Chapter # 01
Data Basics

Q. Define data and information. Explain with examples.

Data

Data is a collection of facts, figures and statistics related to an object. An object can be a person, event or anything about which data is gathered. Data can be processed to create useful information.

Data is a valuable asset for an organization. Data can be used by the managers to perform effective and successful operations of management. It provides a view of past activities related to the rise and fall of an organization. It also enables the user to make better decisions for future. Data is very useful for generating reports, graphs and statistics etc.

Example

Students fill an admission form when they get admission in college. The form consists of raw facts about the students. These raw facts are student's name, father name, address etc. The purpose of collecting this data is to maintain the records of the students during their study period in the college.

Information

The processed data is called information. Information is meaningful, useful and organized. It is used for making decisions. Data is used as input for processing and information is the output of this processing.

Example

Data collected from census is used to generate different types of information. The government can use it to determine the literacy rate in the country. Government can use the information in important decisions to improve literacy rate.

Q. Define data processing. What activities are involved in data processing? Discuss in detail.

The process of manipulating data to achieve the required objectives and results is called data processing. The software is used to process data. The software converts data into meaningful information. A series of actions or operations are performed on data to get the required output or result.

Activities in Data Processing

Different activities involved in data processing are as follows:

- Data capturing
- Data manipulation
- Managing output results

1. Data Capturing

The process of recording the data in some form is called **data capturing**. Data is captured before it can be processed. Data may be recorded on source documents. Data can also be given directly to the computer through input devices.

2. Data Manipulation

The process of applying different operations on data is called **data manipulation**. The following operations can be performed on data:

- **Classifying:** A process of organizing data into classes or groups is called classifying. For example, the data in a college can be classified in two groups. The data of students may be in one group and data of teachers may be in second group.
- **Calculation:** A process of applying arithmetic operations on data is called calculation. The common calculations are addition, subtraction, multiplication and division etc.
- **Sorting:** The process of arranging data in a logical sequence is called sorting. The data can be sorted numerically or alphabetically.
- **Summarizing:** The process of reducing a large amount of data in a more concise and usable form is called summarizing. For example, people deposit money in banks daily. The data of bank can be summarized to show the total money deposited in a particular month instead of showing all deposits.

3. Managing Output Result

The following activities can be performed on data after the data has been captured and manipulated:

- **Storage:** The process of retaining data for future use is called data storage. Different storage medium are used to store the data such as hard disks and tapes etc.
- **Retrieval:** The process of accessing or fetching the stored data is called data retrieval. The data can be retrieved as and when required. The retrieved data can be displayed in different forms such as reports, graphs and charts etc.
- **Communication:** The process of transferring data from one location to another is called data communication. The data may be transferred to different locations for further processing. For example, the result can be sent to the students via email.
- **Reproduction:** The process of copying or duplicating data is called reproduction of data. Data can be reproduced if different users need data at different locations.

Q. What is difference between data and information?

The difference between data and information is as follows:

Data	Information
1. Data is unprocessed raw facts about a particular entity.	1. Information is processed form of data.
2. Data is used as input in the computer.	2. Information is the output of computer.
3. Data is normally huge in its volume.	3. Information is normally short in its volume.
4. Data is the asset of organizations and is not available to people for sale.	4. Information is normally available to people for sale.
5. Data is difficult or even impossible to reproduce. For examples, if Government lose the data of census, if will be almost impossible to reproduce it.	5. Information is easier to reproduce if lost. For example, if the list of illiterate citizens is lost, it can be reproduced easily because the data is still stored.
6. Data is used rarely.	6. Information is used frequently.
7. Data is an independent entity.	7. Information depends on data.

Q. Define file, record and field in detail.

Field

A combination of one or more characters is called field. It is the smallest unit of data that can be accessed by the user. The name of each field in a record is unique. The data type of a field indicates the type of data that can be stored in the field. Each field contains one specific piece of information. A field size defines the maximum number of characters that can be stored in a field.

For example, Employee Number, Employee Name, Grade and Designation are fields.

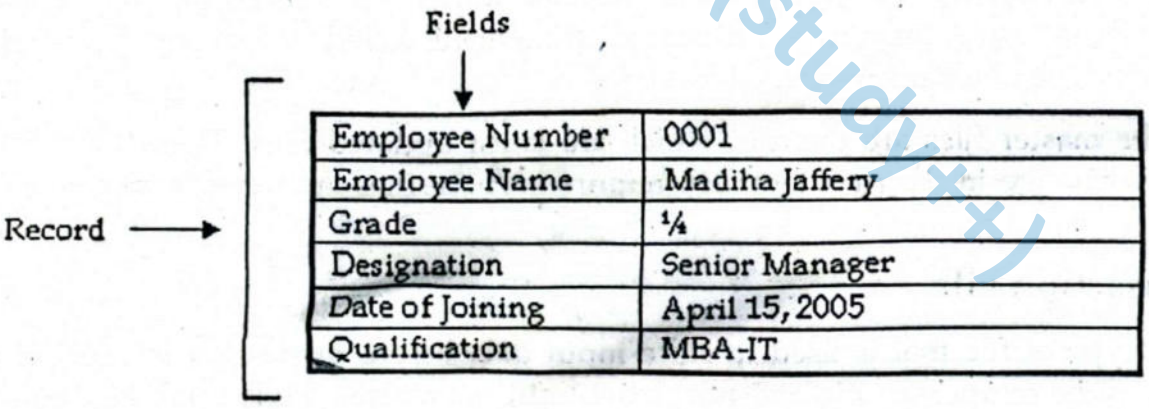

Figure 1.1: Fields and Record

Record

A collection of related fields treated as a single unit is called a record. For example, a employee's record includes a set of fields that contains Employer Number, Employee Name, Grade and Designation etc.

File

A collection of related records treated as a single unit is called file. File is also known as data set. Files are stored in disk like hard disk, CD-ROM or DVD-ROM etc. A Student file may contain the records of hundreds of students. Each student's record consists of same fields but each field contains different data.

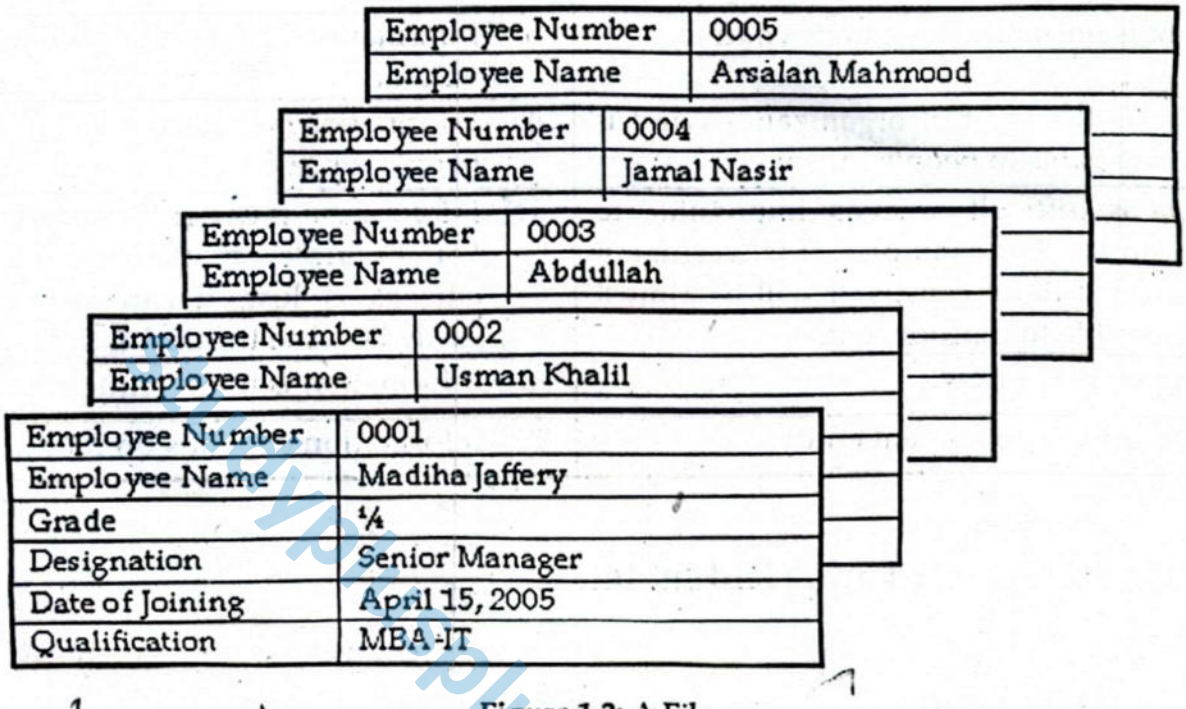

Figure 1.2: A File

Q. Describe the file types from usage point of view.

The types of files from usage point of view are as follows:

1. Master File

Master files are used to store the information that remains constant for a long period of time. For example, the address file for all students enrolled in a college is a master file. It may contain the fields like name, address, phone and email etc.

The master files are the files which are latest updated files. These files are updated when any change in their contents are required. These files are never empty since they are created.

2. Transaction File

A type of file that is used to store input data before processing is called transaction file. It may be temporary file and may exist until the master file is updated. It may also be used to maintain a permanent record of the data about a transaction.

3. Backup File

A type of file that is used to take backup of important data is called backup file. It is a permanent file. It is used to store an additional copy of data. The data can be recovered from this file if the original file is lost or damaged. Backup files are mostly created by using specific software utilities.

Q. Describe the file types from functional point of view.

The files are given proper names from functional point of view. It consists of file name and file extension. The name and extension of a file is separated by dot. The extension is normally assigned by the software in which the file is created.

File Types from Functional Point of View

The types of files from functional point of view are as follows:

1. Program File

A type of file that contains the software instructions is called **program file**. The source program files and executable files are examples of program files. The source program file may have an extension of .com. The executable files may have an extension of .exe.

2. Data File

A type of file that contains data is called **data file**. Data files are created by the software being used. Some examples of data files and software in which these files are created are as follows: => Notepad, Word Processor

Software	File Type	File Extension
Notepad	Text File	.txt
Word Processor	Document	.doc, .rtf
Spreadsheet	Worksheet	.xls, .wks
Database	Data File	.dat, .dbf, .mdb
Image Processor	Image File	.tif, .jpg, .eps, .gif, .bmp
Audio Software	Audio File	.wav, .mid
Video Software	Video File	.avi, .mpg

Q. How do we organize files on storage media?

A technique for physically arranging the records of a file on secondary storage devices is called file organization. The files are organized on storage media in the following ways:

1. Sequential Files

The records in **sequential file organization** are stored in sequence. A sequence means the records are stored one after the other. The records can be retrieved only in the sequence in which they were stored. The principal storage media for sequential files is magnetic tape.

The major disadvantage of sequential access is that it is very slow. If the the last record is to be retrieved, all preceding record are read before reaching the last record.

2. Direct or Random Files

The records in direct file organization are not stored in a particular sequence. A key value of a record is used to determine the location to store the record. Each record is accessed directly without going through the preceding records.

This file organization is suitable for storing data on disk. Direct file organization is much faster than sequetinal file organization for finding a specific record.

A problem may occur in this type of files known as **synonym**. The problem occurs if the same address is calculated to store two or more records.

3. Indexed Sequential Files

In **indexed sequential file organization**, records are stored in ascending or descending order. *The order is* based on a value called **key**. Additionally, indexed file organization maintains an **index** in a file.

An **index** consists of key values and the corresponding disk address for each record in the file. Index refers to the place on a disk where a record is stored. The index file is updated whenever a record is added or deleted from the file.

The records in indexed file organization can be accessed in sequential access as well as **random access** or **direct access**. The records in this file type require more space on storage media. This method is slower than direct file organization as it requires to perform an index search.

Q. What is file processing system? Which problems were faced in traditional file approach in processing information?

File Processing System

Traditional file processing is the first computer-based method to handle business application. In the past, many organizations stored data in files on tape or disk. The data was managed using file processing system. In a typical file processing system, each *department in* an organization has its own set of files. The files are designed specially *for their own* applications. The records in one file are not related to the records in any other file.

Problems in Traditional File Approach

The business organizations faced the following problems in traditional file approach:

1. Data Redundancy means the duplication of data in multiple files

In file processing system, the same data may be duplicated in several files. Suppose there are two files "Students" and "Library". The file "Students" contains the data Roll No, name, address and telephone number and other details of all students in a college. The file "Library" contains the Roll No and name of those students who get a book from library along with the information about the book. The data of one student appears in two files. This is known as data redundancy. This redundancy causes higher storage.

2. Data Inconsistency

File system approach can also result in data inconsistency. Inconsistency means that two files may contain different data of the same student. For example, if the address of a student is changed, it must be changed in both files. There is a possibility that it is changed in *the "Students" file* and not from "Library" file. The data becomes inconsistent in this situation.

3. Program Data Dependency

Program data dependency is a relationship between data in files and program required to update and maintain the files. Application programs are developed according to a particular file format in file processing system. If the format of file is changed, the application

program also needs to be changed accordingly. For example, if there is a change in the length of postal code, it requires change in the program. The changes may be costly to implement.

4. Lack of Flexibility

The file processing system is not very flexible. It cannot easily generate the information that is different from routine. It can take a lot of time to collect the data from different files and write programs to produce the desired information.

5. Integrity Problems

Integrity means reliability and accuracy of data. The stored data must satisfy certain types of consistency constrains. For example, Roll No and Marks of students should be numeric value. It is very difficult to apply these constrains on files in file processing system.

6. Security Problems

File processing system does not provide adequate security on data. In some situations, it is required to provide different types of access to data for different users. For example, a data entry operator should only be allowed to enter data. The chairman of the organization should be able to access or delete the data completely. Such types of security options are not available in file processing system.

Q. What is database? Explain with the help of an example. Which facilities are provided by database system?

Database

A database is a collection of logically related data sets or files. Each file may contain different type of information and are used for specific purposes. The files may be organized in different ways to meet different processing and retrieval requirements of the users.

Example

A bank may have separate files for the clients as follows:

- Saving Accounts
- Current Accounts
- Automobile Loan
- Personal Loan
- Clients Information etc.

The client database of bank will consist of the records from all of the above files. The data of any client can be added, retrieved or updated by using database programs.

Facilities of Database System

A database system normally provides the following facilities to the user:

- Adding new files to database
- Inserting new data in existing files
- Retrieving data from existing files
- Updating data in existing files
- Deleting data from existing files
- Removing existing files from database

Q. Define database system. What are four major components of database systems? Write in details.

Database System

A database system is a collection of data as well as programs required to manage that data. A database system is a computerized record-keeping system. The main purpose of this system is to maintain data and provide it to the user when it is required.

Components of Database Systems

The four major components of database system are as follows:

1. Data

Data is the most important component of database system. Data is a collection of facts stored in the database. The basic purpose of a database system is to store, maintain and process data for the user.

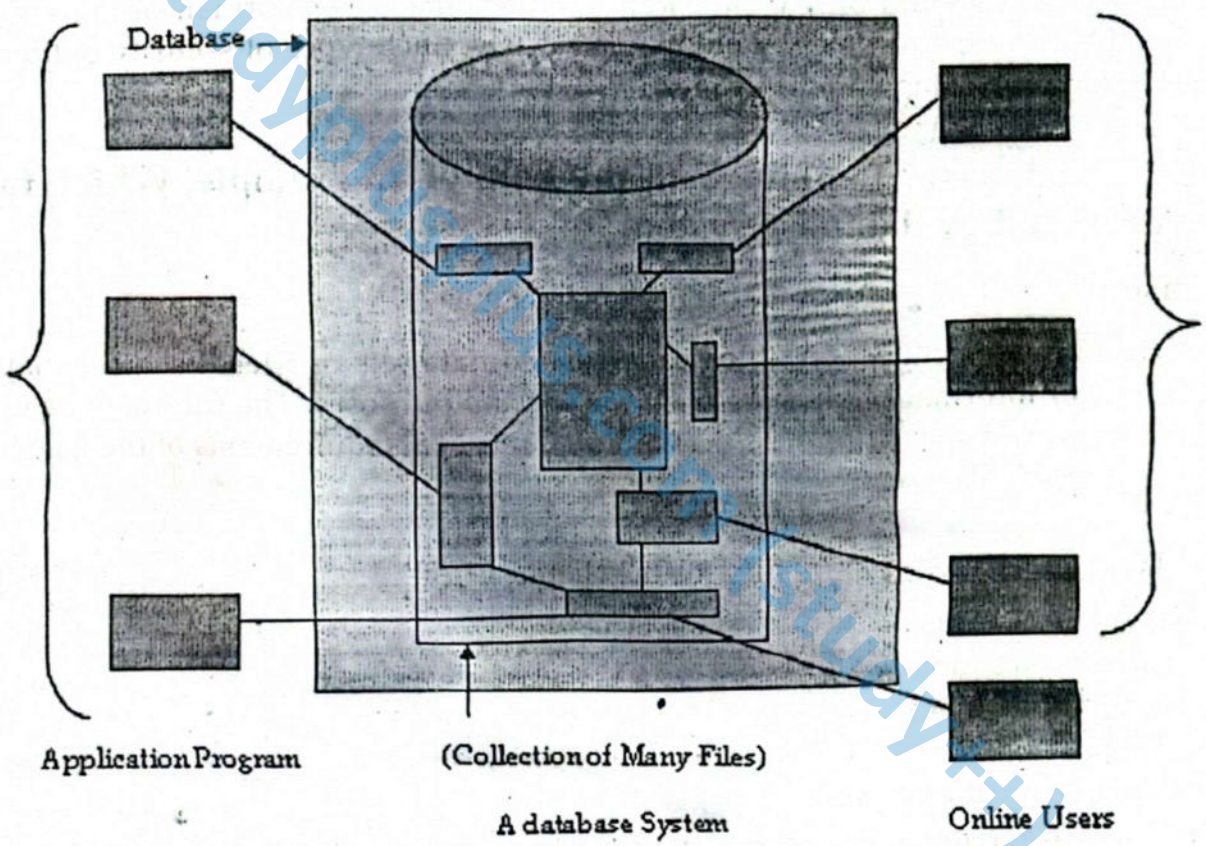

Figure 1.3: Major Components of Database System

2. Hardware

The physical components of a computer system are called hardware. The hardware is used to perform different tasks such as input, output, storage and processing.

Some important hardware components are as follows:

- Secondary Storage
- I/O Devices
- Processors
- Main Memory

3. Software

Software is a collection of programs used by computer within a database system. The most important software is DBMS itself. It uses three types of software to enable the database system work properly. These are as follows:

- **Operating System Software:** It manages all hardware components. It also enables all other software to run on the computer.
- **DBMS Software:** It manages the database in the database system.
- **Application Programs and Utilities:** These are used to access and process the data stored in the database.

4. Personnel

The people related to the database system are called personnel. Different types of persons in a database system are as follows:

- **Database Administrators:** They manage DBMS's use and ensure that database is functioning properly. They are responsible of the whole database system. They authorize access to database and monitor the use and working of database system.
- **Application Programmers:** The application programmers write application programs to access data from database.
- **End Users:** End users are the people who use application programs to perform different tasks on database. They include clerks, managers and directors etc.

Q. Explain different objectives of the databases.

The main objectives of using databases are as follows:

1. Data Integration

The data in file system is stored in separate files. It is very difficult to access data stored in separate and independent files. An important objective of databases is to solve this problem. The data in database may be located at different computers physically but it is connected through data communication links. In this way, data appears centralized logically.

2. Data Integrity

Data integrity means the reliability and accuracy of data. Integrity rules are designed to keep the data consistent and correct. These rules act like a check on the incoming data. It is very important that a database maintains the quality of the data stored in it. DBMS provides several methods to enforce integrity of the data in a database.

Enforcing data integrity ensures the quality of data in the database. For example, if an employee ID is entered as "123", this value should not be entered again. The same ID should not be assigned to two or more employees.

3. Data Independence

Database approach provides the facility of **data independence**. It means that the data and the application programs are separate from each other. The user can change data storage structures and operations without changing the application programs. The user can also modify programs without reorganization of data.

Q. Define database model. Describe different types of database models.

A set of rules and standards that define how the database organizes data is called database model. It also defines how users view the organization of data.

Types of Database Models

There are three types of database models. These models are as follows:

1. Hierarchical Model

The **hierarchical model** arranges records in hierarchy like an organizational chart. Each record type in this model is called a **node** or **segment**. A node represents a particular entity. The top-most node is called **root**. Each node is a subordinate of the node that is at the next higher level. A higher level node is called **parent** and lower level node is called **child**. A parent node can have one or many child nodes. A child node can have only one parent node. This kind of structure is often called **inverted tree**.

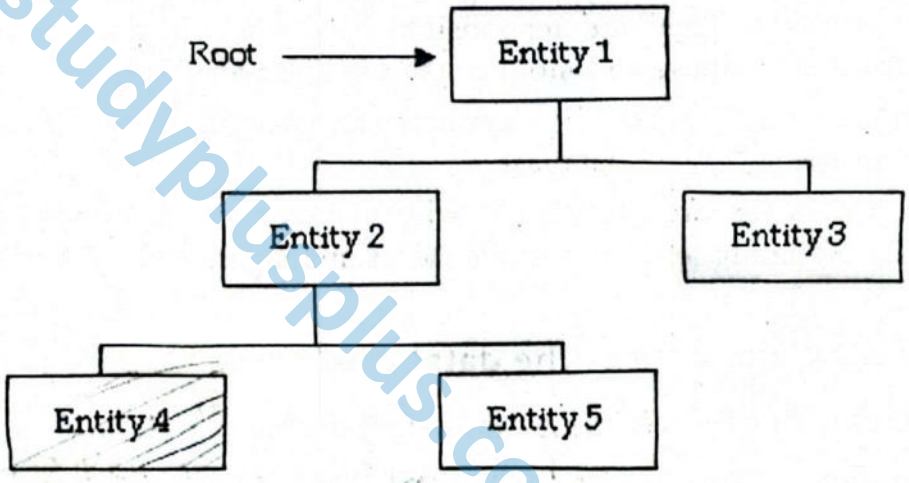

Figure 1.4: Hierarchical Model

2. Network Model

The **network model** is similar to hierarchical model. The difference is that child node can have more than one parent nodes. The child nodes are represented by arrows in network model. It requires more complex diagram to represent a database. It also provides more flexibility than hierarchical model.

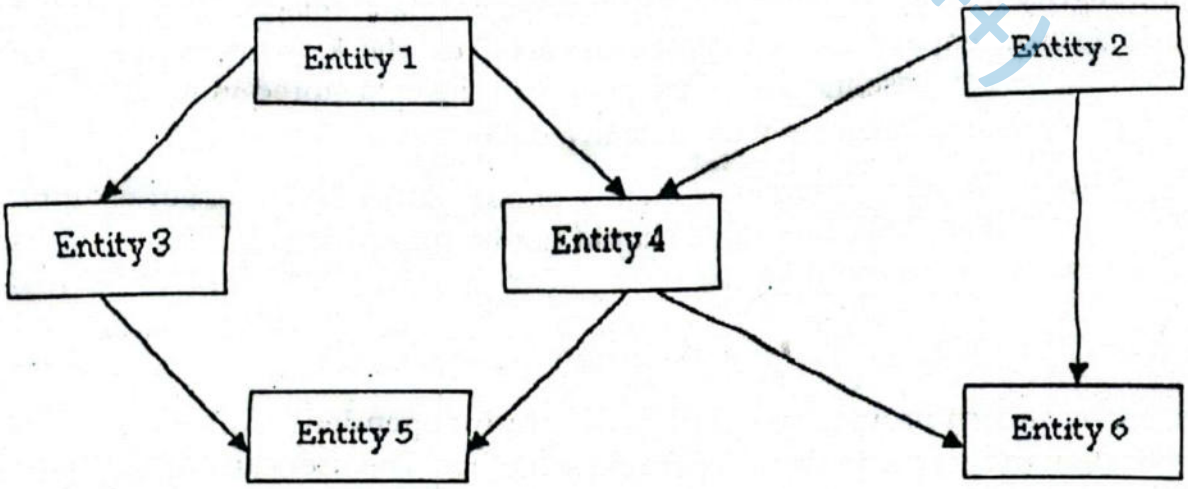

Figure 1.5: Network Model

3. Relational Model

Relation model is the most commonly used database model. It is more flexible than hierarchical and network database models. The relational model consists of simple relations. A relation is a term used for table. A relation represents a particular entity. It is used to store information about the entity. The relationships are based on the data of the entities. The relationship between entities is represented by the following diagram:

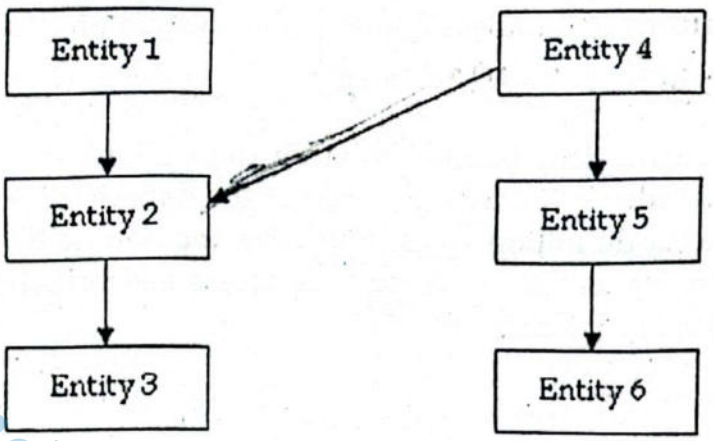

Figure 1.6: Relational Model

Q. What is database management system (DBMS)? Discuss its objectives.

Database Management System

A database management system (DBMS) is a collection of programs that are used to create, maintain and access database in a convenient and efficient manner. DBMS is an improvement over the traditional file management system. DBMS uses database manager software to control the overall structure of a database. DBMS is normally used by large or medium organizations for different purposes.

Objectives of Database Management System

Some important objectives of database management system are as follows:

1. Shareability

An ability to share data is a fundamental objective of database management system. Shareability means that data is shared by different people at the same time. Data is stored at a central place. Different users can access this data from different locations. It reduces storage cost and provides data consistency.

2. Availability

Availability means that the users must be able to access data and DBMS easily. The data should be available when and where it is required.

3. Evolvability

Evolvability means that DBMS should provide the facility to change the database due to increase in user requirements or change in the technology. The change in database may occur in two ways. It may change by contents or by structure. For example, the structure of the database may be changed if the operation of an organization is expanded.

4. Database Integrity

Database integrity ensures that data entered in the database is accurate and consistent. Database is normally shared among different users. Some measures must be maintained to ensure database integrity.

Q. Discuss different features of database management system.

Some important features of database management system are as follows:

1. Data Dictionary

Data dictionary is a file that is used to store data definitions or description of the structure of data used in database. It may also monitor the data that is entered. It ensures that data is according to the data definition rules. The rules include field names, field sizes and data types etc. Data dictionary is also used for data access authorization for database users. Data dictionary is also called repository.

2. Utilities

DBMS utilities are programs that are used to maintain database by processing data, records and files. Some utilities are used for backup and recovery procedures of databases.

3. Query Language

A query language is used to perform different operations on the databases. The most popular query language is SQL. It stands for Structure Query Language. The statements of SQL are written in simple English-like sentences. Some important commands of SQL are SELECT, DELETE, CREATE, MODIFY, UPDATE and INSERT etc. SQL can be used for the following purposes:

- Create table structures
- Enter data in tables
- Retrieve data from tables
- Update data in tables
- Delete data from tables etc.

4. Report Generator

A report generator is a program that is used to produce reports. It retrieves data from database and displays it in different formats. The user can use report generator to format page number, dates, titles and column headings etc. The users can produce useful and attractive reports by using report generator. Report generator is also called report writer.

5. Access Security

Access security refers to the protection of database from unauthorized access. The database management system provides several procedures to maintain data security. The security is maintained by allowing access to the database through the use of passwords.

6. Backup and Recovery

Database management system provides the facility for backup and recovery. Backup facility is used to store an additional copy of data. The data can be recovered from backup file if original file is lost or damaged.

Q. Briefly describe the advantages and disadvantages of DBMS.

Some important advantages of database management system are as follows:

1. Data Independence

Database approach provides the facility of **data independence**. It means that the data and the application programs are separate from each other. The user can change data storage structures and operations without changing the application programs. The user can also modify programs without reorganization of data.

2. Redundancy Control

Redundancy means duplication of data in multiple files. The data in database appears only once. It is not duplicated. If the data is required at many places, the same data is used at all places.

Suppose we want to manage the information of students who study in college and the student who live in hostel. The data of students is stored in college database. If a student also lives in hostel, his record will be stored only in college database. It will not be duplicated.

3. Consistency Constraints

Consistency constraints are the rules that must be followed to enter data in the database. If a data does not fulfill these constraints, it cannot be entered in the database.

Database management systems provide an easy way for applying different consistency constraints that ensure the consistency of the data in databases. For example, a constraint can be applied to ensure that the marks of a student are always between 0 and 100 etc.

4. Support Complex Data Relationships:—

Database management system allows the user to design complex data structures. It enables the user to logically view and access data in different ways.

5. Data Security Features

Data security is the protection of the database from unauthorized access. The database management system provides several procedures to maintain data security. The security is maintained by allowing access to the database through the use of passwords. Every user of database system should not be able to access all the data.

6. Database Backup & Recovery

Database management system provides the facility for backup and recovery. Backup facility is used to store an additional copy of data. The data can be recovered from this file if the original file is lost or damaged. A utility in DBMS is used to recover data. The utility rebuilds the data by using the backup copy of the data.

7. Advanced Capabilities

DBMS provides advanced access capabilities for online reporting. It uses Internet for advanced capabilities.

Disadvantages of Database Management System

Some important disadvantages of database management system are as follows:

1. High Cost of DBMS

A complete database management system is very large and complex software. It is expensive to purchase database management software.

2. Higher Hardware Cost

Database management system is complicated and heavy software. Additional memory and processing power is required to run the DBMS. It may require more powerful hardware.

3. Additional Training (User Training)

The application programmers require additional training for developing efficient programs. The untrained and less-experienced developers may develop inefficient database system. An inefficient system may create many problems for the users and the organization.

4. Problems in Case of Wrong Database Environment (SQL)

A database system may be changed later due to change in requirements. The change can be costly due to conversion and testing of existing programs. Hierarchical system is more difficult to change than other systems. It requires a lot of cost to implement changes. The changes in relational databases are easier to implement. It means that problems may increase later if wrong type of database environment is selected.

5. Need of Data Dictionary (Data Mining)

Another disadvantage of DBMS is the need of data dictionary. Data dictionary stores data definitions or description of the structure of data used in database. It ensures that data is according to the data definition rules. The rules include field names, field sizes and data types etc. Data dictionary is also used for data access authorization for database users.

Data dictionary is very useful tool but it is also expensive. It requires installation costs as well as hardware requirements.

Done ✓

Short Questions

Q.1. Define data.

Data is a collection of facts, figures and statistics related to an object. Data can be processed to create useful information.

Q.2. How is data useful?

Data is a valuable asset for an organization. It is very useful for generating reports, graphs and statistics etc.

Q.3. Define information.

The manipulated and processed form of data is called **information**. It is more meaningful than data. It is used for making decisions. Data is used as input for processing and information is the output of this processing.

Q.4. Differentiate between data and information.

Data is a set of raw facts and information is the processed form of data. Data is used as input in the computer and information is the output of the computer.

Q.5. What is data processing?

The process of manipulating data to achieve the required objectives and results is called **data processing**. The software is used to process raw data. The software converts raw data into meaningful information.

Q.6. List different activities in data processing.

Different activities involved in data processing are data capturing, data manipulation and managing output results.

Q.7. What do you mean by data manipulation?

The process of applying different operations on data is called **data manipulation**. It includes the operations of classifying, calculations, sorting and summarizing.

Q.8. Define field.

A **field** is a combination of one or more characters. It is the smallest unit of data that can be accessed by the user. Each field contains one specific piece of information For example, Employee_No, Designation and Salary are fields.

Q.9. Why is it important to specify data type and size of a field?

The data type of a field specifies the type of data that can be stored in the field. A **field size** defines the maximum number of characters that can be stored in a field.

Q.10. What is meant by a record?

A collection of related fields treated as a single unit is known as **record**. For example, a employee's record includes a set of fields about the employee such as Employer Number, Employee Name, Grade and Designation etc.

Q.11. What is a file?

A collection of related records treated as a single unit is known as **file**. Files are stored in disk like hard disk, CD-ROM or DVD-ROM etc. File is also know as data set

Q.12. List the file types from usage point of view.

- **Master File:** Master files are used to store the information that remains constant for a long period of time. → permant data store
- **Transaction File:** A type of file that is used to store input data before processing is called transaction file.
- **Backup File:** A type of file that is used to take backup important data is called backup file. It is used to store an additional copy of data.

Q.13. List the file types from function point of view.

The file types from function point view are program files and data files.

Q.14. Define program file.

A type of file that contains software instructions is called program file. The source program files and executable files are examples of program files.

Q.15. Define data file.

A file that contains data is called data file. It is created by the software being used.

Q.16. Define file organization.

A technique for physically arranging the records of a file on secondary storage devices is called file organization.

Q.17. Name different types of file organizations.

Different types of file organizations are sequential files, direct or random access files and indexed sequential files.

Q.18. How does the sequential file organization store files?

The records in sequential file organzation are stored in sequence. The sequence means the recrods are stored one after the other. The records can be retrieved only in the sequence in which they were stored. The principal storage media for sequential files is magnetic tape.

Q.19. How the records are stored in direct or random file organization?

Records in direct file organization are not stored in a particular sequence. A key value of a record is used to determine the location to store the record. Each record is accessed directly without going through the preceding records. This file organization is suitable for storing data on disk.

Q.20. Which file organization stores records in an order?

In indexed sequential file organization, records are stored in ascending or descending order based on a value called key. The index file organization maintains an index in a file.

Q.21. Is there any disadvantage of sequential access?

The major disadvantage of sequential access is that it is very slow. If the last record is to be retrieved, all preceding records are read before reaching the last record.

Q.22. Define database.

A database is a collection of logically related data sets or files. Each file may contain different type of information and are used for specific purposes. The files may be organized in different ways to meet different processing and retrieval requirements of the users.

Q.23. Which problems occur due to data redundancy?

Redundant data uses additional storage space and makes it difficult to maintain the accuracy of database when changes are made.

Q.24. What is a database system?

A database system is a collection of data as well as programs required to manage that data. It is a computerized record-keeping system.

Q.25. What is the main purpose of a database system?

The main purpose of database system is to maintain data and provide it to the user when it is required.

Q.26. Name four major components of database systems.

The four major components of database system are data, hardware, software and personnel.

Q.27. What are different objectives of the databases?

- **Data Integration:** It means that data is logically centralized even if it may be located at different locations.
- **Data Integrity:** Data integrity means the reliability and accuracy of data.
- **Data Independence:** It means that the data and the application programs are separate from each other.

Q.28. What is the purpose of database integrity?

Database integrity ensures that data entered in the database is accurate and consistent.

Q.29. Distinguish between data dependence and data independence.

In data dependence, the description of data is included in the application programs that use data. In data independence, the description of data is separated from the application programs.

Q.30. Write the functions of database application.

The functions of database application are creation and execution of queries, creation and execution of forms, creation and execution of reports and controlling application.

Q.31. Define DBMS.

A database management system is a collection of programs that are used to create, maintain and access database in a convenient and efficient manner.

Q.32. How is the structure of a database controlled?

Database management system uses database manager software to control the overall structure of a database.

Q.33. Write some important objectives of database management system.

Some important objectives of database management system are shareability, availability, evolvability and database integrity.

Q.34. Name different features of database management system.

Some important features of database management system are data dictionary, utilities, query language, report generator, access security, backup and recovery.

Q.35. Are there any rules to enter data in the database?

Consistency constraints are the rules that must be followed to enter data in the database. If a data does not fulfill these constraints, it cannot be entered in the database.

Q.36. Is data dictionary useful?

Data dictionary is used to store data definitions or description of the structure of data used in database. It may also monitor the data that is entered. It ensures that data is according to the data definition rules.

Q.37. Do you know the most popular query language?

The most popular query language is SQL. It stands for **Structure Query Language**. The statements of SQL are written in simple English-like sentences.

Q.38. What is the use of query language?

A query language is used to perform different operations on the databases.

Q.39. List some important purposes of query language.

- Create table structures
- Enter data in tables
- Retrieve data from tables
- Update data in tables
- Delete data from tables etc.

Q.40. Why is report generator used in database system?

A report generator is used to produce reports. It retrieves data from database and displays it in different formats. The report format can be specified in advance such as row heading, column heading and page heading etc.

Q.41. What is meant by access security?

Access security refers to the protection of database from unauthorized access. It may include password and access restrictions. DBMS provides the facility of access security.

Q.42. State the purpose of backup and recovery.

Backup is used to store a copy of important data in database. The recovery process uses the data in backup if the original data is damaged due to any reason.

Q.43. Name some advantages of DBMS.

Some important advantages of database management system are data independence, support of complex data relationships, data security, backup and recovery.

Q.44. What do you mean by the term data independence?

Data independence means that data and application programs are separate from each other. The user can change data storage structures and operations without changing the application programs. The user can also modify programs without reorganization of data.

Q.45. List three examples of database systems.

- A system that maintains component part details for a car manufacturer
- A training company keeping course information and participants' details
- An organization maintaining all sales order information

Q.46. Define data redundancy.

Data redundancy means the duplication of data in multiple files.

Q.47. Define data inconsistency.

Data inconsistency means two files may contain different data of the same entity.

Q.48. Differentiate between database and DBMS.

Database is a collection of related data. DBMS is a collection of program to create and maintain databases.

Q.49. Distinguish between file processing and database approach.

In file processing, data may be duplicated in different files that cause data redundancy. It is difficult to apply data integrity checks on files. The programs and data are inter-dependent. In database approach, data is not duplicated and appears only once. It provides many constraints for data integrity. The programs and data are independent of each other.

Q.50. What are the advantages of database approach over traditional file approach?

The advantages of database approach over traditional file approach include data consistency, data independence, data security & integrity and data sharing.

Q.51. What is the use of DDL?

DDL stands for Data Definition Language. It consists of SQL commands used to define a database, creating tables, indexes and views. Some important commands of DDL include CREATE/DROP TABLE, ALTER TABLE, CREATE/DROP VIEW etc.

Q.52. What is the use of DML?

DML stands for Data Manipulation Language. It consists of SQL commands used to load, update and query the database using SELECT command. INSERT, UPDATE and DELETE are some examples of DML commands.

Q.53. What is the use of DCL?

DCL stands for Data Control Language. It consists of SQL commands used to establish user access to the database. Some important commands of DCL are GRANT, ADD USER and REVOKE etc.

Q.54. Differentiate between data redundancy and data inconsistency.

Data redundancy means the duplication of data in multiple files. Data inconsistency means two files may contain different data of the same entity.

Multiple Choice

1. **A collection of raw facts and figure is called:**
   a. Data b. Information c. Processing d. None
2. **The manipulated and processed data is called:**
   a. Object b. Information c. Data d. None
3. **Manipulation of data to achieve the required objectives and result is called:**
   a. Data processing b. Operation c. a and b d. None
4. **Which of the following activities involved in data processing?**
   a. Data capturing b. Data Manipulation
   c. Managing output d. All
5. **The process of arranging data in a logical sequence is called:**
   a. Sorting b. Summarizing
   c. Data capturing d. Classifying
6. **Which of the following operations can be performed on data after the data has been captured and manipulated?**
   a. Storage and Retrieval b. Communication
   c. Reproduction d. All
7. **Storage and retrieval of data is related to:**
   a. Data capturing b. Data Manipulation
   c. Managing output result d. None
8. **A person's account, car, and house are considered:**
   a. Object b. Table c. Data processing d. None
9. **A collection of related fields is:**
   a. File b. Record c. Database d. None
10. **All records in a file have the same:**
    a. Contents b. Structure c. Both a and b d. None
11. **A collection of data that consists of name, address and email of a person is called:**
    a. Byte b. Record c. Character d. Field
12. **A record in a database is the information referring to a:**
    a. person b. Product c. Event d. All
13. **Each item of information within a record is called:**
    a. File b. Field c. Both a and b d. Byte
14. **A logical grouping of characters is a:**
    a. Field b. Record c. File d. All
15. **A field is to a record as:**
    a. Data are to files b. A column is to a row
    c. Files are to tables d. Attributes are to columns
16. **A database containing all students in a class would store basic data of students in:**
    a. Record c. Field d. Cell d. File
17. **A database containing all students in a class would store RollNo of a student in:**
    a. Record b. Field d. Cell d. File
18. **A database containing all students in a class would store the information of individual students in:**
    a. Record c. Field d. Cell d. File

19. Which of the following is also known as data set?
a. Record b. Field c. File d. All

20. A set of related files created and managed by a (DBMS) is called:
a. Field b. Record c. Database d. None

21. Which of the following is an example of a database?
a. Phone book b. Library catalog
c. Student records d. All

22. Which of the following represents a collection of concepts that are used to describe the structure of a database
a. Data warehouse b. Data model
c. Data structure d. Data type

23. Which of the following data model is more flexible?
a. Network model b. Hierarchical model
c. Relational model d. a and b

24. Which of the following type of file require largest processing time?
a. Sequential file b. Random file
c. Indexed sequential file d. Direct access file

25. Which of the following may be a temporary file?
a. Master file b. Transaction file
c. Backup file d. None

26. SQL is a(n):
a. Unstructured language b. Structured Language
c. Object oriented language d. Software

27. SQL stands for:
a. Sort-Query-List b. Self-Quantifying-Language
c. Seek-Qualify-Label d. None

28. SQL can be used to:
a. Create database structures only b. Query database data only
c. Modify database data only d. All

29. The type of files from usage point of view include:
a. Master file b. Transaction file
c. Backup file d. All

30. The type of files from functional point of view include:
a. Program files b. Data files c. a and b d. None

31. Which file is used to store information that remains constant for a long time?
a. Master file b. Transaction file c. Backup file d. None

32. Which of the following file is used to store additional copy of data?
a. Master file b. Transaction file c. Backup file d. None

33. Which of the following file contains the software instructions?
a. Program file b. Data file c. both a and b d. None

34. The extension of program file is:
a. .com b. .exe c. both a and b d. None

35. Which of the following is an image file extension?
a. .tif b. .jpg c. .bmp d. All

36. The extension for a video file is:
a. .avi b. .mpg c. .txt d. both a and b

37. **The extension of audio files is:**
a. .wav b. .mid c. Both a and b d. None

38. **The file name and the extension are separated by a:**
a. dot b. Semicolon c. Colon d. None

39. **In "program.txt", "txt" is the:**
a. File name. b. Directory
c. Extension d. File location

40. **Word processor can save documents using the following extensions EXCEPT:**
a. txt b. rtf c. doc d. mdb

41. **The type of files from storage point of view include:**
a. Sequential files b. Direct or Random file
c. Index sequential d. All

42. **The problem synonym occurs in:**
a. Sequential files b. Direct or random files
c. Indexed sequential files d. None

43. **The process in which the contents are accessed in sequence is called:**
a. Random Access b. Sequential Access-
c. Read-only Access d. None

44. **The file organization that is appropriate for magnetic tape storage media is:**
a. Sequential b. Random
c. Indexed sequential files d. None

45. **A database is an organized collection of _____ related data.**
a. Logically b. Physically c. Loosely d. Badly

46. **The objectives of database management systems include:**
a. Database Integrity b. Data Integration
c. Availability d. All

47. **The most commonly used database model in current database development is:**
a. Relational b. Hierarchical
c. Network d. Object-oriented

48. **Which of the following data model has general shape of an organization chart?**
a. Hierarchical Model b. Network Model
c. Relational Model d. Object Model

49. **Information sharing means that:**
a. Information can be stored once and retrieved any number of times.
b. The same information can be shared by different applications.
c. Both a and b
d. Neither a nor b

50. **Which of the following is handled by DBMS?**
a. Data integrity b. Data security
c. Data independence d. All

51. **A program whose job is to store and retrieve user data in the database is called:**
a. Database Modeling System
b. Database Management System
c. Data Business Model System
d. Data Business Management Service

52. **The database system is composed of four major parts:**
    a. Hardware, Hard drive, Monitor, Data, User
    b. Hardware, Software, People and Data.
    c. Software, You, Me, DBA, Client
    d. DBMS, Hardware, User, Programmer, Engineer
53. **In a database processing system:**
    a. The database application(s) interact with the DBMS
    b. The database application(s) access the database data
    c. The DBMS accesses the database data
    d. a and c
54. **Data that causes inconsistency lacks:**
    a. Good data
    b. Data integrity
    c. Data redundancy
    d. Data anomaly
55. **Which one of the following is an advantage of database management approach?**
    a. Programs are independent of data format
    b. Reduced security and the control of data
    c. Increased duplication of data.
    d. All
56. **DBMS stands for:**
    a. Database Modeling System
    b. Database Management System
    c. Data Business Model System
    d. Data Business Management Service
57. **Which of the following is NOT an advantage of database systems?**
    a. Redundant data
    b. Data independence
    c. Backup /Recovery
    d. Better data quality
58. **Which of the following enables the user to modify data structures without affecting existing programs that use them?**
    a. Data dependence
    b. Data independence
    c. Data integration
    d. Data relationships
59. **Database application contain procedures for:**
    a. Adding records
    b. Deleting records
    c. Processing queries
    d. All
60. **Which of the following is related to a component of DBMS known as personnel?**
    a. Application programmer
    b. End users
    c. Database administrator
    d. All
61. **Which of the following database models is often referred to as an Inverted Tree?**
    a. Hierarchical
    b. Network
    c. Relational
    d. Object-oriented
62. **The major component of DBMS is called:**
    a. Database Manager
    b. File Manager
    c. Data Manager
    d. All
63. **Duplicate data in multiple data files is:**
    a. Data redundancy
    b. Data multiplication
    c. Data Integrity
    d. None
64. **The description of structure and organization of data in a database is contained in:**
    a. Data dictionary
    b. Data mine
    c. Structured query language
    d. None

65. **A request for information from a database in database terminology is called:**
a. Report b. Letter c. Table d. Query

66. **A printed or onscreen display of data or information in the database is called a(n):**
a. Entity b. Report c. Query d. Screen

67. **The name for software to build reports that summarize data from a database is:**
a. Report writer b. Reporter
c. Report Builder d. Report Generator

68. **Which of the following is not a database management system?**
a. MS Access b. MS SQL c. Oracle d. None

Answers

1. a	2. b	3. c	4. d	5. a	6. d	7. c
8. a	9. b	10. b	11. b	12. d	13. b	14. a
15. b	16. d	17. b	18. a	19. c	20. c	21. d
22. b	23. c	24. a	25. b	26. b	27. d	28. d
29. d	30. c	31. a	32. c	33. a	34. c	35. d
36. d	37. c	38. a	39. c	40. d	41. d	42. b
43. b	44. a	45. a	46. d	47. a	48. a	49. c
50. d	51. b	52. b	53. d	54. b	55. a	56. b
57. a	58. b	59. d	60. d	61. a	62. a	63. a
64. a	65. d	66. b	67. d	68. b		

Fill in the Blanks

1. The manipulated and processed form of data is called_____.
2. A _____is a combination of one or more characters.
3. A collection of related fields treated as a single unit is known as_____.
4. A collection of related records treated as a single unit is known as_____.
5. A _____is a collection of logically related data sets or files.
6. The people related to the database system are called_____.
7. The _____has the general shape of an organizational chart.
8. Hierarchical structure is often called_____.
9. The _____consists of a collection of simple files/relations (Entities).
10. A_____ is a collection of programs that are used to create, maintain and access database in a convenient and efficient manner.
11. A file is a collection of_____.
12. Hierarchical data model has the general shape of a_____.
13. Data is a collection of _____,_____ and _____.
14. Before processing the data is recorded in _____.

Answers

1. Information	2. Field	3. Record
4. File	5. Database	6. Personnel
7. Hierarchical model	8. Inverted tree	9. Relational model
10. Database management system	11. records	12. Organizational chart
13. Facts, figures, statistics	14. Transaction files.	

True / False

1. Data can only be processed through computers.
2. The traditional file system approach has many advantages over DBMS approach.
3. Data dictionary is used to view the meaning of database terminology
4. Master file is the latest updated file that never becomes empty since it is created.
5. SQL is used to retrieve information from the database based on certain criteria.
6. Network data model is more popular and widely used than relational data model.
7. Indexed Sequential files can be processed sequentially as well as randomly.
8. Backup files store data prior to its processing.
9. Microsoft Access is a relational database management system
10. A report generator is used to produce a printed document from the database.
11. Another name for database program is DBMS.

Answers

1. F	2. F	3. F	4. T	5. T	6. F
7. T	8. F	9. T	10. T	11. T	