11

DATABASE CONCEPTS

A database is a collection of logically related information in an organised way so that it can be easily accessed, managed and updated.

Some other operations can also be performed on database such as adding, updating and deleting data.

Fundamentals of Database

1. **Data** These are raw and unorganised facts that need to be processed such as digital representation of text, numbers, graphical images or sound.

   e.g. A student's test score is one piece of data.

2. **Information** When data is processed, organised, structured or presented in a given context to make it useful or meaningful, it is called information.

   e.g. The class's average score is the information that can be concluded from the given data.

Types of Database

1. **Network Database** In this type of database, data is represented as a collection of records and relationships among data that are represented as links.

2. **Hierarchical Database** In this type of database, data is organised in the form of tree with nodes. Nodes are connected *via* links.

3. **Relational Database** This database is also known as **structured database** in which data is stored in the form of tables. Where, columns define the type of data stored in the table and rows define the information about the data.

Components of Database

A database consists of several different components. Each component listed, is called an **object**.

Database components are described below

1. **Tables** These are the building blocks or relation of any relational database model where all the actual data is defined and entered.

   Different types of operation are done on the tables such as storing, filtering, retrieving and editing of data. Tables consist of cells at the intersection of records (rows) and fields (columns), *which are described below*

   (i) **Field** It is an area (within the record) reserved for a specific piece of data.

   e.g. Customer number, Customer name, Street address, City, State, Phone number, Current address, etc. Field of a table is also known as **column**.

(ii) **Record** It is the collection of data items of all the fields pertaining to one entity, i.e. a person, company, transition, etc. Record of a table is also known as **row** or a **tuple** and the number of records in a relation is called the **cardinality** of that relation.

2. **Queries** These are basically questions based on the data available in a database. A query consists of specifications indicating which fields, records and summaries a user wants to fetch from the database.

   Queries allow you to extract data based on the criteria that you define.

3. **Forms** Although you can enter and modify data in datasheet view of tables but you neither control the user's action very well nor you can do much to facilitate the data-entry process. To overcome this problem, forms are introduced.

   Like tables, forms can be used to view and edit your data. However, forms are typically used to view the data in an underlying table, one record at a time.

   e.g. A user can create a data entry form that looks exactly like a paper form. People generally prefer to enter data into a well-designed form, rather than a table.

4. **Reports** When you want to print those records which are fetched from your database, design a report. Access even has a **wizard** to help produce mailing labels.

Database Management System (DBMS)

A DBMS is a collection of inter-related data and a set of programs to retrieve data from the database.

It is an organised collection of data viewed as a whole, instead of a group of separate unrelated files.

The primary goal of DBMS is to provide an environment that is both convenient and efficient for user to store and retrieve database information.

e.g. MySQL, Oracle, FoxPro, dBASE, SyBase,

MS-Access. The purpose of database management system is to bridge the gap between information and data.

The basic processes that are supported by DBMS are as follows

(i) Specification of data types, structures and constraints to be considered in an application.

(ii) Storing the data.

(iii) Manipulation of the database.

(iv) Querying the database to retrieve desired information.

(v) Updating the content of the database.

Architecture of DBMS

The architecture of DBMS is divided into three levels are as follows

1. **Internal Level** It is the lowest level of data abstraction that deals with the physical representation of the database on the computer. It is also known as **physical level**.

   It defines how the data is actually stored and organised in the storage medium.
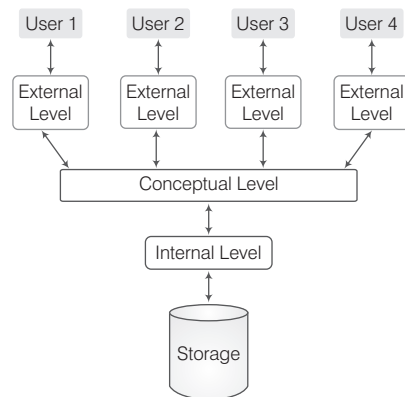
Architecture of DBMS

2. **Conceptual Level** It is the overall view of the database and includes all the information that is going to be represented in the database.

   It describes what type of data is stored in the database, the relationship among the data without effecting to the physical level. It is also known as **logical level**.

3. **External Level** This is the highest level of data abstraction which describes the interaction between the user and the system.

It permits the users to access data in a way that is customised according to their needs, so that the same data can be seen by different users in different ways, at the same time. It is also known as **view level**.

Advantages of DBMS

1. **Reduction in Data Redundancy** The duplication of data refers to data redundancy. DBMS cannot make separate copies of the same data. All the data is kept at a place and different applications refer to data from centrally controlled system.

2. **Better Interaction with Users** In DBMS, the availability of uptodate information improves the data to be access or respond as per user requests.

3. **Improvement in Data Security** DBMS can allow the means of access to the database through the authorised channels.

To ensure security, DBMS provides security tools, i.e. username and password.

4. **Maintenance of Data Integrity** Data integrity ensures that the data of database is accurate. In DBMS, data is centralised and used by many users at a time, it is essential to enforce integrity controls.

5. **Backup and Recovery** The DBMS provides backup and recovery sub-system that is responsible to recover data from hardware and software failures.

Disadvantages of DBMS

1. **Cost of Hardware and Software** A processor with high speed of data processing and memory of large size is required to run the DBMS software. It means that you have to upgrade the hardware used for file based system. Similarly, database software is also very costly.

2. **Complexity** The provision of the functionality that is expected from a good DBMS makes the DBMS an extremely complex piece of software. Failure to understand the system can lead to bad design decisions, which can have serious consequences for an organisation.

3. **Cost of Staff Training** Mostly DBMSs are often complex systems, so the training for user to use the database is required. The organisation has to pay a lot of amount for the training of staff to run the DBMS.

4. **Appointing Technical Staff** The trained technical persons such as database administrator, application programmers, etc., are required to handle the database. You have to pay a lot of amount to these persons. Therefore, the system cost increases.

5. **Database Failure** In most of the organisations, all data is integrated into a single database. If database is corrupted due to power failure or it is corrupted on the storage media, then our valuable data may be lost or whole system stop.

Applications of DBMS

1. **Banking** For customer information, accounts, loans and other banking transactions.

2. **Reservation** For reservation and schedule information.

3. **Universities** For student information, course registration, grades, etc.

4. **Credit Card Transaction** For purchase of credit cards and generation of monthly statements.

5. **Tele-communication** For keeping records of calls made, generating monthly bill, etc.

6. **Finance** For storing information about holdings, sales and purchase of financial statements.

7. **Sales** For customer, product and purchase information.

Relational Database

In a relational database, data is stored in different tables with relationships to each other. In the case of relational database, a Relational Database Management System (RDBMS) performs this task.

An important feature of this database system is that a single database can be spread across several tables.

e.g. Base, Oracle, DB2, SyBase, Informix, etc.

Terms Related to Relational Database

1. **Relation** It is a table with columns and rows which represent the data items and relationships among them. It has three important properties, a name, cardinality and a degree.

   These properties help us to further define and describe relations

   (i) **Name** The first property of a relation is its name, which is represented by the tide or the entity identifier.

   (ii) **Cardinality** The second property of a relation is its cardinality, which refers to the number of tuples (rows) in a relation.

   (iii) **Degree** The third property of a relation is its degree, which refers to the number of attributes (columns) in each tuple.

2. **Domain** It is a collection of all possible values from which the values for a given column or an attribute is drawn. A domain is said to be atomic, if elements are considered to be indivisible units.

3. **Attributes** The heading columns of a table are known as **attributes**. Each attribute of a table has a distinct name.

4. **Tuples** The rows in a relation are also known as **tuples**. Each row or tuple has a set of permitted values for each attribute.

Key

A key is defined as the column or set of columns in a table that is used to identify either row of data in a table or establish relationship with another table.

If a table has Id, name and address as the column names, then each one is known as the **key** for that table. The keys are also used to uniquely identify each record in the database table.

Types of Key

1. **Primary Key** It is a set of one or more attributes that can uniquely identify tuples (rows) within the relation. The primary key should be chosen in such a way, i.e. its value must not be changed.

   There should not be duplicacy in the record of primary key. Primary key can be atomic or composite. The field chosen as primary key, cannot accept null value.

2. **Candidate Key** The set of all attributes which can uniquely identify each tuple (row) of a relation, is known as **candidate key**. Each table may have one or more candidate keys and one of them will become the primary key. The candidate key of a relation is always a minimal key.

3. **Alternate Key** From the set of candidate keys after selecting one of the keys as primary key, all other remaining keys are known as alternate keys.

4. **Foreign Key** It is a non-key attribute whose value is derived from the primary key of the same or some another table. The relationship between two tables is established with the help of foreign key.

   A table may have multiple foreign keys and each foreign key can have a different referenced table.

Database Languages

1. **Data Definition Language** (DDL) It is used to define structure of your tables and other objects in database. In DBMS, it is used to specify a database schema as a set of definitions.

2. **Data Manipulation Language** (DML) It provides various commands used to access and manipulate data in existing database. This manipulation involves inserting data into database tables, retrieving existing data,

deleting data from existing tables and modifying existing data.

3. **Data Control Language** (DCL) These commands are used to assign security levels in database which involves multiple user setups. They are used to grant defined role and access privileges to the users.

Entity-Relationship Model (E-R Model)

It represents the entities contained in the database. It is a diagrammatically representation of entities and relationship between them. It is also known as **E-R Diagram**.

E-R Diagram
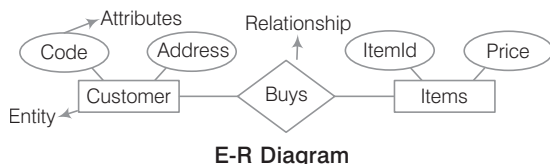
Some terms related to E-R Model are described below

Entity

It is an object that has its existence in the real world. It includes all those things about which the data are collected. "Entities are represented in rectangles". e.g. Customer buys items, it means Customer and Items are entities.

Attributes

It describes the characteristics or properties of entity. In tables, attributes are represented by columns. Attributes are drawn in elliptical shapes. e.g. Items entity may contain ItemId and Price.

Entity Set

It is a set of entities of the same type that shares same properties or attributes. e.g. Students are an entity set of all student entities in the database.

Entity set is of two types which are as follows

1. **Strong Entity Set** It has a primary key or can be easily distinguishable each attribute.

2. **Weak Entity Set** It does not posses sufficient attributes to form a primary key.

Relationship

It is an association among several entities. A relationship describes how two or more entities are related to each other. It is represented by diamond shape.

Relationship can be divided into three parts

(i) **One-to-One** This relationship tells us that a single record in Table A is related to a single record in Table B and *vice-versa.*

(ii) **One-to-Many** This entails one data in Table A to have links to multiple data in Table B. However, a single data in Table B, will have link to a single data in Table A.

(iii) **Many-to-Many** Each data in Table A is linked to all the data in Table B and *vice-versa.*

🔺 Tit-Bits

- Dr. EF Codd represented 12 rules for Relational Database Management System (RDBMS) in 1970.
- **Schema** is a logical structure of the database.
- **Instances** are the actual data contained in the database at a particular point of time.
- **Data Duplication** wastes the space, but also promotes a more serious problem called data inconsistency.
- **Data Mining** is the process of sorting through large data sets to identify patterns and establish relationships to solve problems through data analysis.

QUESTION BANK

1. A .... is a collection of data that is stored electronically as a series of records in a table.
(1) Spreadsheet   (2) Presentation
(3) Database   (4) Web page

2. A collection of interrelated records is called a   **[RBI Grade B 2012]**
(1) utility file
(2) management information system
(3) database
(4) spreadsheet
(5) datasheet

3. Which of the following is the organised collection of large amount of interrelated data stored in a meaningful way used for manipulation and updation?
     [IBPS Clerk Mains 2017]
(1) Database   (2) File
(3) Folder   (4) Data-mining
(5) None of these

4. Items such as names and addresses are considered as
(1) input   (2) data
(3) output   (4) records

5. Which type of database, organised the data in the form of tree with nodes?
(1) Network Database
(2) Hierarchical Database
(3) Relational Database
(4) Multiple Database

6. The database stores information in
(1) rows and columns
(2) blocks
(3) tracks and sectors
(4) All of the above

7. To locate a data item for storage is
(1) field   (2) feed
(3) database   (4) fetch

8. Devices that could be used to input data into a database are
(1) keyboard, fax roller ball
(2) mouse, keyboard, monitor
(3) mouse, keyboard, touch screen
(4) All of the above

9. In a relational database, a data structure that organises the information about a single topic into rows and columns, is
(1) block   (2) record   (3) tuple   (4) table

10. The smallest unit of information about a record in a database is called a
(1) cell   (2) field   (3) record   (4) query

11. .... are distinct items that do not have much meaning to you in a given context.
     [SBI PO 2012]
(1) Fields   (2) Data
(3) Queries   (4) Properties
(5) None of these

12. A collection of related files is called a
     [SBI PO 2011, IBPS Clerk 2013]
(1) character   (2) field
(3) database   (4) record
(5) None of these

13. Which of the following contains information about a single 'entity' in the database like a person, place, event or thing?
(1) Query   (2) Form
(3) Record   (4) Table

14. DBMS is comprised of tables that made up of rows called .... and columns called ....
(1) fields, records   (2) records, fields
(3) address, fields   (4) ranges, sheet

15. What is a stored question about information in a database?   **[SBI Clerk 2015]**
(1) Query   (2) Sort
(3) Report   (4) Record
(5) Field

16. A program that generally has more user-friendly interface than a DBMS is called a
(1) front end   (2) repository
(3) back end   (4) form

17. Which of the following object(s) is/are contained in database?
(1) Table   (2) Query
(3) Form   (4) All of these

18. Which of the following places the common data elements in order from smallest to largest?
(1) Character, File, Record, Field, Database
(2) Character, Record, Field, File, Database
(3) Character, Field, Record, File, Database
(4) Bit, Byte, Character, Record, Field, File, Database

19. What is the overall term for creating, editing, formatting, storing, retrieving a text document?   **[IBPS PO 2012]**
(1) Word processing   (2) Spreadsheet design
(3) Web design   (4) Database management
(5) Presentation generation

20. The database administrator's function in an organisation is
(1) to be responsible for the technical aspects of managing the information contained in organisational databases
(2) to be responsible for the executive level aspects of decision regarding the information management
(3) to show the relationship among entity classes in a data warehouse
(4) to define which data mining tools must be used to extract data

21. The code that relational database management systems use to perform their database task is referred to as
(1) QBE   (2) SQL
(3) OLAP   (4) Sequel Server

22. DBMS helps to achieve
(1) data independency
(2) centralised control of data
(3) selection of data
(4) Both (1) and (2)

23. Which out of the following is not a DBMS software?
(1) dBASE   (2) FoxPro
(3) Oracle   (4) Database 2000

24. In which, the database can be restored up to the last consistent state after the system failure?
(1) Backup   (2) Recovery
(3) Redundancy   (4) Security

25. .... provides total solutions to reduce data redundancy, inconsistency, dependency and unauthorised access of data.
(1) DBMS   **[IBPS Clerk 2012]**
(2) Tables
(3) Database
(4) Protection password
(5) Centralisation of data

26. Periodically adding, changing and deleting file records is called .... file.
(1) updating   (2) upgrading
(3) restructuring   (4) renewing

27. Architecture of database management can be viewed as
(1) two levels   (2) four levels
(3) three levels   (4) one level

28. A collection of conceptual tools for describing data, relationships, semantics and constraints is referred to as
          [IBPS Clerk 2012]
(1) E-R model   (2) database
(3) data model   (4) DBMS
(5) None of these

29. .... is one reason for problems of data integrity.   **[IBPS Clerk 2012]**
(1) Data availability constraints
(2) Data inconsistency
(3) Security constraints
(4) Unauthorised access of data
(5) Data redundancy

30. .... means that the data contained in a database is accurate and reliable.
(1) Data redundancy
(2) Data integrity
(3) Data reliability
(4) Data consistency

31. Which of the following contains data descriptions and defines the name, data type and length of each field in the database?
(1) Data dictionary
(2) Data table
(3) Data record
(4) Data filed

32. An advantage of the database management approach is
(1) data is dependent on programs
(2) data redundancy increases
(3) data is integrated and can be accessed by multiple programs
(4) All of the above

33. Which of the following is the drawback of DBMS?
(1) Improvement in data
(2) Backup and recovery
(3) Complexity
(4) Maintenance of data integrity

34. In which of the following, database is used?
(1) Banking
(2) Finance
(3) Sales
(4) All of these

35. A database that contains tables linked by common fields is called a
(1) centralised database
(2) flat file database
(3) relational database
(4) All of the above

36. Oracle is a(n) **[IBPS Clerk Mains 2017]**
(1) hardware
(2) high level language
(3) operating system
(4) system software
(5) RDBMS

37. The cardinality property of a relation, refers to the
(1) number of database
(2) number of columns
(3) number of rows
(4) number of tables

38. Rows of a relation are called
(1) relation
(2) tuples
(3) data structure
(4) entities

39. A set of possible data values is called
(1) attribute
(2) degree
(3) tuple
(4) domain

40. The purpose of the primary key in a database is to **[IBPS Clerk 2015]**
(1) unlock the database
(2) provide a map of the data
(3) uniquely identify a record
(4) establish constraints on database operations
(5) None of the above

41. In case of entity integrity, the primary key may be
(1) not null
(2) null
(3) Both (1) and (2)
(4) any value

42. In files, there is a key associated with each record which is used to differentiate among different records. For every file, there is atleast one set of keys that is unique. Such a key is called
(1) unique key
(2) prime attribute
(3) index key
(4) primary key

43. Which of the following types of table constraints will prevent the entry of duplicate rows?
(1) Primary key
(2) Unique
(3) Null
(4) Foreign key

44. The particular field of a record that uniquely identifies each record is called the **[SBI PO 2012]**
(1) key field
(2) primary field
(3) master field
(4) order field
(5) None of these

45. .... is a primary key of one file that also appears in another file. **[IBPS Clerk 2013]**
(1) Physical key
(2) Primary key
(3) Foreign key
(4) Logical key
(5) None of these

46. .... is an invalid type of database key.
(1) Structured primary key **[IBPS Clerk 2013]**
(2) Atomic primary key
(3) Primary key
(4) Composite primary key
(5) None of the above

47. Key to represent relationship between tables is called **[SBI Clerk 2010]**
(1) primary key
(2) secondary key
(3) foreign key
(4) composite key
(5) None of these

48. Which database language is used to access data in existing database?
(1) DDL   (2) DML
(3) DCL   (4) None of these

49. An E-R diagram is a graphic method of presenting   **[IBPS Clerk 2011]**
(1) primary keys and their relationships
(2) primary keys and their relationships to instances
(3) entity classes and their relationships
(4) entity classes and their relationships to primary keys
(5) None of the above

50. In an E-R diagram, an entity set is represented by
(1) rectangle   (2) square
(3) ellipse   (4) triangle

51. In an E-R diagram, attributes are represented by
(1) rectangle   (2) square
(3) ellipse   (4) circle

52. In E-R diagram, relationship type is represented by   **[IBPS Clerk 2012]**
(1) ellipse   (2) dashed ellipse
(3) rectangle   (4) diamond
(5) None of these

53. An entity set that does not have sufficient attributes to form a primary key, is a
          [IBPS Clerk 2011]
(1) strong entity set   (2) weak entity set
(3) simple entity set   (4) primary entity set
(5) None of these

54. Relationship can be divided into
(1) One-to-one
(2) Many-to-one
(3) One-to-many
(4) All of the above

55. Dr. E F Codd represented .... rules that a database must obey if it has to be considered truly relational.   **[IBPS Clerk 2012]**
(1) 10   (2) 8
(3) 12   (4) 6
(5) 5

56. A logical schema
(1) is the entire database
(2) is a standard way of organising information into accessible part
(3) describes how data is actually stored on disk
(4) All of the above

57. Data duplication wastes the space, but also promotes a more serious problem called
(1) isolated   **[IBPS PO 2015]**
(2) data inconsistency
(3) other than those given as options
(4) program dependency
(5) separated data

58. When data changes in multiple lists and all lists are not updated. This causes
          [RBI Grade B 2012]
(1) Data redundancy
(2) Information overload
(3) Duplicate data
(4) Data consistency
(5) Data inconsistency

ANSWERS

1. *(3)*	2. *(3)*	3. *(1)*	4. *(2)*	5. *(2)*	6. *(1)*	7. *(4)*	8. *(3)*	9. *(4)*	10. *(2)*
11. *(1)*	12. *(4)*	13. *(3)*	14. *(2)*	15. *(1)*	16. *(4)*	17. *(4)*	18. *(3)*	19. *(4)*	20. *(1)*
21. *(2)*	22. *(4)*	23. *(4)*	24. *(2)*	25. *(4)*	26. *(1)*	27. *(3)*	28. *(3)*	29. *(1)*	30. *(2)*
31. *(1)*	32. *(3)*	33. *(3)*	34. *(4)*	35. *(3)*	36. *(5)*	37. *(3)*	38. *(2)*	39. *(4)*	40. *(3)*
41. *(1)*	42. *(4)*	43. *(1)*	44. *(2)*	45. *(3)*	46. *(1)*	47. *(3)*	48. *(2)*	49. *(3)*	50. *(1)*
51. *(3)*	52. *(4)*	53. *(2)*	54. *(4)*	55. *(3)*	56. *(2)*	57. *(2)*	58. *(5)*		