

Database Management Systems

Lecture # 1 & 2
Subject Teacher: Zartasha Baloch

Data and Database

Lecture # 1 & 2

Disclaimer: The material used in this presentation to deliver the lecture i.e., definitions/text and pictures/graphs etc. does not solely belong to the author/presenter. The presenter has gathered this lecture material from various sources on web/textbooks. Following sources are especially acknowledged:

1. Connolly, Thomas M., and Carolyn E. Begg. *Database systems: a practical approach to design, implementation, and management*. Pearson Education, 2005.
2. Hoffer, Jeffrey A., Venkataraman Ramesh, and Heikki Topi. *Modern database management*. Upper Saddle River, NJ: Prentice Hall,, 2011.

In this Lecture you will Learn about:

- ▶ **Data, Information, Metadata**
- ▶ **Traditional File Processing System**
- ▶ **Examples of Database Systems**
- ▶ **Database**
- ▶ **Database Management Systems**
- ▶ **Database application programs, Views**
- ▶ **Database environment**
- ▶ **Functions of a DBMS**
- ▶ **Advantages and Disadvantages of DBMS**

Data

- ▶ Data is a plural form of ***datum***, which means a piece of information.
- ▶ Data is a collection of facts & figures, such as values or measurements.
- ▶ Data can be,
 - ▶ Numbers
 - ▶ Words
 - ▶ Measurements
- ▶ A collection of data which conveys some meaningful idea is called information.

Data (cont'd)

- ▶ Another definition of data can be a representations of objects and events that have meaning and importance in the user's environment and stored on computer media.
- ▶ For example, in a salesperson's database, the data would include facts such as customer name, address and telephone number.
- ▶ This type of data is called **Structured Data**.
- ▶ The most important structured data types are numeric, character and dates.
- ▶ Structured data are stored in tabular form (in tables, relations, arrays, spreadsheets, etc) and are most found in traditional databases.

Data (cont'd)

- ▶ Today's databases are used to store objects such as documents, maps, photographic images, sound and video segment in addition to structured data.
- ▶ For example, the salesperson's database might include a photo image of the customer contact. It might also include a sound recording or video clip about the most recent product. This type of data is referred to as **Unstructured Data**, or as **Multimedia Data**.
- ▶ Multimedia data are most often found on Web servers and on Web-enabled databases.

Information & Metadata

- ▶ **Information** is data that have been processed in such a way as to increase the knowledge of the person. OR
- ▶ **Information** is facts about situation, person, event, etc.
- ▶ When data is collated or organized into meaningful form, it gains significance, and that meaningful data is called information.
- ▶ **Metadata** is a data about data.
- ▶ **Metadata** is data that describe the properties or characteristics of end-user data, and the context of that data. Some of the properties that are typically described include data names, definitions, size and allowable values.
- ▶ For example, if **ID** is a data then the **ID** must have data type.

Data Vs Information

- ▶ The terms data and information are closely related and in fact are often used interchangeably.
- ▶ However, it is useful to distinguish between data and information.
- ▶ For example
 - ▶ 221120 is a data;
 - ▶ it can be information if 22/11/20 date of your exam.
 - ▶ 221120 can be a zip code of any city.
 - ▶ PKR 221,120 can be Salary of any Person

Class Roster

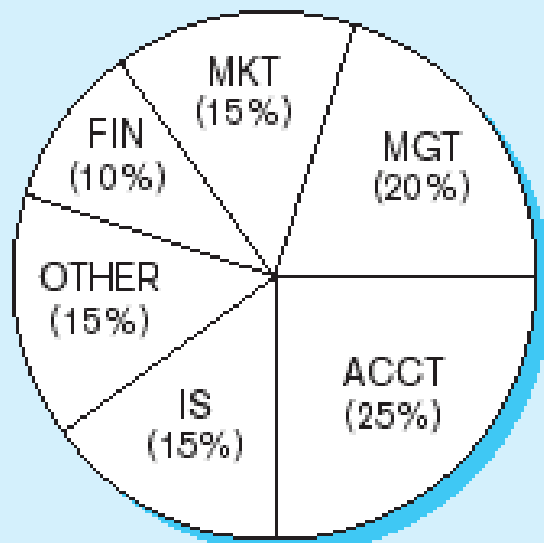
Course: MGT 500 Semester: Spring 200X
Business Policy

Section: 2

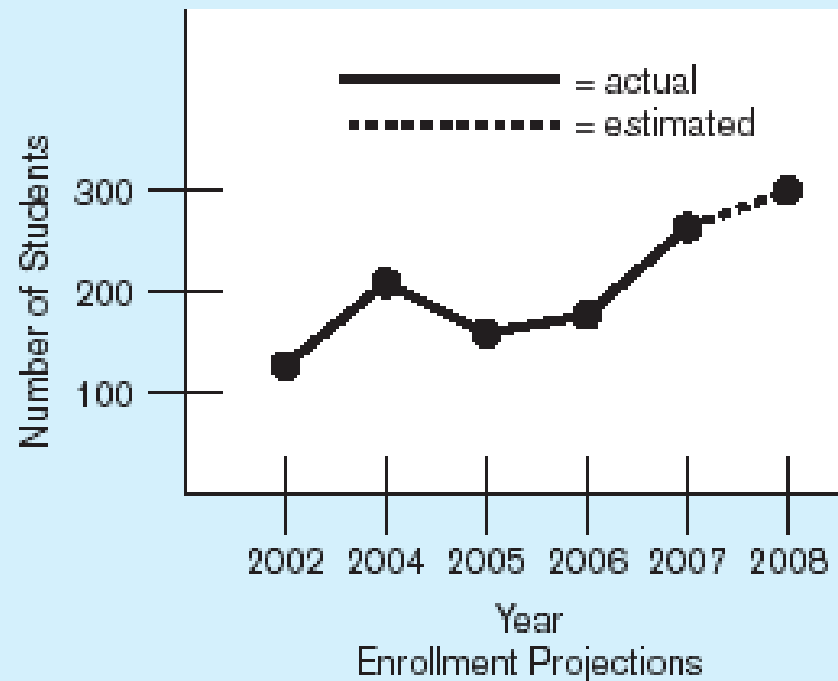
<u>Name</u>	<u>ID</u>	<u>Major</u>	<u>GPA</u>
Baker, Kenneth D.	324917628	MGT	2.9
Doyle, Joan E.	476193248	MKT	3.4
Finkle, Clive R.	548429344	PRM	2.8
Lewis, John C.	551742186	MGT	3.7
McFerran, Debra R.	409723145	IS	2.9
Sisneros, Michael	392416582	ACCT	3.3

Data in context

Summarized data



Percent Enrollment by Major (200X)



Graphical displays turn data into useful information that managers can use for decision making and interpretation

<i>Data Item</i>			<i>Value</i>			
Name	Type	Length	Min	Max	Description	Source
Course	Alphanumeric	30			Course ID and name	Academic Unit
Section	Integer	1	1	9	Section number	Registrar
Semester	Alphanumeric	10			Semester and year	Registrar
Name	Alphanumeric	30			Student name	Student IS
ID	Integer	9			Student ID (SSN)	Student IS
Major	Alphanumeric	4			Student major	Student IS
GPA	Decimal	3	0.0	4.0	Student grade point average	Academic Unit

Descriptions of the properties or characteristics of the data, including data types, field sizes, allowable values, and data context

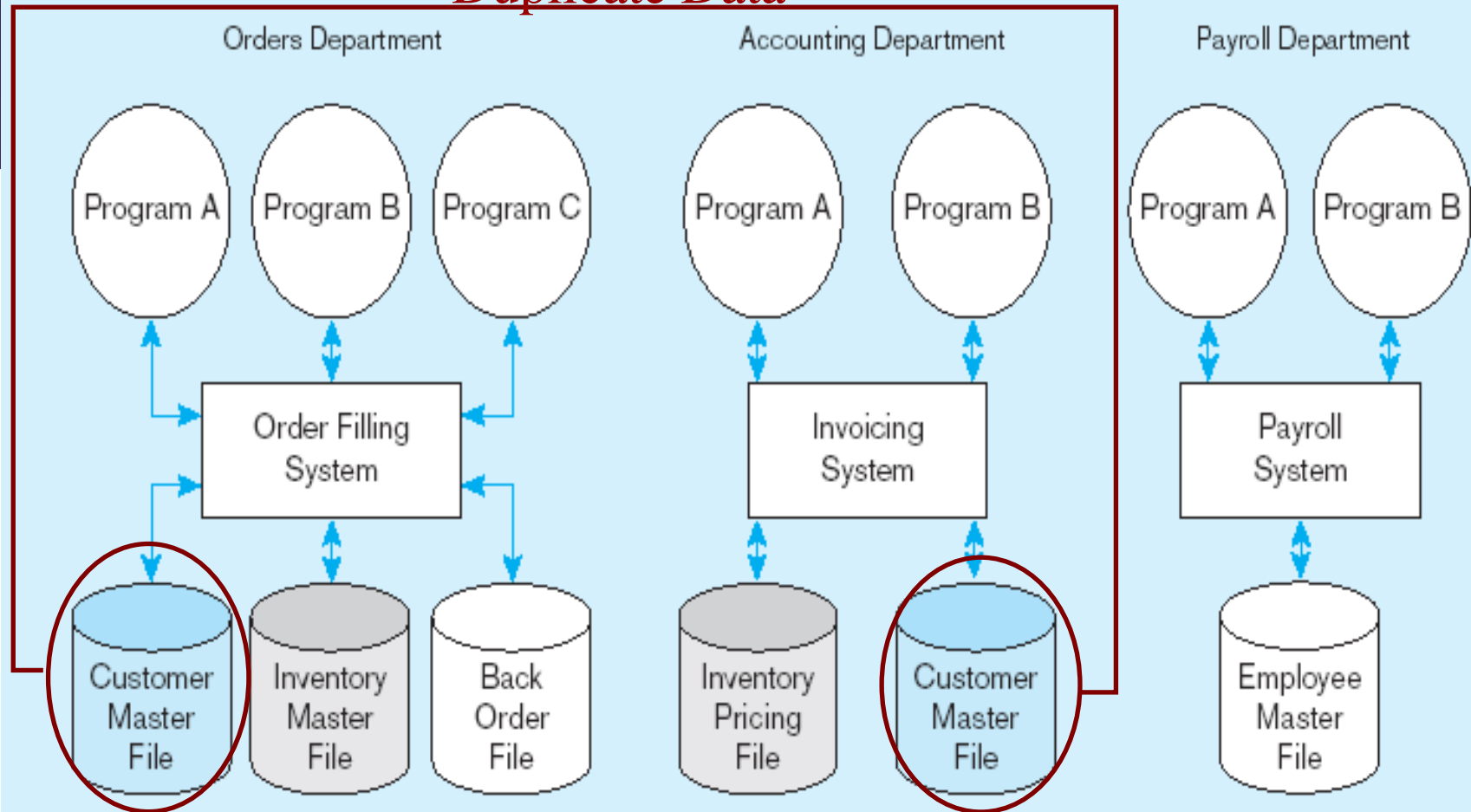
Traditional File Processing System

- ▶ Traditional file processing system or simple file processing system refers to the first computer-based approach of handling the commercial or business applications.
- ▶ That is why it is also called a replacement of the manual file system.
- ▶ While computers were introduced into the business world there is no any concept of DBMS, therefore computer file processing systems were developed to store, maintain and retrieve large files of data.
- ▶ In Traditional File Processing System each application is designed with its own set of data files.

Traditional File Processing System

- ▶ The data may already be present in files for other applications but would need to be restructured to meet the needs of the new application. When a new application is required it is simpler to develop new file structures along with the new application.
- ▶ Different application programs are written to extract data from these files and to add record to these files.
- ▶ But, There are Number of disadvantages in using File Processing System, to store the data.

Duplicate Data



Traditional File Processing System

Disadvantages of Traditional File Processing System

Program-Data Dependence, The reports produced by the file processing system are program dependent, which means if any change in the format or structure of data and records in the file is to be made, the programs must be modified correspondingly. Also, a new program will have to be developed to produce a new report.

Duplication of Data, Since each application has its own data file, the same data may have to be recorded and stored in many files. For example, personal file and payroll file, both contain data on employee name, designation etc. The result is unnecessary duplicate or redundant data items.

Disadvantages of Traditional File Processing System

- Limited Data Sharing,** There is limited data sharing possibilities with the traditional file system. Each application has its own private files and users have little choice to share the data outside their own applications.
- Lengthy Development Times,** each new application requires that the developer essentially start from scratch by designing new file formats and descriptions and then writing the file access logic for each new program.
- Excessive Program Maintenance,** all above factors combined to create a heavy program maintenance load in organization.

Problems with Data Dependency

- ▶ Each application programmer must maintain his/her own data
- ▶ Each application program needs to include code for the metadata of each file
- ▶ Each application program must have its own processing routines for reading, inserting, updating, and deleting data
- ▶ Lack of coordination and central control
- ▶ Non-standard file formats

Problems with Data Redundancy

- ▶ Waste of space to have duplicate data
- ▶ Causes more maintenance headaches
- ▶ The biggest problem:
 - ▶ **Data changes in one file could cause inconsistencies**
 - ▶ Compromises in ***data integrity***

SOLUTION: The DATABASE Approach

- ▶ Central repository of shared data
- ▶ Data is managed by a controlling agent
- ▶ Stored in a standardized, convenient form

Requires a Database Management System (DBMS)

Examples of Database Systems

- ▶ Purchases from the supermarket
- ▶ Purchases using your credit card
- ▶ Booking a holiday at the travel agents
- ▶ Using the local library
- ▶ Renting a video
- ▶ Using the Internet

Database

- ▶ A shared collection of logically related data (and a description of this data), designed to meet the information needs of an organization.

Database

- ▶ Shared collection - can be used simultaneously by many departments and users.
- ▶ Logically related - comprises the important objects and the relationships between these objects.
- ▶ Description of the data - the system catalogue (meta-data) provides description of data to enable data independence.

Database

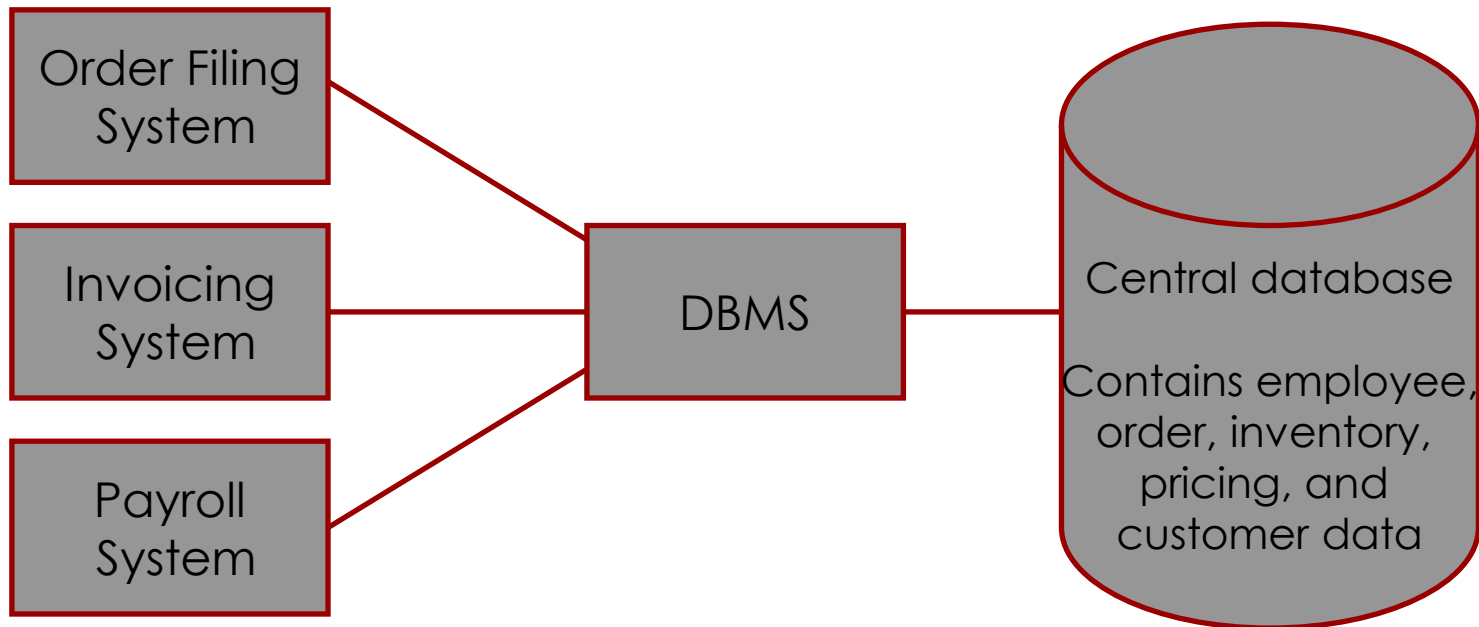
- ▶ In computing, databases are sometimes classified according to their organizational approach.
 - ▶ The most prevalent approach is the **Relational Database**, a tabular database in which data is defined so that it can be reorganized and accessed in a number of different ways.
 - ▶ A **Distributed Database** is one that can be dispersed or replicated among different points in a network.
 - ▶ An **Object-Oriented Database** is one that is congruent with the data defined in object classes and subclasses.

Database Management Systems

- ▶ A software system that enables users to define, create, and maintain the database and that provides controlled access to this database.

DATABASE MANAGEMENT SYSTEMS

A software system that is used to create, maintain, and provide controlled access to user databases



DBMS manages data resources like an operating system manages hardware resources

Database application program

- ▶ A software program that interacts with the database by issuing an appropriate request (typically an SQL statement) to the DBMS.

Views

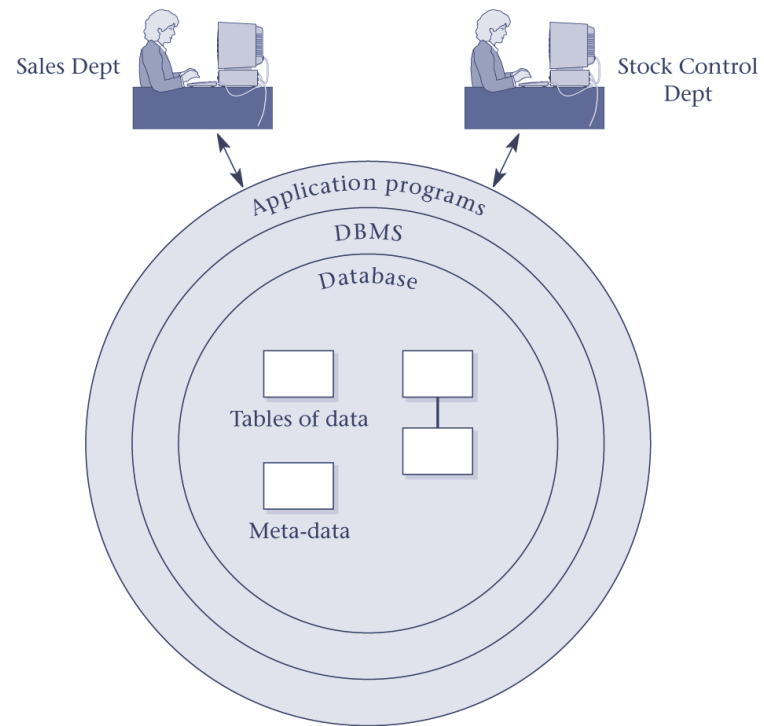
- ▶ Allows each user to have his or her own view of the database.
- ▶ A view is essentially some subset of the database.

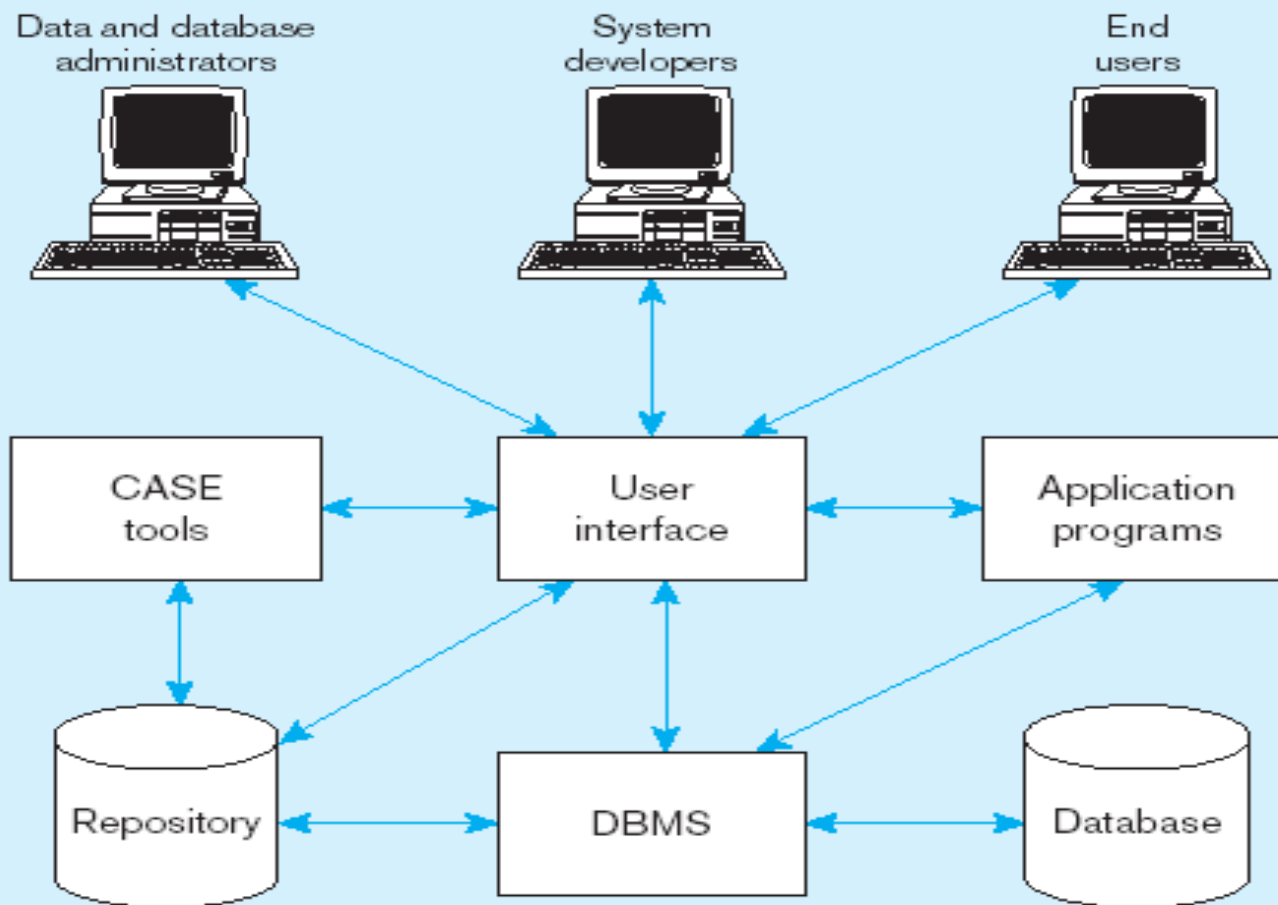
Views

► Benefits include:

- Provide a level of security;
- Provide a mechanism to customize the appearance of the database;
- Present a consistent, unchanging picture of the structure of the database, even if the underlying database is changed.

DBMS





Components of the Database Environment

Components of DBMS Environment

- **CASE Tools**—computer-aided software engineering
- **Repository**—centralized storehouse of metadata
- **Database Management System (DBMS)** —software for managing the database
- **Database**—storehouse of the data
- **Application Programs**—software using the data
- **User Interface**—text and graphical displays to users
- **Data/Database Administrators**—personnel responsible for maintaining the database
- **System Developers**—personnel responsible for designing databases and software
- **End Users**—people who use the applications and databases

Functions of a DBMS

- ▶ Data Storage, Retrieval, and Update.
- ▶ A User-Accessible Catalog.
- ▶ Transaction Support.
- ▶ Concurrency Control Services.
- ▶ Recovery Services.

Functions of a DBMS

- ▶ **Authorization Services.**
- ▶ **Support for Data Communication.**
- ▶ **Integrity Services.**
- ▶ **Services to Promote Data Independence.**
- ▶ **Utility Services.**

Advantages of DBMSs

- ▶ Control of data redundancy
- ▶ Data consistency
- ▶ Sharing of data
- ▶ Improved data integrity
- ▶ Improved maintenance through data independence

Advantages of DBMSs

- ▶ Improved security
- ▶ More information from same amount of data
- ▶ Enforcement of standards
- ▶ Increased Concurrency
- ▶ Improved backup and recovery services

Disadvantages of DBMS

- ▶ Complexity
- ▶ Cost of DBMS
- ▶ Cost of conversion
- ▶ Performance
- ▶ Higher impact of a failure