

# Experiment No.01

## PART A

(PART A: TO BE REFERRED BY STUDENTS)

**A.1 Aim:** Explore various fundamental concepts of DBMS and its future prospective.

1. Write briefly about the need of an effective DBMS in today's world of data explosion
2. List four applications you have used that most likely employed a database system to store persistent data
3. List four significant differences between a file-processing system and a DBMS.
4. In a particular organization, customer names and addresses are kept in one file for the sales department and another file for the billing department. What inconsistency might result from this redundancy?
5. Define the following terms: tuple, attributes, entity, query, schema
6. Describe at least 3 tables that might be used to store information in a social networking system such as Facebook.
7. List any three database software and compare them in terms of licensing, ease of use, scalability, performance, security.

**A.2 Prerequisite: -**

**A.3 Outcome:**

After successful completion of this experiment students will be able to

1. Explain fundamental terminologies of DBMS.
2. Enlist basic components of DBMS.
3. Describe the need of DBMS in software development.
4. Differentiate among various database packages and its usage in real life for solving the problems.
5. Enlist the real life applications of DBMS.
6. Correlate the usage of DBMS for prospective areas such as Data science.

**A.4 Theory:**

To define the database system, the following terminologies must be clarified:

1. File
2. Database
3. Database Management System (DBMS)

File consists of only one file or table, with each entry containing all the required data defined within it. It is like a cabinet containing only one folder which has many pages in it, each page containing all the information for that specific entry. This makes it easy for the user to know where to find requested entries and all the data associated with them. As entire data are at one

place, retrieval of data is easy and fast. But the flat file suffers from the problem of data redundancy. Suppose there is an order processing application which includes the name, address, and phone number of a customer. If this customer has placed order for two items then there could be two entries for this customer with his name, address and phone number. In other words, the file contains redundant data, i.e. twice the name, address, and the phone number of the customer. If his address is to be changed then it is to be changed at two places.

Database is a collection of interrelated files, while database management system is a collection of database, database utilities (DBMS software) operated by user groups and administered by database administrator.

Database is a collection of interrelated files. The conventional file processing system will lead to excessive data redundancy and associated difficulties. But, in a database system which has an integrated single database, all the problems of the conventional file processing system will be minimized, even though the time taken to design and implement a robust database will be relatively longer when compared to the design of files in the conventional file processing system. There will be greater payoff by using the robust database which will offset the cost towards longer duration of designing and implementing the database.

A database is a collection of interrelated data stored together with controlled redundancy to serve one or more applications in an optimal way. The data are stored in such a way that they are independent of the programs used by the people for accessing the data. The approach used in adding the new data, modifying and retrieving the existing data from the database is common and controlled one.

It is also defined as a collection of logically related data stored together that is designed to meet information requirements of an organization. We can also define it as an electronic filing system.

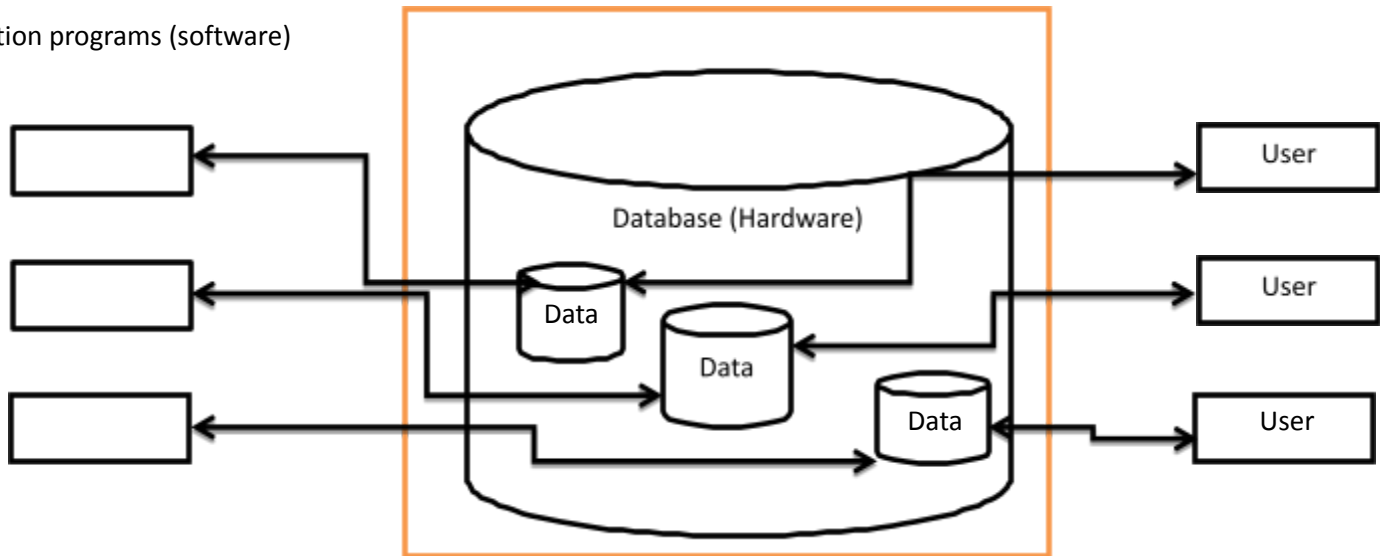
A database management system is a software package that controls all access to database. When a request to access database comes, it examines the request interprets it using some commands. There are many different types of database users, out of which database administrator is responsible for managing overall system and all other users of the system.

The DBMS performs the following five primary functions:

1. Define, create and organize a database: The DBMS establishes the logical relationships among different data elements in a database.
2. Input data: It performs the function of entering the data into the database through an input device with the help of the user.
3. Process data: It performs of manipulation and processing of the data stored in the database.
4. Maintain data integrity and security: It allows limited access of the database to authorized users to maintain data integrity and security.
5. Query database: It provides information to the decision makers that they need to make important decisions. The information is provided by querying the database using SQL.

## Database system environment:

Application programs (software)



## DBMS Users:

The users of a database system are:

1. End users or Naïve Users
2. Online users
3. Application programmers
4. Database administrator

## Role of Database Administrator:

1. Deciding contents of database
2. Deciding structure of internal storage of database
3. Deciding users
4. Granting authorities
5. Deciding rules and constraints
6. Managing security for database
7. Monitoring the performance of the database
8. Taking backup
9. Removal of dump and maintain free space

## PART B

(PART B: TO BE  
COMPLETED BY  
STUDENTS)

**(Students must submit the soft copy as per following segments within two hours of the practical. The soft copy must be uploaded on the portal provided by the instructor)**

Roll No. I066	Name: Srihari Thyagarajan
Program : B Tech AI	Division: I
Batch: B3	Date of Experiment: 15/ 07/ 2022
Date of Submission: 15/ 07/ 2022	Grade :

### B.1 Tasks given in PART A to be completed here

Solutions:

1. A DataBase Management System (DBMS), is a program that manages details related to storage and access for a database.  
The necessity of DBMS in today's world:
  - a. Creation of Databases.
  - b. Maintaining Data.
  - c. Efficient retrieval of data from the database.
  - d. Helps in specifying access and pose restrictions to view the database.
2. Applications which effectively use the DBMS:
  - a. Restaurant Management System (Grofers, Dunzo) - storing food items of restaurants and customer information.
  - b. Food ordering apps (eg. Swiggy, Zomato, etc.) - storing customer information.
  - c. Airplane/ Railways industries (eg, Indigo, GoAir apps) - storing passenger details

3.

File Processing System	DBMS
Less Secure.	More Secure.
Results in Data Redundancy.	Reduces Data Redundancy.
Has multiple formats for which different software are required.	Has standard file format, easy to open and access data from the respective databases.
Less structures/ organized.	More organized which helps in efficient retrieval of data from the database.

4. Data Redundancy occurs when the same data is stored in multiple files/ places in different formats. This results in Data Inconsistency, which leads to false/ in-efficient retrieval of data. Updating the data in certain files does not update the same data in other files/ formats.

From the question, if the address of the customers are updated in the Billing department file, the same data which exists in the Sales department file will not be updated.

5.

a. Tuples - The Rows in the database are called tuples.

b. Attributes - The Columns in the database are called tuples.

c. Entity - is an object that exists and is separate from other objects (eg. a person, company, etc).

d. Query - is a statement which requests data from the database, for viewing, modification, or manipulation of data.

e. Schema - is the logical structure of the database (how data is organized in the database). It's categories include Physical Schema, Logical Schema, View Schema.

6. Examples of three tables

- Customer Login Information - Social media platforms like Facebook, store the customer login information such as Usernames, Email IDs and passwords (security questions).
- User profile information - Name, D.O.B, Address, Profile picture, posts, etc. are stored in databases to recognize/ identify the user ( unique identity).
- Pages/ people followed - Likes and Dislikes of users can be stored in various tables/ databases to help improve their feed recommendation.

7.

	MySQL	Postgre SQL	MongoDB
Licensing	GNU	Open Source	SSPL
Ease of Use	Simple	Moderate to Hard, configuration, setup and support/ resources available.	Easy to learn, similar to MySQL.
Scalability	Vertically scalable, configuration is difficult when tables get too large.	Vertically scalable, fast with large databases as well.	Horizontal (sharding)
Performance	Moderate, gets slow with very large databases.	Good, gets slow with very large databases.	Good, relatively faster with large databases as well.
Security	Provides basic security tools for robust data encryption.	Provides good security, more secure than MySQL.	Moderate security.

## B.2 Observations and Learning:

From the experiment, we learnt:

- Difference between File system and DBMS, pros of DBMS compared to File System.
- The various terminologies and components of DBMS
- Various DBMS software,
- The differences between the DBMS software
- Real life applications of DBMS.

## B.3 Conclusion:

*(Students must write the conclusive statements as per the attainment of individual outcomes listed above and learning/observation noted in section B.2)*

Learning the pros and cons of DBMS, the necessity of it in storage systems in day-to-day activities, terminologies relevant to DBMS, software relevant in the industry.

#### **B.4 Question of curiosity:**

Justify the following statements: “*The knowledge of DBMS is essential for becoming the Data scientist* ”

Justification -

The role of a Data Scientist encompasses collecting, analyzing and interpreting large amounts of data. It requires one to develop hypotheses, make inferences, and analyze customer and market trends based on the data collected. Hence, handling of data is a crucial part of this job and hence the knowledge of DBMS and the various software used in this industry such as MySQL, Postgre SQL, MongoDB, etc. is necessary.

Data scientists are also responsible for setting best practices for collecting data, using analysis tools and interpreting data.

\*\*\*\*\*