

Database Concepts

STUDY NOTES

- **Database:** A database is an organized collection of data.
- **DBMS (Database Management System):** A collection of programs that enables you to store, modify and extract information from a database.
- **Relational Database:** A database in which the data is stored in the form of relations (also called tables) is called a Relational Database.
- **RDBMS (Relational Database Management System):** A DBMS used to manage Relational Databases is called an RDBMS (Relational Database Management System). Some popular RDBMS software available are Oracle, MySQL, Sybase, Ingress.
- **Relation:** A relation is two dimension table.
- **Attributes:** The columns of a table are called attributes.
- **Tuples:** The rows of a table are called tuples.
- **Degree:** The number of attributes in a relation is called the degree of the relation.
- **Cardinality:** The number of tuples (rows) in the relation is called its cardinality.
- **Primary Key:** The group of one or more column used to uniquely identify each row of a relation is called its Primary Key.
- **Candidate Key:** A column or a group of columns which can be used as the primary key of relation is called a candidate key.
- **Alternate key:** A candidate key of table which is not made its primary key is called its Alternate key.
- **Foreign Key:** A foreign key is a column in a table where that column is a primary key of another table.
- **Referential Integrity:** This property of a relational database which ensures that no entry in a foreign key column of a table can be made unless it matches a primary key value in the corresponding related table is called Referential Integrity.
- **File System vs DBMS: Key Differences:** A file system is a technique of arranging the files in a storage medium like a hard disk, pen drive, DVD, etc. It helps you to organize the data and allows easy retrieval of files when they are required. Database Management System (DBMS) is a software for storing and retrieving user's data while considering appropriate security measures. It consists of a group of programs that manipulate the database.
- **Categories of SQL:**
 - (i) **DDL (Data Definition Language):** This is the category of SQL commands. All the commands which are used to create, destroy or restructure databases and tables come under this category. Examples of DDL commands are -CREATE, DROP and ALTER.
 - (ii) **DML (Data Manipulation Language):** This is a category of SQL commands. All the commands which are used to manipulate data within tables comes under this category. Examples: Insert, Update and Delete.
 - (iii) **DCL (Data Control Language):** All the commands which are used to control the access to databases and tables fall under this category. Examples -Grant and Revoke.
- **Benefits of using a DBMS are:** Redundancy can be controlled, Inconsistency can be avoided, Data can be shared, Security restrictions can be applied.

- **Relational algebra:**

- ❖ **Selection:** Operation upon a relation to select a horizontal subset of the relation.

- ❖ **Projection:** Operation upon a relation to select a vertical subset of the relation.

Example:

Table: Employees

ENO	ENAME	DOJ	DNO
E1	NUSRAT	2001-11-21	D3
E2	KABIR	2005-10-25	D1

A selection upon Employees for tuples whose DOJ is in the year 2005 will result in

Table: Employees

ENO	ENAME	DOJ	DNO
E2	KABIR	2005-10-25	D1

A projection upon Employees for ENAME and DOJ of all Employees will result into

Table: Employees

ENAME	DOJ
NUSRAT	2001-11-21
KABIR	2005-10-25

- **UNION:** UNION is used to combine the results of two or more Select statements. However it will eliminate duplicate rows from its result set. In case of union, number of columns and datatype must be same in both the tables.

Example of UNION

The First table,

Rollno	Name
1	abhi
2	adam

The **Second** table,

Roll	Name
2	adam
3	Chester

Union SQL query will be,

```
select * from First
UNION
select * from second
```

The result table will look like,

Roll	NAME
1	abhi
2	adam
3	Chester

The UNION operator is used to combine the result-set of two or more SELECT statements.

- Each SELECT statement within UNION must have the same number of columns.
- The columns must also have similar data types.
- The columns in each SELECT statement must also be in the same order.

- **MINUS:** The MINUS compares the results of two queries and returns distinct rows from the result set of the first query that does not appear in the result set of the second query.

The following illustrates the syntax of the MINUS operator:

```
SELECT select_list1
FROM table_name1
MINUS
SELECT select_list2
FROM table_name2;
```

The basic rules for a query that uses MINUS operator are the following:

- (i) The number and order of columns in both select_list1 and select_list2 must be the same.
- (ii) The data types of the corresponding columns in both queries must be compatible.

- **Cartesian Product:** A Cartesian join or Cartesian product is a join of every row of one table to every row of another table. This normally happens when no matching join columns are specified. For example, if table A with 100 rows is joined with table B with 1000 rows, a Cartesian join will return 100,000 rows.

Note: A Cartesian product may indicate a missing join condition.

Syntax:

The basic syntax of the CARTESIAN JOIN or the CROSS JOIN is as follows –

```
SELECT table1.column1, table2.column2...
FROM table1, table2 [, table3 ]
```

Example

Consider the following two tables.

Table 1: CUSTOMERS table is as follows.

ID	NAME	AGE
1	Ramesh	32
2	Khilan	25
3	kaushik	23

Table 2: PRODUCT Table is as follows –

PID	CUSTOMER_ID	PRICE
100	3	150
101	2	200

Now, let us join these two tables using INNER JOIN as follows –

```
SQL> SELECT ID, NAME, PRICE FROM CUSTOMERS, PRODUCT;
```

This would produce the following result –

ID	NAME	PRICE
1	Ramesh	150
1	Ramesh	200
2	Khilan	150
2	Khilan	200
3	kaushik	150
3	kaushik	200