Chapter 8
# Database Management System

- **Q) What is a database?**

- **ANS:** Database is an organized collection of data related to a particular enterprise. It may contain different types files each one containing many records.

- **Q) What is DBMS?**

- Database Management System is a collection of inter related data and a set of programs to store, modify and access those data

Plus Two Result

Database
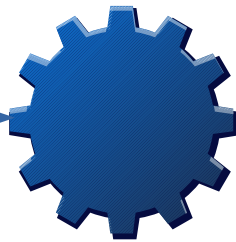
DBMS

User

END

What are the advantages of DBMS ?

- a). **Controlling Data redundancy**:-Duplication of data is known as data redundancy. In DBMS data is kept in one place in centralized manner and the users can access this centrally maintained date for their purpose. So data redundancy can control.

- b). **Data consistency** :- Data redundancy leads data inconsistency (different copies of same data hold different values because the updation of data may not occur in all the copies. In DBMS it avoided by eliminating redundancy.

- c). **Efficient data access**: DBMS utilizes a variety of techniques to store and retrieve data efficiently.

- d). **Data can be shared**:- The data stored in the Database can be shared among many users and new programs can be developed to share the existing data.

- e). **Data Integrity**: Integrity refers to the overall completeness, accuracy and consistency of data in the database. It can be achieved by use of error checking , validation, avoiding duplication etc.

- f). **Security** :- Security refers to accidental or intentional disclosure or unauthorized access, modification or destruction. Through the use of passwords, information in the database is made available only to authorized person. .

- g). **Sharing of data** : The data stored in the database can be shared among several users or programs even simultaneously and each may use it for different purposes.

- h). **Enforcement of Standards**:- The database Administrator defines standards like display formats, report structure, update procedures, access rules etc.. for the DBMS. It is helpful when data transfer occurs between systems.

- i). **Crash recovery** : If the system crashes data in the database may become unusable. DBMS provides some mechanism to recover data from the crashes.

What are the components of a database?

- The components of a database are

- a). **Hardware** : It include actual computer system used for storage and retrieval of Database. It include computers, storage devices, network devices, and other supporting devices.

- b). **Software** : It consists of DBMS, application programs and utilities. DBMS consists of different components that handles tasks like data definition, data manipulation

- c). **Data** :The database should contain all the data needed by the organization. For effective storage and retrieval of information , data is organized as fields, records and files.
  1) Field : A field is a smallest unit of stored data. Each field has a specific type.(eg. Name,- Char, Mark,-num etc).
  2) Record : A record is a collection of related fields. A record store an instant of the relation or the table.
  3) File : A file is a collection of all occurrence of one type of record.

- d). **Users** : The users of database can be classified depending on the mode of their interactions with DBMS. The different categories of users are Database Administrator, Application Programmer, Sophisticated users, and naïve users.

- e). **Procedures** : Procedures refers to the instructions and rules that govern the design and use of the database.

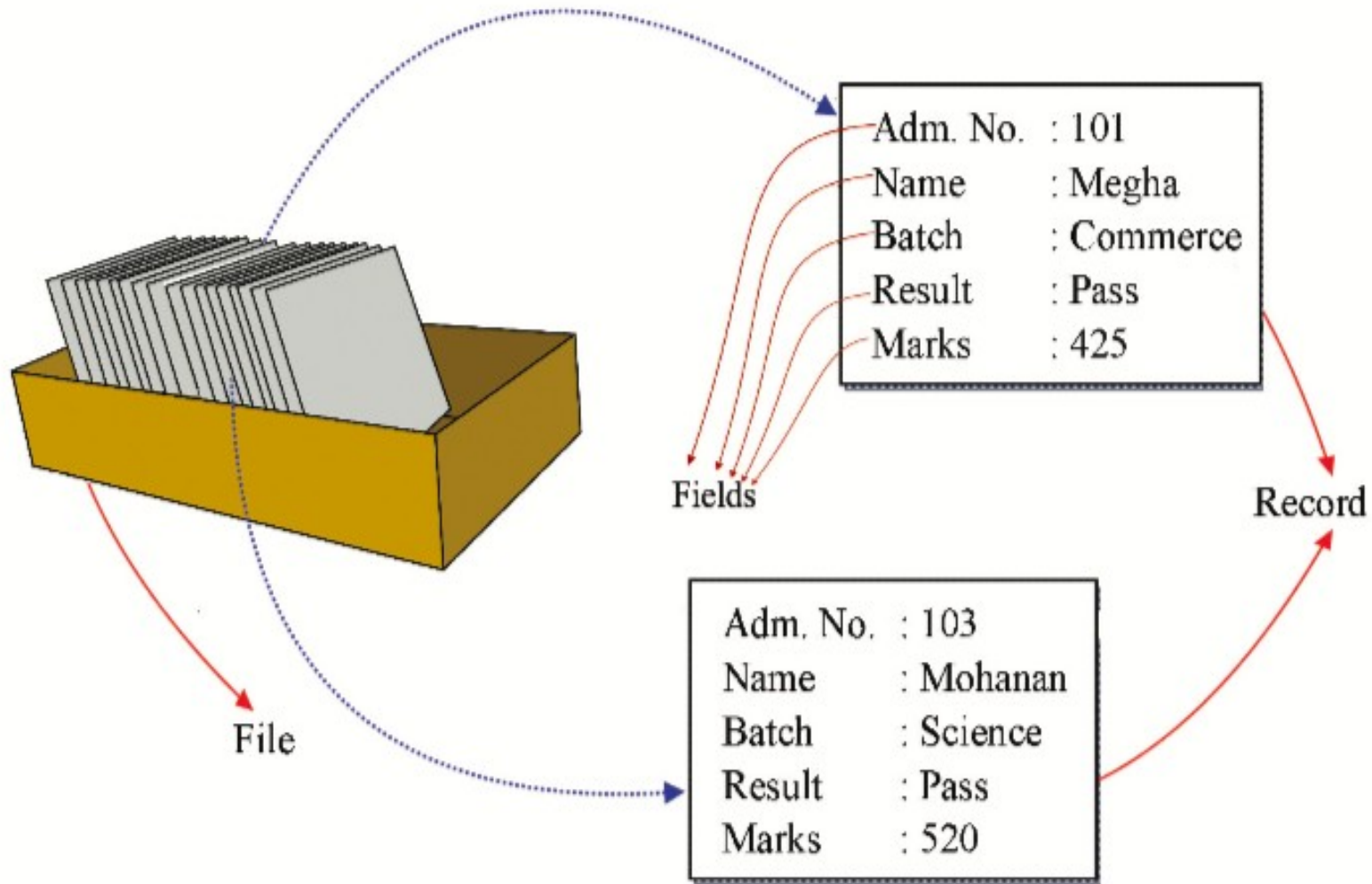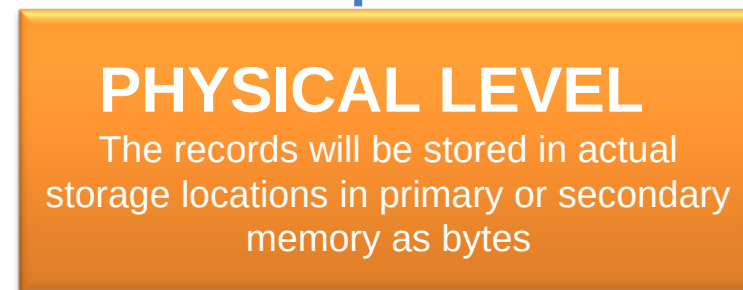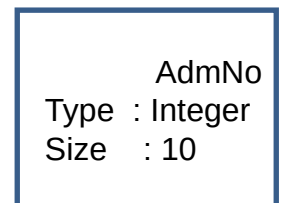Fig. 8.2: Concept of data organisation

END

What is database abstraction?

- The developers in a database system hide the complexity from users through different levels of abstraction. There are three types of abstraction

- a). **Physical Level Abstraction**-It is the lowest level of abstraction. Physical level describes how data is actually stored in the storage medium like disk, tape etc.. The DBMS hide its details from the programmers.

- b). **Logical (conceptual) Level Abstraction**- It is the next high level of abstraction. Logical level describes what data are stored in the database and what relationship among them. In this level the records and its types are defined. The logical level abstraction is done by the Data Base Administrator(DBA).

- c). **View Level Abstraction**- This is the highest level of abstraction. It concerned with the way in which the individual users view the data. Most of the users may not need all the data stored in the database. They need only a part of it. So all other data are hide from the user

Data Abstraction

VIEW LEVEL

Through each view users see only part of the database.

| View 1 | View 2 | View N |

LOGICAL LEVEL
The type of data and their relationships are specified

AdmNo
Type : Integer
Size : 10

PHYSICAL LEVEL
The records will be stored in actual storage locations in primary or secondary memory as bytes

010111
10

END
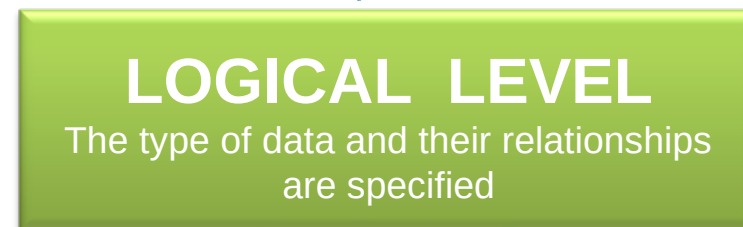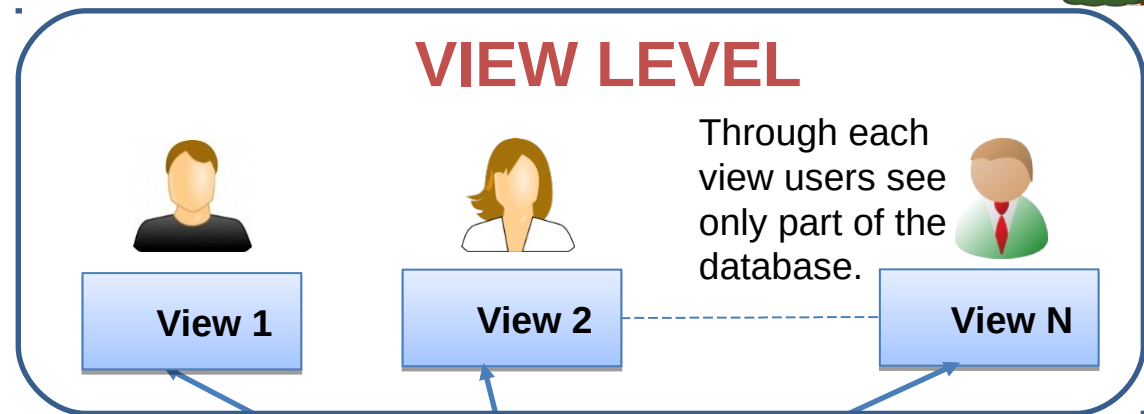
What is meant by data independence?.What are different types of data independence?

- The ability to modify a schema definition at one level (Physical, Logical, View levels) without affecting the schema definition at the next higher level is called data independence. The two types of data independence are ,

- a). **Physical data independence**: It is the ability to modify the physical schema without affecting the conceptual(logical) schema. So the application programs remain same even though the physical schema get modified

- b). **Logical data independence**: It is the ability to modify the logical schema without affecting the view level. So the application programs remains same.. It is more difficult to achieve because the application programs are heavily depend on logical structure.

Data Independence

VIEW LEVEL

Select * From Student;

Select Name, lAdmno from student';

LOGICAL INDEPENDENCE

No Change

LOGICAL LEVEL

Name: IAdmNo
Name: IAdmNo
Type : Integer
Type : Integer
Size : 10
Size : 10
Default :1001

Change in field definition

PHYSICAL INDEPENDENCE

No Change

PHYSICAL LEVEL

01011110

Storage location altered

END
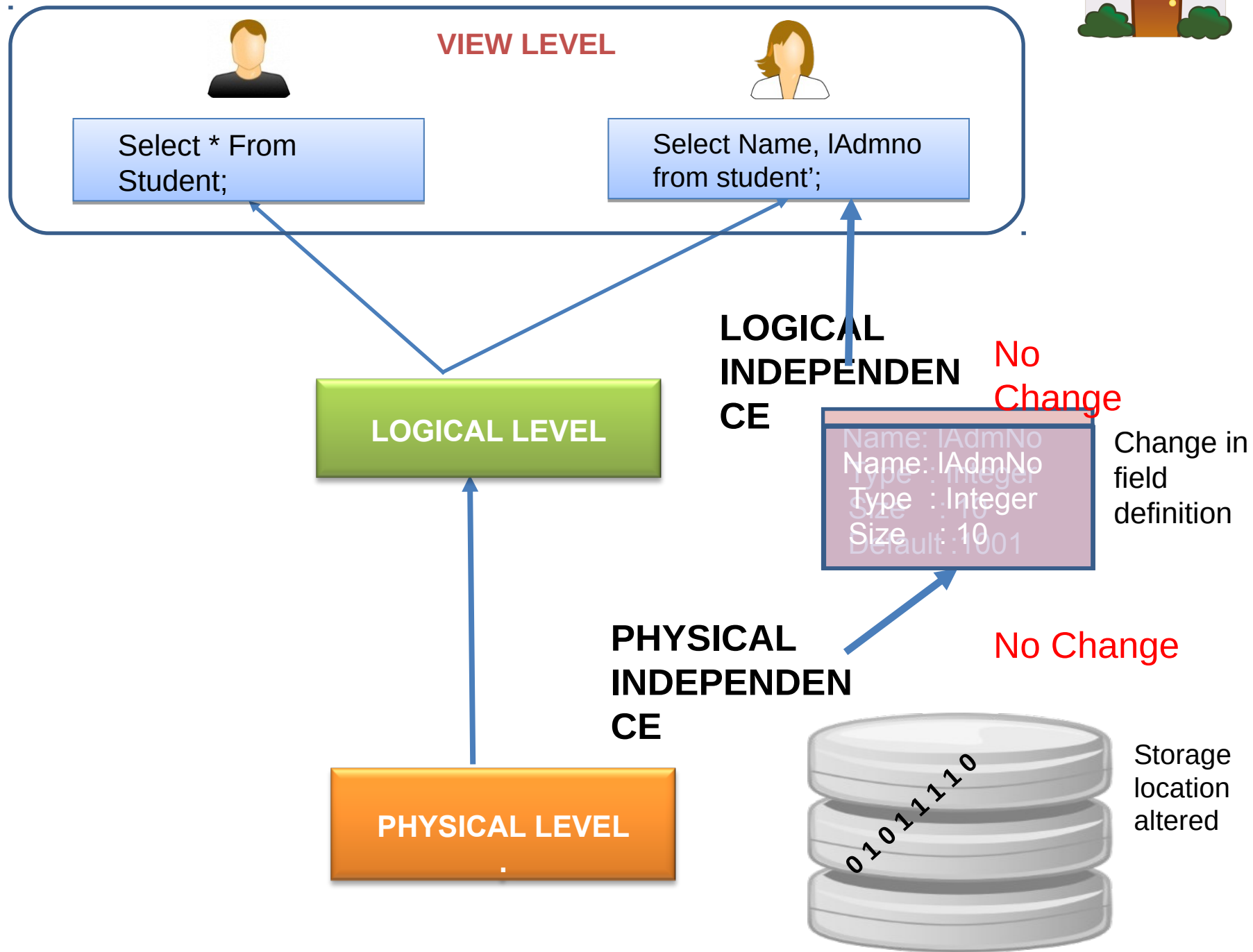
Who are the users of database?

- Users of the database can be classified in to the following groups,

- a) **Database administrator (DBA)**: The person has control over the centralized and shared data in a DBMS. He is the in charge of creating, modifying and maintaining 3 levels of DBMS.

- b) **Application Programmer**: They are computer professionals and interacts the DBMS through application programs. Application programs are programs written in any host languages like Java, C++ etc.. and interact with DBMS through Data Manipulation Language (DML).

- c). **Sophisticated Users** : The users are familiar with DBMS like scientist, business analyst etc. The interact with DBMS by their own Query or request

- d). **Naive Users/ End users**: The uses don't know about the physical and logical structures of DBMS like clerical staff in an office.

- **What is schema?**

- The description or structure of a database is called database schema.

What are the duties of a DBA?

- The person has control over the centralized and shared data in a DBMS.

- He is responsible for,

- **a). Design of the physical and conceptual Schemas.**

- **b). Security and authorization.**

- **c). Data availability and recovery from failures.**

RDBMS

What is a RDBMS or Relational Model ?

Relational Database Management System (RDBMS) represents the database as a collection of relations, each of which is assigned a unique name.

The relation resembles a table.

Relation 1

Data table 1: Project table

Project number	Description	Dept. number
155	Payroll	257
498	Widgets	632
226	Sales Manual	598

Relation 2

Data table 2: Department table

Dept. number	Dept. name	Manager SSN
257	Accounting	005-10-6321
632	Manufacturing	549-77-1001
598	Marketing	098-40-1370

Relation 3

Data table 3: Manager table

SSN	Last name	First name	Hire date	Dept. number
005-10-6321	Johns	Francine	10-07-1997	257
549-77-1001	Buckley	Bill	02-17-1979	632
098-40-1370	Fiske	Steven	01-05-1985	598

END

What are the various terminologies in RDBMS

1)**Entity:** An entity is a person or a thing that can be distinguishable from others. Eg. Student, Book etc..

2)**Relation** : Relation is a collection of data elements organized in rows and columns. It is also called a table.

3)**Tuple** : A row in a relation is called a tuple. It consists of a complete set of values used to represent a particular entity. Values of each attribute are taken from its domain.

4)**Attribute** : The columns of a relation are called attributes. Eg. Ad.no, name, age etc are the attributes of the relation Student.

5)**Degree** : The no of attributes(columns) in a relation is called degree of that relation.

6)**Cardinality** : The no of rows (tuples) in a relation is called cardinality of that relation

7)**Domain** :Domain is a range of values from which a particular value of an attribute for a can be drawn. Eg. 0 to 100 may be domain of values for the attribute mark in student relation.

8)**Instance** : An instance of a relation is a set of tuples has the same number of fields as the relational scheme.

Student

AdmNo	Roll	Name	Batch	Marks	Result
101	24	Sachin	Science	480	EHS
102	14	Rahul	Commerce	410	EHS
103	4	Fathima	Humanities	200	NHS
104	12	Mahesh	Commerce	180	NHS
105	24	Nelson	Humanities	385	EHS
106	8	Joseph	Commerce	350	EHS
107	24	Shaji	Humanities	205	NHS
108	2	Bincy	Science	300	EHS

Relation – Student
Attribute – AdmNo, RollNo,Name, Batch, Marks, Result
Degree-- 6
Cardinality – 8
Domain of Result – EHS, NHS
Domain of RollNo – Positive integer number

END

What are the different Keys in a relation?

- Tuples in a relation must be distinct. So there should be a way to identify each tuples in a relation. Keys allow us to make such distinctions.

- **a). Candidate Key:**-Candidate key for a relation is the minimal set of attributes that uniquely identifies each tuple (row) of the relation. There may more than one candidate key in a relation.

- **b). Primary Key:**-It is one of the candidate keys chosen to be unique identifier of the relation by the database designer. If there is more than one candidate key, then key with least no. of attributes is chosen as Primary key

- **c). Alternate Key:-**The candidate keys that are not chosen as primary key is known as alternate keys. In the above

- **d). Foreign key:-**An attribute or set of attributes in a table which is a primary key in another table is called a foreign key. It helps to link two or more tables. So it is also called Reference key.

What is relational algebra?

- The collection of operations that is used to manipulate the entire relations of a database is known as relational algebra.

- The relational operations takes one or two relation as input and produce a new relation as output.

What are the fundamental operations in Relational Algebra ?

- 1) **Select operation:** SELECT operation is used to select rows from a relation that satisfies a given predicate. The predicate is a user defined condition to select rows. It is denoted by the symbol σ (sigma)

- 2) **Project operation**: The project operation retrieves the columns from a relation. It is denoted by the symbol Π (pie)

- 3) **Union Operation**: It combines two table and forms a new relation. It is denoted by $\cup$ Duplication of tuples are eliminated.

- 4). **Intersection Operation**: The Intersection operation combines the common rows from two tables and forms a new relation. It is denoted by the symbol ∩

- 5) **Set Difference Operation**: The Set difference operation find tuple that are in one relation but are not in another relation. It is denoted by the symbol -

- 6)**Cartesian Product(Cross Product**): The Cartesian Product operation combines tuples from two relations. It is denoted by the symbol X

SELECT Operation

SELECT operation is used to select rows from a relation that satisfies a given predicate.

The predicate is a user defined condition to select rows.

It is denoted by the symbol σ (sigma)

Syntax : σ $_{condition}$ (Relation)

SELECT operation Examples

Student

AdmNo	Roll	Name	Batch	Marks	Result
101	24	Sachin	Science	480	EHS
102	14	Rahul	Commerce	410	EHS
103	4	Fathima	Humanities	200	NHS
104	12	Mahesh	Commerce	180	NHS
105	24	Nelson	Humanities	385	EHS
106	8	Joseph	Commerce	350	EHS
107	24	Shaji	Humanities	205	NHS
108	2	Bincy	Science	300	EHS

σ Batch="Commerce" (Student)

ANSWER

$$\sigma \text{ Batch="Commerce" (Student)}$$

AdmNo	Roll	Name	Batch	Marks	Result
102	14	Rahul	Commerce	410	EHS
104	12	Mahesh	Commerce	180	NHS
106	8	Joseph	Commerce	350	EHS

SELECT operation Examples

Student

AdmNo	Roll	Name	Batch	Marks	Result
101	24	Sachin	Science	480	EHS
102	14	Rahul	Commerce	410	EHS
103	4	Fathima	Humanities	200	NHS
104	12	Mahesh	Commerce	180	NHS
105	24	Nelson	Humanities	385	EHS
106	8	Joseph	Commerce	350	EHS
107	24	Shaji	Humanities	205	NHS
108	2	Bincy	Science	300	EHS

σ Admno<105 (Student)

ANSWER

σ Admno<105 (Student)

AdmNo	Roll	Name	Batch	Marks	Result
101	24	Sachin	Science	480	EHS
102	14	Rahul	Commerce	410	EHS
103	4	Fathima	Humanities	200	NHS
104	12	Mahesh	Commerce	180	NHS

Project Operation

The project operation retrieves the columns from a relation.

It is denoted by the symbol Π (pie)

Π column1, column2……. (Relation)

Example Project

Student

AdmNo	Roll	Name	Batch	Marks	Result
101	24	Sachin	Science	480	EHS
102	14	Rahul	Commerce	410	EHS
103	4	Fathima	Humanities	200	NHS
104	12	Mahesh	Commerce	180	NHS
105	24	Nelson	Humanities	385	EHS
106	8	Joseph	Commerce	350	EHS
107	24	Shaji	Humanities	205	NHS
108	2	Bincy	Science	300	EHS

Π **RollNo,Name (Student)**

Example Project

Π **RollNo,Name (Student)**

Roll	Name
24	Sachin
14	Rahul
4	Fathima
12	Mahesh
24	Nelson
8	Joseph
24	Shaji
2	Bincy

Example Project

1. Π Roll, Name (σ Admno < 105 (Student))

Roll	Name
24	Sachin
14	Rahul
4	Fathima
12	Mahesh

Union Operation

The Union operation combines two table and forms a new relation.

It is denoted by the symbol U

The union operation is valid only when the two relations are in union compatible.

Two relations are union-compatible, then they have the same number of attributes, and corresponding attributes have the same domain.

Syntax :   Relation1 U Relation2

Example Union

Arts

AdmNo	Name	BatchCode
101	Sachin	S2
103	Fathima	H2
106	Joseph	C2
110	Niktha	S1
132	Vivek	C1
154	Nevin	C1

Sports

AdmNo	Name	BatchCode
102	Rahul	C2
103	Fathima	H2
105	Nelson	H2
106	Joseph	C2
108	Bincy	S2
132	Vivek	C1
164	Rachana	S1

Example Union

Details students who participated in arts, sports or both

Arts U Sports

AdmNo	Name	BatchCode
101	Sachin	S2
103	Fathima	H2
106	Joseph	C2
110	Niktha	S1
132	Vivek	C1
154	Nevin	C1
102	Rahul	C2
105	Nelson	H2
108	Bincy	S2
132	Vivek	C1
164	Rachana	S1

Intersection Operation

The Intersection operation combines the common rows from two tables and forms a new relation.

It is denoted by the symbol ∩

The Intersection operation is valid only when the two relations are in union compatible.

Syntax :   Relation1 ∩ Relation2

Example Intersection

Arts

AdmNo	Name	BatchCode
101	Sachin	S2
103	Fathima	H2
106	Joseph	C2
110	Niktha	S1
132	Vivek	C1
154	Nevin	C1

Sports

AdmNo	Name	BatchCode
102	Rahul	C2
103	Fathima	H2
105	Nelson	H2
106	Joseph	C2
108	Bincy	S2
132	Vivek	C1
164	Rachana	S1

Example Intersection

Details students who participated in both arts and sports

- Arts ∩ Sports

AdmNo	Name	BatchCode
103	Fathima	H2
106	Joseph	C2
132	Vivek	C1

Table. 8.13: Relation of ARTS SPORTS

Set Difference Operation

The Set difference operation find tuple that are in one relation but are not in another relation.

It is denoted by the symbol -.

The union operation is valid only when the two relations are in union compatible.

Syntax :   Relation1 - Relation2

Example set difference

Details students who participated in arts but not in sports

Arts - Sports

AdmNo	Name	BatchCode
101	Sachin	S2
110	Nikitha	S1
154	Nevin	C1

Table. 8.14: Relation of ARTS - SPORTS

Example set difference

Details students who participated in sports but not in arts

Sports - Arts

AdmNo	Name	BatchCode
101	Rahul	C2
105	Nelson	H2
108	Bincy	S2
164	Rachana	S1

Table. 8.15: Relation of SPORTS - ARTS

Cartesian Product

The Cartesian Product operation combines tuples from two relations. It is denoted by the symbol X.

Syntax :   Relation1 X Relation2

If Relation1 has m rows and Relation2 has n rows then Relation1 X Relation2 has m x n rows

Cartesian Product Example

Student

AdmNo	Roll	Name	Batch	Marks	Result
101	24	Sachin	Science	480	EHS
102	14	Rahul	Commerce	410	EHS
103	4	Fathima	Humanities	200	NHS
104	12	Mahesh	Commerce	180	NHS
105	24	Nelson	Humanities	385	EHS
106	8	Joseph	Commerce	350	EHS
107	24	Shaji	Humanities	205	NHS
108	2	Bincy	Science	300	EHS

Teacher

TeacherId	Name	Dept
1001	Viswesaran	English
1002	Meenakshi	Computer

Cartesian Product Example

Student X Teacher

AdmNo	Roll	Name	Batch Code	Marks	Result	TeacherId	Name	Dept
101	24	Sachin	S2	480	EHS	1001	Viswesaran	English
101	24	Sachin	S2	480	EHS	1002	Meenakshi	Computer
102	14	Rahul	C2	410	EHS	1001	Viswesaran	English
102	14	Rahul	C2	410	EHS	1002	Meenakshi	Computer
103	4	Fathima	H2	200	NHS	1001	Viswesaran	English
103	4	Fathima	H2	200	NHS	1002	Meenakshi	Computer
104	12	Mahesh	C2	180	NHS	1001	Viswesaran	English
104	12	Mahesh	C2	180	NHS	1002	Meenakshi	Computer
105	24	Nelson	H2	385	EHS	1001	Viswesaran	English
105	24	Nelson	H2	385	EHS	1002	Meenakshi	Computer
106	8	Joseph	C2	350	EHS	1001	Viswesaran	English
106	8	Joseph	C2	350	EHS	1002	Meenakshi	Computer
107	24	Shaji	H2	205	NHS	1001	Viswesaran	English
107	24	Shaji	H2	205	NHS	1002	Meenakshi	Computer
108	2	Bincy	S2	300	EHS	1001	Viswesaran	English
108	2	Bincy	S2	300	EHS	1002	Meenakshi	Computer

END

Previous Questions

- 1. Explain any three advantages of DBMS

- 2. What is relational algebra ?Explain any three relational algebra operations.

- 3. The number of attributes in a relation is called ....
  (a) tuple (b) degree (c) cardinality (d) domain

- 4. Explain the components of DBMS.

- 5. Define the following
  a) Field
  b) Record

- 6. In RDBMS a relation contains 10 rows and 5 columns. What is the degree of the relation?

- 7. Explain the different levels of data abstraction in DBMS.

- 8. .... symbol is used for select operation in relational algebra.
  a). σ b). π c). υ d).ñ

- 9. Explain advantages of DBMS over conventional file system.

- 10.Define the term Data independence. Explain different levels of data independence

- 11. Explain any three fundamental operations in relational algebra

- 12. In DBMS
  a) Explain classification of database users
  b) Define the terms -Cardinality, Schema and Alternate key

- 13. What is key? Explain any two keys in a relational database management system