**DATA SCIENCE QUERY**

**Data science :**

**Data Science** is the process of extracting useful insights from raw data using statistics, machine learning, and programming.

**DATABASE MANAGEMENT SYSTEM (DBMS):**

**DDL (Data Definition Language)** – Defines and modifies database structure.

Commands: CREATE, ALTER, DROP, TRUNCATE

**DML (Data Manipulation Language)** – Handles data within tables.

Commands: INSERT, UPDATE, DELETE, SELECT

**DCL (Data Control Language)** – Controls access to the database.

Commands: GRANT, REVOKE

**TCL (Transaction Control Language)** – Manages database transactions.

Commands: COMMIT, ROLLBACK, SAVEPOINT

**DQL (Data Query Language)** in DBMS is used to retrieve data from the database.

**Command:** SELECT

**Example :**

**Database creation:**

postgres=# create database datascience;

**Connect to database :**

postgres=# \c datascience

**Table creation :**

datascience=# create table student(rollno int,name varchar(50),age int,dept varchar(30),year int,gender varchar(20));

**select the table :**

datascience=# select\*from student;

rollno | name | age | dept | year | gender

--------+------+-----+------+------+--------

(0 rows)

**Alter table name :**

datascience=# alter table student rename stud\_details;

datascience=# alter table student rename to stud\_details;

ALTER TABLE

datascience=# select \*from stud\_details;

rollno | name | age | dept | year | gender

--------+------+-----+------+------+--------

(0 rows)

**Add column :**

datascience=# alter table stud\_details add column emailid varchar(30);

ALTER TABLE

datascience=# select \*from stud\_details;

rollno | name | age | dept | year | gender | emailid

--------+------+-----+------+------+--------+---------

(0 rows)

**Drop column :**

datascience=# alter table stud\_details drop emailid;

ALTER TABLE

datascience=# select \*from stud\_details;

rollno | name | age | dept | year | gender

--------+------+-----+------+------+--------

(0 rows)

**Insert the values:**

datascience=# insert into stud\_details values(1,'meerab',23,'AI&DS',2020,'female'),(2,'murtasim',27,'AI&DS',2020,'male'),(3,'haya',23,'IT',2020,'fem

ale');

INSERT 0 3

datascience=# select \*from stud\_details;

rollno | name | age | dept | year | gender

--------+----------+-----+-------+------+--------

1 | meerab | 23 | AI&DS | 2020 | female

2 | murtasim | 27 | AI&DS | 2020 | male

3 | haya | 23 | IT | 2020 | female

**Update the values:**

datascience=# update stud\_details set name='marium' where rollno=3;

UPDATE 1

datascience=# select \*from stud\_details;

rollno | name | age | dept | year | gender

--------+----------+-----+-------+------+--------

1 | meerab | 23 | AI&DS | 2020 | female

2 | murtasim | 27 | AI&DS | 2020 | male

3 | marium | 23 | IT | 2020 | female

**Delete the values(row):**

datascience=# delete from stud\_details where rollno=3;

DELETE 1

datascience=# select \*from stud\_details;

rollno | name | age | dept | year | gender

--------+----------+-----+-------+------+--------

1 | meerab | 23 | AI&DS | 2020 | female

2 | murtasim | 27 | AI&DS | 2020 | male

**Drop the values(column):**

datascience=# alter table stud\_details drop column year;

ALTER TABLE

datascience=# select \*from stud\_details;

rollno | name | age | dept | gender

--------+----------+-----+-------+--------

1 | meerab | 23 | AI&DS | female

2 | murtasim | 27 | AI&DS | male

**Order by :**

datascience=# select \*from stud\_details order by rollno;

rollno | name | age | dept | gender

--------+----------+-----+-------+--------

1 | meerab | 23 | AI&DS | female

2 | murtasim | 27 | AI&DS | male

3 | haya | 23 | IT | female

**Descending order:**

datascience=# select \*from stud\_details order by name desc;

rollno | name | age | dept | gender

--------+----------+-----+-------+--------

2 | murtasim | 27 | AI&DS | male

1 | meerab | 23 | AI&DS | female

3 | haya | 23 | IT | female

**Ascending order:**

datascience=# select \*from stud\_details order by name asc;

rollno | name | age | dept | gender

--------+----------+-----+-------+--------

3 | haya | 23 | IT | female

1 | meerab | 23 | AI&DS | female

2 | murtasim | 27 | AI&DS | male

**Limits :**

datascience=# select \*from stud\_details limit 2;

rollno | name | age | dept | gender

--------+----------+-----+-------+--------

2 | murtasim | 27 | AI&DS | male

1 | meerab | 23 | AI&DS | female

**Check whether values are null:**

datascience=# select \*from stud\_details where year isnull;

rollno | name | age | dept | gender | year

--------+----------+-----+-------+--------+------

2 | murtasim | 27 | AI&DS | male |

**Particular column:**

datascience=# select name from stud\_details;

name

----------

murtasim

meerab

haya

**Like keyword:(starting letter)**

datascience=# select \*from stud\_details where name like 'm%';

rollno | name | age | dept | gender | year

--------+----------+-----+-------+--------+------

1 | meerab | 23 | AI&DS | female | 2020

2 | murtasim | 27 | AI&DS | male | 2020

**Like keyword:(ending letter)**

datascience=# select \*from stud\_details where name like '%a';

rollno | name | age | dept | gender | year

--------+------+-----+------+--------+------

3 | haya | 23 | IT | female | 2020

**Like keyword:(middle letter)**

datascience=# select \*from stud\_details where name like '%e%';

rollno | name | age | dept | gender | year

--------+--------+-----+-------+--------+------

1 | meerab | 23 | AI&DS | female | 2020

**In keyword :**

datascience=# SELECT \* FROM stud\_details WHERE dept IN ('AI&DS');

rollno | name | age | dept | gender | year

--------+----------+-----+-------+--------+------

1 | meerab | 23 | AI&DS | female | 2020

2 | murtasim | 27 | AI&DS | male | 2020

(2 rows)

**Between Keyword:**

**Where :**

datascience=# SELECT \* FROM stud\_details WHERE age between 20 and 25;

rollno | name | age | dept | gender | year

--------+--------+-----+-------+--------+------

1 | meerab | 23 | AI&DS | female | 2020

3 | haya | 23 | IT | female | 2020

(2 rows)

**Having:**

datascience=# select dept,count(\*)as count from stud\_details group by dept having count(\*)>1;

dept | count

-------+-------

AI&DS | 2

**(1 row)**

**Aggregate Functions :**

**Sum:**

datascience=# select sum(score)from student;

sum

-----

250

(1 row)

**Average:**

datascience=# select avg(score)from student;

avg

---------------------

83.3333333333333333

**Count:**

datascience=# select count(score)from student;

count

-------

3

(1 row)

**Min and Max:**

datascience=# select min(score)from student;

min

-----

60

(1 row)

datascience=# select max(score)from student;

max

-----

100

(1 row)

**Inner join :**

datascience=# select \*from stud\_details s inner join student k on s.rollno =k.rollno;

rollno | name | age | dept | gender | year | rollno | name | markid | score

--------+----------+-----+-------+--------+------+--------+-------+--------+-------

1 | meerab | 23 | AI&DS | female | 2020 | 1 | abi | 1 | 90

2 | murtasim | 27 | AI&DS | male | 2020 | 2 | efa | 2 | 60

3 | haya | 23 | IT | female | 2020 | 3 | zinho | 3 | 100

(3 rows)

**Left join & right join:**

datascience=# select \*from stud\_details s left join student k on s.rollno =k.rollno;

rollno | name | age | dept | gender | year | rollno | name | markid | score

--------+----------+-----+-------+--------+------+--------+-------+--------+-------

1 | meerab | 23 | AI&DS | female | 2020 | 1 | abi | 1 | 90

2 | murtasim | 27 | AI&DS | male | 2020 | 2 | efa | 2 | 60

3 | haya | 23 | IT | female | 2020 | 3 | zinho | 3 | 100

(3 rows)

datascience=# select \*from stud\_details s right join student k on s.rollno =k.rollno;

rollno | name | age | dept | gender | year | rollno | name | markid | score

--------+----------+-----+-------+--------+------+--------+--------+--------+-------

1 | meerab | 23 | AI&DS | female | 2020 | 1 | abi | 1 | 90

2 | murtasim | 27 | AI&DS | male | 2020 | 2 | efa | 2 | 60

3 | haya | 23 | IT | female | 2020 | 3 | zinho | 3 | 100

| | | | | | 4 | virat | 3 | 90

| | | | | | 5 | enanya | 5 | 80

| | | | | | 6 | zanya | 6 | 100

**Ranking :**

datascience=# select name,age ,rank() over(order by age desc) from stud\_details;

name | age | rank

----------+-----+------

murtasim | 27 | 1

meerab | 23 | 2

haya | 23 | 2

**Dense Ranking :(does not skip numbers)**

datascience=# select name,age ,dense\_rank() over(order by age desc) from stud\_details;

name | age | dense\_rank

----------+-----+------------

murtasim | 27 | 1

meerab | 23 | 2

haya | 23 | 2

(3 rows)

Since **Meerab and Haya** have the same age, they get **rank 2**, and the next rank is **4 (not 3)**.

DENSE\_RANK() **does not skip numbers**, so the next rank is **3 (not 4)**.

**Unique Numbers :**

datascience=# SELECT name, age, ROW\_NUMBER() OVER (ORDER BY age DESC) AS row\_num FROM stud\_details;

name | age | row\_num

----------+-----+---------

murtasim | 27 | 1

meerab | 23 | 2

haya | 23 | 3