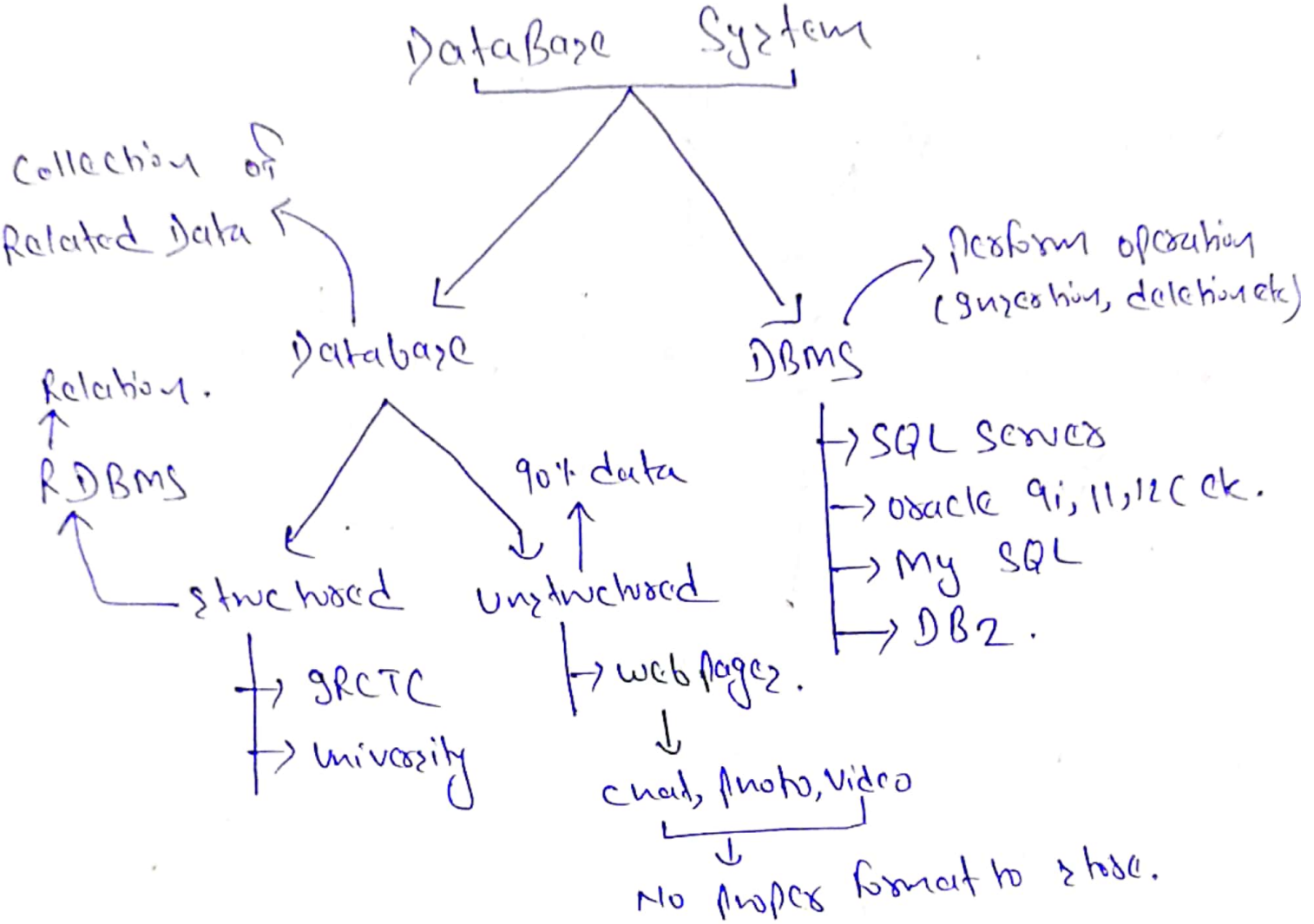
Data Base System



=> File System Vs DBMS

(i) Data (Data store in Server)

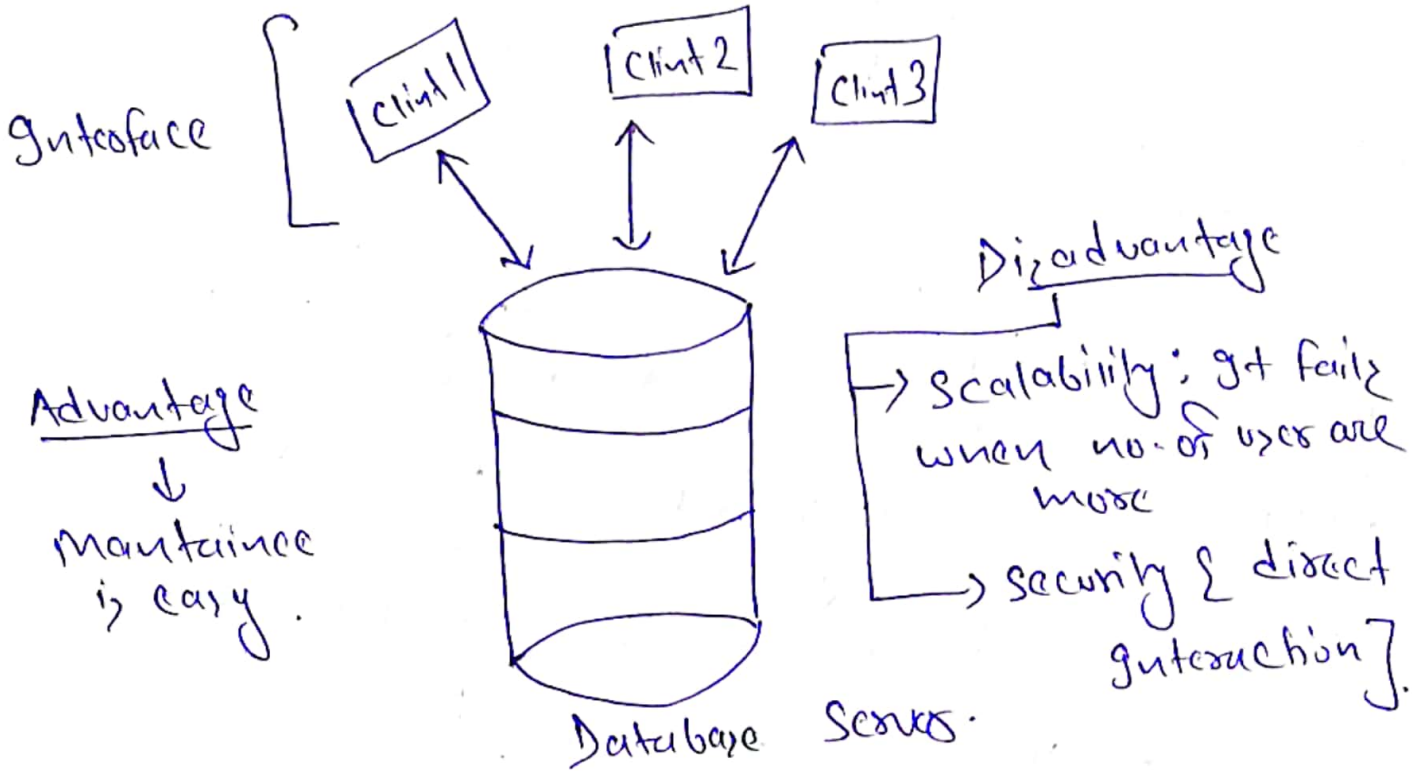
(ii) Attributes (file location)

(iii) Concurrency (No. of people accessing at time)

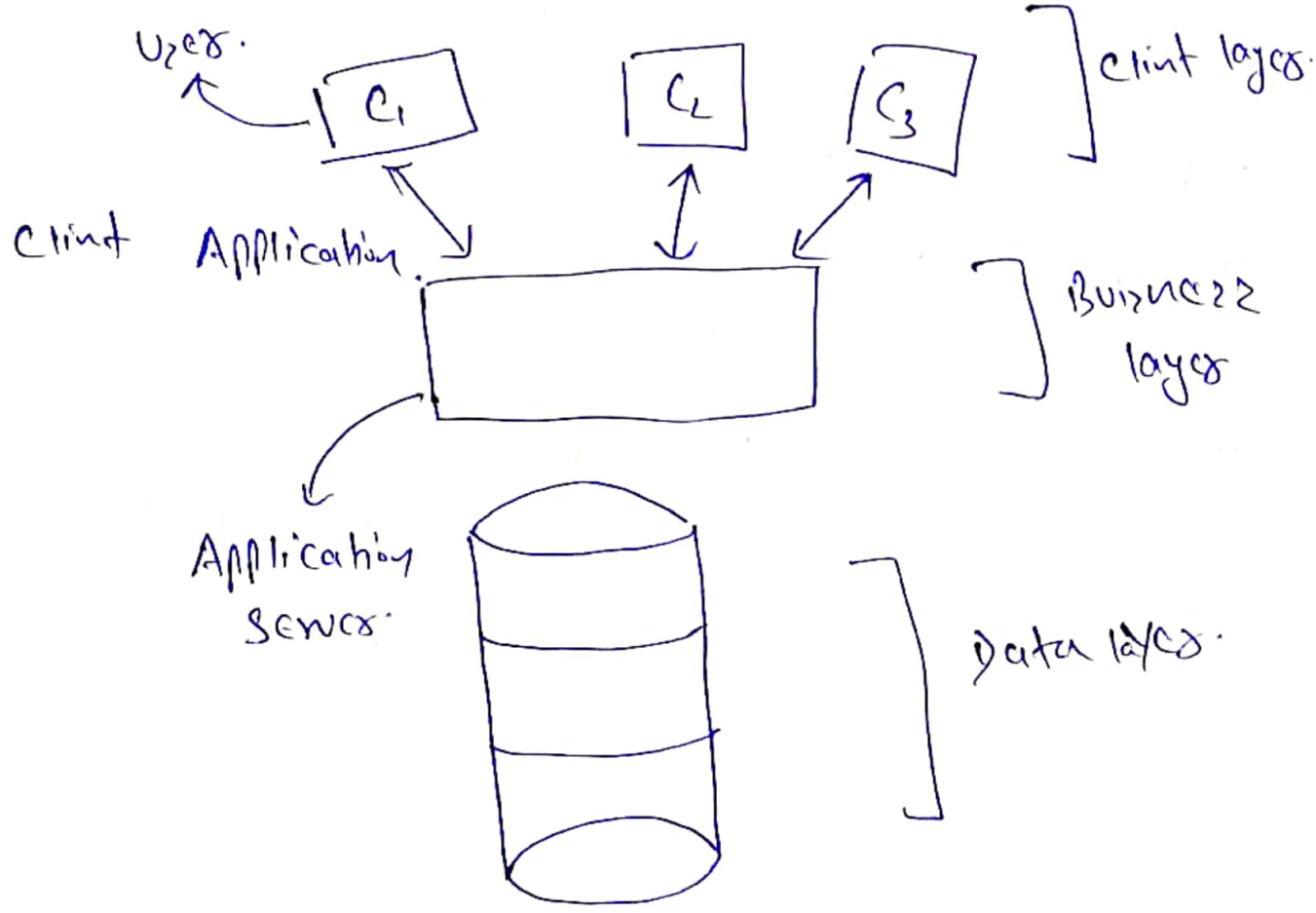
(iv) Security [Role base access control]

(v) Redundancy [No duplication - Store unique Data]

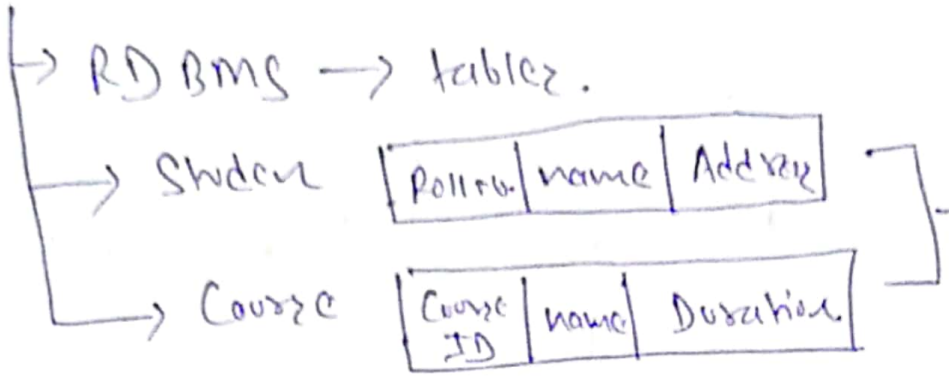
* 2-tier Architecture :



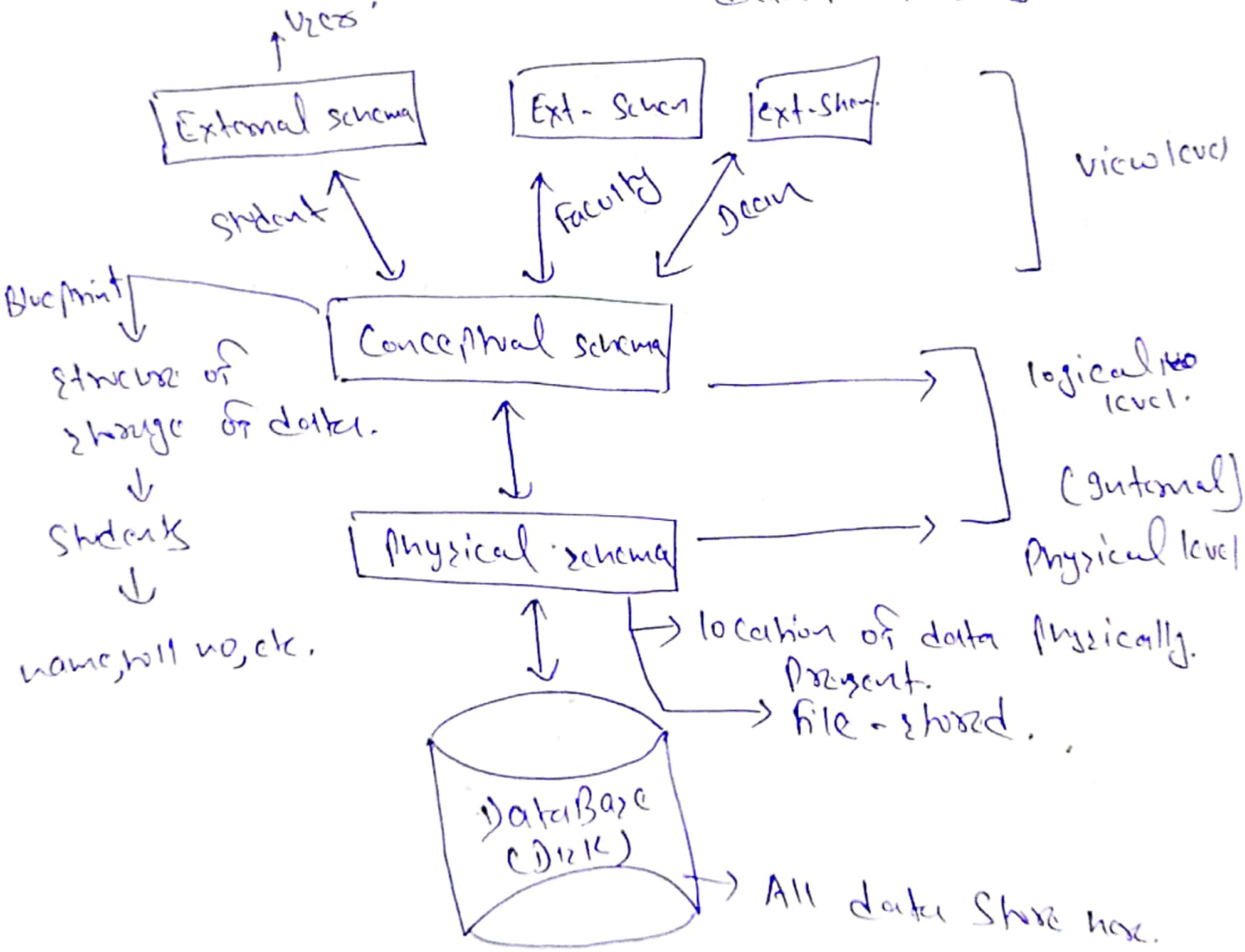
* 3-tier Architecture



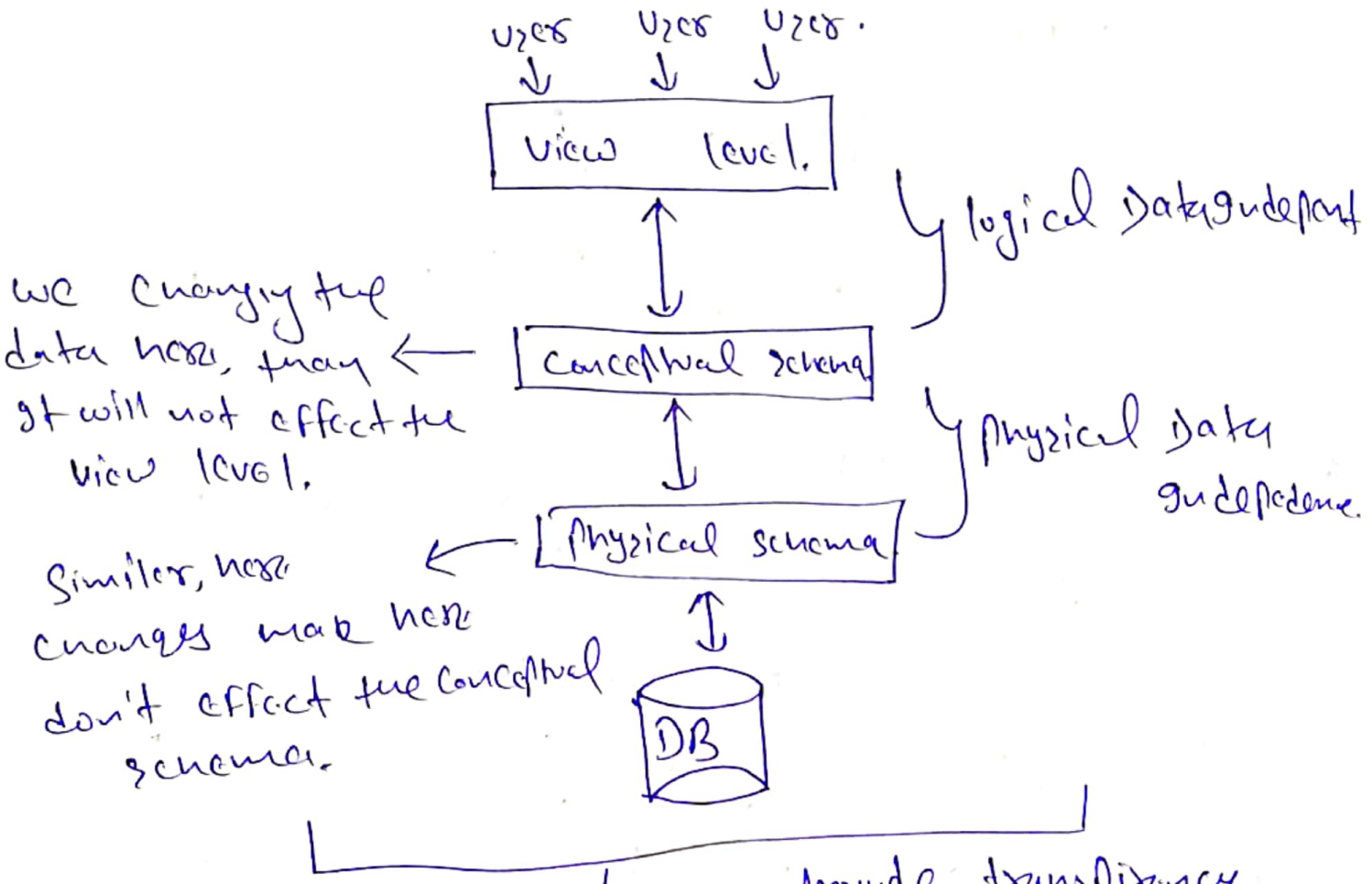
* **Schema** : logical Representation.



* **3-Schemas Architecture** : [aim to hide actual storage of data from users]



Data Independence



⇒ Concept of Candidate key:

→ what is key → Attribute.

→ Use of key → uniquely identify. → key which we use to identify is called primary key. and Remaining called Alternative key.

Student tab.

S-name	City	Age.
Reddy	hydrabad	20
Rajwal	kanpur	21
Reddy	hydrabad	20

1) Adhar Card
2) Roll No.
3) Reg. No.
4) Licence no.
5) Voter ID
6) Phone No.
7) Email ID

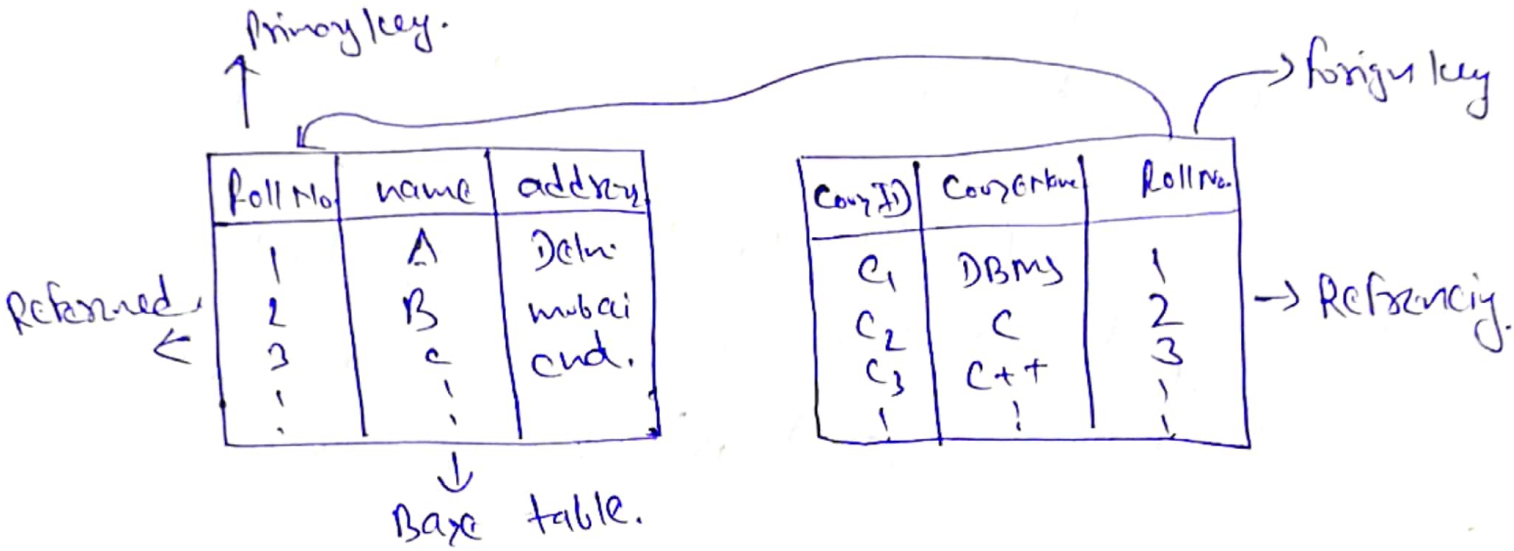
This all are unique for everyone. and this called Candidate key.

⇒ Primary key: (unique + Not Null)

↳ We can't leave this coloumn blank.

⇒ foreign key: It is an attribute or set of attribute that references to primary key of some table or another (relation).

→ Maintains referential integrity



foreign key → Referential integrity.

↳ value should be same.

Referenced table:

i) Insert: No violation.

ii) Delete: may cause violation → best case.

→ Solution: on delete Cascade. [use as part of syntax]

→ solution: on delete set Null.

→ on delete No action. [By default]

iii) Updation: May cause violation.

→ on updation Cascade

→ On updation set Null

→ on updation set NULL.

→ Referencing table:

1) Insert → May cause violation.
2) Delete → will not cause any violation.
3) Updation → May cause violation.

⇒ Super key: A super key is a combination of all possible attributes which can uniquely identified two tuples in a table.

→ Super set of any candidate key is super key.

Entity - Relationship model (ER-Model)

Eg → Student ←→ Anything | which has physical Appearance

Attribute ← (Roll no, age, addr) Detecting the relationship.

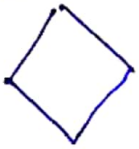
Eg. Course : [CSC, IT]

→ Relation between student and course is called study.

Representation

Entity → 

Attributes → 

Relationship → 

* Types of Attributes:

1) Single vs Multivalued Attributes.
   - Single → Reg. No.
   - Multivalued → More than one Attribute
     - eg. → Phone No] → Single Student can have 2 Phn. no.
     - eg → Address [Permanent + correspondence]

2) Simple vs Composite Attributes:
   - It can be further divided.
     - eg → age
   - It can be divided further. {name: first name + middle name + last name}

3) Stored vs Derived Attributes:
   - It is fixed, it can be derived.
     - eg: Date of birth.
   - We don't know Actual Attribute but deriving the attribute from some other attribute.
     - eg → age, Course End date.
   - It is denoted by (dotted ellipse) → dotted ellipse

4) Key vs Non-key Attribute:
   - It is used to uniquely identify the data.
     - No redundancy.
     - eg → Reg No.
   - Non-unique value is also possible.

(v) Required vs optional Attribute:

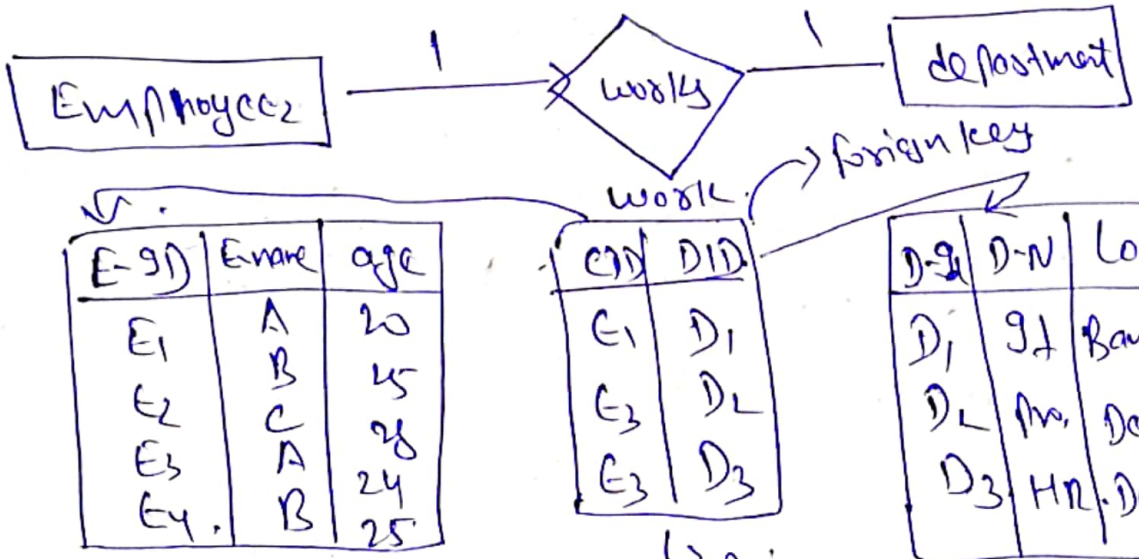
- It is must, necessary.
  - eg -> Name.
- It is optional
  - eg -> passing marks.

---

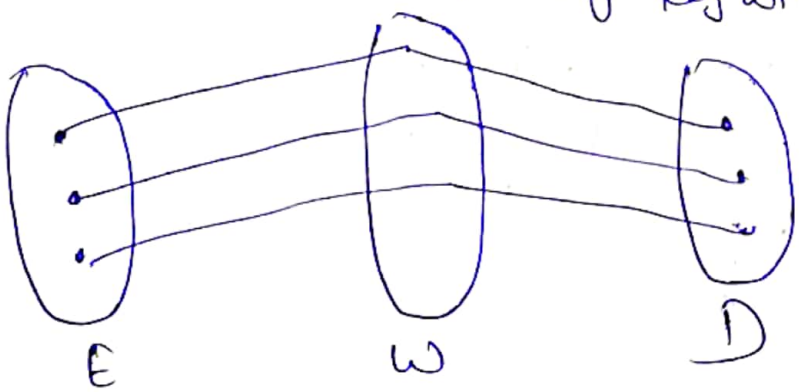
* Various types of relationship:

- 1-1 (one to one)
- 1-M (one to many)
- M-1 (many to one)
- M-M (M-N) (many to many).

=> one to one Relationship:



↳ Primary key will either (E-ID) and (D-ID).

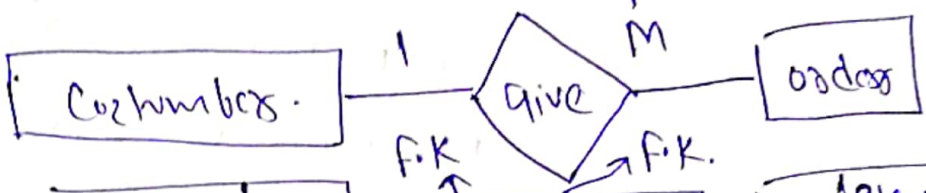


→ We can combine the two table to make the single table: make one primary key.

Primary key.

EID	E-name	age	DID
E1	A	20	D_1
E2	B	25	D_3
E3	C	28	D_2
E4.	A	24	— (NULL)

⇒ One- Many - Relationship:



P.K. ←

I.D	Name	City
C_1	A	Jal.
C_2	B	Delhi
C_3	C	Mum
C_4	A	Mum

F.K ↑ F.K.

S.D	O.N	Date
C_1	O_1	
C_1	O_2	
C_2	O_3	
C_2	O_4	

O.N	Ste-N.	Cost
O_1	Shirt	100
O_2	Shoes.	200
O_3	Shirt.	1500
O_4	Jeans	1200

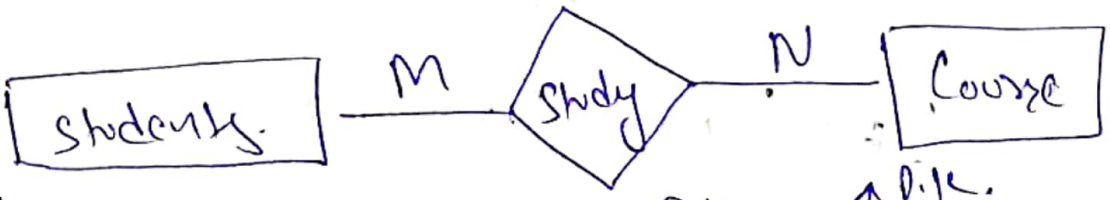
↳ O.N act as primary key here.

{ Many side primary key will decide the P.K for give. }

⇒ Merging the table:

ID	O-N	Cost.
C_1	O_1	100
C_2	O_2	200
C_2	O_3	100
C_1	O_4	1500

⇒ Many to many Relationship:



P.K.

Roll.	name	age
1	A	16
2	B	17
3	A	16
4	C	17
5	D	15

F.K F.K

Roll-No	C-ID
1	C_1
2	C_2
1	C_2
2	C_1
3	C_3

P.K.

C-ID	name	credit
C_1	Maths	4
C_2	Phy	4
C_3	Chem	4
C_4	Hindi	4

↳ (Rool No + C-ID) → Composite key will make Primary key.

In this we can't reduce the table.

DBMS

Data :- Data is the raw material that can be processed for any computing machine eg :- employee name, product name, name of the student, marks of the student, any number, image

Information :- It is the data that has been converted into more useful or intelligable form eg report card sheet

Why we need information

1) To gain knowledge about the surroundings
2) To keep the system upto date
3) To know about the rules and regulation of the society

Knowledge :- Human mind purposefully organise the information and evaluate it to produce knowledge eg 238 is a data and Marks of student is information and the hard work require to get mark is knowledge

Knowledge

Fact Based Heuristic Based

1) Fact Based :- The knowledge gain from fundamental & through experiment

2) Heuristic Based :- It is the knowledge of good practice

and good judgement like hypothesis

Difference between data and information

	Data	Information
①	Data is the raw fact	It is the processed form of data
②	It is not significant to a buisness	It is significant to a buisness
③	Data are Atomic level piece of information	It is a collection of data
④	Data doesnot help in decision making	It help in decision making
⑤	eg:- product name, name of student	eg report Card Sheet

Database:- The related information when placed in an organised form makes a database or an organised collection of related information is known as database eg Dictionary, Telephone Directory, Mobile Contact

Operation perform on Database

1) Insertion
2) Updation
3) Deletion
4) Retrive
5) Sorting

Difference between Computerised databases and manual databases

Traditional File System :-

File System :- A file system is the method of storing and organising the computer files and the data they contain to make it easy to find and access them

Characteristic of files System :-

1) It is a group of files for storing the data of an organisation
2) Each file is independent from one another
3) Each file is called a flat files
4) Files are design by using the program written in programing language such as c, c++

Limitation / Disadvantages of File processing System :-

1) Seperated and isolated data
2) Duplication of data :- 1) It cost time and money
   2) It takes up additional storage space
   3) It can lead to loss of Data Integrity
3) Data dependencies :- files and record were describe by specific physical format that were code in the application program by the programs
4) Difficulty in representing the data from the User point of view
5) Data security :- The security of data is low in the file based system because the data is maintain in a flat file is easy accessible

6) Transactional problems :- This System does not satisfy transactional properties called ACID properties A → Atomicity, C → Consistency I → Insolation, D → Durability

7) Concurrency problems :- When multiple user access a same peace piece of data at a same interval of time then it is called as Concurency of System. When two or more user read the data simultaneously then their is no problem but when they like to update the file simultaneously It may result in a problem

Building Block of Database :-

1) Column / fields
2) Rows / tulple / Record
3) Tables

DBMS (Data Base Management System) :- It is the Software System that allow the User to define, to create and maintain the database and provide control access to the data

Application of Database :-

i) Library System
2) Banking System
3) ATM

Database :- Mysql, oracle, sql server, DB2, Microsoft Access

Components of DBMS :-

1) Hardware :- The hardware is the actual computer system used for keeping and accessing the database Conventional DBMS hardware consist of secondary storage devices such as harddisk. Database run on the range of machine from micro computers to main frames

2) Software :- Software is the actual DBMS between the physical Database and the users of the system All the request from the user for accessing the database are handled by DBMS

3) Data :-

4) Users :- There are no of users who can access or retrieve the data on demand using the application and the interfaces provided by the DBMS
The Users of the database can be classified into the following groups.

1) Naive Users
2) Online users
3) Sophisticated users
4) Specialized Users
5) Application programmers
6) DBA - Database Administrator

① Naive Users :- Those user who need not be aware of the presence of the database system. They are the end users of the database who

work through a menu driven application programs, where the type and range of response is always indicated to the users

2) Online Users:- Those users who may communicate with database directly through an online terminal or indirectly through user interface and application program.

3) Sophisticated User:- They are those user who interact with the system without writting the program. Instead they form their request in database query language.

4) Specialized User:- Those users who write specialized database application that donot fit into the fractional database processing framework.

5) Application Programmer:- Those users who are responsible for developing the application programs or user interface. The application programs could be written in high level language.

6) DBA - Database Administrator:- It is a person or the group incharge for implementing the database system within the organisation. The DBA has all the privilage allowed by the DBMS and can assign or remove the privilages from the users

5) Procedure :-

Disadvantages of DBMS :-

1) Complexity :-
2) Size
3) Performance
4) Higher impact of failure
5) Cost of DBMS

Differentiate between File Management System & DBMS.

* Master file :- Master file are those file which remain static. There is no change

- Transaction File :- Transaction file are those file which is dynamic in nature. We can made changes

- Instances :- The situation data in the database at a particular moment of time is called an instance

- Schema :- The overall design of the database is called Schema OR Description of database.

- Subschema :- It is the subset of the schema and inherit the same property that a schema has. It gives the users a window through which he/she can view only that part of database which is of interest to him.

Passion
Page No.
Date

Architecture of DBMS :-

There is 3 level

1) External level
2) Conceptual Level
3) Internal Level

Objective of three Level Architecture or Sparc 3level Architecture :-

The objective is to seperate each users view of the data from the way the database is physically represented.

There are Several reasons

1) The internal structure of the database should be uneffected while changes to the physical aspects of storage
2) The DBA should be able to change the conceptual structure of the database without affecting all othess users

1) External level | viewlevel :- This level describes that part of the database that is relavent to each users. This level insulates the users from the details of conceptual and the internal level.

2) Conceptual level | logic level :- This level describe what data is stored into the database and the relationship among the data.

It represents :-

(a) All the Entities, attributes and there relation -ships
(b) The constraints on the data

c) Security and integrity information

/Storage level

3) Internal level:- It is the physical representation of the database on the computer. This level describe How the data is stored in the database. It covers the data structures and file organisation used to store the data on storage devices.

Schemas:-

1) External Schema
2) Conceptual Schema
3) Internal Schema

1) External Schema:- The external view is described by means of a schema called External Schema. That corresponds to differents view of the data

2) Conceptual Schema:- The conceptual view is defined by conceptual schema, which describes all the entities, attributes and their relationship with the integrity constraints

3) Internal Schema:- Internal level is defined by internal schema, which is a complete description of the internal model

There is only 1 Conceptual Schema and 1 internal schema per database and more than 1 external Schema

Schema is also known as Intension

NOTE:- Instance / Extension of database

Mapping between the levels:-
1) External / Conceptual Mapping
2) Conceptual / Internal Mapping

1) External / Conceptual Mapping:- Each external Schema is related to the conceptual Schema by external conceptual Mapping. This Mapping gives the correspondance among the records and the relationships of the external & Conceptual Views. There is a mapping from a particular logical record in the external View to one or more conceptual record in the Conceptual View.

2) Conceptual / Internal Mapping:- Conceptual Schema is related to Internal Schema by Conceptual Internal Mapping. Mapping between the Conceptual and Internal level Specify the method of deriving the Conceptual Record from physical database

Data Independence:-
1) Logical data independency
2) Physical data independency

1) Logical data independency:- It indicates that the Conceptual Schema can be changed without effecting the existing external schema. The changes would be absorbed by the mapping

between external and conceptual level

2) Physical data independency:- It indicates that the physical storage structure or devices can be changed without effecting the conceptual schema. The change would be absorded by the conceptual internal mapping.

→ Logical data independency is much more difficult to achieve than Physical data independency as it requires the flexibility in the design of the database and programer has to see the future requirement or modification in the design

- Limitation of file processing system:-

1) Separated and Isolated Data:- To make a decision, a user might need data from two separate files. First the files were evaluated by analysts and programmers to determine the specific data required from each file and the relationship between the data and then application could be written in a programming language to process and extract the needed data.

2) Difficulty in representing data from the user's View:- To create useful application for the user, often data from various files must be combined. In file processing it was difficult to determine relationships between isolated data in order to meet user application.

Components of DBMS:-

- Data:- It is the most important component of DBMS environment from the end users point of view. One of the major features of database is that actual data are separated from the programs that use the data. A database should always be designed, built and populated for a particular audience and for a specific purpose.

- Procedures:- Procedures refer to the instructions and rules that govern the design and use of the database. The user of the system and the staff that manage the database require documented procedures on how to use or run the system.

Disadvantages of DBMS:-

1) Complexity:- The provision of the functionality that is expected of a good DBMS makes the DBMS an extremely complex piece of software. Database designers, developers, database administrators and end-users must understand this functionality to take full advantage of it. Failure to understand the system can lead to bad design decisions, which can have serious consequences for an organization.

2) Size:- The complexity and breadth of functionality makes the DBMS an extremely large piece of Software, occupying megabytes of disk space and requiring substantial amount of memory to run efficiently.

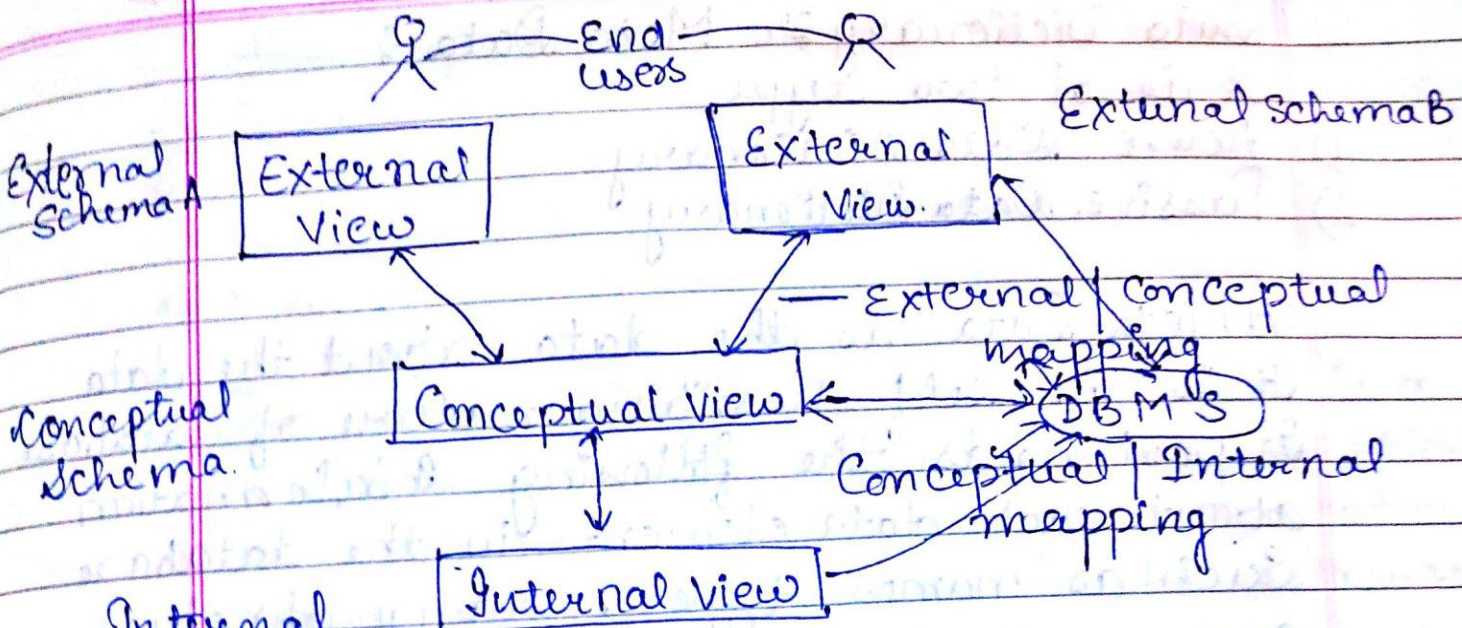
3) Performance:- A file Based System is written for a specific application such as invoicing. As result performance is generally very good. However the DBMS is written to be more general, to cater for many application rather than just one.

4) Higher impact of a failure:- The centralization of resource increases the vulnerability of the system. Since all user and applications rely on the availability of the DBMS, the failure of any component can bring operation to a halt.

5) Cost of DBMS:- The cost of DBMS varies significantly depending on the environment and functionality provided. There is also the recurrent annual maintenance cost.

Passion
Page No.
Date

File Management eg C++ or COBOL Program	Database Management eg oracle or Sybase
1) Small System	1) Large System
2) Relatively cheap	2) relatively expensive
3) few files	3) many files
4) files are files	4) files not necessarily files
5) Simple Structure	5) Complex Structure
6) little preliminary design	6) Vast preliminary design.
7) integrity left to application programmer	7) rigorous inbuilt integrity checking
8) no security	8) rigorous Security
9) Simple, primitive backup/recovery	9) Complex & Sophisticated backup/recovery
10) often Single User	10) multiple user



Role of DBA (Database Administrator.) :- It is a person or group incharge for implementing DBMS in an organisation. The DBA job requires high degree of technical expertise DBA consist of a team of people rather than just one person

Responsibilities of DBA :-

1) Makes the decision concerning the content of the database
2) Plans the storage structure and access strategy
3) Provides the support to the users
4) Defines the security and integrity checks
5) Interprets backup and recovery strategies
6) Monitoring the performace and responding to the changes in the requirements

Data Dictionary or Meta Data :-

It is of two type

1) Active Data Dictionary
2) Passive Data Dictionary

A Meta Data is the data about the data It is the self describing nature of database. It part holds the following information about each data element in the database such as names, types, range of values access authorization, inolicate which application program uses the data

Meta Data is used by the developers to develop the program, queries to manage and manipulate the data

Dictionary:-

1) Active Data Dictionary:- It is manage automatically by the data management software. It is always consistent with the current structure of the Database.

2) Passive Data Dictionary:- It is the one use for documentation purposes. It is managed by the user of the system and is modified manually by the users.

Database Languages :-

(1) Data Defination Language (DDL) :- It is a language that allows the user to define the data and there relationship to

other type of data. Command are
1) Create
2) Alter
3) Rename
4) Drop

2) Data Manipulation language (DML) :- It is a language that provides a set of operation to support the basic data manipulation, operation on the data held in the database. Command used are :-
1) Insert
2) Delete
3) Select
4) Update

3) Data Control Languages (DCL) :- Command are :-
1) GRANT
2) REVOKE

Data Models :- There are three type of data Models.
1) Object based
2) Record based logical models
3) Physical Based

It can be define as an integrated collection of concepts for describing and manipulating the data, relationships between the data and Constraint on the data in an organisation.
It comprises of three component
1) Structural part :- There are rule which help in designing model

2) Manipulative part :- which type of operation is apply on model

3) Integrity Rules :-

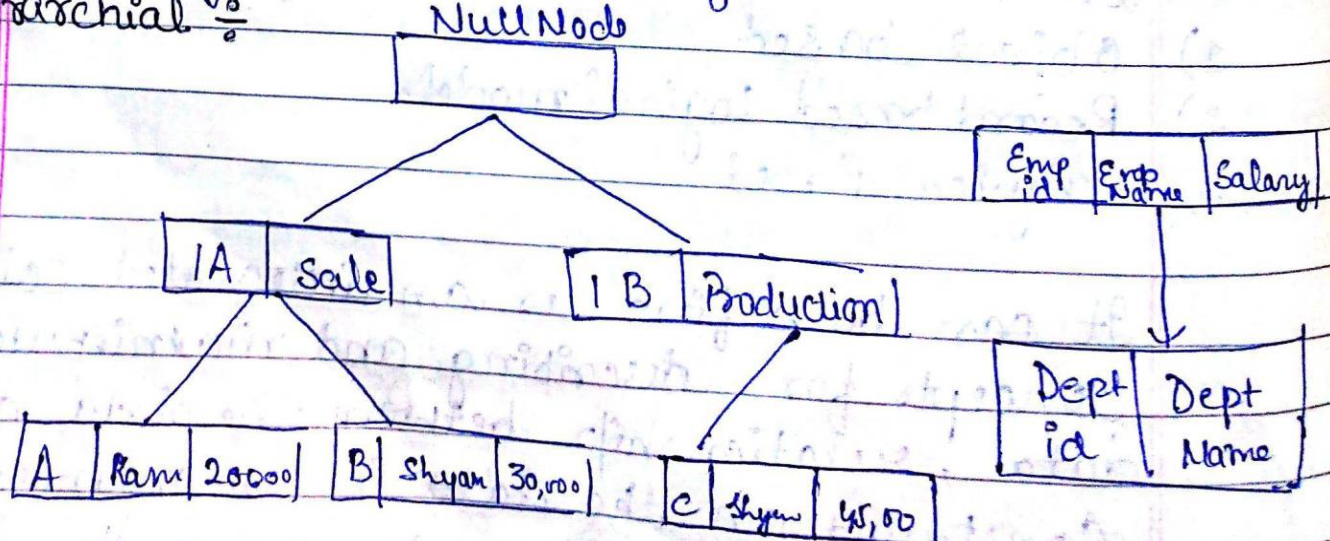
Data model are divided into three category

1) object based :- It uses a concept such as entities, attribute and there relationship. This model can be used to describe the data at the Conceptual and External level Eg E-R models

2) Physical based :- These model describe how the data is stored in the Computer. This model is used to describe the data at the internal level.

3) Record based logical models :- These models are used in describing the data at the logical and the view level. These models are used to Specify the over all logical structure of the database

Eg of Hierarchial :-



1) Hierarchial model:-

It is based on tree structure. It consist of collection of records that are connected to each other by links. The tree structure use in a Hierarchial model is known as routed tree. The root node of that tree is an empty node. So Hierarchial model is a collection of routed trees and the relationship exists in the Hierarchial model is one to many and ~~many~~ one to one.

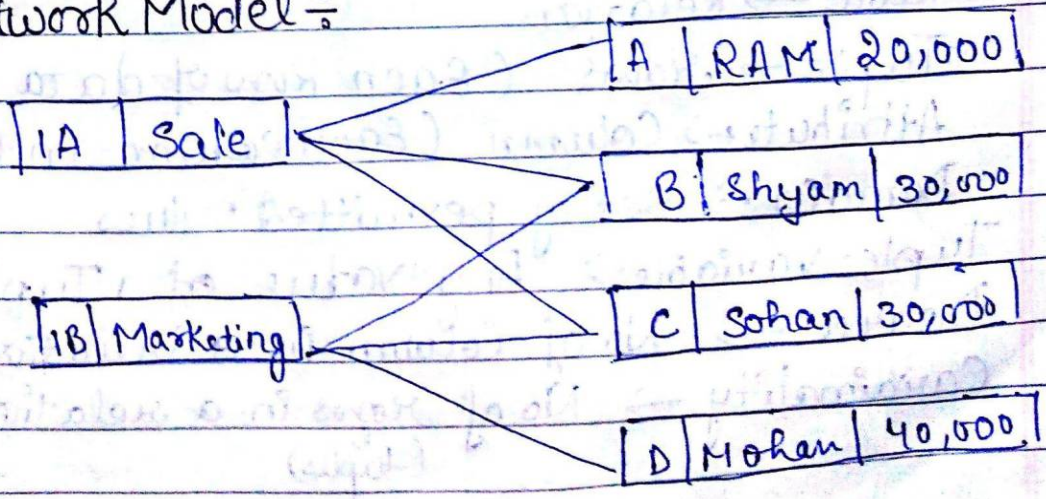
Advantages:-

1) It is easy to understand.
2) More efficient than ER model

Disadvantage:-

1) Data inconsistency occur when the parent node is delete that result in the deletion of the child node.
2) Wastage of storage space
3) Complex to design
4) Absence of structural independency.

2) Network Model:-



for one to many →
one to one ↔
many to many ←

It is based on Graph Structure. It consist of collection of records, which are connected to each other by links.

Advantage:-
1) It is easy to design than the Hierarichial model
2) Data access is easy in the network model

Disadvantages:-
1) It is complex to design than a relational model
2) Efficiency are less than the relational model
3) Absence of structural independency

3) Relational Model:- Relational model stores data in form of tables.

Employee

Emp id	Name	Salary
1	A	20,000
2	B	30,000
3	C	40,000
4	D	50,000

→ Domain

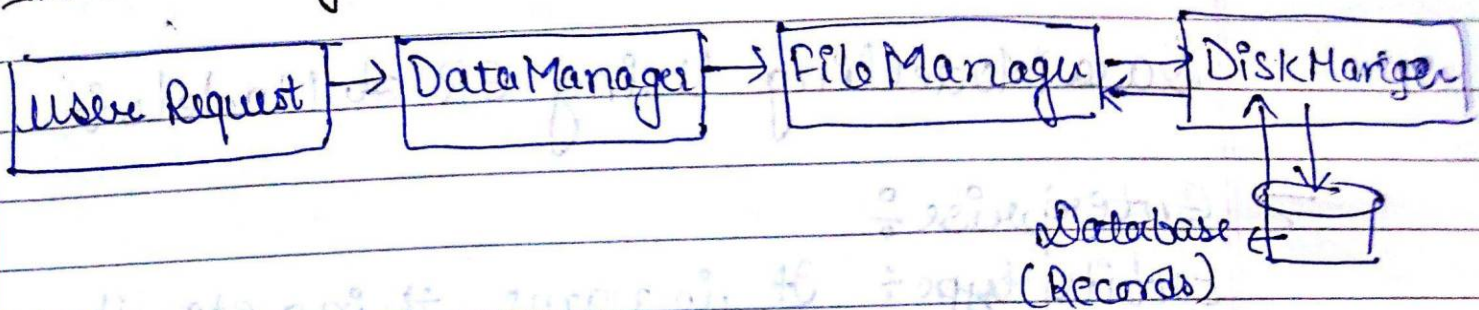
Tuple Variable

Department

Dep id	Name
A	Sales
B	Marketing

Table → Relation.
Tuple → rows (Each Row of data)
Attributes → Column. (Each Column in the tuple
Domain → Set of permitted values
Tuples variable → Any value of a Tuple
Degree → No of Column in a relation → 3 in emp.
Cardinality → No of rows in a relation. (4 tuples)

Data Manager :-



UNIT-2

- In Relational model many to many relationship can be easily implemented.
- It is useful for representing most of the real world object and relationship among them.
- Relational model doesnot maintain physical connection among records. Data is organized logically in the forms of rows and columns and stored in table.

UNIT - 2

Data Modelling Using ER & Models :-

→ Enterprise :-
Entity type :- It is name, thing etc. These are the data object about which information is to be collected
Attributes :- characteristic of entity. OR Data elements & Data fields.

Type of Attributes :-

1) Single Value Attributes :- Those attribute which contain a single value. for eg :- Age., Salary etc

2) Multivalued Attribute :- That contain more than one value of a single entity for eg Phone no.

3) Composite Attributes :- Those attribute which can be further divided. for eg name → First name Last name, Date of Birth etc.

4) Simple or Atomic attributes :- Those attributes which can not be further divided for eg Age.

5) Stored Attributes :- Attributes which can not be derived from another attribute Date of Birth

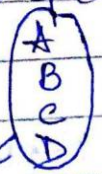
6) Null value :- Empty.

Entity Types :- Collection of Entity that Share the same attribute.

eg :- Employee → Name, Age, salary.
↓
Entitytype

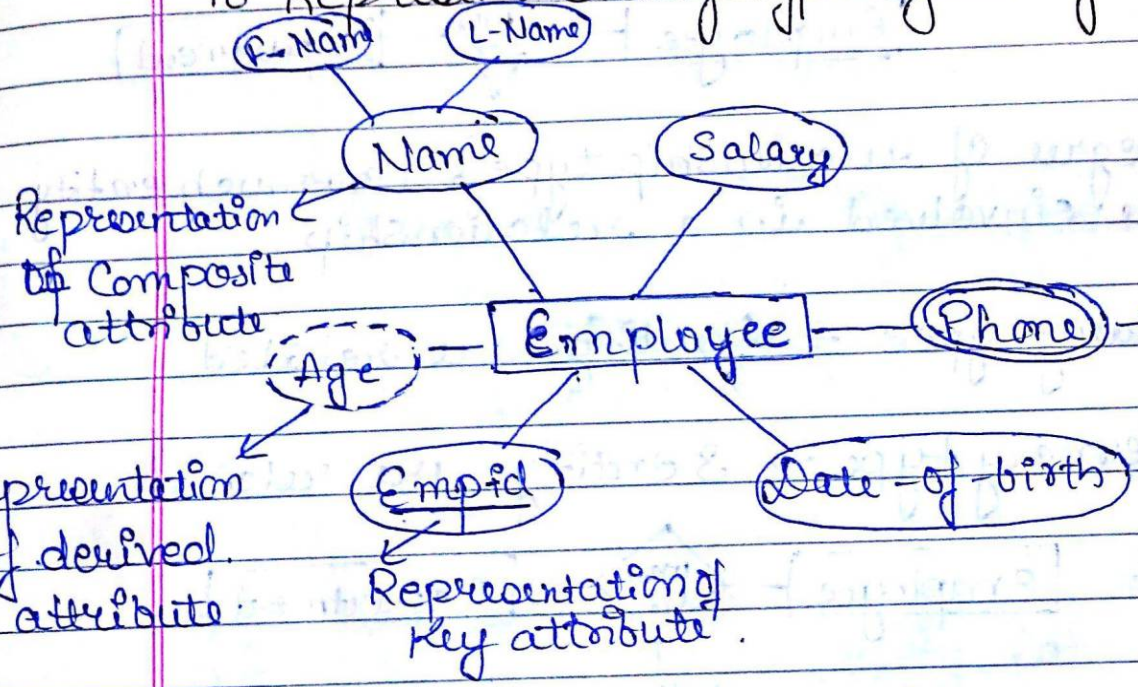
Passion
Page No.
Date

Entity Set :- Employee — Name, Age, Salary



Entity Set

To Represent Entity Type by using ER diagram.

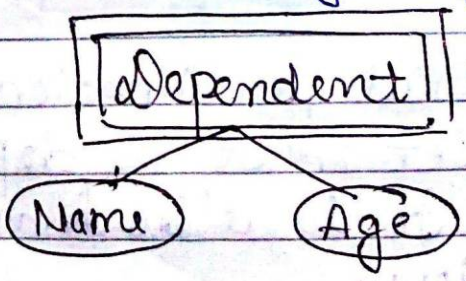


Key attribute :- Key attribute are those attribute which is uniquely identify the record.

Weak Entity type :- Those entity type in which no key attributes is present.

Strong Entity type :- Those Strong Entity type which consist of key attributes.

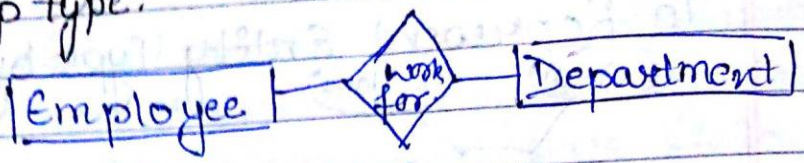
Eg of weak Entity type



Domain of the attributes: Set of possible values.

To join two entities → Relationship: Representation of Relationship

Relationship type:

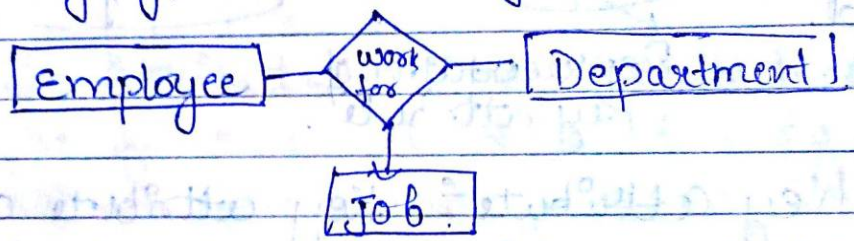


Degree of relationship type: How much entity are involved in a relationship.

1) Binary type → 2 entity are related

2) Ternary type → 3 entity are related

Eg:



Relationship Constraints:
- Participation Constraints
- Cardinality ratio

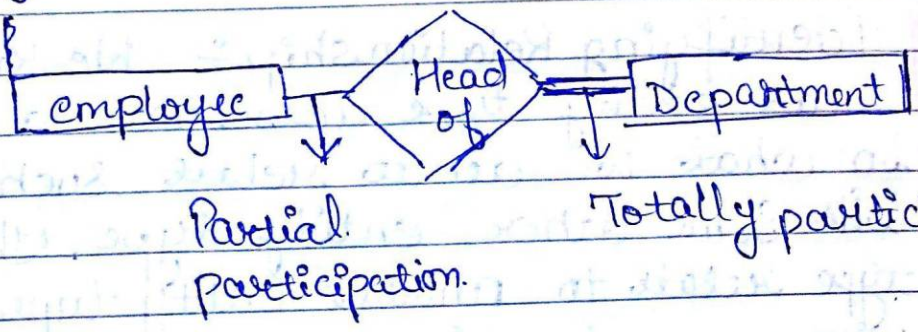
① Participation Constraints

Total participation Partial Participation

① Total Participation: In total participation every entity in the empty set must be depend on another entity. It is also known as existence dependency.

In E-R diagram it is represented as a double line connecting the participating Entity type two the relationship.

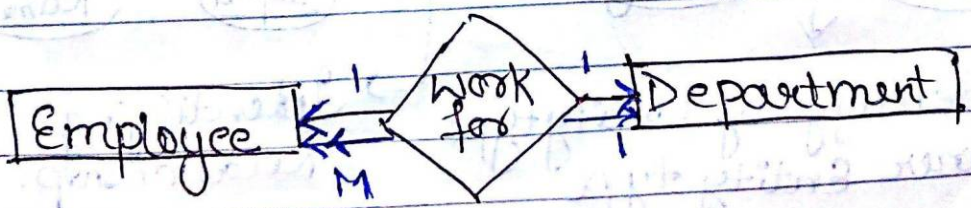
(2) Partial participation: In partial participation some entities in the entity set are depend upon another entity.

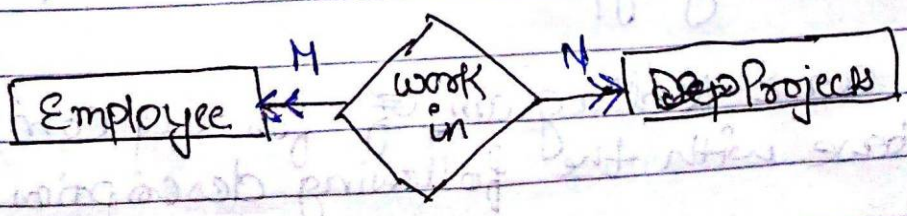


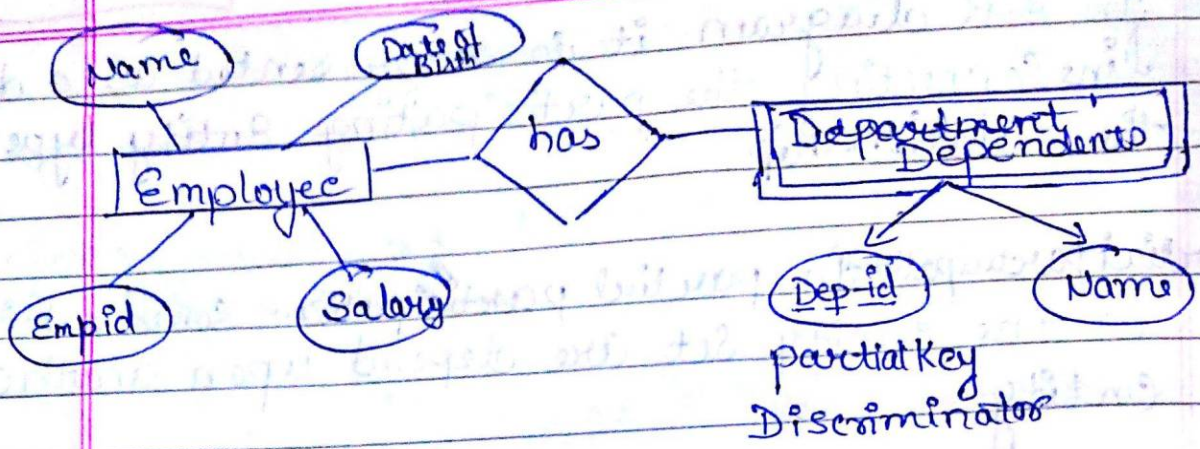
(2) Cardinality ration: Cardinality ratio for binary relationship specifies the no. of relation ship instance that an entity can participate in a relation set.

Relationship exists are:
one to one → (1:1)
one to many → (1:N)
Many to one → (N:1)
Many to Many → (M:N)

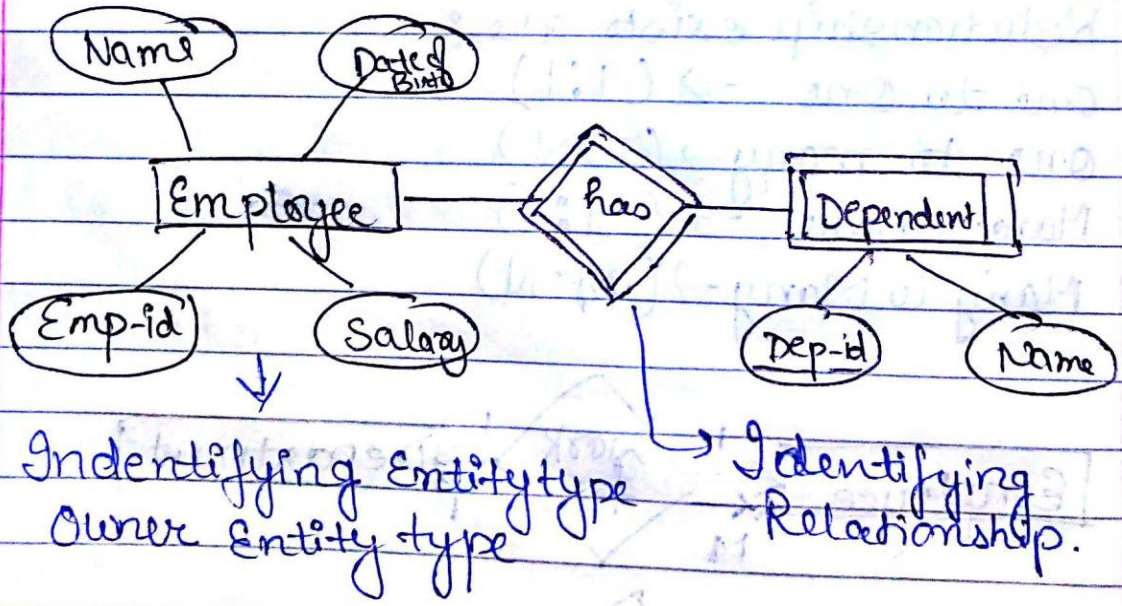
eg:-







→ Identifying Relationship :- We know that a weak entity type does have a key attribute, so what we do to relate such entity type with some other entity type. The weak entity type relate to another Entity type in combination with some of their attributes value. We call this other entity type the identifying or the owner entity type and we call the relationship type that relates to its owner the identifying relationship.



Q Make an E-R diagram of for the company database with the following description

(1) The company is organised into departments. Each department has a unique name and unique no. A department may have several locations

(2) A department controls a no of projects, each of which has a unique name, unique no and a single location.

(3) We store each employee name, social security no, address and salary an employee is assign to one deparment but may work on several projects, which are not necessarily control by the same department. We want to keep track of the dependen -ce of each Employee for insurance purposes.

Construct ER diagram for Teacher Student database