

# CACS 201: Computer Fundamentals and Applications

PREPARED BY KRISHNA PD. ACHARYA  
ASST. PROFESSOR

(MECHI MULTIPLE CAMPUS)

# Database Management System

- Introduction to Database Management
- Database Model
- SQL
- Database Design and Security
- Data warehouse and data mining.
- Database Administrator

Copyright@Krishna Pd. Acharya  
Mechi Multiple Campus  
BCA Programme

# Database Management System

- **Data** is a collection of unprocessed items, which can include text, numbers, images, audio, and video.
- **Information** is processed data. it is organized, meaningful, and useful.
- The collection of interrelated data of an organization usually referred to as the **database**, which organize in a manner that allows access, retrieval, and use of that data.
- **A database management system (DBMS)** is a collection of interrelated data and a set of programs to access those data. The primary goal of a DBMS is to provide a way to store and retrieve database information that is both convenient and efficient.
- Database systems are designed to manage **large** amount of information. Management of data involves both **defining structures** for storage of information and providing mechanisms for the **manipulation of information**. In addition, the database system must **ensure the safety** of the information stored, despite system crashes or attempts at **unauthorized access**. If data are to be **shared** among several users, the system must avoid possible **inconsistent results**. Ex. Oracle, MS-SQL server, MySQL, Sybase, IBM DB2 MS-Access etc.

## Purpose of Database Systems(Advantage)

- **Controlling Redundancy**: Data redundancy is a condition created within a database in which the same piece of data is held in two separate places. This can mean two different fields within a single database, which is controlled by dbms.
- **Integrity can be enforced**: Integrity of data means that data in database is always accurate, such that incorrect information cannot be stored in database. In order to maintain the integrity of data, some integrity constraints are enforced on the database.
- **Inconsistency can be avoided** : When the same data is duplicated and changes are made at one file, which is not propagated to the other file, it gives rise to inconsistency and the two entries regarding the same data will not agree. At such times the data is said to be inconsistent.

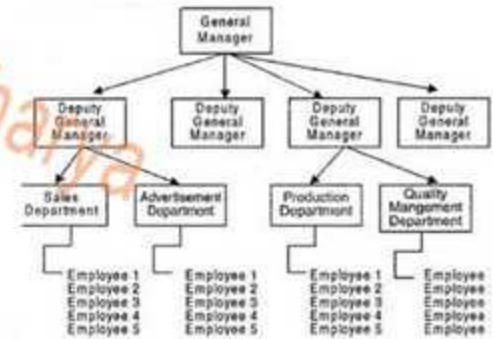


# Database Management System

- **Data can be shared:** data can be shared in different sites among multiple user with use of network.
- **Standards can be enforced :** Since DBMS is a central system, so standard can be enforced easily. The standardized data is very helpful during migration or interchanging of data.
- **Restricting unauthorized access:** When multiple users share a database, it is likely that some users will not be authorized to access all information in the database.
- **Concurrency Control :** DBMS systems provide mechanisms to provide concurrent access of data to multiple users.
- **Providing Backup and Recovery:** A DBMS must provide facilities for recovering from hardware or software failures. The backup and recovery subsystem of the DBMS is responsible for recovery.

## Database model

- database model refers to the logical structure, representation or layout of a database and how the data will be stored, managed and processed within it. It helps in designing a database and serves as blueprint for application developers and database administrators in creating a database.
- **Types of Database model**
  - Hierarchical database model
  - Network model
  - Entity-relationship model
  - Relational model
  - Object-oriented database model



## Hierarchical database model

Data is represented in the form of hierarchical **or tree-like structure**. The root may have any number of **dependents**, each of these dependents may have any number of lower level dependents. A parent record can have several children but a **child can have one parent**. So it maintain **one to one or one to many** relationship

# Database Management System

## Hierarchical database model

Data is represented in the form of hierarchical **or tree-like structure**. The root may have any number of **dependents**, each of these dependents may have any number of lower level dependents. A parent record can have several children but a **child can have one parent**. So it maintain **one to one or one to many** relationship.

## Advantages(pros)

1. It promotes data sharing.
2. Parent/child relationship promotes conceptual simplicity.
3. Database security is provided and enforced by DBMS.
4. Parent/child relationship promotes data integrity.
5. It is efficient with 1:1 or 1:M relationships.

## Disadvantages(pros)

1. Complex implementation requires knowledge of physical data storage characteristics.
2. Navigational system yields complex application development, management, and use; requires knowledge of hierarchical path.
3. Changes in structure require changes in all application programs.
4. There are implementation limitations (no multiparent or M:N relationships).
5. There is no data definition or data manipulation language in the DBMS.
6. There is a lack of standards.



# Database Management System

## The network model

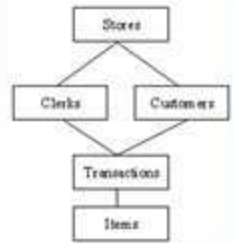
It is very similar to the hierarchical model actually. In fact, the hierarchical model is a subset of the network model. However, instead of using a single-parent tree hierarchy, the network model uses set theory to provide a tree-like hierarchy with the exception that child tables were allowed to have more than one parent. This allowed the network model to support many-to-many relationships.

## Advantage of Network model

1. Conceptual simplicity is at least equal to that of the hierarchical model.
2. It handles more relationship types, such as M:N and multiparent.
3. Data access is more flexible than in hierarchical and file system models.
4. Data member relationship promotes data integrity.
5. There is conformance to standards.
6. It includes data definition language (DDL) and data manipulation language (DML) in DBMS.

## Disadvantage of Network model

1. Increase System complexity due to many to many relationship.
2. Navigational system yields complex implementation, application development, and management.
3. Structural changes require changes in all application programs. (no data independence)

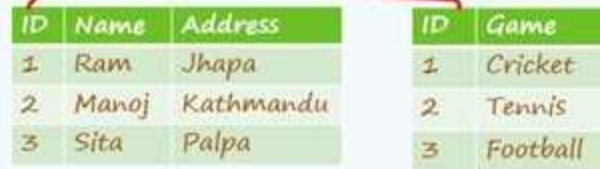


# Database Management System

## Relational Model.

The relational model uses a collection of tables to represent both data and the relationships among those data. Each table has multiple columns, and each column has a unique name. Tables are also known as relations.

- Single row of table is called tuple or record
- Each column of table is field or attribute
- Single table is relation



ID	Name	Address
1	Ram	Jhapa
2	Manoj	Kathmandu
3	Sita	Palpa

ID	Game
1	Cricket
2	Tennis
3	Football

## Advantage:

**Structural independence** is promoted by the use of independent tables. Changes in a table's structure do not affect data access or application programs.

2. Tabular view substantially improves conceptual **simplicity**, thereby promoting easier database design, implementation, management, and use.

3. Ad hoc query capability is based on SQL. (Ad hoc means dynamic query which means query is formed based on condition selected by user)

4. Powerful RDBMS isolates the end user from physical-level details and improves implementation and management simplicity.

## Disadvantage

1. The RDBMS requires substantial hardware and system software overhead.

2. Conceptual simplicity gives relatively untrained people the tools to use a good system poorly, and if unchecked it may produce the same data anomalies found in file systems.

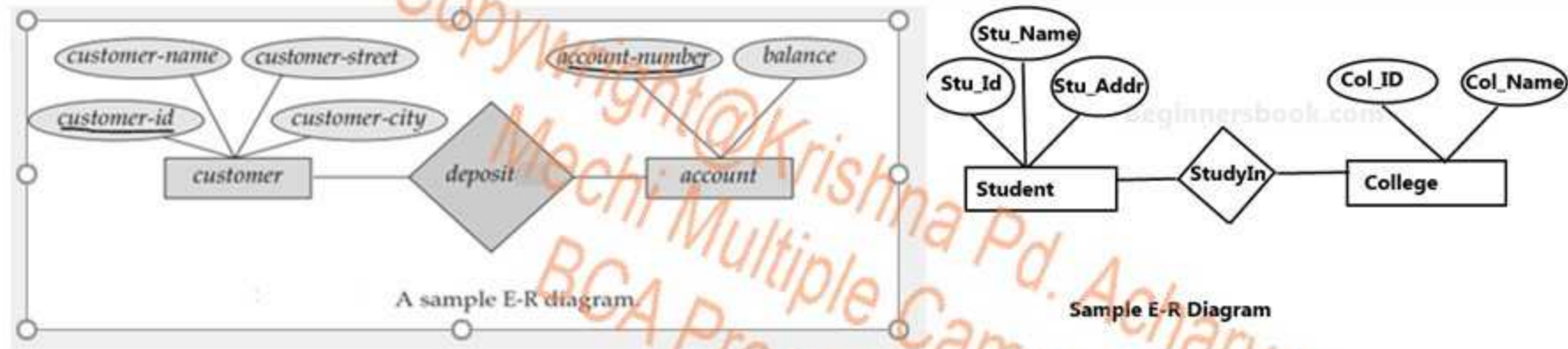
3. It may promote islands of information problems as individuals and departments can easily develop their own applications.



# Database Management System

## The entity-relationship (E-R) data model:

it uses a collection of basic objects, called entities, and relationships among these objects. An entity is a "thing" or "object" in the real world that is distinguishable from other objects. The entity-relationship model is widely used in database design. For example, each person is an entity, and bank accounts can be considered as entities. Entities are described in a database by a set of attributes. For example, the attributes account-number and balance may describe one particular account in a bank, and they form attributes of the account entity set. Similarly, attributes customer-name, customer-street address and customer-city may describe a customer entity.



## Advantage:

1. Visual modeling yields exceptional conceptual simplicity.
2. Visual representation makes it an effective communication tool.
3. It is integrated with the leading relational model.
4. Implement real word object as entity.



# Database Management System

## Disadvantage:

1. There is limited constraint representation.
2. There is limited relationship representation.
3. There is no data manipulation language.
4. Loss of information content occurs when attributes are removed from entities to avoid crowded displays. (This limitation has been addressed in subsequent graphical versions.)

An **object-oriented database management system** (OODBMS) is a database management system that supports the creation and modeling of data as objects. OODBMS also includes support for classes of objects and the inheritance of class properties, and incorporates methods, subclasses and their objects. Most of the object databases also offer some kind of query language, permitting objects to be found through a declarative programming approach.

A database system that incorporates all the important object-oriented concepts that are includes **Abstract Data Types** (Class definition, provides extension to complex attribute types) **Encapsulation** (data and functions combine into single unit) **Inheritance** (transfer the certain attribute from its parent or root) **Method overloading** (one function can work differently based on input parameter)

Class	Objects	Attributes/Data
Employee	Deepak Reddy	Age: 25
		Designation: TL
		Salary: 50,000 pm
	Vamshi Krishna	Age: 23
		Designation: SSE
		Salary: 25,000 pm



# Database Management System

## **Advantage:**

1. Semantic(structure) content is added.
2. Visual representation includes semantic content.
3. Inheritance promotes data integrity.
4. class, overloading, encapsulation and inheritance are available.

## **Disadvantage:**

1. Slow development of standards caused vendors to supply their own enhancements, thus eliminating a widely accepted standard.
2. It is a complex navigational system.
3. There is a steep learning curve.
4. High system overhead slows transactions.
5. Required more memory to represent object.

## **SQL:**

Structured Query Language (SQL) is a standard computer language for relational database management and data manipulation. SQL is used to query, insert, update and modify data. Most relational databases support SQL, which is an added benefit for database administrators (DBAs), as they are often required to support databases across several different platforms.

## **Data Definition Language (DDL)**

A data definition language (DDL) is a Query language used to create and modify the structure of database objects in a database. These database objects include views, schemas, tables, indexes, etc.

**Data manipulation language (DML)** is a family of query languages including commands permitting users to manipulate data in a database. This manipulation involves inserting data into database tables, retrieving existing data, deleting data from existing tables and modifying existing data.



# Database Management System

## Database Administrators

One of the main reasons for using DBMS is to have central control of both the data and the programs that access those data. A person who has such central control over the system is called a database administrator (DBA). The functions of a DBA include

**Schema definition.** The DBA creates the original database schema by executing a set of data definition statements in the DDL.

**Storage structure and access-method definition.**

**Schema and physical-organization modification.** The DBA carries out changes to the schema and physical organization to reflect the changing needs of the organization, or to alter the physical organization to improve performance.

**Granting of authorization for data access.** By granting different types of authorization, the database administrator can regulate which parts of the database various users can access. The authorization information is kept in a special system structure that the database system consults whenever someone attempts to access the data in the system.

**Routine maintenance.** Examples of the database administrator's routine maintenance activities are:

- Periodically **backing up** the database, either onto tapes or onto remote servers, to prevent loss of data in case of disasters such as flooding.
- Ensuring that enough **free disk space** is available for normal operations, and upgrading disk space as required.
- **Monitoring jobs running** on the database and ensuring that performance is not degraded by very expensive tasks submitted by some users.



# Database Management System

## Database Security:

It protection from malicious attempts(unauthorized user) to steal (view) or modify data.

**Secrecy:** Users should not be able to see things they are not supposed to.

E.g., A student can't see other students' grades.

**Integrity:** Users should not be able to modify things they are not supposed to.

E.g., Only instructors can assign Internal marks.

**Availability:** Users should be able to see and modify things they are allowed to.

There are three main **security objectives**. First, information should not be disclosed to unauthorized users (**secrecy**). Second, only authorized users should be allowed to modify data (**integrity**). Third, authorized users should not be denied access (**availability**).

## Authentication

Authentication is the process of confirming that a user logs in only in accordance with the rights to perform the activities he is authorized to perform. User authentication can be performed at operating system level or database level itself.

**Authorization** is the process of giving someone permission to do something. In multi-user database management systems, a database administrator defines some privileges. Based on those privilege, users are allowed/disallowed to access the resources of database.



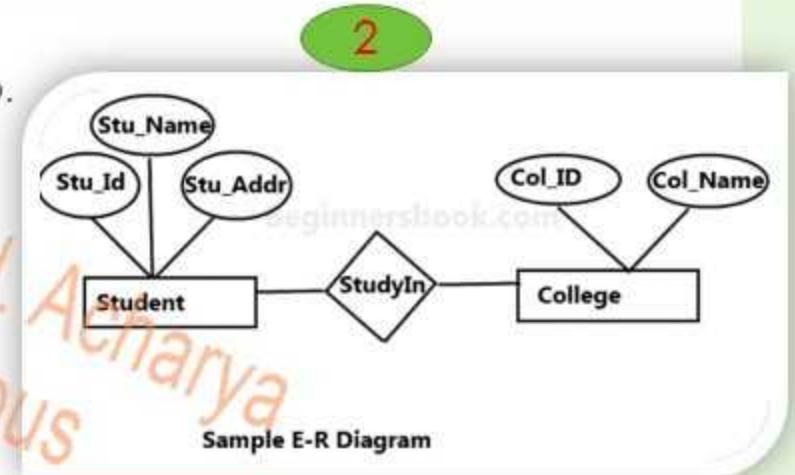
# Database Management System

**Database design** is the process of producing a detailed data model of a database. This data model contains all the needed logical and physical design choices and physical storage parameters needed to generate a design in a data definition language, which can then be used to create a database. A fully attributed data model contains detailed attributes for each entity.

The process of doing database design generally consists of a number of steps which will be carried out by the database designer. Usually, the designer must:

- Determine the data to be stored in the database.
- Determine the relationships between the different data elements.
- Apply to a logical structure upon the data on the basis of these relationship.

Business situation: A College has many students in different faculties.



3

**create table STUDENT**

(*Stud\_id* **integer primary key**,

*STU\_NAME* **varchar(10)**,

*Stud\_addr* **varchar(25));**

4

Stud_id	Stud_name	Stud_addr

# Database Management System

## Data warehouses (DW)

- DW are centralized data repositories that integrate data from various transactional, legacy, or external systems, applications, and sources.
- The data warehouse provides an environment separate from the operational systems and is completely designed for decision-support, analytical-reporting, ad-hoc queries, and data mining.
- This isolation and optimization enables queries to be performed without any impact on the systems that support the business' primary transactions (i.e transactional and operational systems).

The data characteristics/Feature of a Data Warehouse are:

- Subject-oriented
- Time-variant
- Non-volatile
- Integrated

Advantage of DW.

- Allows end users to perform extensive analysis.
- Allows a consolidated view of corporate data.
- Better and more timely information.
- Enhanced system performance.
- Simplification of data access.
- Enhance business knowledge
- Present competitive advantage
- Facilitate decision making
- Help in reforming business processes

## Disadvantage of DW

- Data warehouses are not the optimal environment for unstructured data. Because data must be extracted, transformed and loaded into the warehouse.
- Data warehouses can have high maintenance costs.
- Data warehouses can get outdated relatively quickly.
- Required skilled person to generate DW.
- Longer time to developed.



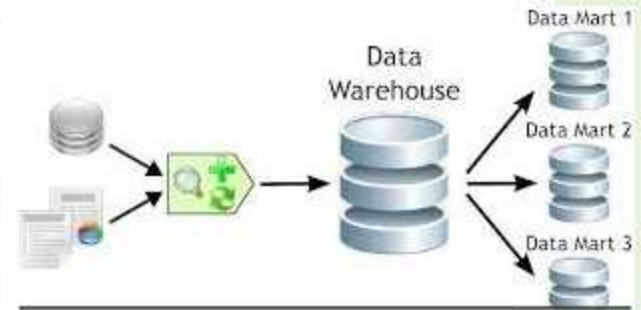
# Database Management System

## Data warehouses (DW)

- DW are centralized data repositories that integrate data from various transactional, legacy, or external systems, applications, and sources.
- The data warehouse provides an environment separate from the operational systems and is completely designed for decision-support, analytical-reporting, ad-hoc queries, and data mining.
- This isolation and optimization enables queries to be performed without any impact on the systems that support the business' primary transactions (i.e transactional and operational systems).

The data characteristics of a Data Warehouse are:

- Subject-oriented
- Time-variant
- Non-volatile
- Integrated



## DATA MART :

A data mart is a subset of a data warehouse in which a summarized or highly focused portion of the organization's data is placed in a separate database for a specific population of users. For example, a company might develop marketing and sales data marts to deal with customer information. A data mart typically focuses on a single subject area or line of business, so it usually can be constructed more rapidly and at lower cost than an enterprise-wide data warehouse.

# Database Management System

## Data Mining:

- Data Mining is defined as extracting the information from the huge set of data. In words, data mining is mining the knowledge from data.
- It can find out hidden patterns and relationships in large databases and deducing rules from them to predict future behavior. The patterns and rules are used to guide decision making and forecast the effect of those decisions.
- The types of information obtainable from data mining include associations, sequences, classifications, clusters, and forecasts.

**Associations:** Determine occurrences linked to a single event.

**Sequences:** Determine events that are linked over time.

**Classification:** Discover characteristics of customers and make predictions about their behavior.

**Clustering:** Discover groups within data.

**Forecasting:** Use existing values to forecast what other values will be.



Figure 1: Data mining process [1,2]



# Assignment

1. Define Database and DBMS with their advantages and Disadvantages.
2. Explain the different database models with their merits and demerits.
3. Write down the steps to design the database with example.
4. What do you mean by data security. Explain authorization and authentication with example.
5. Define Data Warehouse with its pros and cons.
6. What is data mart? Differentiate between data warehousing and data mart
7. Who is DBA? Write the responsibilities of DBA?
8. What is data mining? Explain the data mining process.
9. What is SQL? Explain types of Query language.
10. What is flat file? Discuss its limitation.

## Disadvantage of DBMS

1. Training required for normal user:
2. Increase software and hardware cost
3. Database Administrator required
4. Technical Problems.
5. Time-Consuming Design
6. Compatibility issue
7. Complexity increased