Database Management System

Chapter 1

Computer science class 12

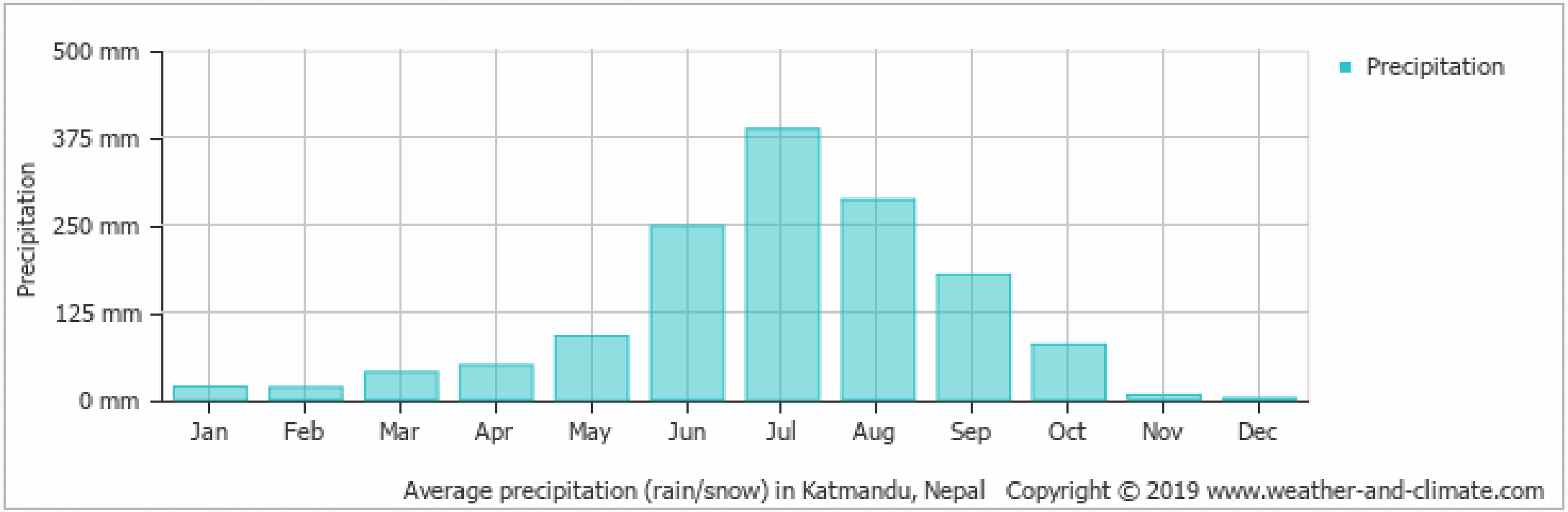
Data

Data is a collection of raw facts, such as numbers, **words**, measurements, observations or just descriptions of things. When activities in the organization take place , the effect of these activities need to be recorded which is also known as data

ID	NAME	CLASS	MARK	SEX
1	John Deo	Four	75	female
2	Max Ruin	Three	85	male
3	Arnold	Three	55	male
4	Krish Star	Four	60	female
5	John Mike	Four	60	female
6	Alex John	Four	55	male

Information

- Data after processing is known as information. The objective of data processing is to generate the information required for carrying out the business activities.



Here we have the annual precipitation record of Kathmandu. It is data represented by a bar graph. After processing or analyzing it **we get the information that highest precipitation was in the month of July and the lowest was in December**.

Database

Order	
int	OrderID
Int	ClientID
datetime	Created
datetime	Modified

ProductProcess	
int	ProductProcessID
Int	OrderID
datetime	TimeStamp
datetime	State
int	ConfigID

Product	
int	Id
Int	Productid
datetime	Productname
datetime	ProductDesc
int	State

Config	
int	ConfigID
Int	PiConfig
Datetime	PsConfig

- A database is an organized collection of data, generally stored and accessed electronically from a computer system. Where databases are more complex they are often developed using formal design and modeling techniques. In other words we can also say database as the collection of related data. Simple example of database is telephone diary.

Function of database

i. Storing manipulating and managing information
ii. Reducing repeated storage of data
iii. Keeping latest and correct information
iv. Providing data security
v. Improvising disk optimization

Advantages of database

i. Sharing data
ii. Reduced data redundancy
iii. Data backup and recovery
iv. Inconsistency avoided
v. Data integrity
vi. Data security
vii. Data independence
viii. Multiple user interfaces
ix. Process complex query

Database Management system (DBMS)

A **Database Management System** (**DBMS**) is software designed to store, retrieve, **define**, and manage data in a database. It is general purpose software system that facilities the process of defining , constructing and manipulating data in the database for various application. DBMS handles user request for database action and allows for control of security and data integrity requirement. MS SQL server, MySQL, MS Access are some popular DBMS software.

Objective of DBMS

i. Providing very easy access to data for user.
ii. Providing very fast response to user requested for any data
iii. Removing any duplicate data
iv. To provide huge space for storage of relevant data
v. To facilitate more than 1 user to use same data at a same time
vi. To protect data against unwanted and unauthorized access.

vii. To maintain data integrity

viii. To give user the latest modification of database.

Function of DBMS

i. DBMS have to supply data for various application programs
ii. Creating and manipulating data dictionary
iii. Making proper arrangement and facility for sharing database
iv. DBMS should allow the user to store, retrieve and update information as easy as possible
v. Securing data
vi. Recover data in case of failure.

Different between Database and DBMS

Database	DBMS
1. it is a collection of related information about a subject organized in a useful manner that provides a base or foundation for procedure, such as retrieving information, drawing conclusion and make decision.	1. A DBMS is a set of programs that manages the database files. It allows accessing the files, updating the records and retrieving data as requested.
2. The technique for designing a database using top-down methods is to write a main database parts that names all the major storage and retrievals it will need. Later, the programming team looks at the requirements of each of those database components and the process is repeated.	2. the top-down method starts from the general and moves to the specific. Basically, you start with a general idea of what is needed for the system and then ask the end-users what data they need to store. The analyst will then work with the users to determine what data should be kept in the database. Using the top-down method requires that the analyst has a detailed understanding of the system. The top-down method also can have shortcomings. In some cases, top-down design can lead to unsatisfactory results because the analyst and end-users can miss something that is important and is necessary for the system.

What is database object

- A database object in a database is a data structure used to either store or reference data. The most common object that most people interact with is the table. Other objects are indexes, stored procedures, sequences, views and many more.
- in simple words we can refer database objects to anything we create in any database engine.

Fields and records in database table

- Fields and records are two basic components of a database, which is an organized collection of information, or data. The term "fields" refers to columns, or vertical categories of data; the term "records" refers to rows, or horizontal groupings of unique field data. Examples of both database components are seen in older, printed databases and also in modern, computer-based databases.

field

First Name	Last Name	Address	City	Age
Mickey	Mouse	123 Fantasy Way	Anaheim	73
Bat	Man	321 Cavern Ave	Gotham	54
Wonder	Woman	987 Truth Way	Paradise	39
Donald	Duck	555 Quack Street	Mallard	65
Bugs	Bunny	567 Carrot Street	Rascal	58
Wiley	Coyote	999 Acme Way	Canyon	61
Cat	Woman	234 Purrfect Street	Hairball	32
Tweety	Bird	543	Itotltaw	28

Record

Keys in DBMS

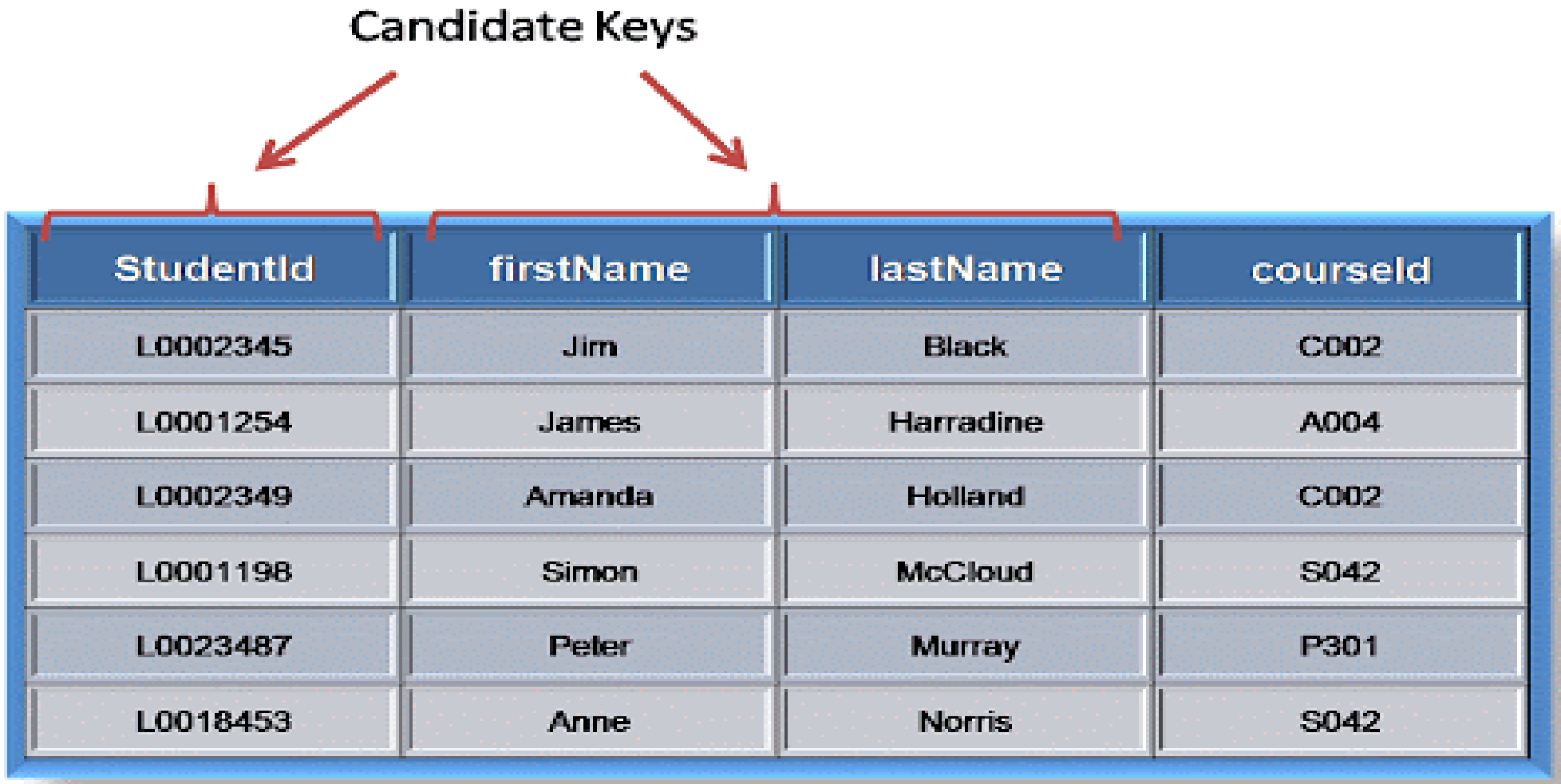
A key in DBMS is an attribute or a set of attributes that help to uniquely identify a tuple (or row) in a relation (or table). Keys are also used to establish relationships between the different tables and columns of a relational database. Individual values in a key are called key values. Following are the types of keys.

i. Candidate key
ii. Primary key
iii. Foreign key
iv. Alternate key

Candidate key

A candidate key is a specific type of field in a relational database that can identify each unique record independently of any other data.

Experts describe a candidate key of having "no redundant attributes" and being a "minimal representation of a tuple" in a relational database table.


Candidate Keys
StudentId
firstName
lastName
courseId
L0002345
Jim
Black
C002
L0001254
James
Harradine
A004
L0002349
Amanda
Holland
C002
L0001198
Simon
McCloud
S042
L0023487
Peter
Murray
P301
L0018453
Anne
Norris
S042

Primary key

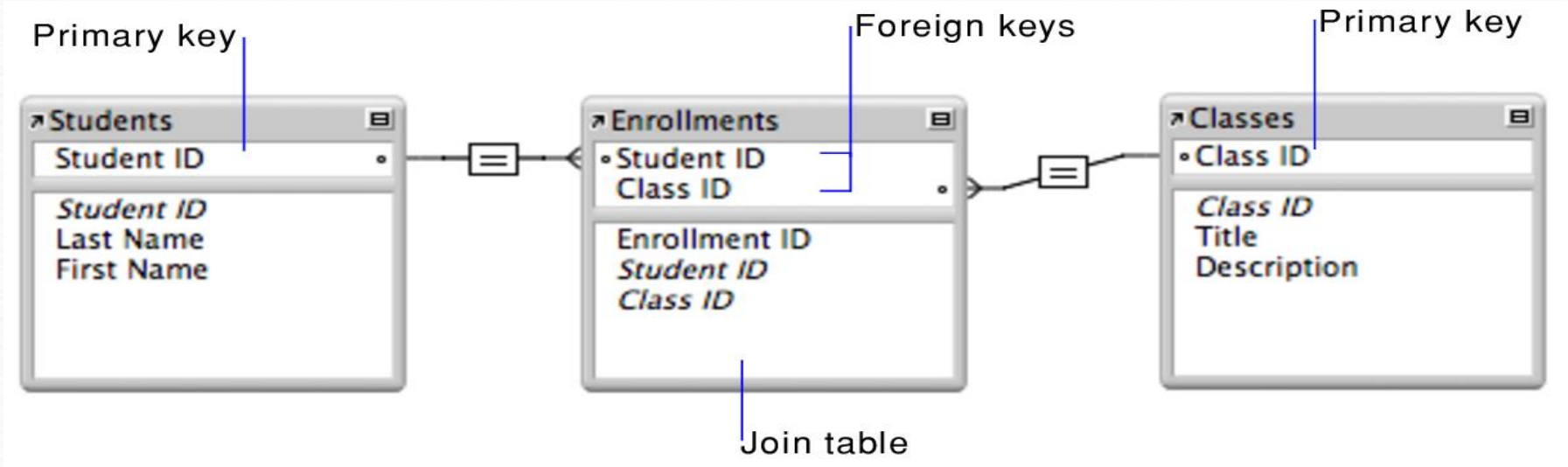
In the relational model of databases, a primary key is a specific choice of a minimal set of attributes that uniquely specify a tuple in a relation. Informally, a primary key is "which attributes identify a record", and in simple cases are simply a single attribute: a unique id.

Primary Keys

StudentId	firstName	lastName	courseId
L0002345	Jim	Black	C002
L0001254	James	Harradine	A004
L0002349	Amanda	Holland	C002
L0001198	Simon	McCloud	S042
L0023487	Peter	Murray	P301
L0018453	Anne	Norris	S042

Foreign key

Foreign keys are the columns of a table that points to the **primary key** of another table. They act as a cross-reference between tables.



ALTERNATE KEYS

- Alternate key is a column or group of columns in a table that uniquely identify every row in that table. A table can have multiple choices for a primary key but only one can be set as the primary key. All the keys which are not primary key are called an Alternate Key.

StudID	Roll No	First Name	LastName	Email
1	11	Tom	Price	abc@gmail.com
2	12	Nick	Wright	xyz@gmail.com
3	13	Dana	Natan	mno@yahoo.com

In this table, StudID, Roll No, Email are qualified to become a primary key. But since StudID is the primary key, Roll No, Email becomes the alternative key.

What is SQL

- SQL or **Structured Query Language** is the most common language for inserting, extracting and organizing data that is stored in a relational database. A database is a table that consists of rows and columns. SQL is the language of databases. It facilitates storing and retrieving specific information from databases that are further used for analysis.
- There are two types of SQL they are data definition language (DDL) and data manipulation language (DML)

Data definition language (DDL)

- Data Definition Language (DDL) is a language used to define data structures and modify data. For example, DDL commands can be used to add, remove, or modify tables within in a database. DDLs used in database applications are considered a subset of SQL, the Structured Query Language. However, a DDL may also define other types of data, such as XML.
- A Data Definition Language has a pre-defined syntax for describing data. For example, to build a new table using SQL syntax, the CREATE command is used, followed by parameters for the table name and column definitions. The DDL can also define the name of each column and the associated data type. Once a table is created, it can be modified using the ALTER command. If the table is no longer needed, the DROP command can be used to delete the table

Data manipulation language(DML)

- DML is Data Manipulation Language which is used to manipulate existing data. For example: INSERT statement is used to insert new record in the table, UPDATE statement is used for updating record , DELETE statement is used for deleting a record present in table of a database.

Difference between DDL and DML:

DDL	DML
It stands for Data Definition Language.	It stands for Data Manipulation Language.
It is used to create database schema and can be used to define some constraints as well.	It is used to add, retrieve or update the data.
It basically defines the column (Attributes) of the table.	It add or update the row of the table. These rows are called as tuple.
It doesn't have any further classification.	It is further classified into Procedural and Non-Procedural DML.
Basic command present in DDL are CREATE, DROP, RENAME, ALTER etc.	BASIC command present in DML are UPDATE, INSERT, MERGE etc.
DDL does not use WHERE clause in its statement.	While DML uses WHERE clause in its statement.

DDL examples

Here are some commands that come under DDL:

- CREATE
- ALTER
- DROP

a. CREATE It is used to create a new table in the database.

Example:
CREATE TABLE EMPLOYEE(Name VARCHAR2(20), Email VARCHAR2(100), D
OB DATE);

b. DROP: It is used to delete both the structure and record stored in the table.

Example:

DROP TABLE EMPLOYEE;

c. ALTER: It is used to alter the structure of the database. This change could be either to modify the characteristics of an existing attribute or probably to add a new attribute.

Example:

ALTER TABLE table_name ADD column_name COLUMN-definition;

DML examples

Here are some commands that come under DML:

- INSERT
- UPDATE
- DELETE

a. **INSERT:** The INSERT statement is a SQL query. It is used to insert data into the row of a table.

example:

```
INSERT INTO javatpoint (Author, Subject) VALUES ("Sonoo", "DBMS");
```

---

b. UPDATE: This command is used to update or modify the value of a column in the table.

example:

```
UPDATE students
SET User_Name = 'Sonoo'
WHERE Student_Id = '3‘;
```

c. DELETE: It is used to remove one or more row from a table.

example:

```
DELETE FROM javatpoint
WHERE Author="Sonoo";
```

Database models

- A database model is a type of data model that determines the logical structure of a database and fundamentally determines in which manner data can be stored, organized and manipulated. The most popular example of a database model is the relational model, which uses a table-based format

ER model

Entity-Relationship database model: this model is based on perception of a real world that contains a collection of basic objects, called entities and of relationship among these objects and characteristics of an entity. It shows relationship between different entities.

The Element of ER model are as follows

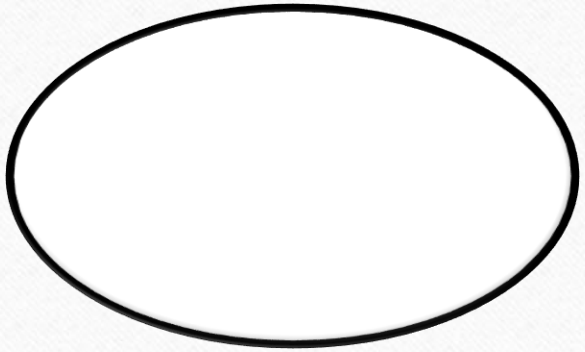
1. Entity 2. Attributes 3. Relationship

1. Entity

An entity is a thing of interest to an organization about which data is to be held. Example is customer, employee, suppliers. It is represented by rectangle.

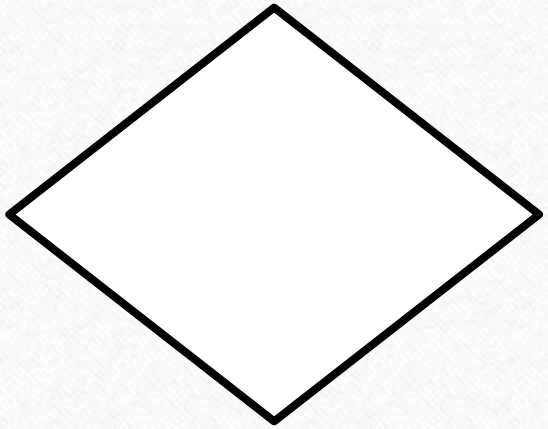
2. Attribute

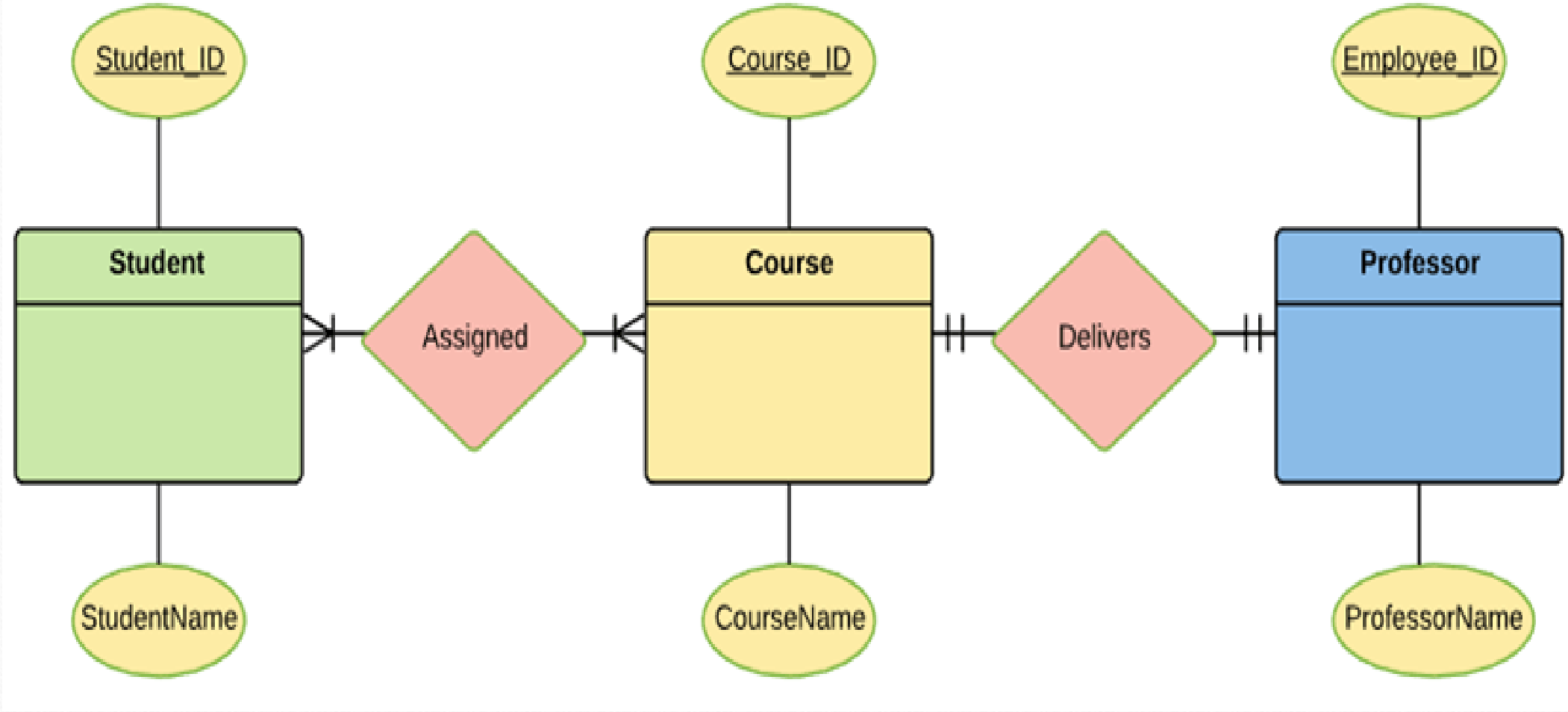
- Attribute is a property or characteristic of an entity. Example of attribute associated with a customer include customer id, customer name, title, address etc. it is denoted by an oval shape.



3. Relationship

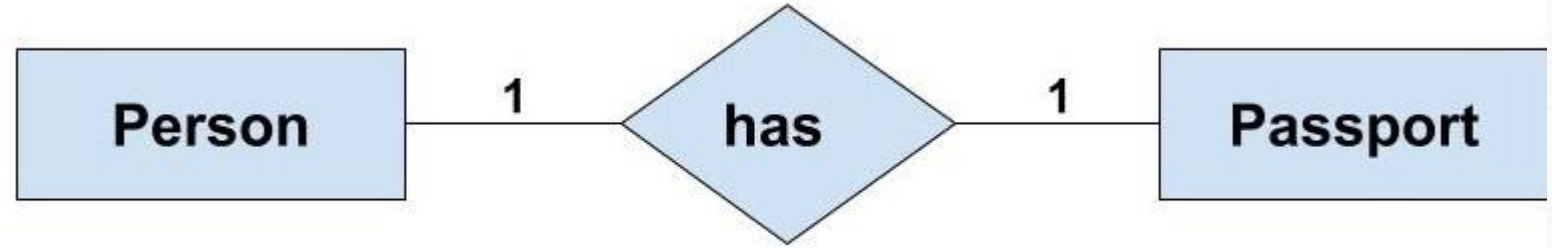
A relationship is a link of association between entities. Example is link between customer and supplier. It is represented by diamond shape




Student_ID
Course_ID
Employee_ID
Student
Assigned
Course
Delivers
Professor
StudentName
CourseName
ProfessorName

There are following types of relationship

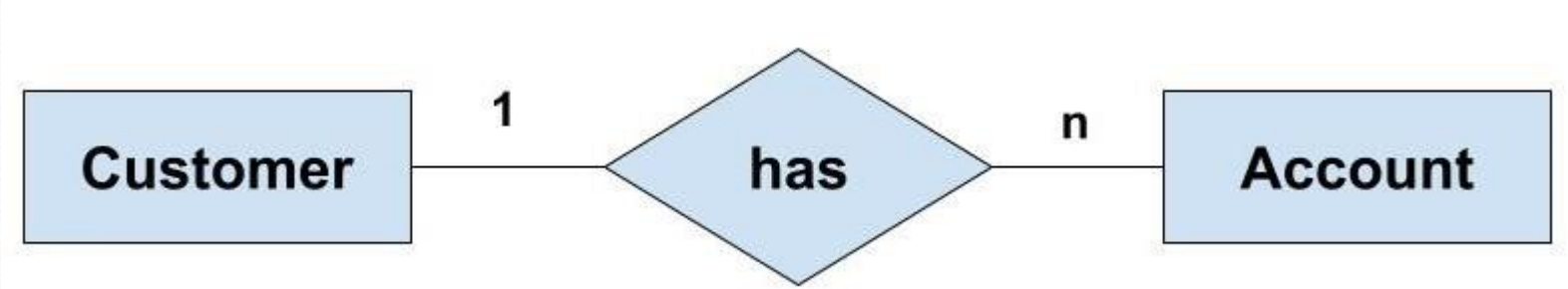
One to one relationship: Such a relationship exists when an instance of an entity is associated with one instance of another entity.



For example, If there are two entities 'Person' (Id, Name, Age, Address)and 'Passport'(Passport_id, Passport_no). So, each person can have only one passport and each passport belongs to only one person.

One-to-Many or Many-to-One Relationship

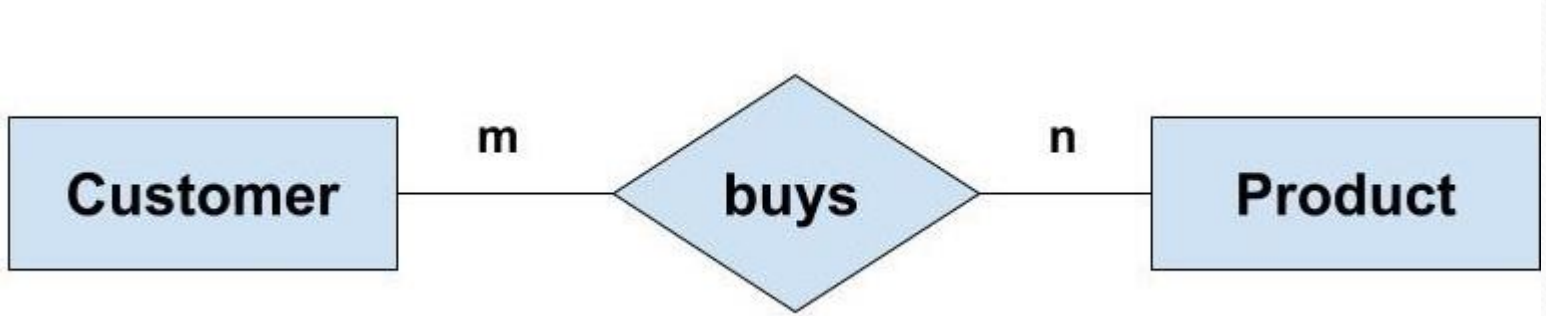
Such a relationship exists when each record of one table can be related to one or more than one record of the other table. This relationship is the most common relationship found. A one-to-many relationship can also be said as a many-to-one relationship depending upon the way we view it.



example, If there are two entity type 'Customer' and 'Account' then each 'Customer' can have more than one 'Account' but each 'Account' is held by only one 'Customer'. In this example, we can say that each Customer is associated with many Account. So, it is a one-to-many relationship. But, if we see it the other way i.e. many Account is associated with one Customer then we can say that it is a many-to-one relationship.

Many-to-Many Relationship

Such a relationship exists when each record of the first table can be related to one or more than one record of the second table and a single record of the second table can be related to one or more than one record of the first table. A many-to-many relationship can be seen as a two one-to-many relationship which is linked by a 'linking table' or 'associate table'. The linking table links two tables by having fields which are the primary key of the other two tables.



Example: If there are two entity type 'Customer' and 'Product' then each customer can buy more than one product and a product can be bought by many different customers.

Advantages of ER model

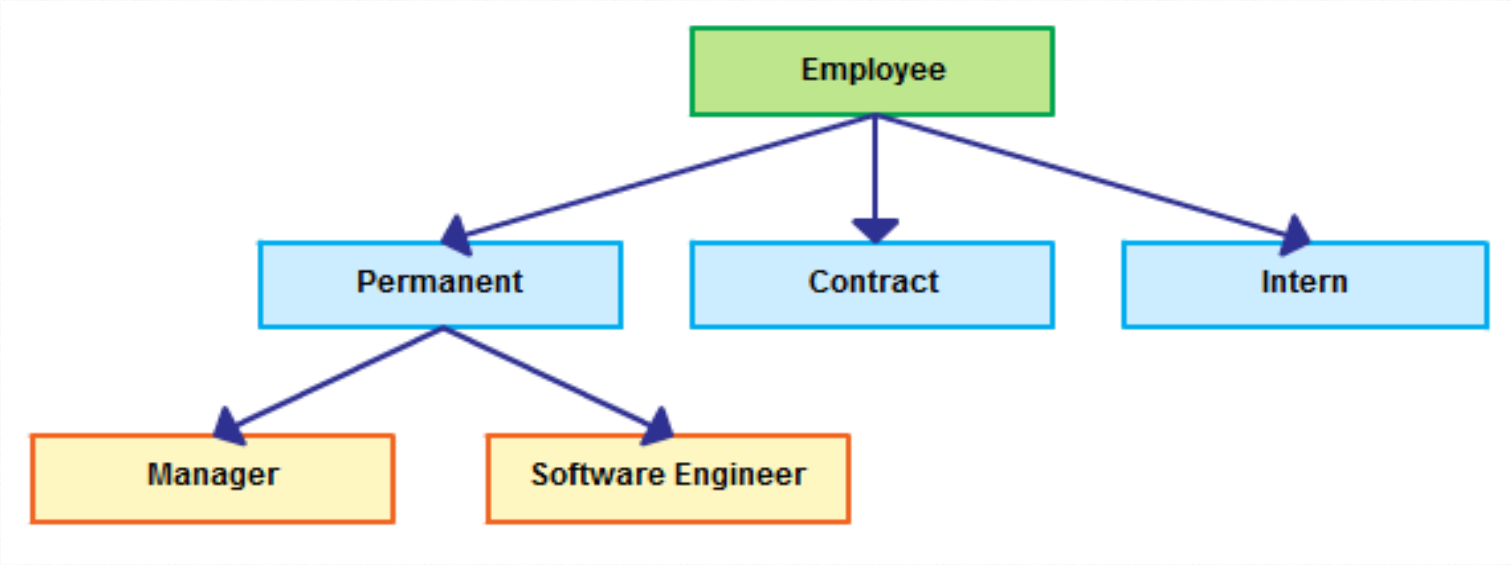
i. ER model is very well integrated with relational database model
ii. Database's main entity and their relationship are easily viewed and understood through ER model

Disadvantages of ER model

i. No data manipulation language or commands are available in ER model
ii. This model becomes crowded due to huge presence of entity

2. Hierarchical database model

It is the database model in which records are grouped in such a way that there relationship form a branching, tree like structure. In very simple language this type of database model organizes data in top down structure that resembles a tree. Top node of the hierarchy is called root and end node is called leaves. The top node is also called parent and below node is also called child.


Employee
Permanent
Contract
Intern
Manager
Software Engineer

Advantages of Hierarchical database model

i. It is the most simplest and easiest database model
ii. If parent is known then searching is easy and fast
iii. Database security is said to be good because we can't modify, delete a child without consulting its parent
iv. It is efficient to handle 1 to many relationship

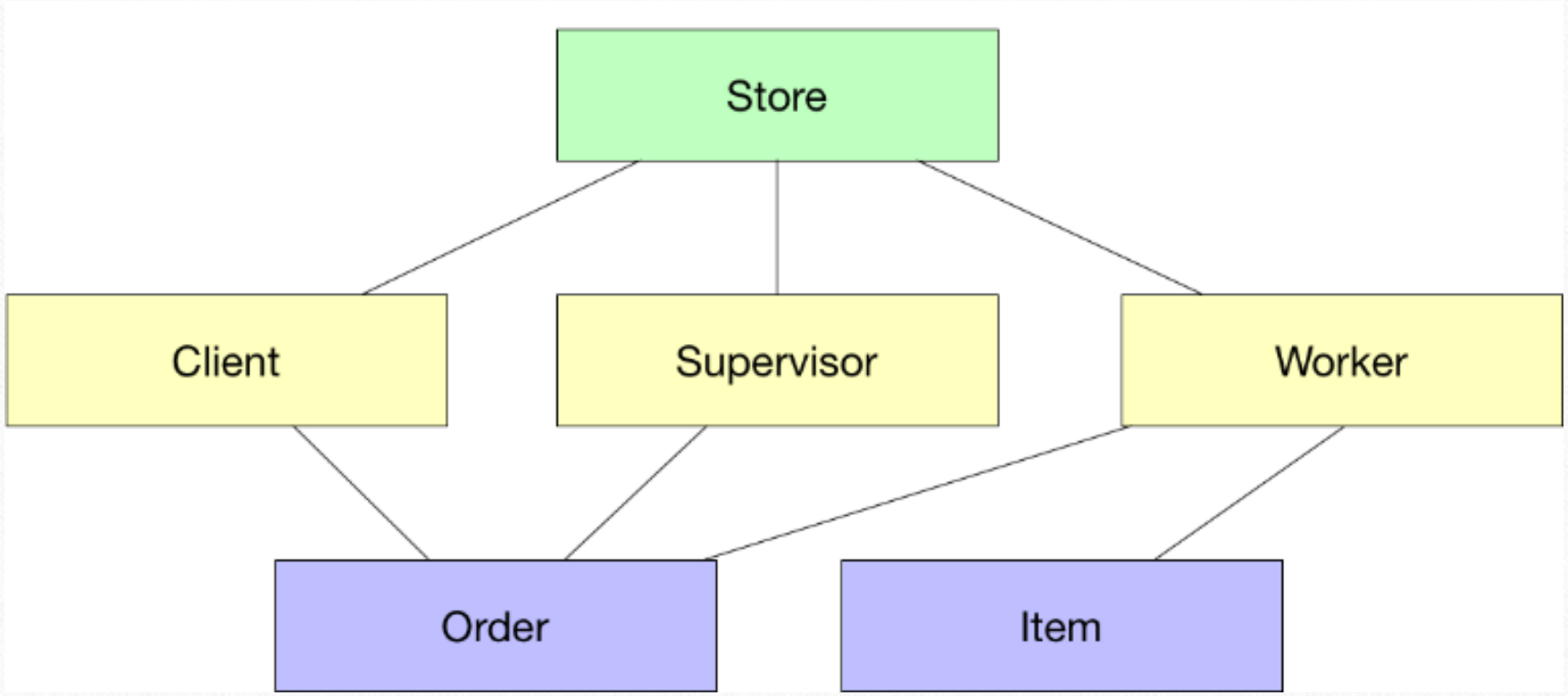
Disadvantages to hierarchical database model

i. This model nowadays has become an outdated model
ii. This model cannot handle many to many relationship
iii. When parent node is deleted, child are automatically deleated

Network database model

Network database model is replaced hierarchical database model due to some limitations on the model. Suppose, if an employee relates to two departments, then the hierarchical database model cannot able to arrange records in proper place. So network, database model was emerged to arranged non-hierarchical database. The structure of database is more like graph rather than tree structure.

Here nodes are associated with each other forming network database model



Advantages of Network database model

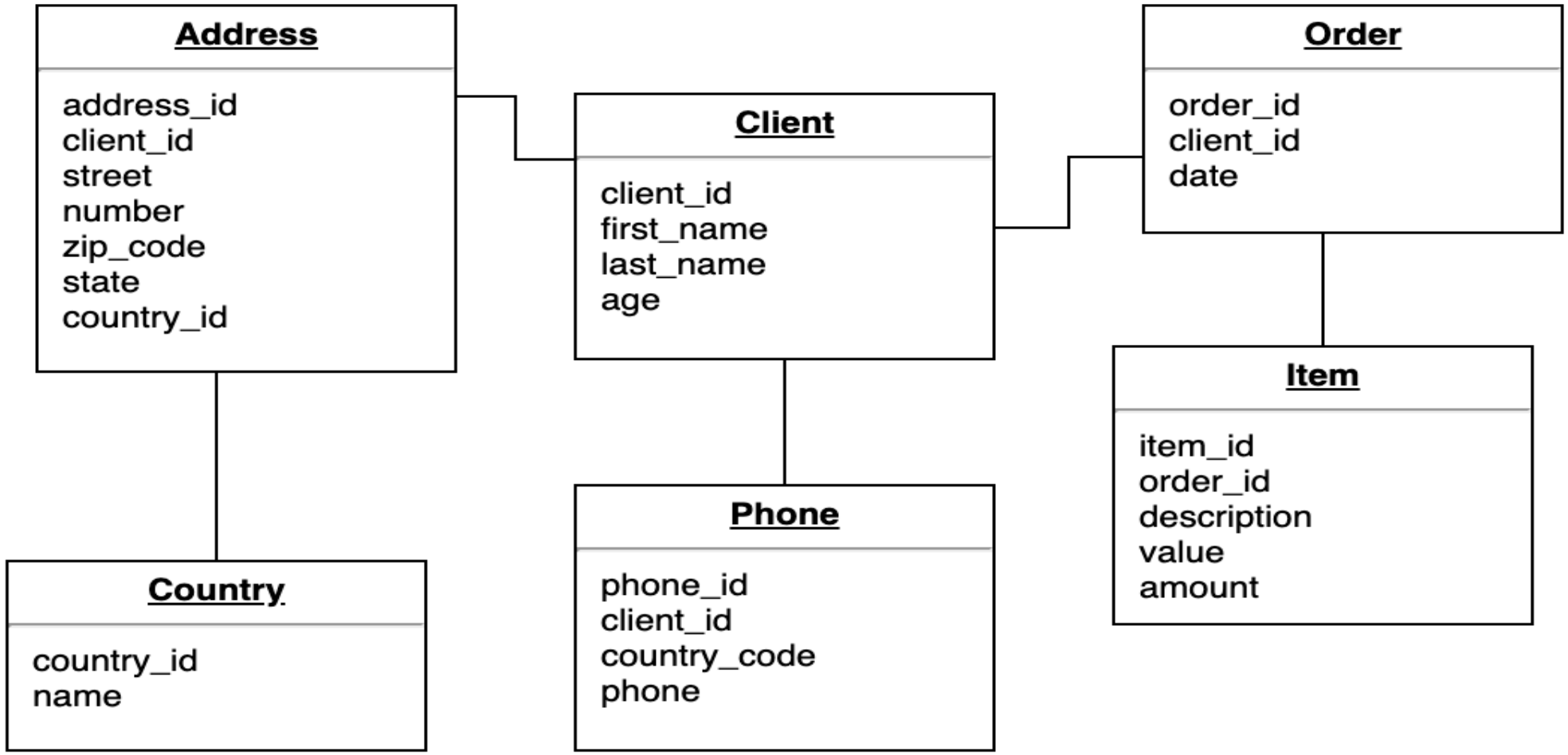
i. It is able to handle many to many relationship
ii. Searching is generally fast due to presence of multidirectional pointer
iii. Data redundancy is reduced because same data is not needed again and again

Disadvantages of network database model

i. It is too complex, because database administrator, programmer must be familiar with inter structure in order to access the database
ii. There is need of long programs to handle relationship

Relational database model

In this model, the data is organized into tables which contain multiple rows and columns. These tables are related with each other. A row in a table represents a relationship among a set of values. Since a table is a collection of such relationships, it is generally referred to the mathematical term relation, from which the relational database model derives its name. It uses 2 dimensional table which are made up of rows and columns. Each row represents record and each column represents a field. Record is also referred as tuples and field is referred as attributes.


Address
address_id
client_id
street
number
zip_code
state
country_id
Client
client_id
first_name
last_name
age
Order
order_id
client_id
date
Item
item_id
order_id
description
value
amount
Country
country_id
name
Phone
phone_id
client_id
country_code
phone

Advantages of Relational database model

i. There are few redundancy in this database model
ii. The normalization of database is possible
iii. The database processing is also very fast

Disadvantages of relational database model

- It is the most complex database model
- Here are too many rules which makes this database model non user friendly

Object oriented database model

- Object oriented data model is based upon real world situations. These situations are represented as objects, with different attributes. All these object have multiple relationships between them.
- In this model both data and relationship are contained in a single structure known as object.

Elements of Object oriented data model

Objects

The real world entities and situations are represented as objects in the Object oriented database model.

Attributes and Method

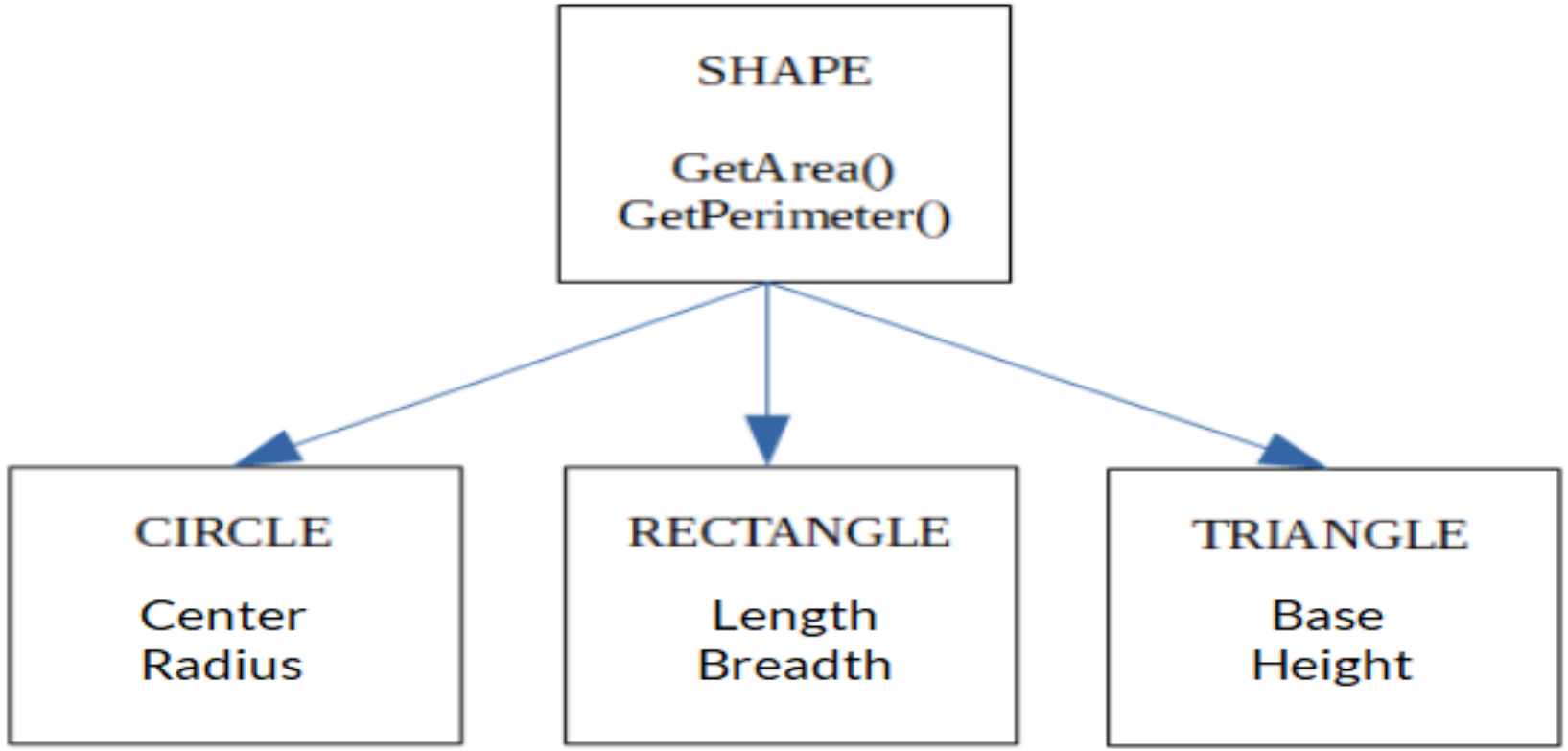
Every object has certain characteristics. These are represented using Attributes. The behavior of the objects is represented using Methods.

Class

Similar attributes and methods are grouped together using a class. An object can be called as an instance of the class.

Inheritance

A new class can be derived from the original class. The derived class contains attributes and methods of the original class as well as its own.



Shape, Circle, Rectangle and Triangle are all objects in this model. Circle has the attributes Center and Radius. Rectangle has the attributes Length and Breath, Triangle has the attributes Base and Height. The objects Circle, Rectangle and Triangle inherit from the object Shape.

Advantages of object oriented database model

Sematic(correct and acceptable) content is added

Visual representation contains sematic content

Inheritance promotes data integrity

Works well with object-oriented programming languages.

disadvantages of object oriented database model

- Object databases are not widely adopted.
- In some situations, the high complexity can cause performance problems.

Normalization

- **Normalization** is the process of minimizing **redundancy** from a relation or set of relations. Redundancy in relation may cause insertion, deletion and updating anomalies (abnormal way). So, it helps to minimize the redundancy in relations. **Normal forms** are used to eliminate or reduce redundancy in database tables.

There are three type of normalization

- 1NF (First Normal Form)
- 2NF (Second Normal Form)
- 3NF (Third Normal Form)

Full Names	Physical Address	Movies Rented	Salutation
Janet Jones	First Street Plot No 4	Pirates of the Caribbean, Clash of the Titans	Ms.
Robert Phil	3rd Street 34	Forgetting Sarah Marshal, Daddy's Little Girls	Mr.
Robert Phil	5th Avenue	Clash of the Titans	Mr.

Here you see **Movies Rented column has multiple values.** Now let's move into 1st Normal Forms:

1NF (First Normal Form) Rules

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

Full Names	Physical Address	Movies rented	Salutation
Janet Jones	First Street Plot No 4	Pirates of the Caribbean	Ms.
Janet Jones	First Street Plot No 4	Clash of the Titans	Ms.
Robert Phil	3rd Street 34	Forgetting Sarah Marshal	Mr.
Robert Phil	3rd Street 34	Daddy's Little Girls	Mr.
Robert Phil	5th Avenue	Clash of the Titans	Mr.

2NF (Second Normal Form) Rules

- Rule 1- Be in 1NF
- Rule 2- Single Column Primary Key

Membership ID	Full names	Physical Address	Salutation
1	Janet Jones	First Street Plot No 4	Ms.
2	Robert Phil	3rd Street 34	Mr.
3	Robert Phil	5th Avenue	Mr.

Membership ID	Movies rented
1	Pirates of the Caribbean
1	Clash of the Titans
2	Forgetting Sarah Marshal
2	Daddy's Little Girls
3	Clash of the Titans

It is clear that we can't move forward to make our simple database in 2nd Normalization form unless we partition the table above. We have introduced a new column called Membership_id which is the primary key for table 1. Records can be uniquely identified in Table 1 using membership id

3NF (Third Normal Form) Rules

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

What are transitive functional dependencies?

- A transitive functional dependency is when changing a non-key column, might cause any of the other non-key columns to change
- Consider the table 1. Changing the non-key column Full Name may change Salutation.

MEMBERSHIP ID	FULL NAMES	PHYSICAL ADDRESS	SALUTATION
1	Janet Jones	First Street Plot No 4	Ms.
2	Robert Phil	3rd Street 34	Mr.
3	Robert Phil	5th Avenue	Mr.

Change in Name

May Change Salutation

3NF Example

Membership ID	Full names	Physical Address	Salutation ID
1	Janet Jones	First Street Plot No 4	2
2	Robert Phil	3rd Street 34	1
3	Robert Phil	5th Avenue	1

Membership ID	Movies rented
1	Pirates of the Caribbean
1	Clash of the Titans
2	Forgetting Sarah Marshal
2	Daddy's Little Girls
3	Clash of the Titans

Salutation ID	Salutation
1	Mr.
2	Ms.
3	Mrs.
4	Dr.

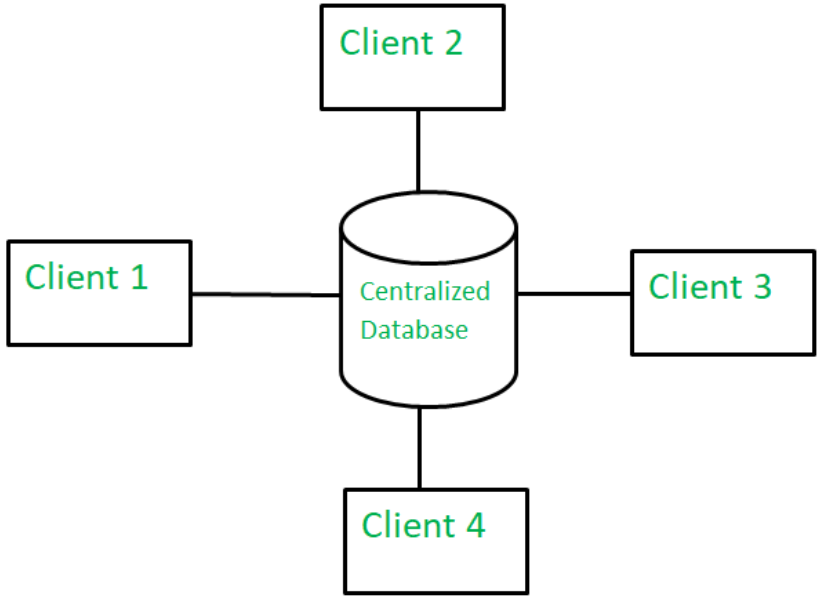
We have again divided our tables and created a new table which stores Salutations.

There are no transitive functional dependencies, and hence our table is in 3NF

In Table 3 Salutation ID is primary key, and in Table 1 Salutation ID is foreign to primary key in Table 3

1. Centralized Database

- A centralized database is basically a type of database that is stored, located as well as maintained at a single location only. This type of database is modified and managed from that location itself. This location is thus mainly any database system or a centralized computer system. The centralized location is accessed via an internet connection (LAN, WAN, etc). This centralized database is mainly used by institutions or organizations.



Advantages –

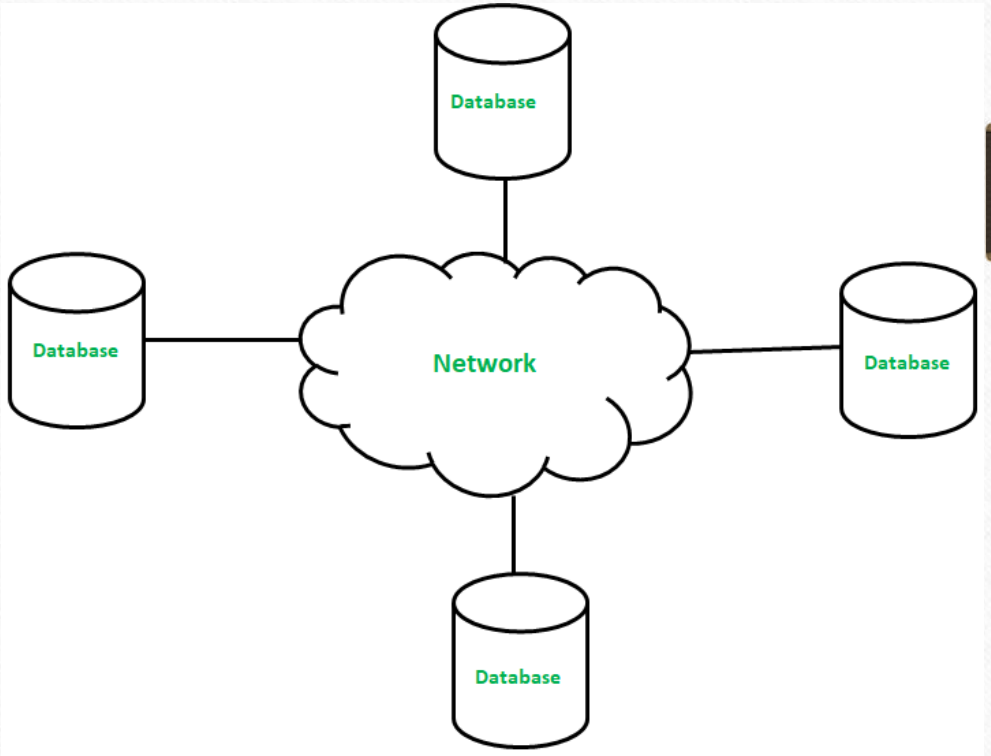
- Since all data is stored at a single location only thus it is easier to access and coordinate data.
- The centralized database has very minimal data redundancy since all data is stored in a single place.
- It is cheaper in comparison to all other databases available.

Disadvantages –

- The data traffic in the case of centralized database is more.
- If any kind of system failure occurs at the centralized system then the entire data will be destroyed.

2. Distributed Database :

- A distributed database is basically a type of database which consists of multiple databases that are connected with each other and are spread across different physical locations. The data that is stored on various physical locations can thus be managed independently of other physical locations. The communication between databases at different physical locations is thus done by a computer network.



Advantages

- This database can be easily expanded as data is already spread across different physical locations.
- The distributed database can easily be accessed from different networks.
- This database is more secure in comparison to centralized database.

Disadvantages

- This database is very costly and it is difficult to maintain because of its complexity.
- In this database, it is difficult to provide a uniform view to user since it is spread across different physical locations.

S.NO.	Centralized database	Distributed database
1.	it is a database that is stored, located as well as maintained at a single location only.	It is a database which consists of multiple databases which are connected with each other and are spread across different physical locations.
2.	The data access time in the case of multiple users is more in a centralized database.	The data access time in the case of multiple users is less in a distributed database.
3.	The management, modification, and backup of this database are easier as the entire data is present at the same location.	The management, modification, and backup of this database are very difficult as it is spread across different physical locations.
4.	This database provides a uniform and complete view to the user.	Since it is spread across different locations thus it is difficult to provide a uniform view to the user.
5.	This database has more data consistency in comparison to distributed database.	This database may have some data replications thus data consistency is less.
6.	The users cannot access the database in case of database failure occurs.	In distributed database, if one database fails users have access to other databases.
7.	Centralized database is less costly.	This database is very expensive.

Terms used in relational database model

- **Data redundancy:** data redundancy means that some data is duplicate in many different table. For example details of sales person's name, address and paid rate might be held on a payroll table for calculating the payroll. The same data may be held on a table in the personnel department along with other personal data and also in sales department.
- **Data inconsistency: Data inconsistency** is a situation where there are multiple tables within a database that deal with the same **data** but may receive it from different inputs. **Inconsistency** is generally compounded by **data** redundancy.
- **Data Integrity:** it is the accuracy of data and its conformity to its expected value. In simple word it means data is accurate and updated.

Domains and Tuples

- In data management and database analysis, a data domain refers to all the values which a data element may contain. The rule for determining the domain boundary may be as simple as a data type with an enumerated list of values.
- Tuple is the collection of information about the attributes of table for single instance. In simple this also can be called as a 'row' in a Table.

Data Security:

- **Database security** refers to the various measures organizations take to ensure their **databases** are protected from internal and external threats. **Database security** includes protecting the **database** itself, the **data** it contains, its **database** management system, and the various applications that access it. **Data security** includes **data** encryption, hashing, tokenization, and key management practices that protect **data** across all applications and platforms. **Data security** is also known as information **security** (IS) or computer **security**.

database security measures

1) Establish strong passwords

This first measure is really easy to put in place. You must put together a combination of capitals, lower-case letters, numbers, and symbols to create a strong password. The more characters you put, the better. With that, you must avoid using your birthday or any personal information and change the password accordingly.

2) Set up a firewall

In order to protect your network, firewalls are an important initiative to consider. They are a must-have for any company, as they control the internet traffic coming and leaving your business.

3) Think of antivirus protection

Antivirus and anti-malware are indispensable to protecting your Data. They are designed to prevent, search for, detect and remove viruses but also adware, worms, trojans, and so on.

4) Updating is important

Your computer must be properly patched and updated. Recent updates allow your Data to be more secured.

5) Secure every laptop

Laptops are portable so there is a higher risk that they can be stolen. As a consequence, it is important to take more security measures in order to protect all laptops. A simple solution is to encrypt them. In doing so, without the right password, your computer's Data is unreadable.

6) Secure mobile phones

Mobile phones are even more easily stolen than laptops but they are as valuable for companies. Equally to laptops, phones can be encrypted- you can put a strong password and enable an automatic lock-out. You can also set up a wiping process if the phone is lost or stolen.

7) Schedule backups

You can schedule backups to external hard drives or in the cloud in order to keep your data stored safely. The right frequency is weekly but you can do incremental backups every few days. You can also use Wimi, which centralizes your documents. With it, you can then share your documents with your team, your clients, and your partners

8) Monitor steadily

Data, Software, technologies, everything is moving so fast. Keep track of them, keep in touch with news to see what is new on the market.

9) Be smart with emails and surfing the web

Downloading apps or files, opening emails and clicking on links can infect your computer and your network. Be careful with the sources you find online or you receive. Take every "warning box" seriously.

10) Educate your employees about Data Security

Prevention is the best way to keep your Data safe. Warned employees will always be more attentive.

Assignment

1. List out advantage and disadvantages of database system
2. Explain data, information, database and database management system.
3. write different between database and database management system
4. what is database object explain
5. what is key? explain types of key in DBMS
6. what is SQL? explain DDL and DML with example
7. write different between DDL and DML
8. Explain ER database model with advantages and disadvantages
9. explain type of relationship in ER database model
10. explain hierarchical database model with its advantages and disadvantages

11. explain network database model with its advantages and disadvantages
12. explain relation database model with its advantages and disadvantages
13. explain object oriented database model with its advantages and disadvantages
14. what is normalization explain with its types
15. write different between centralized and distributed database system
16. what is data and database security explain