Employee Data Analysis

Problem statement:

You are a part of the HR department in a company, and you have been asked to extract, update, and delete the employees' details to maintain the records for further analysis.

Objective:

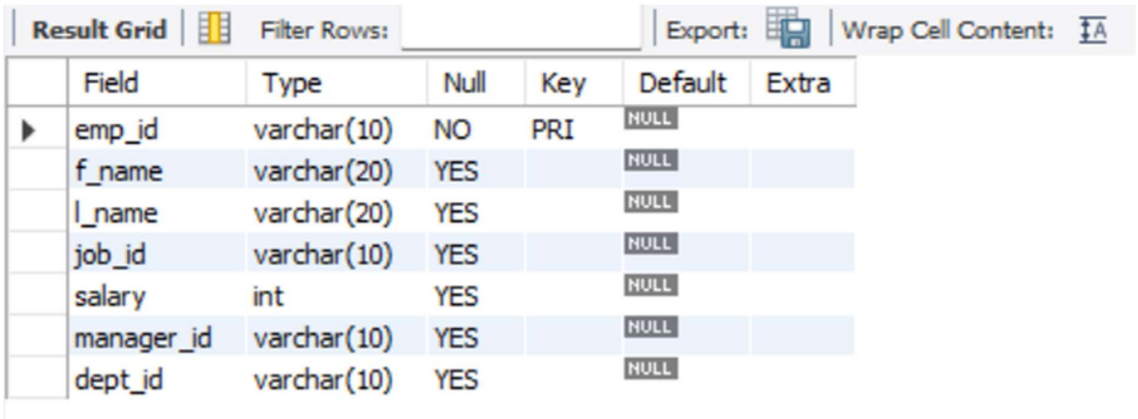
The objective is to design a database to analyze the performance of the employees on a quarterly basis.

1. **-- create a database named Practice:**
   create database practice;

2. **-- select database:**
   use practice;

3. **-- Write a query to create an employee table with employee ID, first name, last name, job ID, salary, manager ID, and department ID fields:**
   create table employee_datasets (
   emp_id varchar(10) not null primary key,
   f_name varchar(20),
   l_name varchar(20),
   job_id varchar(10),
   salary int,
   manager_id varchar(10),
   dept_id varchar(10));

   describe employee_datasets;

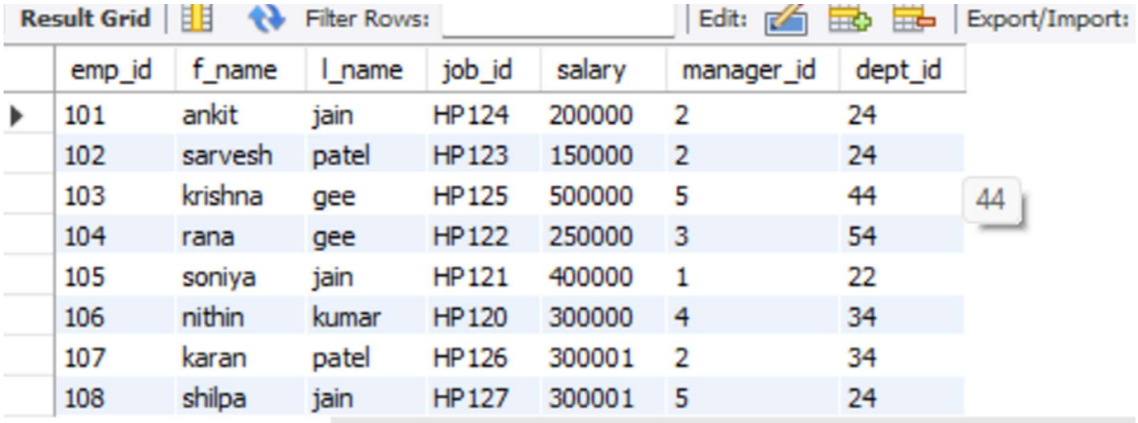


	Field	Type	Null	Key	Default	Extra
▶	emp_id	varchar(10)	NO	PRI	NULL	
	f_name	varchar(20)	YES		NULL	
	l_name	varchar(20)	YES		NULL	
	job_id	varchar(10)	YES		NULL	
	salary	int	YES		NULL	
	manager_id	varchar(10)	YES		NULL	
	dept_id	varchar(10)	YES		NULL	

4. **-- Write a query to insert values into the employee table:**
   insert into employee_datasets (
   emp_id, f_name, l_name, job_id, salary, manager_id, dept_id)
   values

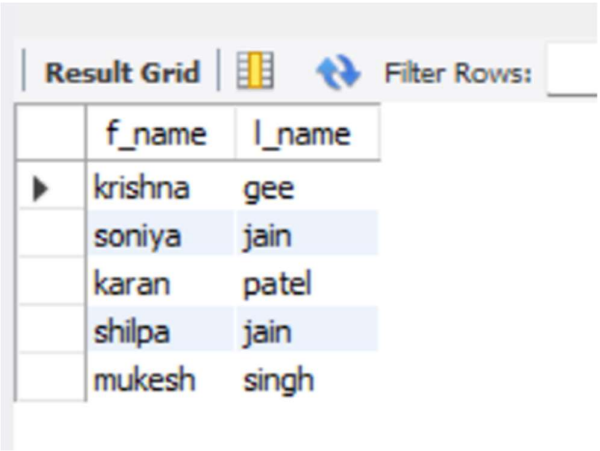
('101', 'ankit', 'jain', 'HP124', 200000, 2, 24),
('102', 'sarvesh', 'patel', 'HP123', 150000, 2, 24),
('103', 'krishna', 'gee', 'HP125', 500000, 5, 44),
('104', 'rana', 'gee', 'HP122', 250000, 3, 54),
('105', 'soniya', 'jain', 'HP121', 400000, 1, 22),
('106', 'nithin', 'kumar', 'HP120', 300000, 4, 34),
('107', 'karan', 'patel', 'HP126', 300001, 2, 34),
('108', 'shilpa', 'jain', 'HP127', 300001, 5, 24),
('109', 'mukesh', 'singh', 'HP128', 300001, 4, 44);

select * from employee_datasets;


Result Grid | Filter Rows: | Edit: | Export/Import:

emp_id	f_name	l_name	job_id	salary	manager_id	dept_id
101	ankit	jain	HP124	200000	2	24
102	sarvesh	patel	HP123	150000	2	24
103	krishna	gee	HP125	500000	5	44
104	rana	gee	HP122	250000	3	54
105	soniya	jain	HP121	400000	1	22
106	nithin	kumar	HP120	300000	4	34
107	karan	patel	HP126	300001	2	34
108	shilpa	jain	HP127	300001	5	24

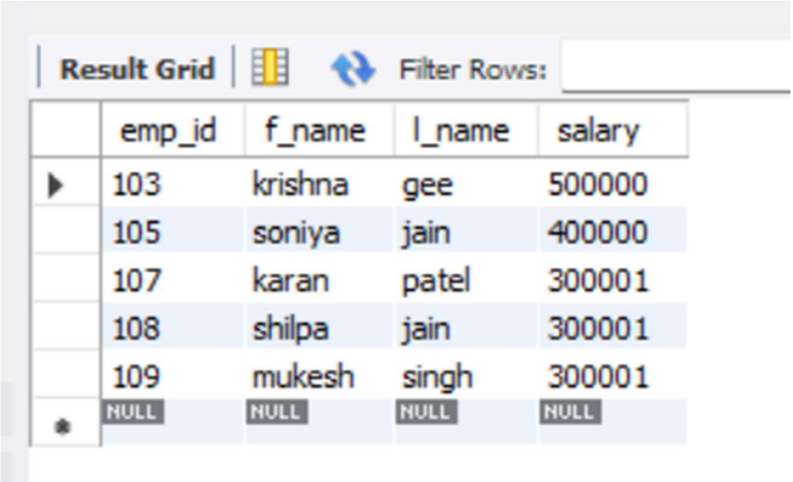
5. **-- Write a query to find the first and last names of every employee whose salary is higher than the employee with the last name Kumar:**
select f_name, l_name from employee_datasets
where salary> (select salary from employee_datasets where l_name ='kumar');


Result Grid | Filter Rows:

f_name	l_name
krishna	gee
soniya	jain
karan	patel
shilpa	jain
mukesh	singh

6. **-- Write a query to display the employee ID, first name, last name and salary of every employee whose salary is greater than the average:**
select emp_id, f_name, l_name, salary from employee_datasets

where salary> (select avg(salary) from employee_datasets);

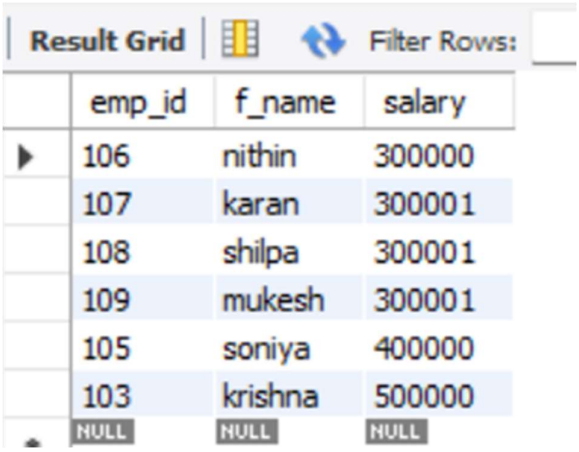


Result Grid | Filter Rows:

emp_id	f_name	l_name	salary
103	krishna	gee	500000
105	soniya	jain	400000
107	karan	patel	300001
108	shilpa	jain	300001
109	mukesh	singh	300001
NULL	NULL	NULL	NULL

7. **-- Write a query to display the employee ID and first name of every employee whose salary is higher than the salary of the JOB_ID = HP122 and sort the results in the ascending order of the salary:**

```
select emp_id, f_name, salary from employee_datasets
where salary > ALL (select salary from employee_datasets where job_id ='HP122')
order by salary;
```

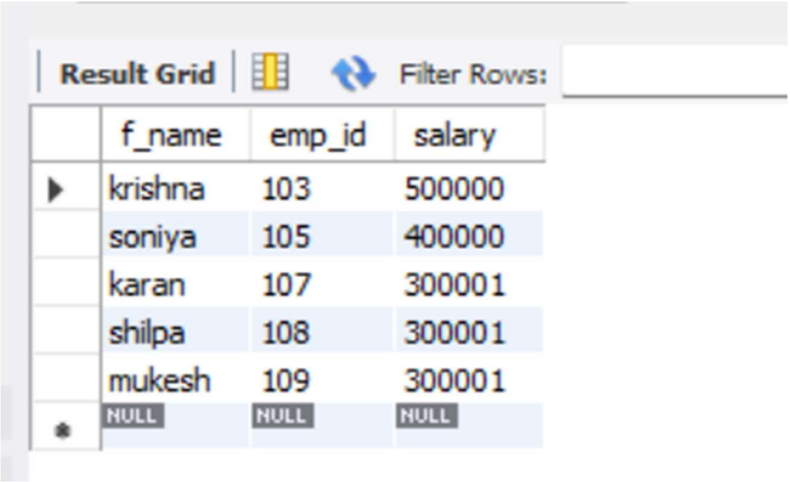


Result Grid | Filter Rows:

emp_id	f_name	salary
106	nithin	300000
107	karan	300001
108	shilpa	300001
109	mukesh	300001
105	soniya	400000
103	krishna	500000
NULL	NULL	NULL

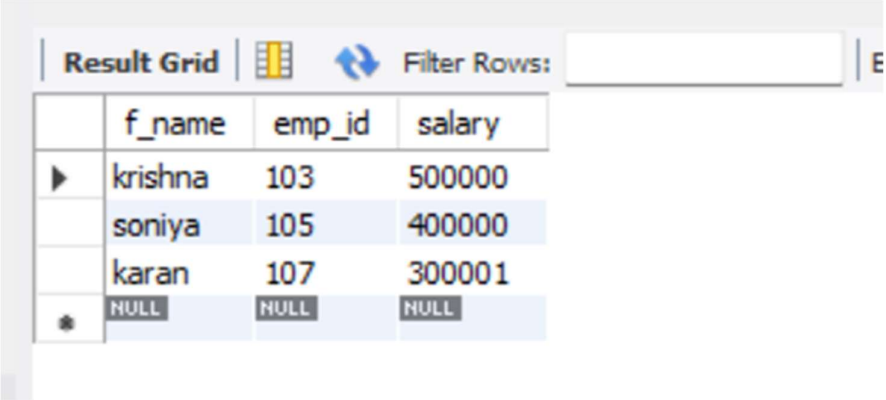
8. **-- Write a query to display the first name, employee ID, and salary of the three employees with the highest salaries:**

```
select distinct f_name, emp_id, salary from employee_datasets a
where 3>= (select count(distinct salary) from employee_datasets b
where b.salary >= a.salary) order by a.salary desc;
```



f_name	emp_id	salary
krishna	103	500000
soniya	105	400000
karan	107	300001
shilpa	108	300001
mukesh	109	300001
NULL	NULL	NULL

-- or
select f_name, emp_id, salary from employee_datasets
order by salary desc
limit 3;



f_name	emp_id	salary
krishna	103	500000
soniya	105	400000
karan	107	300001
NULL	NULL	NULL

9. **-- add 3 new columns to table i.e. role, department and employee rating:**
alter table employee_datasets
Add role varchar(40),
Add dept varchar(40),
Add emp_rating int;

describe employee_datasets;

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Field	Type	Null	Key	Default	Extra
l_name	varchar(20)	YES		NULL	
job_id	varchar(10)	YES		NULL	
salary	int	YES		NULL	
manager_id	varchar(10)	YES		NULL	
dept_id	varchar(10)	YES		NULL	
role	varchar(40)	YES		NULL	
dept	varchar(40)	YES		NULL	
emp_rating	int	YES		NULL	

select * from employee_datasets;

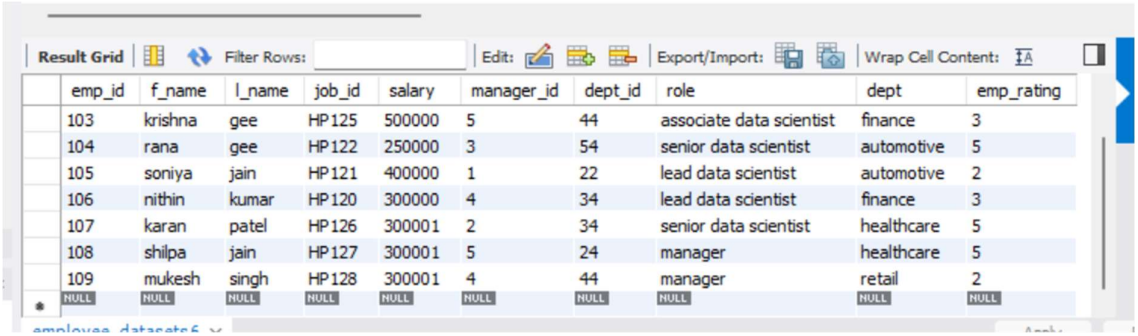
Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Co

emp_id	f_name	l_name	job_id	salary	manager_id	dept_id	role	dept	emp_rating
101	ankit	jain	HP124	200000	2	24	NULL	NULL	NULL
102	sarvesh	patel	HP123	150000	2	24	NULL	NULL	NULL
103	krishna	gee	HP125	500000	5	44	NULL	NULL	NULL
104	rana	gee	HP122	250000	3	54	NULL	NULL	NULL
105	soniya	jain	HP121	400000	1	22	NULL	NULL	NULL
106	nithin	kumar	HP120	300000	4	34	NULL	NULL	NULL
107	karan	patel	HP126	300001	2	34	NULL	NULL	NULL
108	shilpa	jain	HP127	300001	5	24	NULL	NULL	NULL

10. **-- insert values in new columns:**

```
update employee_datasets set role="senior data scientist", dept="retail", emp_rating=3
where emp_id='101';
update employee_datasets set role="junior data scientist", dept="retail", emp_rating=4
where emp_id='102';
update employee_datasets set role="associate data scientist", dept="finance", emp_rating=3
where emp_id='103';
update employee_datasets set role="senior data scientist", dept="automotive",
emp_rating=5 where emp_id='104';
update employee_datasets set role="lead data scientist", dept="automotive", emp_rating=2
where emp_id='105';
update employee_datasets set role="lead data scientist", dept="finance", emp_rating=3
where emp_id='106';
update employee_datasets set role="senior data scientist", dept="healthcare",
emp_rating=5 where emp_id='107';
update employee_datasets set role="manager", dept="healthcare", emp_rating=5 where
emp_id='108';
update employee_datasets set role="manager", dept="retail", emp_rating=2 where
emp_id='109';

select * from employee_datasets;
```

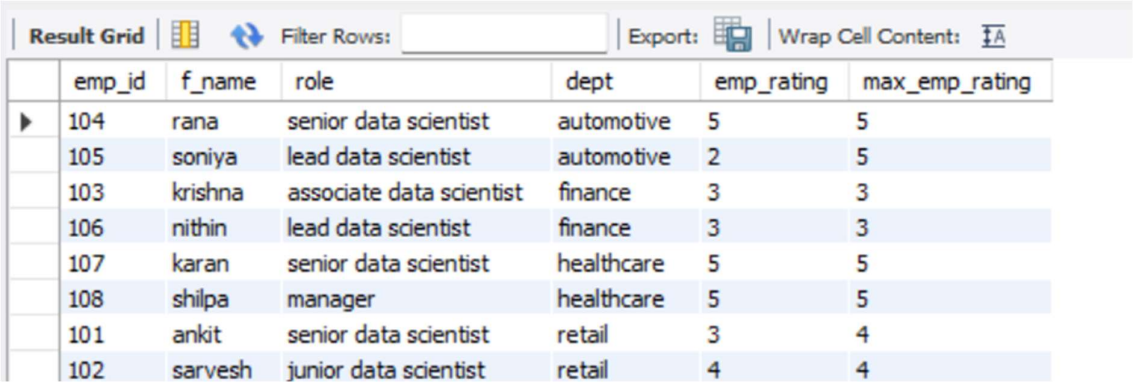


Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content:

emp_id	f_name	l_name	job_id	salary	manager_id	dept_id	role	dept	emp_rating
103	krishna	gee	HP125	500000	5	44	associate data scientist	finance	3
104	rana	gee	HP122	250000	3	54	senior data scientist	automotive	5
105	soniya	jain	HP121	400000	1	22	lead data scientist	automotive	2
106	nithin	kumar	HP120	300000	4	34	lead data scientist	finance	3
107	karan	patel	HP126	300001	2	34	senior data scientist	healthcare	5
108	shilpa	jain	HP127	300001	5	24	manager	healthcare	5
109	mukesh	singh	HP128	300001	4	44	manager	retail	2
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL

11. **-- SELECT EMP_ID, FIRST_NAME,ROLE,DEPT,EMP_RATING and calculate the maximum EMP_RATING in a department from the employee table using partition clause on department, Max function:**

    select emp_id, f_name, role, dept, emp_rating, max(emp_rating) over (partition by dept) as max_emp_rating from employee_datasets;

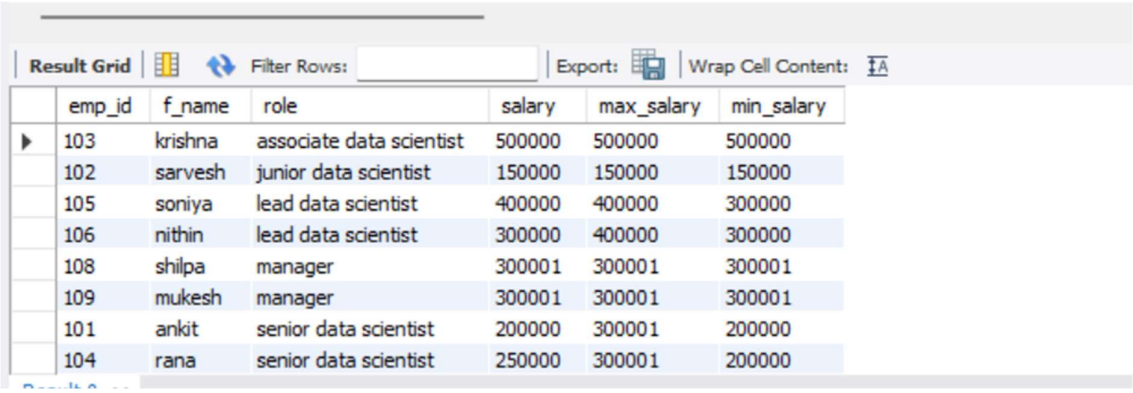


Result Grid | Filter Rows: | Export: | Wrap Cell Content:

emp_id	f_name	role	dept	emp_rating	max_emp_rating
104	rana	senior data scientist	automotive	5	5
105	soniya	lead data scientist	automotive	2	5
103	krishna	associate data scientist	finance	3	3
106	nithin	lead data scientist	finance	3	3
107	karan	senior data scientist	healthcare	5	5
108	shilpa	manager	healthcare	5	5
101	ankit	senior data scientist	retail	3	4
102	sarvesh	junior data scientist	retail	4	4

12. **-- Display the employee's ID, first name, role, and salary by finding the minimum and maximum salary of the employees using PARTITION BY clause, MIN, and MAX functions on role and salary fields:**

    select emp_id, f_name, role, salary, max(salary) over (partition by role) as max_salary, min(salary) over (partition by role) as min_salary from employee_datasets;

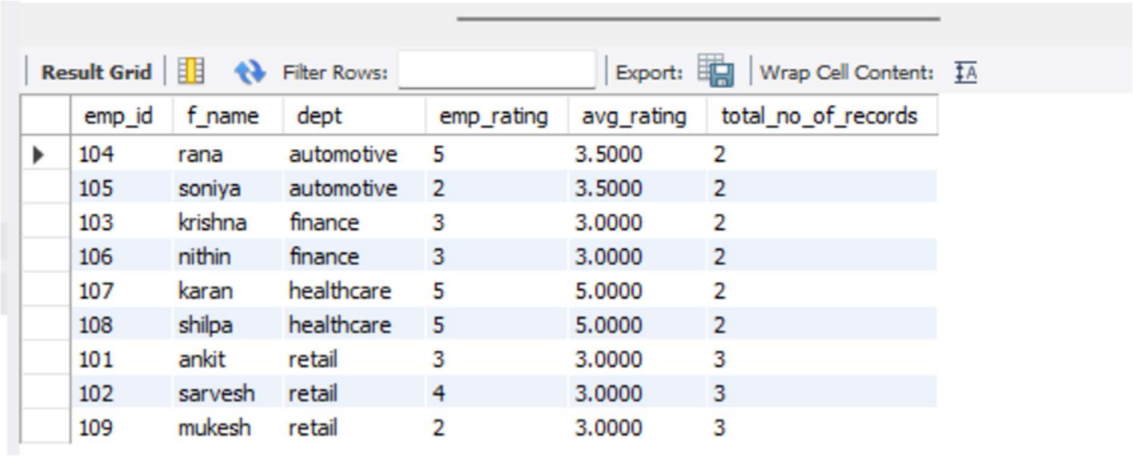


Result Grid | Filter Rows: | Export: | Wrap Cell Content:

emp_id	f_name	role	salary	max_salary	min_salary
103	krishna	associate data scientist	500000	500000	500000
102	sarvesh	junior data scientist	150000	150000	150000
105	soniya	lead data scientist	400000	400000	300000
106	nithin	lead data scientist	300000	400000	300000
108	shilpa	manager	300001	300001	300001
109	mukesh	manager	300001	300001	300001
101	ankit	senior data scientist	200000	300001	200000
104	rana	senior data scientist	250000	300001	200000

13. **-- Display the employee's ID, first name, department, and employee rating by calculating the average employee rating and the total number of records in a department using PARTITION BY clause, AVG, and COUNT functions on department and employee rating fields:**

```
select emp_id, f_name, dept, emp_rating, avg(emp_rating) over (partition by dept) as
avg_rating, count(*) over (partition by dept) as total_no_of_records from
employee_datasets;
```

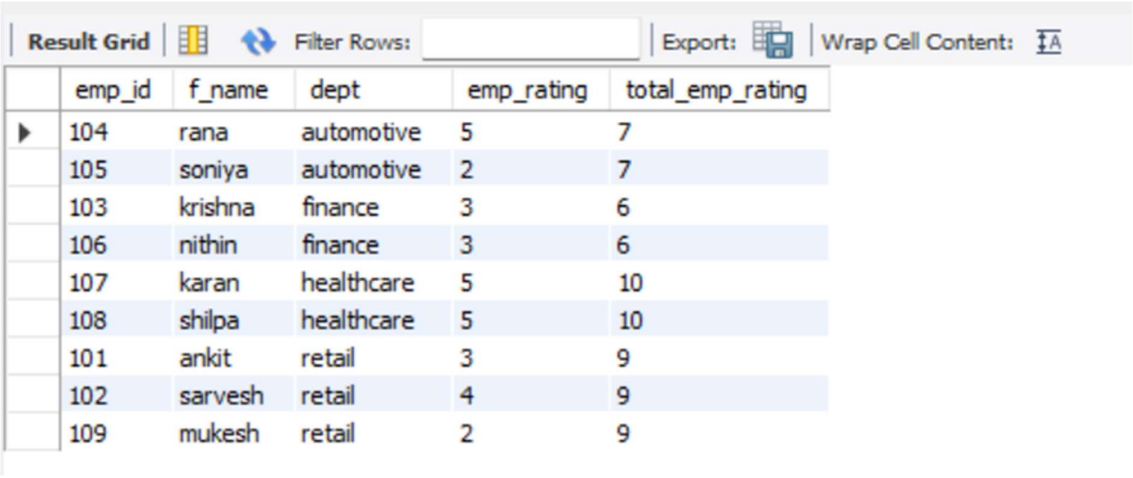


Result Grid | Filter Rows: | Export: | Wrap Cell Content:

emp_id	f_name	dept	emp_rating	avg_rating	total_no_of_records
104	rana	automotive	5	3.5000	2
105	soniya	automotive	2	3.5000	2
103	krishna	finance	3	3.0000	2
106	nithin	finance	3	3.0000	2
107	karan	healthcare	5	5.0000	2
108	shilpa	healthcare	5	5.0000	2
101	ankit	retail	3	3.0000	3
102	sarvesh	retail	4	3.0000	3
109	mukesh	retail	2	3.0000	3

14. **-- Display the employee's Id, first name, department, and employee rating by calculating the total employee rating in a department using PARTITION BY clause and SUM function on the department and the employee rating fields respectively:**

```
select emp_id, f_name, dept, emp_rating, sum(emp_rating) over (partition by dept) as
total_emp_rating from employee_datasets;
```



Result Grid | Filter Rows: | Export: | Wrap Cell Content:

emp_id	f_name	dept	emp_rating	total_emp_rating
104	rana	automotive	5	7
105	soniya	automotive	2	7
103	krishna	finance	3	6
106	nithin	finance	3	6
107	karan	healthcare	5	10
108	shilpa	healthcare	5	10
101	ankit	retail	3	9
102	sarvesh	retail	4	9
109	mukesh	retail	2	9

15. **-- Display the employee's ID, first name, department, and employee rating by assigning a rank to all the employees based on their employee rating using ORDER BY clause, RANK, and DENSE RANK functions on the employee rating field:**

select emp_id, f_name, dept, emp_rating, rank() over (order by emp_rating) as emp_rating_rank, dense_rank() over (order by emp_rating) as emp_rating_dense_rank from employee_datasets;

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

emp_id	f_name	dept	emp_rating	emp_rating_rank	emp_rating_dense_rank
105	soniya	automotive	2	1	1
109	mukesh	retail	2	1	1
101	ankit	retail	3	3	2
103	krishna	finance	3	3	2
106	nithin	finance	3	3	2
102	sarvesh	retail	4	6	3
104	rana	automotive	5	7	4
107	karan	healthcare	5	7	4
108	shilpa	healthcare	5	7	4

16. **-- Display the employee's ID, first name, department, employee rating by assigning a number to each employee in descending order of their employee rating using ORDER BY clause and ROW NUMBER function on the employee rating field:**
select emp_id, f_name, dept, emp_rating, row_number() over (order by emp_rating desc) as emp_id_asc_row_number from employee_datasets;

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

emp_id	f_name	dept	emp_rating	emp_id_asc_row_number
104	rana	automotive	5	1
107	karan	healthcare	5	2
108	shilpa	healthcare	5	3
102	sarvesh	retail	4	4
101	ankit	retail	3	5
103	krishna	finance	3	6
106	nithin	finance	3	7
105	soniya	automotive	2	8
109	mukesh	retail	2	9

17. **-- Display employee's ID, first name, role, department, and employee rating by calculating the percentile of the employee rating in a department using ORDER BY clause and PERCENT RANK function on an employee rating field:**
select emp_id, f_name, dept, emp_rating, percent_rank() over (order by emp_rating) as percentile_emp_rating from employee_datasets;

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

emp_id	f_name	dept	emp_rating	percentile_emp_rating
105	soniya	automotive	2	0
109	mukesh	retail	2	0
101	ankit	retail	3	0.25
103	krishna	finance	3	0.25
106	nithin	finance	3	0.25
102	sarvesh	retail	4	0.625
104	rana	automotive	5	0.75
107	karan	healthcare	5	0.75
108	shilpa	healthcare	5	0.75