

### KEY CONCEPTS

■ Data Base ■ Data Structure ■ Data Base Management System ■ Data Mining ■ Data Warehousing

### Learning Objectives

#### To understand:

- The concept of Database with basics of Data Structure
- Database Abstraction, Implementation and Creation
- DBMS- Users of DBMS and its applications
- The purpose of Data Warehousing and Data Mining
- The usefulness of Data warehousing and Data mining in context to a large retail firm

### Lesson Outline

- Data Base Concepts
- Data Structure
- Data Base Management System
- Data Base Files
- Data Mining and Warehousing
- Lesson Round-Up
- Test Yourself

## DATABASE CONCEPTS

Database concepts are vital to understand how databases function, and databases are a crucial element in modern computer systems. The aim of databases, data abstraction, data models, database management systems, data integrity, data security, and the benefits and drawbacks of databases are just a few of the topics that we'll cover in this chapter.

### Purpose of Database

A database's goal is to have a collection of data serve as many applications as it can. As a result, a database is frequently considered to be a collection of data required to carry out specific tasks inside a business or organisation. It would enable not just data retrieval but also on going data alteration, which is necessary for operation control. You might be able to search the database to get answers to your inquiries or details for planning.

### Database Abstraction

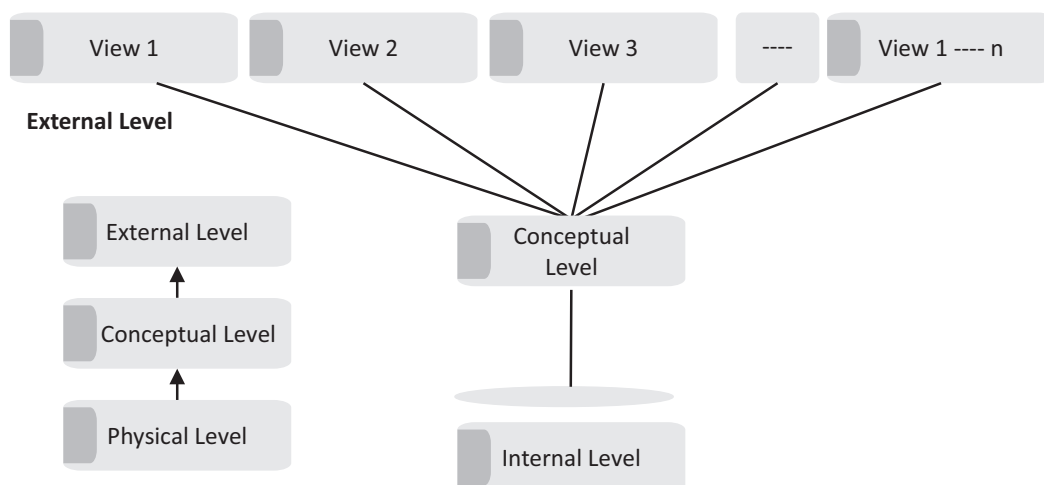
Providing the user with only the information they need is one of a database's main goals. This indicates that the system hides some information about how the data is maintained and kept rather than disclosing all the information about the data. They are shielded from the complexity of databases, which, if necessary, are arranged across a number of layers of abstraction to ease their engagement with the system. Three layers are used to implement the various database levels:

**Physical Level (Internal Level):** Internal level, the lowest level of abstraction, is most analogous to physical storage. It explains the specifics of how the data is kept on the storage media.

**Conceptual Level:** The data that is specifically saved in the database is described at this level of abstraction. Moreover, it describes the connections between the data. At this level, databases are logically explained using straightforward data structures. Users at this level are not concerned with the actual implementation of these logical data structures.

**View Level (External Level):** It relates to how specific users see the data and is the level that is closest to the users.

Each modification at one level can have an impact on plans at other levels since a database can be seen through three different levels of abstraction. Databases may occasionally undergo numerous modifications as they continue to expand. The database shouldn't need to be redesigned and implemented as a result of this. The idea of data independence is useful in this situation.



**Data Models:** A database's data and connections are conceptually represented by a data model. It offers a blueprint that details how information is set up, kept, and used in a database. Hierarchical, network, and relational models are the three basic types of data models.

**Hierarchical Model:** The hierarchical approach organises data in a tree-like form, with a parent-child connection between each item. No of how many kids a parent has, each child can only have one parent. Organizational charts and other types of data with a natural hierarchical structure are best suited for the hierarchical approach.

**Network Model:** As an extension of the hierarchical model, the network model enables records to have numerous parent-child connections. It is helpful for modelling data when a record may include several parent and child connections since it may express complicated relationships.

**Relational Model:** The model that is being utilised the most is the relational model. Data is arranged in tables with rows and columns, where each row denotes a distinct record and each column denotes a field or attribute. A large variety of data kinds and connections may be handled by the relational model, which is adaptable.

**Database Management System (DBMS):** Software that enables users to build, edit, and administer databases is known as a Database Management System (DBMS). To manage the storing, retrieval, and modification of data in a database, it offers a variety of features and capabilities.

**Data Definition Language (DDL)** is used to create and edit database schema. Data Manipulation Language (DML) is used to query and modify data. Transaction Control Language (TCL) is used to govern how transactions are handled.

**Data Integrity:** Any database must maintain data integrity. It speaks of the precision, thoroughness, and consistency of information in a database. By imposing restrictions and guidelines on the data in a database, a DBMS protects data integrity.

**Constraints:** The structure and connections between tables are established by constraints, which are rules. They make sure that linkages between tables are upheld and that data is not duplicated. Each row in a table is guaranteed to be unique by the primary key constraint, and each record in one table is guaranteed to have a matching record in another table by the foreign key constraint.

**Rules:** Data values must fall inside particular ranges or satisfy certain requirements, which is why rules like check constraints are used. A check constraint, for instance, can make sure that a person's birthdate is not later than the present day.

**Data Security:** A crucial idea in databases is data security. It speaks of safeguarding data from unlawful access, usage, disclosure, or obliteration. To guarantee data security, a DBMS offers a variety of security capabilities, including authentication, authorisation, and encryption.

**Authentication:** It includes confirming the user's identity, usually using a username and password.

**Authorization:** It regulates each user's degree of access. It entails allowing or refusing access to particular database sections.

**Encryption:** Data is protected by encryption by being transformed into a secret code that can only be cracked with a special key. It guarantees that the data cannot be read even if an unauthorised person obtains access to the database.

### Advantages of Database

Let's look at some of the advantages offered by database systems and how they solve the issues described above: -

- Significantly reduces database redundancy.
- The database has strong control over inconsistent data.
- Data exchange is made easier by the database.
- Databases impose rules.
- Data security may be ensured via the database.
- Databases may be used to ensure integrity.

Database systems are therefore recommended for systems with superior performance and efficiency.

### Disadvantages of Database

There may be various issues that arise as a result of the database system's need to do sophisticated tasks; these issues can be referred to as the database system's drawbacks. They include:

- Without effective controls, security might be jeopardised.
- Without effective controls, integrity might be jeopardised.
- Further hardware could be needed.
- Possible performance overhead is considerable.

The mechanism is probably complicated.

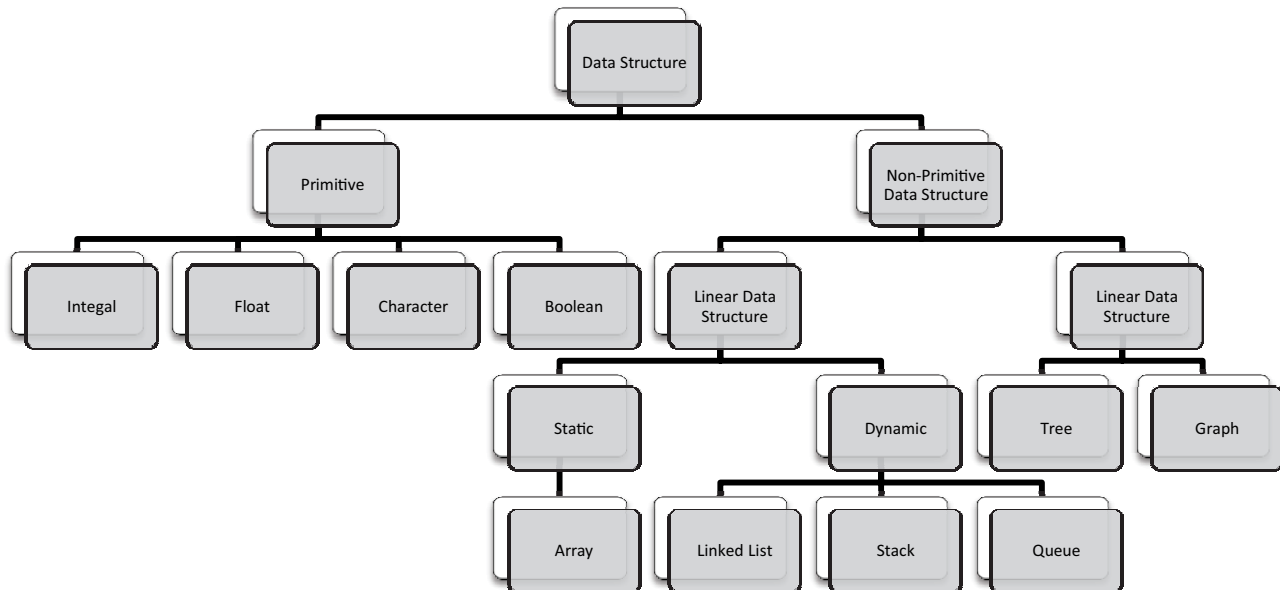
Therefore, understanding database principles is essential to comprehending how databases operate. Anybody dealing with databases has to understand key concepts including data models, DBMS, data integrity, and data security. To guarantee efficient and successful data administration, it is crucial to keep current with the most recent database ideas and technology. Databases may offer a strong foundation for organising and analysing massive volumes of data if the correct skills and resources are applied.

## DATA STRUCTURE

A fundamental idea in computer science, data structures are crucial for effective programming. It is a method of setting up and keeping track of data in a computer so that it may be easily retrieved and used. We will explore the numerous facets of data structures in this chapter, including their types, operations, and implementation.

## Types of Data Structures

The two types of data structures are Primitive and Non-Primitive data structures.



### Primitive Data Structures

Primitive data structures are the character and number-based data structures that are pre-built into programmes. Machine-level instructions might directly alter or use these data structures.

Primitive Data Structures include fundamental data types including Integer, Float, Character, and Boolean.

Due to the fact that they include characters that cannot be further subdivided, these data types are also known as simple data types.

### Non-Primitive Data Structures

Data structures that are descended from primitive data structures are known as non-primitive data structures. Machine-level instructions cannot directly alter or use these data structures.

These data structures' main goal is to create a collection of data pieces that are either homogenous (all of the same kind) or heterogeneous (different data types).

These data structures may be divided into two sub-categories: linear data structures and non-linear data structures, based on the organisation and structure of the data.

**Linear Data Structures:** Each component in a linear data structure is related to the element before it and the element after it in a consecutive way. The most popular linear data structure types are queues, stacks, linked lists, and arrays.

**Arrays:** A distinct index is given to each element, starting at 0. Although they can't be resized in-place, arrays are fixed in size and offer quick access to their elements.

**Linked Lists:** A linked list is made up of a group of objects called nodes, each of which has a reference to the

node after it as well as data. Linked lists make it simple to add and remove components and can be used to construct dynamic data structures.

**Stacks:** It only permits adding and deleting pieces from the top of the stack, which is one end.

**Queues:** The First-In-First-Out (FIFO) principle governs the operation of queues, which are a type of data structure. It enables the addition of elements at the back and the removal of parts from the front.

**Non-Linear Data Structures:** The organisation of data in non-linear data structures is not sequential. Instead, they show intricate connections between different data points. The most popular non-linear data structure types are heaps, trees, and graphs.

**Tress:** Trees are hierarchical data structures in which each node contains zero to more than one offspring and at most one parent. In hierarchical data representations like file systems, organisational charts, and database indexes, trees are employed.

**Graphs:** Graphs are collections of nodes connected by connecting edges. Graphs are used to describe intricate connections between data items, including those seen in social, transportation, and communication networks.

**Heaps:** A unique kind of binary tree that meets the heap condition is called a heap. The element with the greatest (or lowest) priority will always be at the bottom of the heap according to the heap attribute.

### Key Features of Data Structures

Here are some of the Important Characteristics of these Data Structures:

**Robustness:** In theory, all computer programmers strive to create software that produces accurate results for all conceivable inputs and performs effectively across all hardware architectures. Both legitimate and erroneous inputs must be managed by this powerful programme.

**Adaptability:** The development of large-scale software systems is necessary for many years in order for programmes like web browsers, word processors, and internet search engines to function correctly and efficiently. Software also changes as a result of new technology or dynamic market situations.

**Reusability:** Adaptability and Reusability are complementary qualities. It is common knowledge that a programmer need several resources to create any software, making it an expensive endeavour. Yet, if the software is created in a way that makes it adaptive and reusable, it can be used in the majority of future applications. Consequently, it is feasible to create reusable software that looks to be both time and cost-effective by using excellent data structures.

**Operations on Data Structures:** Data structures support a number of operations, including insertion, deletion, traversal, and searching. A data structure gets a new element by insertion.

**Deletion:** Deletion involves removing a data structure piece. A data structure is traversed by going through each element in turn. A data structure element is located using search.

### Implementation of Data Structures:

The use of arrays, linked lists, and pointers can be used to build data structures. While they have a fixed size, arrays are the most straightforward data structure to build. Linked lists provide dynamic data structures but demand more memory for pointer storage. Although pointers are an effective tool for building data structures, they need for a deep understanding of memory management.

Therefore, it can be concluded that a fundamental idea in computer science, data structures are crucial for effective programming. Data structures may be divided into two categories: linear and non-linear. Although non-linear data structures show intricate interactions between data pieces, linear data structures arrange data in a

sequential manner. The numerous operations supported by data structures include insertion, deletion, traversal, and searching. These can be put into practise utilising pointers, linked lists, and arrays. Data structures may offer a potent platform for data organisation and manipulation with the correct skills and resources.

## DATABASE MANAGEMENT SYSTEM

A Database Management System (DBMS) is a piece of software that controls how data is stored, organised, and retrieved from databases. It is a crucial tool for successfully and efficiently handling massive amounts of data. We shall talk about a DBMS's architecture, parts, and features in this chapter, among other things.

### DBMS Architecture:

There are three levels of a DBMS structure: external, conceptual and internal.

**External Level:** It specifies the actions that may be carried out on the data as well as how the data is displayed to the user. The application software that accesses the database is of interest at the external level.

**Conceptual Level:** The database is logically represented at the conceptual level. It describes the connections between the various data pieces as well as the general structure of the database. The conceptual level is unrelated to how the data is physically stored.

**Internal Level:** The database's physical representation is at the internal level. It specifies how information is kept and retrieved on the actual storage hardware.

### DBMS Components:

Data Definition Language (DDL), Data Manipulation Language (DML), query language, and data dictionary are the four parts of a Database Management System (DBMS).

**Data Definition Language (DDL):** DDL, or Data Definition Language, is used to specify the database's structure, including the tables, columns, and connections between them. Also, it outlines the restrictions that apply to database data storage.

**Data Manipulation Language (DML):** The database's data can be changed using the Data Manipulation Language (DML). Operations like insertion, deletion, modification, and retrieval of data are included.

**Query Language:** Data is collected from the database using query language. The database library standard language, SQL (Structured Query Language), is included.

**Data Dictionary:** A database that contains metadata about the data in the database is called a data dictionary. It contains details about data kinds, connections between data items, and access rights.

**DBMS Features:** A DBMS has a number of capabilities that make it an indispensable tool for handling massive amounts of data. The following are some of a DBMS's important characteristics:

**Data Security:** To prevent unauthorised access to data, a DBMS offers a variety of security options. Access control, authentication, and encryption are all part of it.

**Data Integrity:** By imposing constraints like uniqueness, referential integrity, and domain integrity, a DBMS protects the accuracy of the data.

**Concurrency Control:** Concurrency control is a feature of DBMSs that permits concurrent entry to the database without compromising with other users' operations.

**Backup and Recovery:** In the case of a system breakdown, a DBMS offers backup and recovery options to avoid losing data. It has capabilities like recovery and transaction recording.

**Scalability:** A DBMS offers scalability characteristics to meet an organization's expanding data demands. It has capabilities like replication and segmentation.

### DBMS Users

Users with varying access and permissions use a typical DBMS for a variety of objectives. Some people back up their data while others retrieve it. A DBMS's users may be roughly divided into three groups: administrators, designers and end users.

**Administrator:** Administrators are in charge of managing the database and maintaining the DBMS. They are in charge of monitoring how it should be utilised and by whom. In order to enforce security and preserve isolation, they construct access profiles for users and set restrictions. Moreover, administrators take care of DBMS resources including system licences, necessary equipment, and other upkeep-related software and hardware.

**Designers:** A company's DBA staff includes database designers. The individuals who really work on the designing portion of the database are known as designers. They are in charge of comprehending end customers' database requirements. They closely monitor what information needs to be stored and in what format. The whole collection of entities, relations, constraints, and views are identified and designed by them.

**End Users:** End users are the ones who genuinely profit from a DBMS. End users might be as basic as viewers who focus on the logs or market prices or as complex as business analysts. They can also be knowledgeable experts who access data by directly querying the database using a DBMS's query interface. Those that need to use data from a database to power sophisticated applications like knowledge bases, decision systems, intelligent systems, data analytics, or the storage and administration of big data might be considered to be another group of end users.



### Advantages of DBMS

**Controls database redundancy:** Because it keeps all the data in a single database file and stores the recorded data in the database, it is able to manage data redundancy.

**Data sharing:** A DBMS enables an organization's authorised users to distribute data across several users.

**Easy Maintenance:** Because the database system is centralised, it may be simple to maintain.

**Time savings:** It lessens the requirement for maintenance and development time.

**Backup:** It has subsystems for backup and recovery that automatically back up data in the event of hardware or software problems and restore the data as needed.

**Multiple User Interface:** It offers a variety of user interfaces, including graphical user interfaces and application programme interfaces.

### Disadvantages of DBMS

**Cost of Hardware and Software:** To operate DBMS software, a fast data processor and plenty of memory are needed.

**Size:** To operate them effectively, it takes up a lot of memory and storage space.

**Complexity:** The database system adds more requirements and complexity.

**Higher Impact of Failure:** Failure has a greater impact on databases since, in the majority of organisations, all data is kept in a single database. If the database is destroyed due to an electrical failure or database corruption, all data may be permanently lost.

### DATABASE FILES

Data is kept in files in a Database Management System (DBMS). A database file is a group of connected records, each of which has one or more data fields.

The capacity to organise data consistently and logically is one of the main benefits of utilising database files. Intuitive and effective data access and manipulation are made possible for users as a result. In a Relational Database Management System (RDBMS), for instance, data is arranged into tables with columns and rows, making it simple to query for and retrieve certain bits of data. Contrarily, document-oriented databases arrange data as documents with nested fields and arrays, which makes it simpler to manage unstructured, complicated data.

The capacity to maintain data consistency is a key benefit of database files. Database files avoid data duplication or discrepancies by specifying a set of guidelines for how data should be kept and arranged. By ensuring that data is correct and current, this lowers mistakes and enhances the ability to make decisions.

Database files offer better data security in addition to organised data management and data consistency enforcement. Since databases allow for restricted access to data, managers can limit user access to sensitive data. They also provide user authentication and data encryption services to stop unwanted access and security breaches.

Despite the many benefits of database files, there are a few drawbacks to take into account. The expense of developing and maintaining a database is one of the biggest obstacles. Database systems may be pricey, especially for individuals or small organisations. However, setting up and managing databases requires specific technical skills, which can be difficult for someone without a background in database administration.

The systems' intricacy is another possible issue with database files. The setup and configuration of databases may be difficult and time-consuming, requiring extensive planning and design effort. Moreover, they are prone to data corruption and system failures, both of which can result in the loss of crucial information. We will go through the different kinds of database files, how they are organised, and how to access and edit them in this chapter.

### Types of Database Files

Database files come in a variety of forms, each having unique properties and applications. Database files often come in one of these formats:

**Relational Data:** Data is stored in tables with rows and columns in relational files. Each column denotes a field of data, and each row denotes a single record. The relationships between the tables are used to arrange the relational file.

**Hierarchal Files:** A tree-like structure is used to arrange data in hierarchical files. One or more parent records and one or more child records are connected to each record. The relationships between the records are used to structure the records in the hierarchical structure.

**Network Files:** A record may have more than one parent using network files, which are equivalent to hierarchical files. This makes it possible for records to have more intricate associations.

**Object-Oriental Files:** Files that store data as objects with attributes and methods are said to be object-oriented files. The properties of each object stand in for a single record, and the fields of data they represent.

#### **Organization of Database Files:**

To speed up access times and maximise storage space, database files are structured into a variety of forms. These are the most typical database file organisation structures:

**Sequential Files:** Records are kept in a sequential file in a certain order. Data are appended to the file's end and accessed one at a time starting at the starting.

**Indexed Files:** An index is made in an indexed file that associates each entry with a different identifier. This makes it possible to quickly retrieve particular files' records.

**Hashed Files:** Datasets are assigned to specific positions in a hashed file according to a hashing algorithm. This enables quick access to particular records using a search key.

### **Access and Manipulation of Database Files**

The DBMS offers a number of procedures to access and work with the database files. The most frequent operations include:

**Insert:** New records are added to the database file using the insert procedure.

**Delete:** Remove records from the database file using the delete procedure.

**Update:** Modify existing records in the database file using the update operation.

**Retrieve:** To obtain records from the database file, utilise the retrieve procedure.

**Search:** Using a search key, the search method locates particular records in the database file.

Therefore, it can be concluded that any database management system must include database files. They do it in a way that makes it easy to access and manipulate the data they hold. There are several database file formats and organisational patterns, each with certain benefits and drawbacks. Insert, delete, update, retrieve, and search are just a few of the operations that may be used to access and manipulate database files. Organizations may handle their data efficiently and base judgements on the data by comprehending the various database file formats and their organisational structures.

## **DATA MINING AND WAREHOUSING**

Organizations acquire enormous volumes of data daily in today's data-driven environment. Yet, this information is meaningless without a thorough study. Organizations may gain useful insights from their data using two methods: data mining and data warehousing. The ideas of data mining and data warehousing, their distinctions, and their applications in businesses will all be covered below.

### **Data Mining**

The procedure of extracting patterns, trends, and insights from huge databases is known as data mining. By seeing patterns and connections between the data, it includes drawing information from the data. To glean insights from data, data mining employs a variety of methods including machine learning, artificial intelligence, and statistical analysis. Finding hidden patterns and insights that may be utilised to make wise judgements is the main objective of data mining.

### Some key characteristics of data mining include

**Large datasets:** Large datasets with potentially millions or billions of records are frequently analysed using data mining techniques. To manage the data's complexity and volume, specialist tools and approaches are needed.

**Complex data structures:** Data mining frequently works with non-tabular data types including text, photos, and audio. Techniques for assessing unstructured and semi-structured data are necessary for this.

**Multidimensional analysis:** Analyzing data across several dimensions, such as time, place, and client demographics, is known as data mining. As a result, patterns and insights that might not be seen from a single-dimensional perspective of the data might be found.

**Machine learning:** Machine learning techniques are frequently used in data mining to find patterns in the data. These algorithms are capable of automatically learning from the data and progressively improving.

**Predictive modeling:** Building predictive models that may foretell future trends and behaviours through data mining is feasible. These models may be applied to improve corporate operations and make well-informed judgements.

**Data visualization:** Data visualisation methods are frequently used in data mining to portray the data in a form that is simple to grasp and analyse. This makes it simpler for analysts to find patterns and connections in the data.

**Iterative process:** Data processing, model development, and assessment are all steps in the iterative process of data mining. Once the desired outcomes are obtained, this procedure may be repeated several times.

### Data Warehousing

The procedure of gathering, arranging, and storing data from numerous sources in a main repository is known as data warehousing. A data warehouse's organisation of the data makes for effective querying and analysis. Data extraction from diverse sources, including databases, applications, and external sources, is followed by data transformation into a standard format and transferring into a data warehouse.

Some key characteristics of data modelling include:

**Abstraction:** In order to make data simpler to comprehend and handle, data modelling entails building a simplified representation of the real world. This entails defining and portraying the essential entities, connections, and features of the system being represented in a form that is simple to understand.

**Structure:** The categories of data, the connections between them, and the limitations that apply to them are all defined by data modelling, which also identifies the structure of data in a system. With the help of this structure, data may be stored, organised, queried, and subjected to analysis.

**Flexibility:** A reliable data model is adaptable enough to take into account system changes. This indicates that the model should have the flexibility to adapt as the needs of the system do.

**Clarity:** To ensure that the data modelling simple for others to comprehend and utilise, it should be clear and unambiguous. To do this, it is necessary to define the connections and constraints of the data being represented in detail as well as to use standardised notations and language.

**Data integrity:** Data consistency, accuracy, and completeness should be ensured by data modelling. To do this, it is necessary to specify the guidelines and limitations that apply to the data and to apply them by utilizing data validation and error-checking.

**Scalability:** Efficient data modelling will be able to handle massive amounts of data and allow sophisticated searches and analyses. This necessitates building the system to handle massive datasets and performance-enhancing data structure and indexing.

**Iterative process:** As new information becomes available, data modelling is an iterative process that involves adjusting and rewriting the model. To make sure that the model effectively reflects the demands of the system, collaboration is required amongst stakeholders, including developers, analysts, and end users.

### Differences between Data Mining and Data Warehousing

Focus is the main distinction between data mining and data warehousing. Data warehousing focuses on storing and organising data in a way that makes efficient querying and analysis possible whereas data mining focuses on analysing data to find patterns and insights.

Data is analysed using data mining to find hidden patterns and insights that may be utilised to guide decisions. Data warehousing, on the other hand, is a method for storing data in a form that makes efficient querying and analysis possible.

Whereas data warehousing is a systematic process that involves gathering, organising, and storing data, data mining is an exploratory process that involves finding patterns and insights in data.

### Uses of Data Mining and Data Warehousing:

Several businesses employ data mining and data warehousing to glean insightful information from the data. Data warehousing and mining are frequently used for the following purposes:

**Marketing:** To find trends and insights in consumer data, data mining is utilised in marketing. With this data, tailored marketing efforts with a higher chance of success may be created.

**Healthcare:** In the healthcare industry, data mining is used to examine patient data and spot trends that may be utilised to better patient outcomes.

**Retail:** Data collection from several sources, including point-of-sale systems, inventory management systems, and customer loyalty programmes, is done in the retail industry by means of data warehousing. With this information, personalised marketing campaigns are created as well as inventory management and pricing optimisation.

**Finance:** Finance uses data mining and data warehousing to examine market data, spot patterns, and make wise investment decisions.

Data warehousing and data mining are crucial methods that let businesses get the most out of their data. Data warehousing is gathering, organising, and storing data in a way that makes efficient querying and analysis possible. Data mining entails analysing data to find patterns and insights. These methods are applied across a range of sectors to enhance decision-making, streamline operations, and create focused marketing efforts. Organizations may efficiently manage their data and make choices based on the insights gleaned from their data by comprehending the principles of data mining and data warehousing.

### LESSON ROUND-UP

- Any collection of facts, numbers, or statistics that are gathered and examined with a specific goal is data. Data comes in a variety of forms, including semi-structured, unstructured, and structured data. Unstructured data does not adhere to a set format, whereas structured data does. Both structured and unstructured data components can be found in semi-structured data.

- Data structures describe how information is arranged and kept in computer systems. Arrays, linked lists, stacks, queues, and trees are the most popular types of data structures. The type of data and the operations that must be carried out on it determine which data structure should be used.
- A database management system, or DBMS, is a piece of software that controls how data is stored, retrieved, and modified. The DBMS conducts functions like data backup, recovery, and security while also giving users a way to interact with the database through an interface. MySQL, Oracle, SQL Server, and PostgreSQL are a few popular database management system.
- The data kept in a database is physically represented by database files. These documents are arranged in tables with rows and columns. Each column is a data element, and each row denotes a distinct record. The files may be kept on a hard drive or other types of storage.
- The method of obtaining valuable information from huge databases is known as data mining. To find patterns, correlations, and trends in the data, statistical and machine learning approaches are used. Data mining findings may be applied to forecasting, predictions, and decision-making.
- The procedure of gathering, storing, and managing data from many sources in a single repository is known as data warehousing. The relationships across various data pieces are reflected in the schema that was used to arrange the data. Organizations can rapidly and effectively access and analyse vast volumes of data thanks to data warehousing. Business intelligence, reporting, and analytics frequently use it.

#### TEST YOURSELF

*(These are meant for recapitulation only. Answer to these questions are not to be submitted for evaluation.)*

1. What is a database? Explain the different levels of database.
2. What are some common data structures used in computer systems?
3. What is a Database Management System (DBMS)? What are the different categories of Database Management Users?
4. What is a database file? How is data organized within a database file?
5. What is data mining? How can it be used in decision-making?
6. What is data warehousing? What are the characteristics of data warehousing?
7. What is the difference between:
  - a) Data Mining and Data Warehousing
  - b) Administrator versus End User in a Database Management System.
8. How can data warehousing be used for different purposes?

