

|  |
| --- |
| A" textbooks Of the Punjab Curriculum and Textbook Board carry a triangular shaped security sticker on the title page. The sticker exhibits a unique colour shift due to change of angle (orange to green) in the logo of Punjab Curriculum and Textbook Board. Lahore. The name Of this organization can seen around the logo in bright red colour. Moreover. if you Scratch protective coating on the white strips both sides below With a coin. it will reveal the registered trade name "PCTB"\_ This is a hallmark of the authenticity of original textbook Of the Punjab Curriculum and Textbook Board. Look for the security sticker while purchasing the  there is no sticker on the title page or it has been tampered with do not buy such a book. |



A Textbook Of



Computer Science

for class

12

Punjab Curriculum and Textbook Baord, Lahore



All rights are reserved with the Punjab Curriculum & Textbook Board, Lahore. Prepared by: Punjab Curriculum & Textbook Board. Lahore.

 Approved by: Federal Ministry of Education, Curriculum Wing, Islamabad. Vide its Letter No . E.I-10/200S-Maths (Comp. sc.) Dated July 29, 2006.

* Syed Zulqurnain Jaffery

Assistant Professor

Institute of lnrormatio•n Technology Sector 11-8, Islamabad, 

* MS. Shahina Naz



Head or Department Computer Science  IslanE1bad Model College for Girls F-10/2. IslaÆ1abad

* Mr. Asif Ali Magsi

I.ecturer (Computer Science) -College for Boys, Islamabad.

Editor



Mirza Mubåshar Baig

Research Associate 

Department of Computer Science

Lahore University Of Management Sciences (LUMS), I ,ahore.

Supervision: Muhammad Akhtar Shirani

Subject Specialist,

Punjab Curriculum & Textbook Board

Pre-Press: Syed Mubeen mehmond & Co. Urdu Bai9r, Lahore.

Process: LG Scan

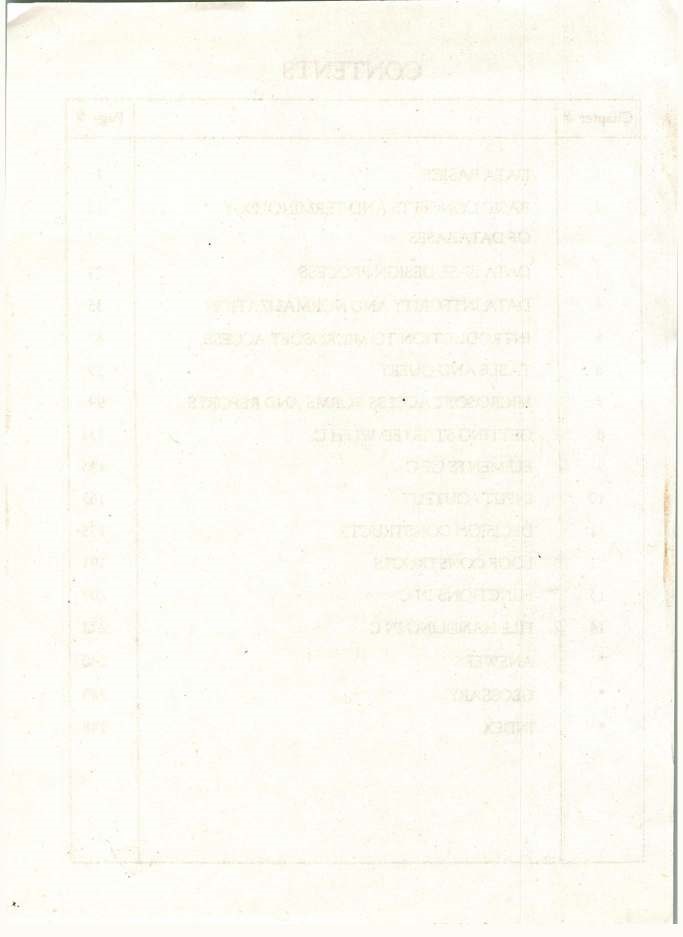
Published by: Frontier Publishing Company I-hr. Printer : Ahmad Tayyab Printer Sagian Lahore.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Date of Printin | Edition | 1m |  | n | No. of Co • s | Price |

May 2016 1st 11 10,000 110.00

CONTENTS

|  |  |  |
| --- | --- | --- |
| Chapter # |  | Page |
| 3  4  5  6  7    8  9  -10  11    12  13 | DATA BASICS    BASIC CONCEPTS AND TEu1NOLOGY  OF DATABASES        DATA BASE DESIGN PROCESS  DATA INTEGRITY AND NORMALIZATION    INTRODUCTION TO MICROSOFT ACCESS  TABLE AND QUERY  MICROSOFT ACCESS-FORMS AND REPORTS  GETING STARTED wrrH C  ELEMENTS OF C    INPUT / OUTPt-rr  DECISION CONSTRUCTS  LOOP CONSTRUCTS    FUNCTIONS IN C  FILE HANDLING IN C    ANSWERS  GLOSSARY    INDEX | 13    23  35  47  59  131  145  163  175  193    207    240  243  248 |



|  |  |  |
| --- | --- | --- |
|  |  | Chaptei |

1.1 OVERVIEW

Data: Data is a collection of åcts, figures and statistics — related to an object, that can be processed to produce a meaningfill information. In an organization, data is an asset that enables the managers to perform an effective and successful operation of management. It gives view of past activities (rise and fall) and enables to make better decisions for future. It is also useful for generating some useful reports, graphs, statistics etc,

Information: The rnanipulated and processed data is called information

e.go the percentage of students results. It refers to all the facts, figures or statistics that are precisely meaningful to the people. So by definition, it is an output Of a certain process,

Operations: Manipulation of data (after capturing from different sources)  to achieve the required objectives and results is called operation. For this purpose, a software (program) is used to process raw data which is converted  to meaningful information. Thus, effectively, a series of actions/operations are performed on raw data to achieve some output or result.



These are-categorized into three basic activities :

 Data Capturing: Data must be recorded or captured in some form before it Can be processed. Data may first be recorded On source documents or given directly through input devices.

Data Manipulation: The following may then be performed on the gathered data.

 ClassiWing: Organizing data into classes /grOups. Items may be assigned predetermined codes, they can be numeric, alphabetic or

 Calculations: Arithmetic manipulation ofthe data.

•;• Sorting• Data is arranged in logical sequence (numerically or alphabeticall').

 Summarizing: Masses Of data are reduced to a more concise and usable form.

Managing the Output Results: Once the data is captured and manipulated, it may be :

 Storing and Retrieval : Data is retained for future reference. Accessing / fetching the stored data and/or information is the Retrieve activity.

 Communication and Reproduction: Data may be transferred from one location or operation to another, for further processing. It is sometimes necessary to copy or to rnake duplicate of data. This activity is called Reproduction.

1.2 TRADITIONAL FILE SYSTEM

Record: A collection of related fields (facts about something) treated as single unit is called a record. Let us assume an employee's biographic information in a bank.

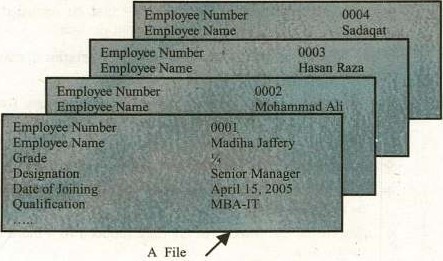


A

"Record"

As it belongs to one employee of the bank, so it is an individual employee's record (Of biographic information).

File: A collection Of related records treated as a single unit is called a file or a data set. If we collect records (as shown below) of all the employees, it becomes a file (bio-information) of all the employees of the bank .



Files are categorized according to different criteria as discussed below :

File Types



+ Master File : These are the latest updated files which never become empty, ever since they are created. They rnaintain information that remains constant over a long period of time. Whenever the information changes in files/records, it is updated.

Methods of updating are adding, deleting or editing records in a

 Transaction File: These are those files in which data prior to the stage of processing is recorded. It may be temporary file, retained till the master file is updated. It may•also be used to maintain a permanent record of transaction data.



of

sc

n

+ Backup File: These are again permanent files and their purpose is the protection of vital files/data of an organization by creating them using some specific software utilities.

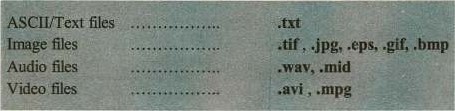


For this purpose, the files are given appropriate names, consisting of two parts i.e., file name and file extension, having a dot "." in between. Normally, the extension is given by the software being used at the time of initial såve. These files are summarized as under:

Program files: These files contain the software instructions i,e. source program files and executable files. The source program files may have the extension as .com and the executable files as . exe.

•9 Data files: These are the files that contain data and are created by the software being used. A few of them are given as under :

|  |
| --- |
| Software |
|  |



+

Some

other

types

are

as

given

below:

File Organization (Storage point of view)

+ Sequential Fires: As the name refers, these files are stored or created on the Storage media in the order the records are entered i.e., one after another in the sequence. They require comparatively more processing time.

 Direct or Random These files reside on the storage media according to the address which is calculated against the value Of the key field of the record. Some times, the same address is calculated, which leads to the concept of synonym.

# Indexed Sequential: The key field of the records (in a file)are stored separately along with the address of each record. These files can be processed sequentially as well as randomly. They require relatively rmre space on the storage media but the processing is as fast as random/direct files.

1.3 DATABASES

A database is a collection of logically related data sets or files. Normally, these files/datasets are of different nature, used for specific purposes. These may be organized in various ways to meet various processing and retrieval requirements of the organizations or users. For example;

A bank may have separate files for its clients Ee.,

> Savings A/C

 Automobile loan

 Personal loan

Clients biographic information etc.

The bank's clients/customer database would include records from each Of these files. Using a series ofprograms, data for any client may be added, retrieved or updated depending upon the activity at a particular time.

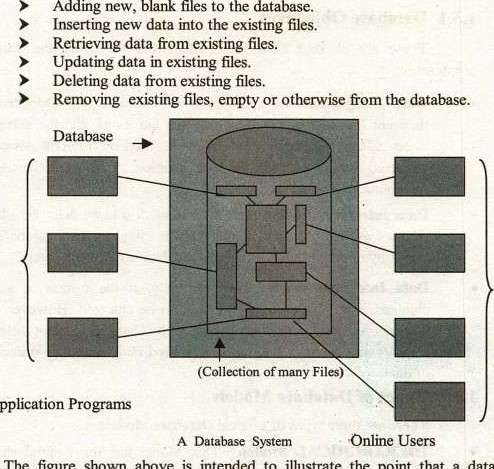


of

12

It is a computerized system whose overall purpose is to maintain information and to make that information available at any time.

A data base system is just a computerized record keeping system, A data base itself can be regarded as a kind of electronic file cabinet, a warehouse or a repository for a collection of computerized data files. The user of the database normally has the following mcilities to enjoy.



to

Application

The

figure

shown

above

is

intended

to

illustrate

the

point

that

a

database

system involves four major components, namely.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | |  |  | | --- | --- | | Data | The Information | | Hardware | The physical components i.e.,   * Secondary Storage * I/O devices * Device controllers * I/O chennels * Processors   Main memory | |
|  |



|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | |  |  | | --- | --- | | Software | User/system software  Set ofprograms  • Utilities | | Personnels | The people   * Programmers/Analysts * End users * Database Administrators | |

13.1 Database Objectives

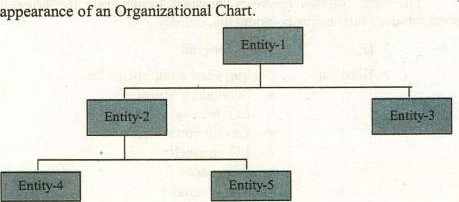
There are at least three main objectives for using the data base organization.

* Data  In a database, information is coordinated from different files and operated on a single file. Logically, the information is centralized, physically data may be located on different devices i.e., scattered around over on different locations, connected through data communication links.
* Data Integrity: If a data item is contained in more than one file, then all files must be updated if that item is changed. In a database, only one copy of data is kept, therefore, the data is more consistent.

 Conventionally, if the format of a file is changed, then all the programs have to be changed. However, a data base allows the organization of data to be changed without the need to re-program. It allows programs to be modified without re-organization of data.

13.2 Types of Database Models.

There are three types ofLogical Database Models.

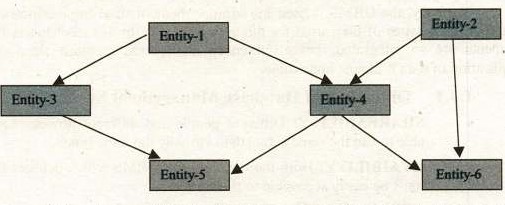
* HIERARCHICAL ModeE This Model has the general shape or appearance of an Organizational Chart.

of 7 n



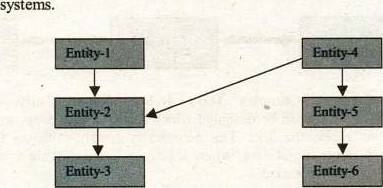
A node on the chart, representing a particular Entity is subordinate at the next highest level, just as on an organizational chart, an employee reports to only boss. This kind of structure is often referred to as an "Inverted Tree", with the top most referred to as the 'Root".

* NETWORK Model: This is somewhat similar to that of the Hierarchical model but has one major difference.



Subordinate entities, depicted by arrows on the network diagram, may participate in as many subordinate relationships as desired. Therefore a much more complex diagrams may be used to represent the structure of the database. Networks provide more flexibility than a simple hierarchical system in the data relationships may be maintained.

• RELATIONAL Model: This system consists ofa collection of simple files/Relations (Entities), each of which has no structural or physical connection such as those typically used in hierarchical or network



The various entities ppssess the interrelationships as depicted by a network like diagram. But these relationships are based on the data content of the entities not by pointer chains or other types of structural connection techniques. 



1.4 DATABASE MANAGEMENT SYSTEM

The data management system (a collection of programs) which is used for storing and manipulating databases is called database management system (DBMS). It is an irnprovement over the traditional file management systems. It uses DBMS software (database manager) which controls the overall structure of a database and access to the data itself

Normally, the DBMS is used for large or medium sized organizations, having heterogeneous types of files, used for different purposes. In this mechanism, the data elements are so integrated, cross referenced and shared amongst them that the duplication of data is almost impossible.

1.4.1 Objectives of Database Management System

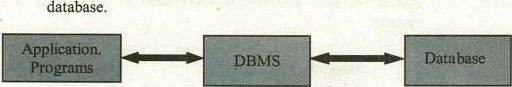
* SHAREABILITY: Different people and different processes must be able to use the same actual data virtually the same time.

 AVAIBILITY: Both the data and the DBMS which delivers the data must be easily accessible to the users.

* EVOLVABILITY: The ability of the DBMS to change in response to growing user needs and advancing technology.
* DATABASE INTEGRITY: Since data is shared among multiple users, adequate integrity control measures must be maintained.

1.4.2 Advantages of Database Systems

* Data Independence: Application prograrns are not aware of the physical implementation of the data sets. The DBMS sits in between the application programs and the actual data sets that make up the



 Support Complex Data Relationships: Fairly complex data structures can be designed which allow various ways to logically view or access the data. This complexity greatly enhances the ability of a designer to put data 'Where it belongs", and provide a path to that data whenever needed.

* Sophisticated Data Security Features: Provide enhanced security mechanisms for access to data. Data base security mechanisms typically go much further in adding more extensive security features. If granted Read access to a file/table, the user may see each record in the file, and every data field it contains. Access intent of each application program (read, write, update, delete) can be specified



explicitly. An application program's view of data records may be controlled to the field level.

Data Base Backup / Recovery: Provide sophisticated backup / recovery mechanism.

Backup / Recovery capabilities onen distinguish between true DBMS and a software package that only claims this facility. A DBMS has a logging or recording mechanism that captures information on changes to data within a data base. In case of data base recovery, a utility within the DBMS rebuilds it by using a backup copy Of the data and log Of changes as input.

Advanced Capabilities: DBMS normally have advance access capability for on-line and ad-hoc reporting capabilities. However, the ability to provide data independence to create complex data structures, to provide security to data access, and to provide backup / recovery capability are the primary requirements of a Database Management

# Systems.

1.4.3 Disadvantages of Database Systems

Although, the DBMS are very powerftl tools to do the job, but they are not more in use. It is because of some disadvantages :

Require additional System Overhead: Additional overhead is required to access data, in case of doing some simple jobs; like reading and processing a tape file, which might take a little time and resources to do the job. If we have to do it on DBMS, it is like "requiring too much to do too little".

Additional Training required for Training of Staff: Application programmers require a sort of precise training to code emcient programs that will run under a DBMS. There is a possibility that inadequate training or experience of application development staff will lead the creation of grossly ineffcient database calls. Quite often, the problem might not be found until the program reaches production status. The typical example is that of using proper and improper indexes for accessing the database.

Problems can multiply in selecting a wrong type of Dbase Environment: A later change in structure, forced by changing requirements, can be costly in terms of conversion and testing of existing programs, Hierarchical data base systems are, in particular, more sensitive than network or relational systems towards this kind of problem, and implementing changes costs a great deal. On the other hand, doing these changes on relational data bases are fairly easy and less costly.



Data must be considered a corporate resource: The data in a company's data base no longer belong to one organization alone. True, One organization normally has the primary responsibility for creating a data base. However, as data base systems mature, more companies or organizations can share the same data across applications.

* A Need of a Dictionary: In order to share data across application systenu, or to simply given end users the ability to identi$' the location of information they need in order to do their jobs, the internal data contents of a company's data bases need to be documented in a consistent manner. For this purpose, they have to install a data dictionary system, which is another overhead on the DBMS.

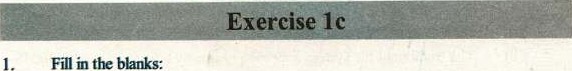
1.4.4 Features of a DBMS

* Data Dictionary: Some databases have a data dictionary, a procedures document or disk file that stores the data definitions or a description of the structure of data used in the database. The data dictionary may monitor the data being entered to make sure it conforms to the data definition rules i.e., file names, field names, field sizes, data types etc. It may be used for data access authorization for the database users.
* Utilities: The DBMS utilities are the software programs that are used to maintain the database by manipulating the data, records and files. Some programs are also used for backup and recovery procedures of the databases.
* Query Language: Normally, SQL (Structured Query Language) is used for creating table structures, entering data into them and retrieving/updating the selected records, based on the particular criteria and format indicated, within the databases. Typically, the query is in the form of a sentence or English-like command i.e., SELECT, DELETE, CREATE, MODIFY, UPDATE and INSERT commands.



 Report Generator: A report generator is a program that is used to produce an on-screen or printed document from the database. The report format can be specified in advance i.e., row headings, column headings, page headers etc. Even the non-experts can create very useful and attractive reports by using this facility.

* Access Security: By using this ficility, the database administrators can assign specific access privileges for the users of the databases.
* Backup and Recovery: It is an important feature available in almost all the DBMS programs. By using this feature, we are able to have the backup of our data and can later, use it to reinstate it in case of data åilure, corruption or loss.



in

(i) DBMS stands for is a collection of related fields

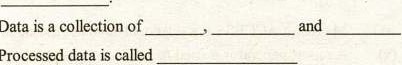
(iii) A file is a collection of

(W) Before processing the data is recorded in is a collection of logically related data

* 1. The data definitions is stored in
  2. SQL stands for

(Viii) Hierarchical data Model has the general shape ofa(n)

(ix)



Data

Processed

(x)

2. Sekct the correct

(i) Which ofthe following represents a collection of concepts that are used to describe the structure of a database? 

a) data warehouse b) data model

|  |  |
| --- | --- |
| Textbook ofComputer Science | Class 12 |

 c) data structure d) data type

(if) Which of the following data model is more flexible?

a) Network data b) Hierarchical data model

c) Relational data model d) object data model

(iii) Which ofthe following type of file require largest processing time?

a) Sequential file b) Random file 

c) Indexed sequential.file d) Direct access file

(W) Which of the following may be a temporary file?

 a) Master file 

c) Backup file

(v) SQL is a(n):

a) Unstructured language

c) Object oriented language

b) Transaction file

d) None of these

b) Structured language

d) Software

12 Cha ter I



1. Write T for true and F for false statement.
   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 meanings of database terminology (iv) Master file is the latest updated file which never becomes empty, ever since it is created.

 SQL is used to retrieve information from the database based on certain

* 1. The Network Data Model is more popular and widely used than Relational Data Model.
  2. Indexed sequential files gan be processed sequentially as well as  randomly.

(Viii) Backup files store data prior to its processing.

1. (ix) Microsoft ACCESS is a relational database management system

(x) A report generator is used to produce a printed document from the

1. a) Differentiate between Data and Infonnation 

b) What activities are involved in data processing? Discuss in details.

6, Define file, record and field in details?

1. Describe the file types from usage point ofview and åanctional point ofview?
2. How do we organize the files on storage media?
3. In general, what activities are to be performed on the databases? Discuss in
4. What are the four major components Ofthe database systems? Write in details,

Discuss the objectives of the databases in your own words

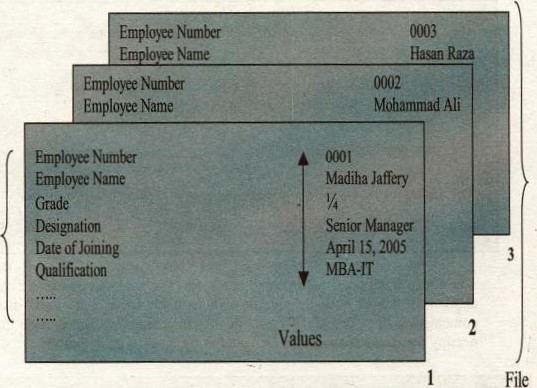
1. Describe the different database models?
2. Discuss the objectives and features of the DBMS?
3. What are the advantages and disadvantages ofthe DBMS?

o r

|  |  |
| --- | --- |
| BASIC CONCEPTS AND  O OG EDATABASES | Chapter |

2.1 OVERVIEW

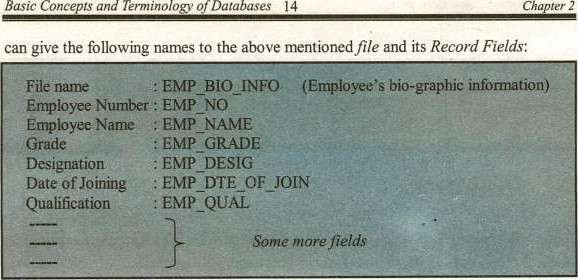
In the previous chapter, we have discussed Files and Records in details. In fact, the concept of databases evolved from the old, traditional working of File Management System (FMS). Let us see how this evolution emerged by considering the same diagram of Record and File.

Fields

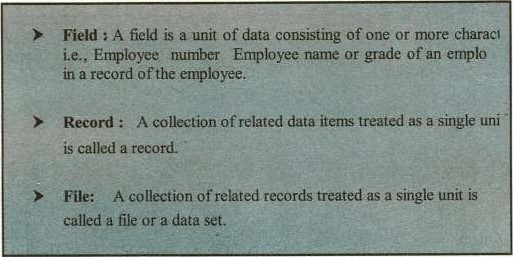


Records

It is worth mentioning here that the file must have a meaningful name and all the field names in each record must be unique and meaningful too. For example, we



0



Basic

Definitions

The above mentioned record can be shown (as it would appear on the hard disk) as follows. Actually, all the data will be saved onto hard disk as a combination of and "l"s i.e„ the machine language form Each field will occupy the space as would be allocated at the time ofits definition.



EMP\_NO EMP\_NAME EMP\_ EMP\_ EMP\_DTE\_ EMP\_ (some more fields)

GRADE DESIG OF\_JOIN QUAL

of



2.1.1 Data Handling in File Management Systems

• The record layouts are properly defined in the file management systems Le. each field is given a fixed- or variable-length sequences of bytes and they are put together contiguously in fixed- or variablelength collections called records. Thus, each field corresponds to a proper starting memory address. As the fields have already been given proper lengths, so the values Of each field are determined within those memory addresses. See the example given below :

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | | | | | |  |
|  | Jaffery |  | Manager | pril  2005 | MBA. |  |
|  | | | | | | |

Thus, the field names are used only as a "names reference" within the programs using them. Their values "flows" with them as the contents ofthe memory spaces they are occupying.

2.2 ATTRIBUTES, ROWS AND TABLES

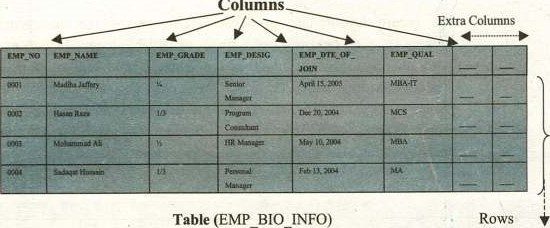
In the late 1960s, the researchers came up with an idea Of using Relational

Databases instead of file structures. They gave the idea of defining a file as a "Two

Dimensional" array (or Table having a unique name), placing all the fields as columns (having unique names) of that table and putting each record as a row into the table. Each row is also known as a tuple or occurrence in that table. Following these concepts, we can see the traditional file structure into a new and easy to manage database structui•e called a Table or Relation.

Concepts of Chapter 2





C

olumns

So, consequently, we summarize the above discussion as follows :

|  |  |  |
| --- | --- | --- |
| |  | | --- | | Data Elements : The fields or data items in databases are termed data iterns, items, attributes or columns in database structures.  Records in file management structures are termed as rous or tup in database structures.  Filesor datasets in databases are termed as tables, relations or data objects in database structures.  The collection oftables-with some traditional files and some other necessary data objects is termed as database. | |  |

2.3 RELATION or TABLE

 A two dimensional array or table of data containing descriptive information about an entity. The entity must have a unique identifier, which is composed of a combination of one or more attributes, and each attribute must have one and only one value. It is appropriate to define the word Entity here. 

Entity: An entity is any thing about which you want to keep information in the database. Let us consider an example of 'Student Information System", which has entities like student, teacher, course lis!, scholarships, time-tabling

Q/



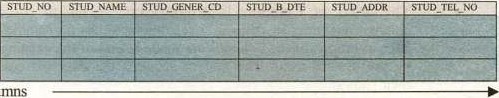
etc. Thus, the entities involved in this case are the same and the entity "student" can be defined in the form of database modeling as follows :

•:• STUDENT (STUD NO, STUD\_NAME, STUD B DTE,

STUD\_ADDR,

From the above given definition ofentity, we can easily construct a twodimensional array or a relation by converting all the attributes in the brackets into columns of the array, as follows :

Rows



Columns

 In the above diagram, all the attributes are called columns ofthe table and all occurrences of the records are called rows.

For example, considering the entity 'STUDENT' as above, following table or relation can be constructed as a part of the database.

STUDENT

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| sno xo | STUD NAME | s•ruo GENER cn | SWO B | STR@ADOR | STUD TEL\_NO |
| fm9•SOA | Sato |  |  |  |  |
|  | Adil |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

A relation STUDENT having rows (records) under columns (attributes)

2.3-1 

A Relation or a Table which is the basis of 'Relational DBMS, by definition must have certain inherent characteristics that form the basic for its underlying strength and flexlbility. Because of these features, an application  implemented by using such a system is much more flexible and can be easily



Conc

of

C

ter2

modified when alterations or enhancements to the underlying data base take place. These characteristics are:

* + - No duplicate rows exist: No two rows can be identical. Why to put two rows (records) for the same entity (e.g., a Person). It will also violate the definition of what a relation represents, as it says by definition that there must be a unique key for each row in a

relation/table.

* + - The order of Rows is insignificant: There is no ordering or sequencing of the rows in the tables. The relational implementation of the tables support all required access mechanism i.e„ it is not necessary to sequence the rows according to the key field.
    - The order of Columns is insignificant: Again, the order of the columns/attributes in defining a relation/table has no significance. The later insertions of the columns are made at end ofthe existing columns by the system itself. The system acquires the data (of columns) by their names.
    - Columns/Attributes are all Elemental or Atomic: All the intersections of Rows and Columns must have a (single) value. The nulls are inserted by the system at the time of later insertion of a column, which should immediately be replaced by zeros/spaces or valid values for that particular column.

2.4 VIEWS

Views are created by using SQL which is a powerful database language, used for data definition and data manipulation purposes. The purpose of using views is purely to keep the data safe and secure from un-authorized and illegal users. The views provide the descriptions of relations that are not stored, but constructed as needed from stored relations.

To create a view, normally the following create sql command is used :

|  |  |
| --- | --- |
| CREATE VIEW | STUDENT VIEW 01 AS |
| SELECT | STUD NO, STUD NAME, STUD ADDR |
| FROM | STUDENT |
| WHERE | STUD GENDER CD = |

This will create a view from the STUDENT table for only male students, which can be used by the users according to the authorization given to them, leaving the original table aside, safe and secure.

2.5 INDEXES

It is another table created by the system developer/DBA containing the key attributes of the table for which the Index is created. It has a very vital role in the data



base managernent systems, especially in RDBMS. The important associations defined in the system make use ofthis. It helps the system run smooth and fist.

2.6 KEYS

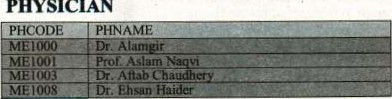
A key is a single or combination of one or more fields and its purpose is to point/retrieve the data rows from the tables, according to the requirement.

Keys are defined in the relations/tables to access or sequence the stored data fast and smooth or to create the links between them

Let us assunr two relations/tables as follows to define different types ofkey.

PATIENT

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | FNAME  ABID+ | \*HCODE  -MEI | •1000 |  |
| 9293 |  | SAL[M FAR | MEIOOI  ME 1003 | 1000102 | 01 |
| 98 |  | KASHIF |  | 10 104 |  |
|  |  | ZAHID | MEIOOO | 1000101 | 09 |
|  |  | NAEEM | MEI | 1000105 |  |



PHYSICIAN

2.6.1 Types or KEYS

* + Primary Key: In a relation, the attribute (column) or a combination of attributes (columns) that uniquely identifies a row or a record. PATNO is the attribute that uniquely identifies each patient and thus can be used as a Primary key for the above defined table 'PATIENT'. On the other hand, PNAME is normally not unique, so it can not be used as a primary key.
  + Secondary Key: A secondary key is non-unique field that is used as a secondary(alternate) key. In the above table, by using PHCODE (physician code), we can scan the records from the table. If the physician leaves, we can change it with a new one by using update statement.
  + Candidate Key/AIternate Key: Sometimes, it is unclear which field to select as the primary key. There might exist some additional field (or combination of fields) that also have the uniqueness property. These keys can be termed as Candidate keys or Alternate keys.



* + In addition for the uniqueness requirement, Candidate keys must possess following two properties:

Concepts Te ofDatahases 2

Composite Key: These keys consists of two or more data elements or attributes. Invariably, these are the same as Candidate/Alternate keys that of uniqueness requirement. In order to make it unique, assign STATUS or another attrll•ute (say patient's ID number).

* + Sort/Controi Key: A sort'Control key is used to physically sequence the stored date according to our need. Multiple attributes can be used as sort fields.
  + Foreign Key: A foreign key is an attribute in a table whose values must match a primary key in another table. The table in which the foreign key is found is called as dependent table and to which it refers is called as parent table.

|  |  |  |  |
| --- | --- | --- | --- |
| |  | | --- | | :  I:M relationships | | |  | | --- | | ré1Åt10kk%1pÅ •are across the Relations/ Tables in a relational data | |

 THE USER 

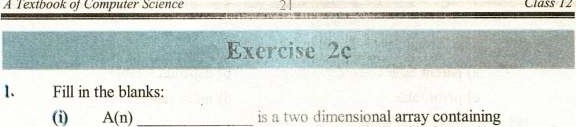
The user or end-user is simply a person who uses the computers for his specific need. He might have a moderate knowledge of computers, computer science and information technology, and his need io use the computers may be entertainment, education, or professional (asks. He does need to know the in-depth knowledge of the computer systems, but instead, he should be aware of the installed software he intends to use.

2.8 THE DATA ADMINISTRATOR

A data administrator (DA) is responsible for the entire data of an organization. He normally develops the overall functional requirements for the databases being used in the office. He shares in developing the logical design for each database. He should control and manage the databases, establish the data standards, supervise the data distribution within the organization and communicate with the users when necessary. He should also participate in developing the data dictionary, prepare documentation and conduct user training where needed. Normally, the Data Administrator serves as a bridge between users and data processing staff.

2.9 THE DATABASE ADMINISTRATOR

A database is responsible for the design, implementation, operation, managernelit maintenance of the database. He/She must be technically expert on the wlrieaeies of the database and DBMS. He is supposed to plan, coordinate and carry variety of jobs during all phases of the database projects. He must possess tcchnical skills because he has to work on the complex software and hardware issues involved and to solve the problems of the system and application experts in [he organizatiom He is also responsible to make sure the database access rights, to safeguard its security and to maintain and fine-tune the database functionality.

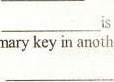


of

descriptive information about an entity.

is any thing about which the information is kept in the database.

1. In a table the order of rows and columns is
2. In a relation, the attribute or a combination of attributes that uniquely identifies a record is called
3. A(n) describes the characteristics of an entity.
4. A(n) is an attribute in a table whose values must match
5. A(n) is responsü»le for the design, implementation, operation, management and maintenance of the database.



a

pnmary

another

lahie.

1. A(n)consists of two or more attributes.
2. A(n)is the dynamic result ofone or more relational operations on the base reintions to produce another relation.
3. Therefers to raw facts and figures.
4. Select the correct option:
   1. Insert command is used to insert:

a) a new table b) a new record

* + 1. a view d) dependencies
  1. CREATE command is used to ereate a:

a) table b) view

* + 1. report d) query
  1. SQL is used for:

a) data definition b) data manipulation

* + 1. data definition and manipulatipn d) searching records



* 1. The foreign key is found in:

|  |  |
| --- | --- |
| a) parent table | b) dependant table |
| c) pivot table  (v) A table must have: | d) index table |
| a) prirnary key | b) secondary key |
| c) composite key | d) sort key |

1. Mark as True or False
   1. The view is not stored in the database,
   2. Two tables can not have the Same name in the database.
   3. Index makes the searching of a record åster. (iv) Secondary key must be unique.

(v) The primary key can not work as a sort key.

 The DBA is responsible for maintaining the database.

* 1. A file is a collection ofrelated fields,
  2. DBMS provide more security to protect data than traditional file management systems.
  3. DBMS is a software used to train database administrator.
  4. A relation is also termed as a tuple in relational database.

 How the Records and Files are constructed in traditional File Management System ?

 How the data is stored and retrieved in FMS (file management system)?

6, How the tables/relations are formed up in DBMS?

7. Discuss the data manipulation in DBMS system.

8• Write down the properties ofrelations in details ,

9. What is a VIEW? How do we create it? 

10. What is usage of indexes in FMSind DBMS?



Of

23

3.1

OVERVIEW

Before we design a database for any Organization, we have to consider many aspects to find out the practical scenario of owning a database. Few of them are discussed as follows :

Feasibility Study: This is also called preliminary investigation Of the required database. It involves the area identification and selection i.e. which area or aspect is to be selected to start with. After the project is selected, it is allocated a specific fund and a proper planning is chalked out for its practical implementation. Side by side, a proper market analysis is also worked out.

Requirements Analysis: During this activity, the requirements are gathered i.e. the possible inputs for the database and the required functionality out of it. The users precisely narrate their needs of the database and the possible domain and restrictions are also chalked out.

Project Planning: A proper schedule is laid down to accomplish this activity. All thc cost factors arc taken into consideration i.e., the salaries of team members, their logistics involved, other trivial expenses (such as marriage gifts, insurances etc) and hardware costs.

Data Analysis: This is an irnportant analysis aspect while designing a database. It involves the following activities :

1. Data Flow Diagrams (DFD)
2. Decision Tables
3. Decision Trees

However, a detailed discussion on these topics is beyond the Scope Of this book.

3.2 DATA MODELING

Data Modeling is the process of identi$ring the data objects and the relationships between them.

Ingredients of I)ata Modeling:

Entities/Objects : A data entity/object is anything that is participating in the System. It is always properly identifiable Le., a TEACHER, a STUDENT, an AEROPLANE.

Database Design Process 3

Attributes: Attributes define the objects, describe their characteristics and in some cases, make references to other objects(s) i.e., attributes for a TEACHER could be :

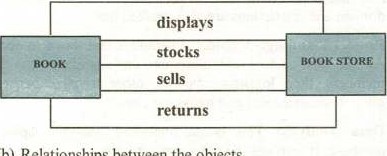
Teacher Name, Gender, Last Degree, Appointment Date, Pay Scale, Nationality, Telephone No. etc.

Relationships: The relationship indicates how the Entities/Objects are Connected or Related to each other.

The Data objects are related/connected to one another in different ways. Consider the data objects BOOK and BOOK STORE in the following diagrams.



(a) A basic connection between the Objects



(b)

Relationships

between

the

objects

orders

Following are the different possible and relevant relationships between them:

A BOOK STORE orders BOOK(s).

 A BOOK STORE displays BOOK(s).

 A BOOK STORE stocks BOOK(s).

•:• A BOOK STORE sells BOOK(s).  A BOOK STORE returns BOOK(s).

It is important to note that :  All the relationships define the relevant connections between both objects.

 All the relationships are bi-directional.

 We have to consider only the relevant relationship (in the context of the requirement)



Cardinality:

* Whether some occurrence(s) of object- I are related to some occurrence(s) of object-2.

 It is expressed as one or many e.g.,

 A husband can have only one wife and

 A Father can have many children.

* The relationships can be

 One to one

 One to many

•9 Many to many

 Recursive  None Modality:

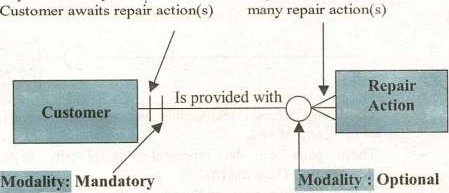
* It defines the nature of the relationship i.e.,

 Optional . . represented by 0

+ Mandatory represented by I

* Consider two objects Customer and Repair Action in a Workshop environment.

Is provided with



Implies

that

only

one

Implies

that

there

may

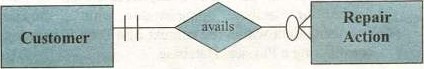
be

Cardinality

Cardinality

|  |  |
| --- | --- |
| Implies that in order to | Implies that there may be |
| have a repair action(s), | a situation in which a repair action |
| we must have a customer | is not necessary |

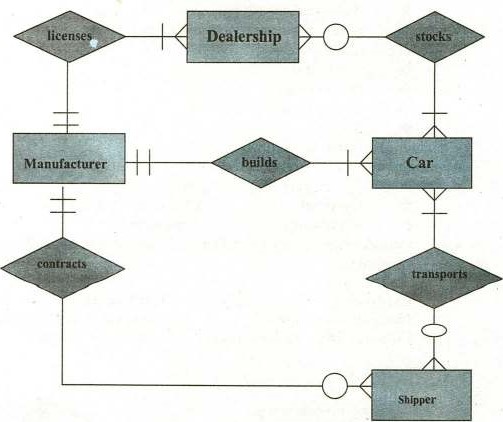
* A simple Data Model can be drawn from the above as follows:





* By connecting all the Data Objects alongwith their Relationships in the above manner, an ERD (Entity Relationship Diagram) is constructed.

Entity-Relation Diagram (An Example)



* The Entity/Object — relationship pair discussed above is the objective ofthe Data Model.

These pairs can be represented graphically using the EntityRelationship Diagram (ERD).

* It was basically proposed/used for design pf a Relational Database System and now is being adopted for other Database types also.
* A set of primary components are identified for the ERD: Data objects, Attn%utes, Relationships, Cardinality and Modality.
* The primary objective of the ERD is to represent Entities/Objects and their relationships / association.
* Data Modeling and the Entity-Relationship diagram provide the Analyst or database administrator with a concise notation for examining data within the context of a Data Processing Application or constructing a Physical Database.

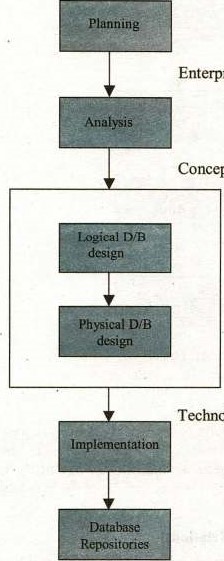
OfCom u'er 27



3.3 DATABASE DESIGN

The major objective of Database design is to map the conceptual data model to an implementation model that a particular DBMS can process with performance that is acceptable to all users throughout the organization. In today's compatible economy, database users require information that is complete and up-to-date and they expect to be able to access this information quickly and easily.

Following is the database development process.

Enterprise Data Model

Conceptual Data Model

Technology Model

Database design can be divided into the following two phases:

33.1 Conceptual (Logical) Database Design:

The process of mapping (lw conceptual data models (from analysis) to structures that are specific to the target DBMS. If the target environment is a relational DBMS, then the conceptual data models are mapped to normalized relations.

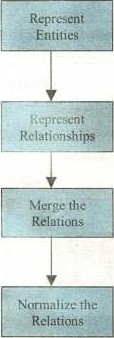
ro Mowing diagram presents an overview of logical design process.

Conceptual Data Model

(E-R Diagram)



l'.•ncess



LOGICAL DATA MODEL (Normalized Relations)

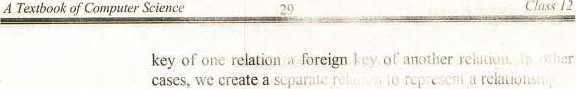
* + 1. Represent Entities:



Each entity type in the E-R diagram is represented as a relation in the Relational View or Data Model. The identifier Of the entity bee"mes the Primary key of the relation, and other attributes of the entity type become non-key attributes of the relation.

* + 1. Represent Relationships:

Each relationship in an E-R diagram must be represented in the relationol model. It depends on its nature. For example, in some cases, we represent a relationship by making the primary



* + 1. Merge the Relations:

In some eases, there  (that is, 'wo or more relations that dco,Tibe [he game entity type). They musi be merged to remove the redundancy. This process is also Isnown as View •m Suppose we have one relation EMPLOYEE' ,

And another relation as:

EMPLOYEE2(EMENO , FNAME. EMP-ADDR,

CODE, EMP-DOB)

Since the two relation; have primary key (EMPNO) they describe the CliE ity ami may be merged into one relation. The result of merging the relations is the following relation.



JOB-CODE,EMP-DOB)

the R&thas:

The relations that are created in step (i) 1 ii) may have unnecessary redundancy and may be Subject to anomalies (or errors) when they are updated. Normalization is the process that refines the relations to avoid these problems. (A detailed  discussion is given in chapter 4)

343.2 Ihtaba\*

It is the last stage of the database design process. The major objective of physical database design is to implement the database as a set of stoled records, files, indexes and other data structures that will provide adequate performance and ensure database integrity, securi(y and recoverability.

There are three major inputs to Physical database design.

Logical database structures (developed during logical database design) i.e., the Normalized Relations.

User processing requirements i.e. size and frequency Of use  ofthe data-Base , response time, security, backup, recovery etc.

Cbaracteristics of the DBMS and other components Of the computer Operating environment.



Components of Physical Database Design:

 Data Volume and Usage Analysis:

To estimate the size or volume and the usage patterns of the database. Estimates of database size are used to select Physical storage devices and estimate the costs of storage. Estimates of usage paths or patterns are used to select the file organization and access methods, to plan for the use ofindexes and to plan a strategy for data distribution.

Data Distribution Strategy.

Many organizations today have distributed computing networks. For these organizations, a significant problem in physical database design is deciding at which nodes (or sites) in the network to physically locate the data.

Basic data Distribution Strategies.

a, Centralized: All data are located at a single site. It is Eiirly easy to do but it has at least three disadvantages.

* data are not readily accessible at remote sites.

Data communication costs may be high.

* The database system fails totally when the central system fails.

Partitioned: The database is divided into partitions (fragments). Each partition is assigned to a particular site. Major advantage of this is that data is moved closer to local users and so is more access-able.

1. Replicated; Full copy of database is assigned to more than one site in the network. This approach maximizes local access but creates update problems, since each database change must be reliably processed and synchronized at all ofthe sites.
2. Hybrid: In this strategy, the database is partitioned into critical and non-critical fragments. Non-critical fragments are stored at only one site, while critical fragments are stored at multiple sites.
3. File Organization:

A technique for physically arranging the records of a file on secondary storage devices. For selecting a file organization, the system designer must recognize several Constraints, including the physical characteristics of the secondary storage devices, available operating systems and file management software, and user needs for storing and accessing data. Following is the criteria for selecting file organizations.

* + - * Fast access for retrieval.



* + - * High throughput for processing transactions.  Effcient use of storage space. 
      * Protection from failure or data loss.

 Minirnizing need for re-organization  Accommodating growth.

Security from un-authorized use.

1. Indexes:

An index is a table that is used to determine the location ofrows in a table ( or tables) that satisW• some condition. They may be created on primary key, secondary key, foreign key etc.

1. Integrity Constraints:

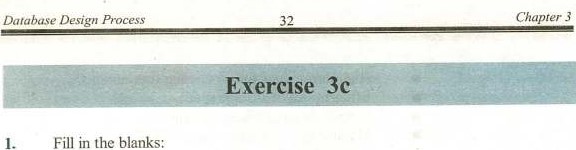
Database integrity refers to the correctness and consistency of data. It is another form of database protection. While it is related to security and precision, it has some broader implications. Security involves protecting the data from unauthorized operations, while integrity is concerned with the quality of data itself. Integrity is usually expressed in terms of certain constraints which are the consistency rules that the database is not permitted to violate. A few of them are discussed in chapter 4.

3.4 IMPLEMENTATION

In database implementation phase, the builder or the database administrator normally requires a server computer which will be linked with hundreds and thousands of computer users who would want to share and interact with the server (database).

For this purpose, the dba might need the services of network administrators to connect the users with the server. The users are normally given the authorizations / permissions defined by their respective managers so that they can perforrn the authorized tasks while using the database facilities. 

In distnhlted computing environment, the database Servers and users might be thousands of kilometers apart, so a lot of expensive telecommunication links are required to perform the designated tasks. NADRA and CRICKINFO are some of the typical examples of this type of databases.



Fill

in

1. Duringphase, the project requirements are gathered and identified.
2. DED stands for
3. The process of identifying data objects and relationship between them is called
4. The number of occurrences of participating entities in a relationship is determined by the ratio.

 Modality determines Whether the participation of an entity in a relationship is or optional.

ERD stands for

1. In ERD model, a(n)is represented by a rectangular box.
2. In database systems, all the data is stored at a single site.

database multiple copies of the same data are stored at different sites on the network.

(x) In distributed databases, the data is among various sites.

2- Select the correct option:

Which of the following keys does not hold uniqueness property.

a) candidate key b) foreign key

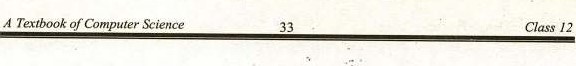
c) primary key d) secondary key

 An entity related to itself in an ERD model refers to:

a) recursive relationship b) one-to-many relationship

e) many-to-many relationship d) one-to-one relationship

1. Database development process involve mapping Of conceptual data model into:
   1. Object oriented data model b) Network data model



c) Implementation model d) Hierarchical data model

1. In ERD model, the relationship txtween two entities is represented by a:
   1. diamond symbol b) rectangular box

c) oval symbol d) kne

(y) In hybrid distribution which kind of fragments are stored at only one

a) critical fragments b) non-critical fragments

c) critical and non-critical fragments d) only large fragments

1. Write T for true and F for false statement. 
   1. In one-to-one relationship only one instance of each entity can participate in the relationship.
   2. The optional modality is represented by 1.

(iii) One-to-many is a uni-directional relationship.

In ERD model, a condition is mentioned in a diamond symbol.

(v) ERD is a physical data model.

 In hybrid distribution the database is partitioned in critical and noncritical fragments.

* 1. In distributed databases, the consistency refers to availability of same data at all sites ofthe network.
  2. Indexing maximizes the time required to search a piece of information from a database.

Analysis is less important activity than coding, so minimum time should be spent over analyzing the system.

(x) Relationship defines the logical connection between entities.

1. Describe different steps involved in analysis stage while designing a database,

5' Explain the following with the help Of figures:

* 1. Entity/Object
  2. Attribute
  3. Relationship
  4. Cardinality 
  5. MOdality

34



6. Draw and explain ER diagram for the system of getting admission in your college.

 Explain the following:

Physical data model

(ii) Conceptual data model

1. What are the cornponents of a logical data model?
2. What elements combined , produce the physical database design? Explain
3. Define and explain the following terms:
   * + 1. Data distribution strategy
       2. File Organization
4. Define the term Analysis. Briefly discuss the following terms:
   * + 1. Feasibility study
       2. Requirement Analysis
       3. Project Planning
       4. Data Analysis

12. Briefly explain the database design process with the help of a diagram.