadlock.

Banker's Algorithm : is resource allocation and deadlock avoidance algorithm which test all the request made by processes for resources, it checks for the safe state, if after granting request system remains in the safe state it allows the request and if there is no safe state it doesn't allow the request made by the process.

Deadlock prevention: it analyses the transaction to determine whether any deadlock situation can arise or not.

- Wait die scheme
- Wound wait scheme

DATABASE RECOVERY:

It means recovering the data when it gets deleted, hacked or damaged accidentally

Purpose:

- To bring the database in the last consistent state which existed prior to any failure
- To preserve transaction properties(ACID).

FAILURE

Transaction failure system crash disk failure

  - Logical errors
  - System errors

FIELDS IN TRANSACTION LOG:

- **TRANSACTION IDENTIFIER:** is the unique identifier of the transaction the performed the write operation
- **DATA ITEM:** unique identifier of the data item written
- **OLD VALUE:** value of data item prior to write
- **NEW VALUE:** value of data item after write operation

LOG BASED RECOVERY:

A log is a sequence of log records, recording all the update activities done on the database by all database users.
Recovery: after a system crashes, the system consults with the log to determine which transactions need to be redone and which need to be undone.
2 approach:

1. Deferred database modification
2. Immediate database modification

SHADOW PAGING:

-Is a recovery technique that is used to recover databases.
-In this database is considered as made up of fixed size logical units of storage which are referred to as pages.
-Pages are mapped into physical blocks of storage, with the help of a page table.
-This method uses two page tables: current & shadow page table.