ScienceQtech Employee Performance Mapping

Problem scenario:

ScienceQtech is a startup that works in the Data Science field. ScienceQtech has worked on fraud detection, market basket, self-driving cars, supply chain, algorithmic early detection of lung cancer, customer sentiment, and the drug discovery field. With the annual appraisal cycle around the corner, the HR department has asked you (Junior Database Administrator) to generate reports on employee details, their performance, and on the project that the employees have undertaken, to analyze the employee database and extract specific data based on different requirements.

Objective:

To facilitate a better understanding, managers have provided ratings for each employee which will help the HR department to finalize the employee performance mapping. As a DBA, you should find the maximum salary of the employees and ensure that all jobs are meeting the organization's profile standard. You also need to calculate bonuses to find extra cost for expenses. This will raise the overall performance of the organization by ensuring that all required employees receive training.

Dataset description:

emp_record_table: It contains the information of all the employees.

- EMP_ID – ID of the employee
- FIRST_NAME – First name of the employee
- LAST_NAME – Last name of the employee
- GENDER – Gender of the employee
- ROLE – Post of the employee
- DEPT – Field of the employee
- EXP – Years of experience the employee has
- COUNTRY – Country in which the employee is presently living
- CONTINENT – Continent in which the country is
- SALARY – Salary of the employee
- EMP_RATING – Performance rating of the employee
- MANAGER_ID – The manager under which the employee is assigned
- PROJ_ID – The project on which the employee is working or has worked on

Proj_table: It contains information about the projects.

- PROJECT_ID – ID for the project
- PROJ_Name – Name of the project
- DOMAIN – Field of the project
- START_DATE – Day the project began
- CLOSURE_DATE – Day the project was or will be completed
- DEV_QTR – Quarter in which the project was scheduled
- STATUS – Status of the project currently

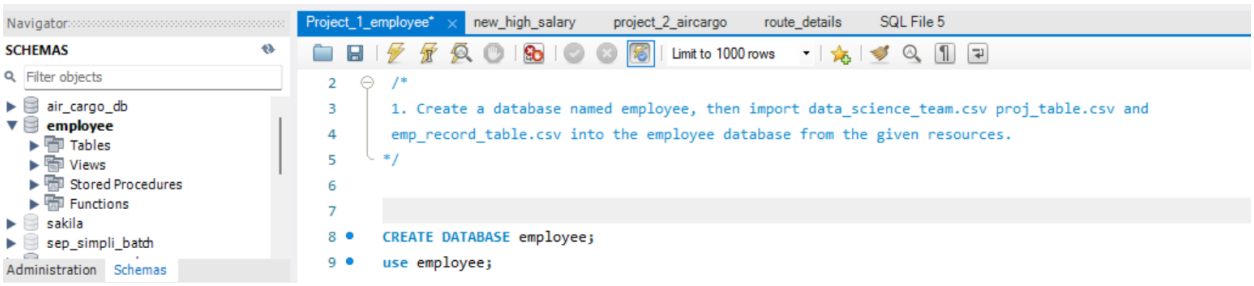
Data_science_team: It contains information about all the employees in the Data Science team.

- EMP_ID – ID of the employee
- FIRST_NAME – First name of the employee
- LAST_NAME – Last name of the employee
- GENDER – Gender of the employee
- ROLE – Post of the employee
- DEPT – Field of the employee
- EXP – Years of experience the employee has
- COUNTRY – Country in which the employee is presently living
- CONTINENT – Continent in which the country is

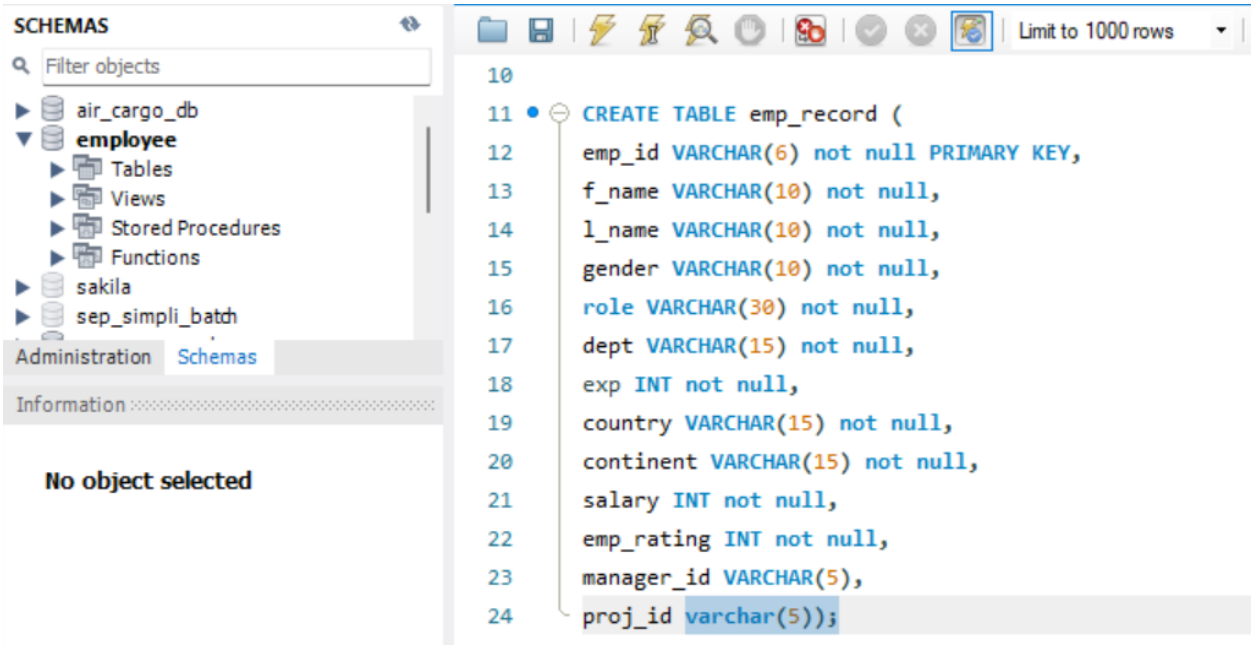
The task to be performed:

1. Create a database named *employee*, then import **data_science_team.csv proj_table.csv** and **emp_record_table.csv** into the **employee** database from the given resources.

- **Sql code:**
  CREATE DATABASE employee;



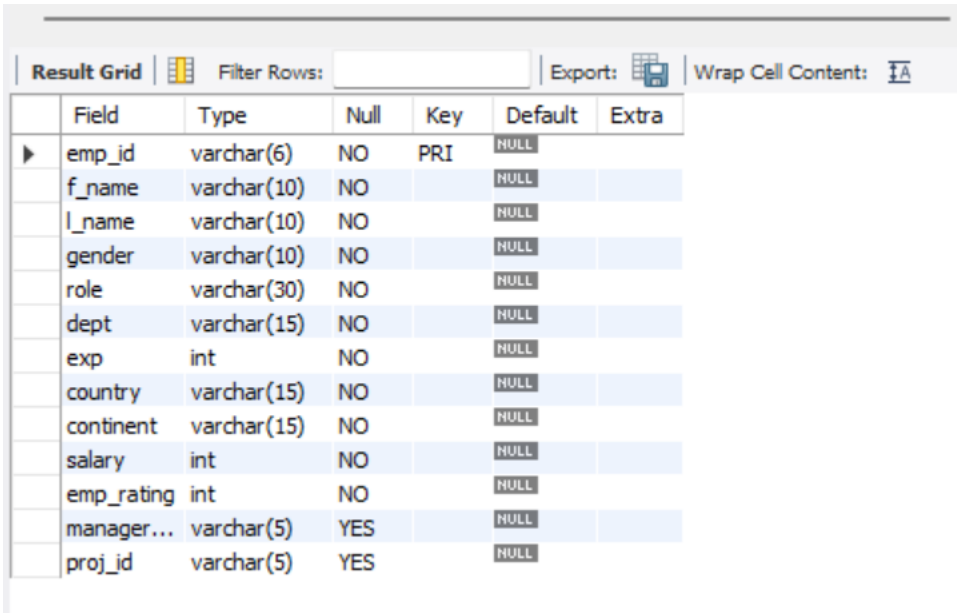
- **Sql code**
  CREATE TABLE emp_record (
  emp_id VARCHAR(6) not null PRIMARY KEY,
  f_name VARCHAR(10) not null,
  l_name VARCHAR(10) not null,
  gender VARCHAR(10) not null,
  role VARCHAR(30) not null,
  dept VARCHAR(15) not null,
  exp INT not null,
  country VARCHAR(15) not null,
  continent VARCHAR(15) not null,
  salary INT not null,
  emp_rating INT not null,
  manager_id VARCHAR(5),
  proj_id varchar(5));



- **Sql code**

  describe emp_record;

  output



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

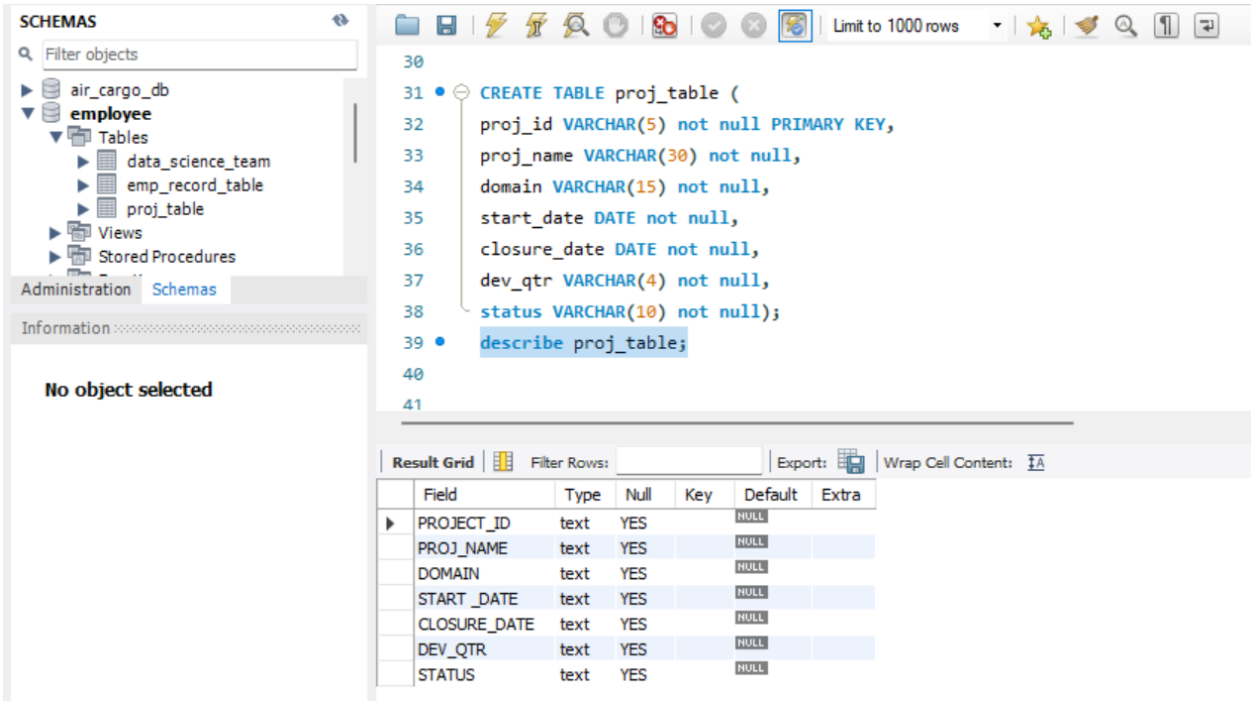
Field	Type	Null	Key	Default	Extra
emp_id	varchar(6)	NO	PRI	NULL	
f_name	varchar(10)	NO		NULL	
l_name	varchar(10)	NO		NULL	
gender	varchar(10)	NO		NULL	
role	varchar(30)	NO		NULL	
dept	varchar(15)	NO		NULL	
exp	int	NO		NULL	
country	varchar(15)	NO		NULL	
continent	varchar(15)	NO		NULL	
salary	int	NO		NULL	
emp_rating	int	NO		NULL	
manager...	varchar(5)	YES		NULL	
proj_id	varchar(5)	YES		NULL	

- **Sql code**

```
CREATE TABLE proj_table (
proj_id VARCHAR(5) not null PRIMARY KEY,
proj_name VARCHAR(30) not null,
domain VARCHAR(15) not null,
start_date DATE not null,
closure_date DATE not null,
dev_qtr VARCHAR(4) not null,
status VARCHAR(10) not null);

describe proj_table;
```

output

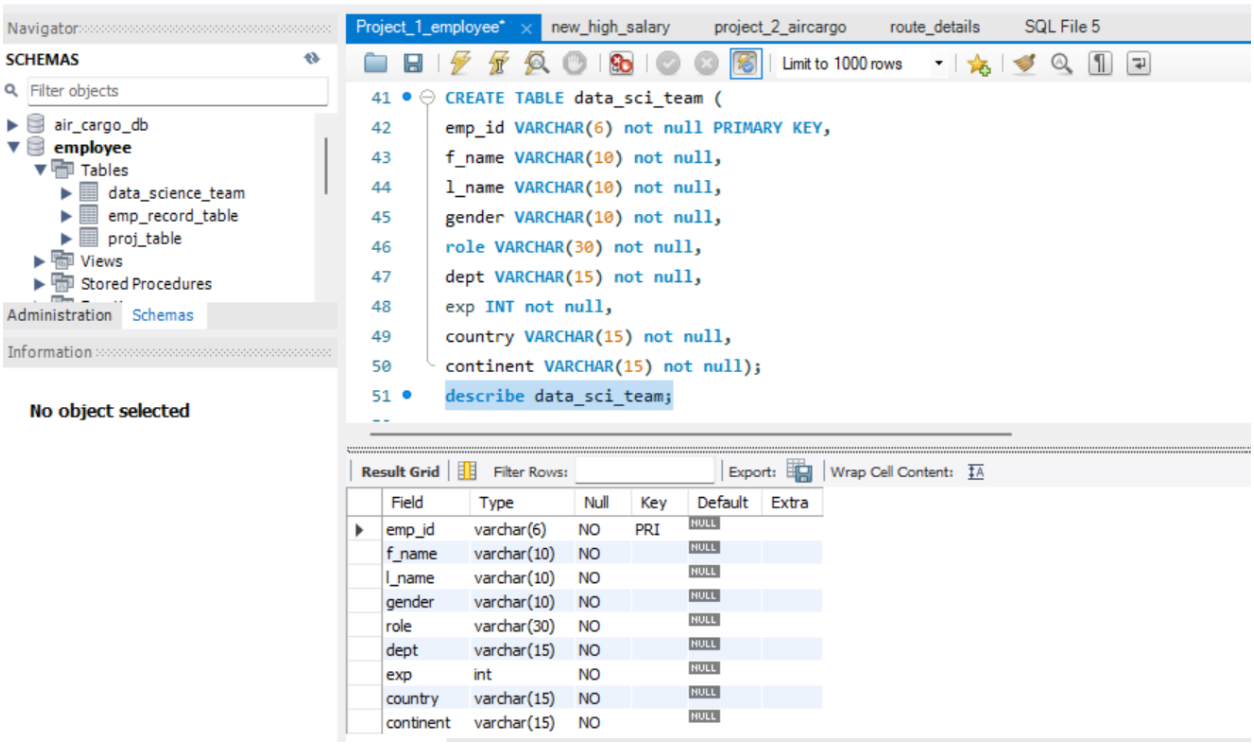


- **Sql code**

```
CREATE TABLE data_sci_team (
emp_id VARCHAR(6) not null PRIMARY KEY,
f_name VARCHAR(10) not null,
l_name VARCHAR(10) not null,
gender VARCHAR(10) not null,
role VARCHAR(30) not null,
dept VARCHAR(15) not null,
exp INT not null,
country VARCHAR(15) not null,
continent VARCHAR(15) not null);

describe data_sci_team;
```

output



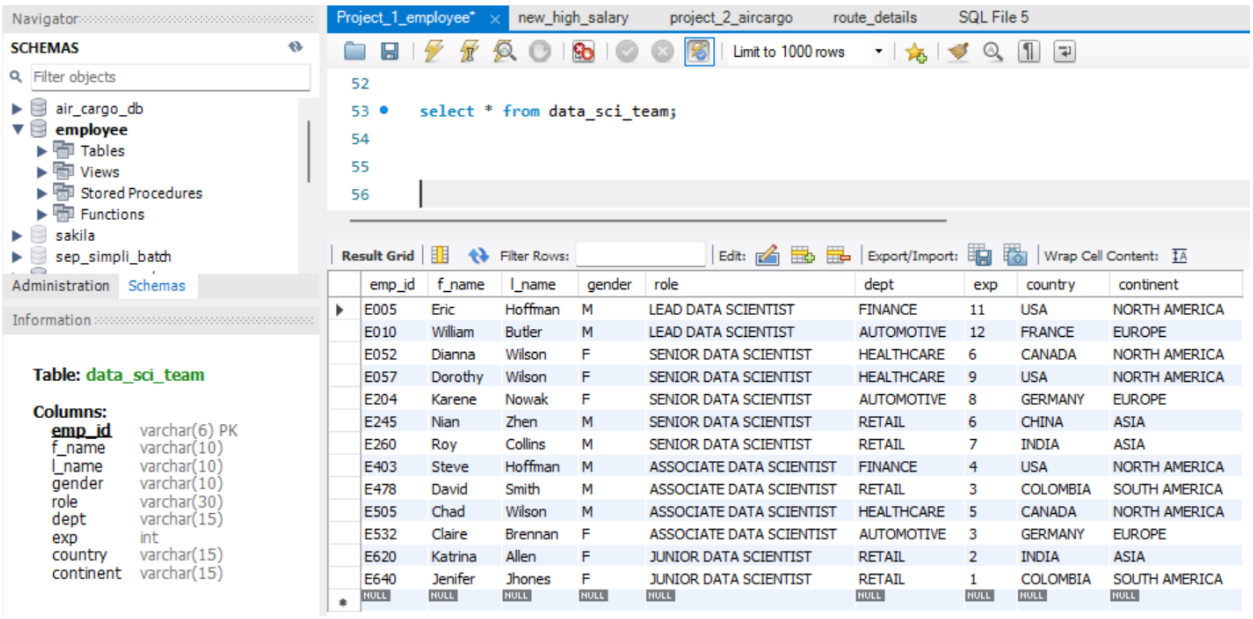
Import data into tables

- **Sql code:**

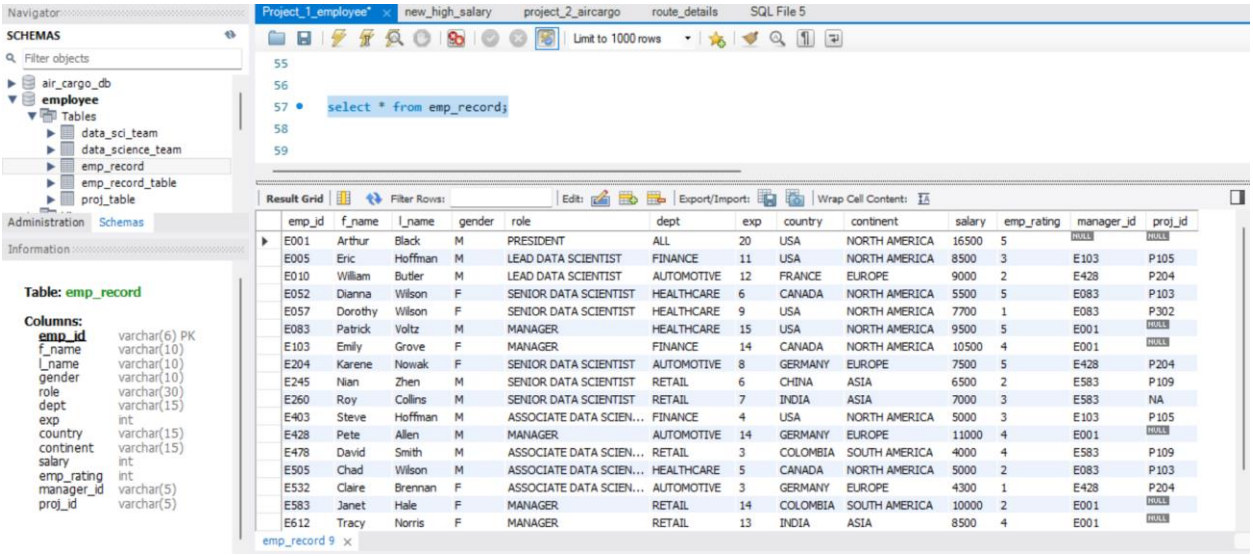
  Import using import function on workbench

Output:

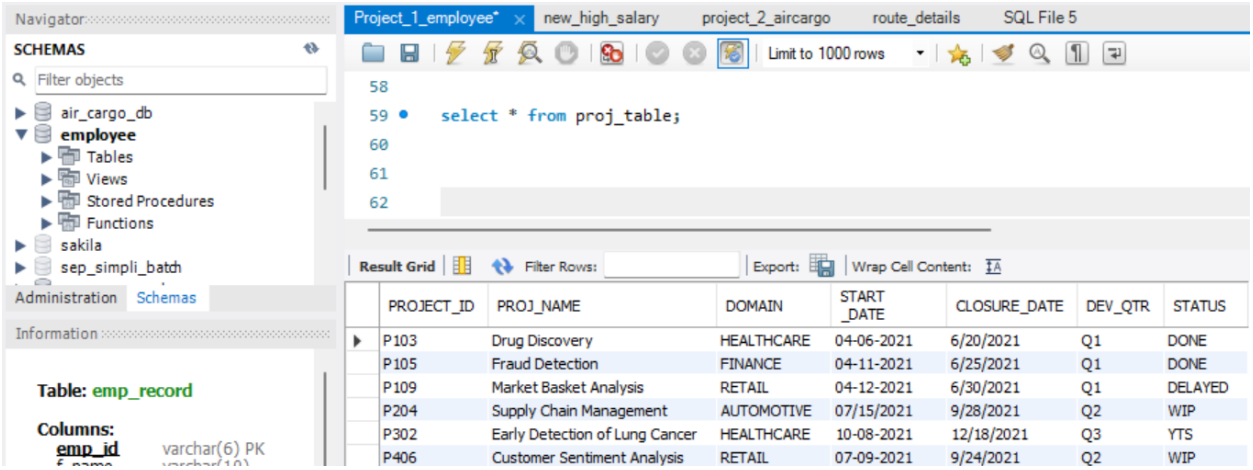
Data_sci_team;



Emp_record



Proj_table



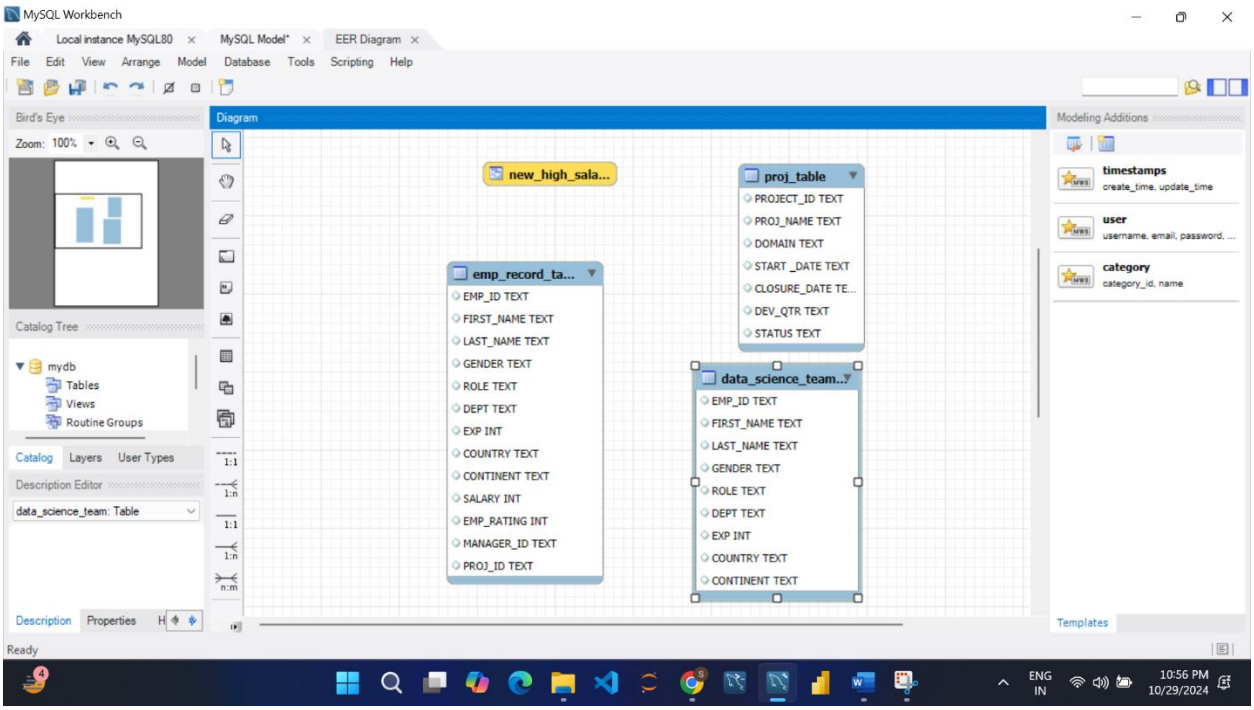
PROJECT_ID	PROJ_NAME	DOMAIN	START_DATE	CLOSURE_DATE	DEV_QTR	STATUS
P103	Drug Discovery	HEALTHCARE	04-06-2021	6/20/2021	Q1	DONE
P105	Fraud Detection	FINANCE	04-11-2021	6/25/2021	Q1	DONE
P109	Market Basket Analysis	RETAIL	04-12-2021	6/30/2021	Q1	DELAYED
P204	Supply Chain Management	AUTOMOTIVE	07/15/2021	9/28/2021	Q2	WIP
P302	Early Detection of Lung Cancer	HEALTHCARE	10-08-2021	12/18/2021	Q3	YTS
P406	Customer Sentiment Analysis	RETAIL	07-09-2021	9/24/2021	Q2	WIP

2. Create an ER diagram for the given **employee** database.

- **Sql code**

  Reverse engineering of employee database

  Output

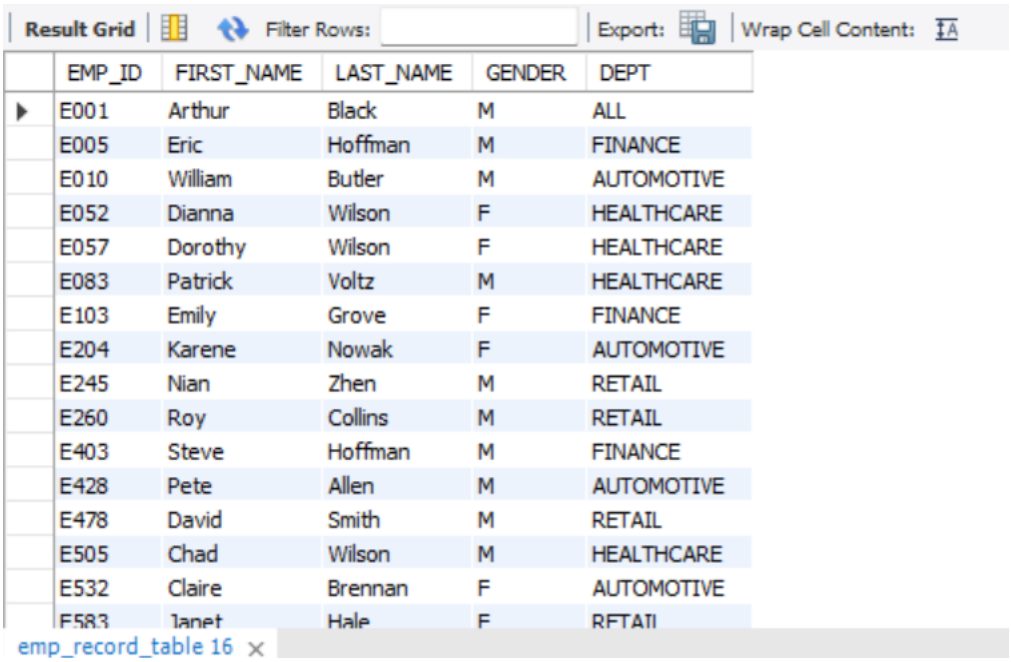


3. Write a query to fetch EMP_ID, FIRST_NAME, LAST_NAME, GENDER, and DEPARTMENT from the employee record table, and make a list of employees and details of their department.

- **Sql code**

```
select EMP_ID,FIRST_NAME,LAST_NAME,GENDER,DEPT
from emp_record_table;
```

output



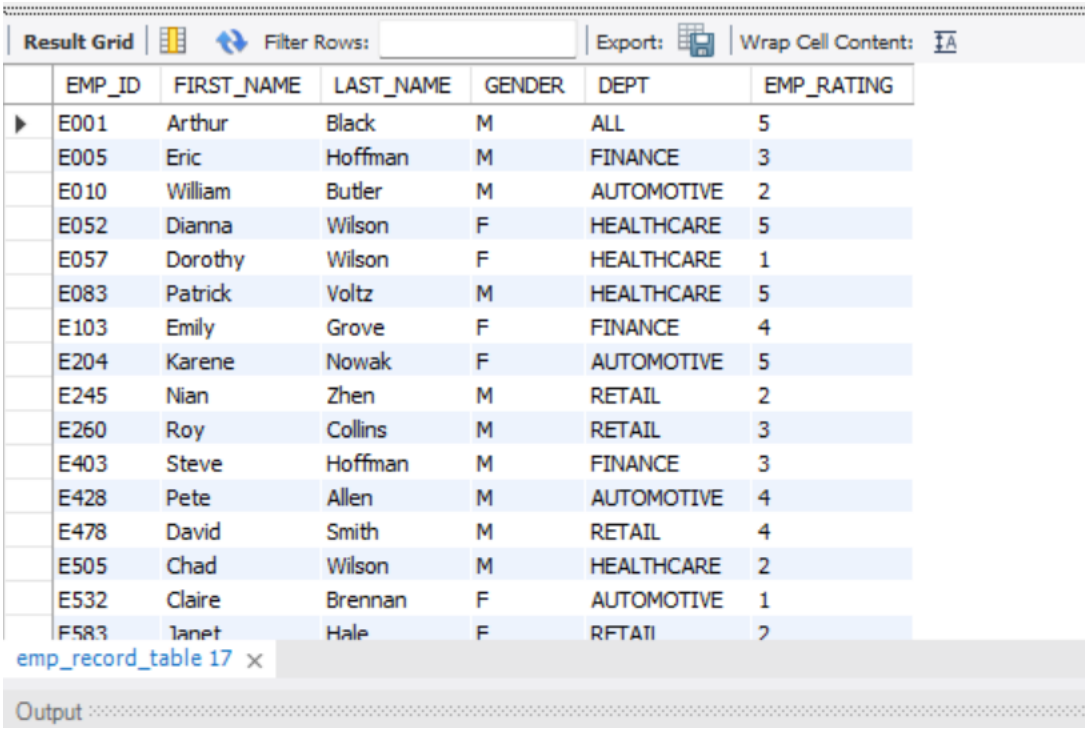
EMP_ID	FIRST_NAME	LAST_NAME	GENDER	DEPT
E001	Arthur	Black	M	ALL
E005	Eric	Hoffman	M	FINANCE
E010	William	Butler	M	AUTOMOTIVE
E052	Dianna	Wilson	F	HEALTHCARE
E057	Dorothy	Wilson	F	HEALTHCARE
E083	Patrick	Voltz	M	HEALTHCARE
E103	Emily	Grove	F	FINANCE
E204	Karene	Nowak	F	AUTOMOTIVE
E245	Nian	Zhen	M	RETAIL
E260	Roy	Collins	M	RETAIL
E403	Steve	Hoffman	M	FINANCE
E428	Pete	Allen	M	AUTOMOTIVE
E478	David	Smith	M	RETAIL
E505	Chad	Wilson	M	HEALTHCARE
E532	Claire	Brennan	F	AUTOMOTIVE
E583	Janet	Hale	F	RETAIL

4. Write a query to fetch EMP_ID, FIRST_NAME, LAST_NAME, GENDER, DEPARTMENT, and EMP_RATING if the EMP_RATING is:
    - less than two
    - greater than four
    - between two and four

- **sql code**

```
select EMP_ID, FIRST_NAME, LAST_NAME, GENDER, DEPT, EMP_RATING
from emp_record_table
where EMP_RATING <2
        or EMP_RATING >4
   or EMP_RATING between 2 And 4;
```

output



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

EMP_ID	FIRST_NAME	LAST_NAME	GENDER	DEPT	EMP_RATING
E001	Arthur	Black	M	ALL	5
E005	Eric	Hoffman	M	FINANCE	3
E010	William	Butler	M	AUTOMOTIVE	2
E052	Dianna	Wilson	F	HEALTHCARE	5
E057	Dorothy	Wilson	F	HEALTHCARE	1
E083	Patrick	Voltz	M	HEALTHCARE	5
E103	Emily	Grove	F	FINANCE	4
E204	Karene	Nowak	F	AUTOMOTIVE	5
E245	Nian	Zhen	M	RETAIL	2
E260	Roy	Collins	M	RETAIL	3
E403	Steve	Hoffman	M	FINANCE	3
E428	Pete	Allen	M	AUTOMOTIVE	4
E478	David	Smith	M	RETAIL	4
E505	Chad	Wilson	M	HEALTHCARE	2
E532	Claire	Brennan	F	AUTOMOTIVE	1
E583	Janet	Hale	F	RETAIL	2

emp_record_table 17 ×

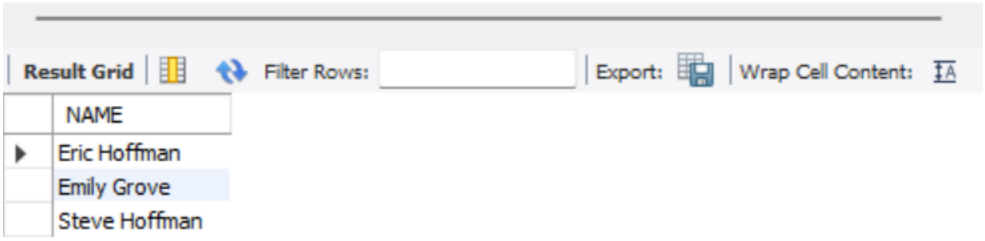
Output

5. Write a query to concatenate the FIRST_NAME and the LAST_NAME of employees in the *Finance* department from the employee table and then give the resultant column alias as NAME.

- **Sql code**

```
select concat(FIRST_NAME,' ',LAST_NAME) as NAME
from emp_record_table
where DEPT = "FINANCE";
```

output

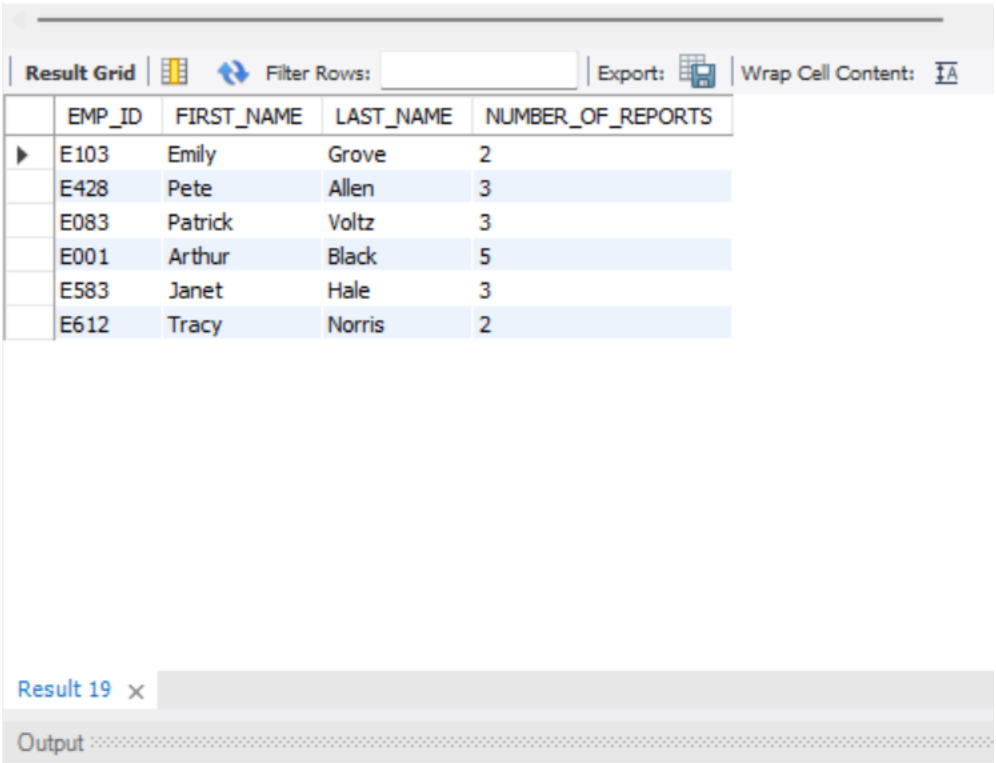


6. Write a query to list only those employees who have someone reporting to them. Also, show the number of reporters (including the President).

- **Sql code**

```
SELECT
   e.EMP_ID,
   e.FIRST_NAME,
   e.LAST_NAME,
   COUNT(r.EMP_ID) AS NUMBER_OF_REPORTS
FROM
   emp_record_table e
JOIN
   emp_record_table r ON e.EMP_ID = r.MANAGER_ID
GROUP BY
   e.EMP_ID, e.FIRST_NAME, e.LAST_NAME
HAVING
   COUNT(r.EMP_ID) > 0;
```

output



7. Write a query to list down all the employees from the healthcare and finance departments using union. Take data from the employee record table.

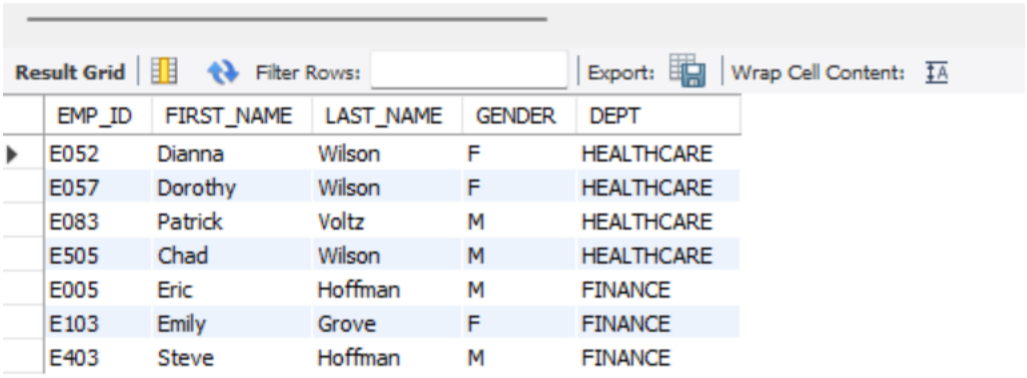
- **Sql code**

```
select EMP_ID, FIRST_NAME, LAST_NAME, GENDER, DEPT
from emp_record_table
where DEPT = 'healthcare'

union

select EMP_ID, FIRST_NAME, LAST_NAME, GENDER, DEPT
from emp_record_table
where DEPT = 'finance';
```

output



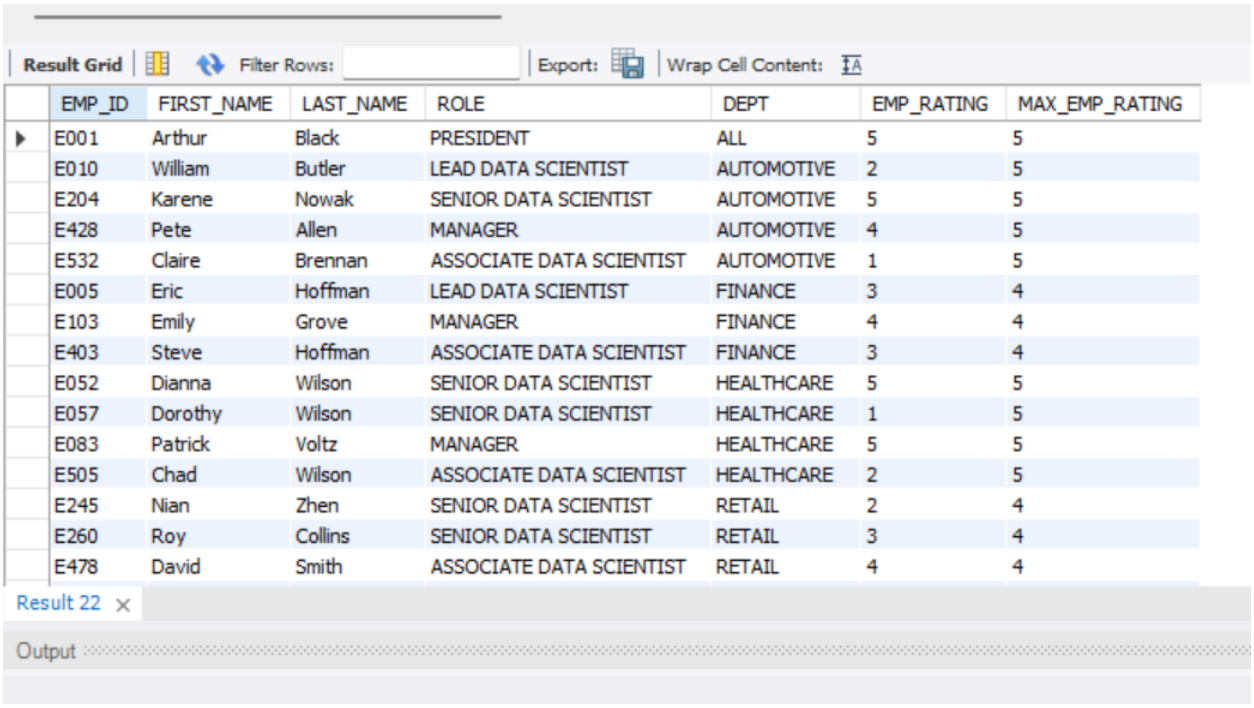
EMP_ID	FIRST_NAME	LAST_NAME	GENDER	DEPT
E052	Dianna	Wilson	F	HEALTHCARE
E057	Dorothy	Wilson	F	HEALTHCARE
E083	Patrick	Voltz	M	HEALTHCARE
E505	Chad	Wilson	M	HEALTHCARE
E005	Eric	Hoffman	M	FINANCE
E103	Emily	Grove	F	FINANCE
E403	Steve	Hoffman	M	FINANCE

8. Write a query to list down employee details such as EMP_ID, FIRST_NAME, LAST_NAME, ROLE, DEPARTMENT, and EMP_RATING grouped by dept. Also include the respective employee rating along with the max emp rating for the department.

- **Sql code**

```
SELECT
   EMP_ID,
   FIRST_NAME,
   LAST_NAME,
   ROLE,
   DEPT,
   EMP_RATING,
   MAX(EMP_RATING) OVER (PARTITION BY DEPT) AS MAX_EMP_RATING
FROM
   emp_record_table
ORDER BY
   DEPT, EMP_ID;
```

output



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

EMP_ID	FIRST_NAME	LAST_NAME	ROLE	DEPT	EMP_RATING	MAX_EMP_RATING
E001	Arthur	Black	PRESIDENT	ALL	5	5
E010	William	Butler	LEAD DATA SCIENTIST	AUTOMOTIVE	2	5
E204	Karene	Nowak	SENIOR DATA SCIENTIST	AUTOMOTIVE	5	5
E428	Pete	Allen	MANAGER	AUTOMOTIVE	4	5
E532	Claire	Brennan	ASSOCIATE DATA SCIENTIST	AUTOMOTIVE	1	5
E005	Eric	Hoffman	LEAD DATA SCIENTIST	FINANCE	3	4
E103	Emily	Grove	MANAGER	FINANCE	4	4
E403	Steve	Hoffman	ASSOCIATE DATA SCIENTIST	FINANCE	3	4
E052	Dianna	Wilson	SENIOR DATA SCIENTIST	HEALTHCARE	5	5
E057	Dorothy	Wilson	SENIOR DATA SCIENTIST	HEALTHCARE	1	5
E083	Patrick	Voltz	MANAGER	HEALTHCARE	5	5
E505	Chad	Wilson	ASSOCIATE DATA SCIENTIST	HEALTHCARE	2	5
E245	Nian	Zhen	SENIOR DATA SCIENTIST	RETAIL	2	4
E260	Roy	Collins	SENIOR DATA SCIENTIST	RETAIL	3	4
E478	David	Smith	ASSOCIATE DATA SCIENTIST	RETAIL	4	4

Result 22 ×

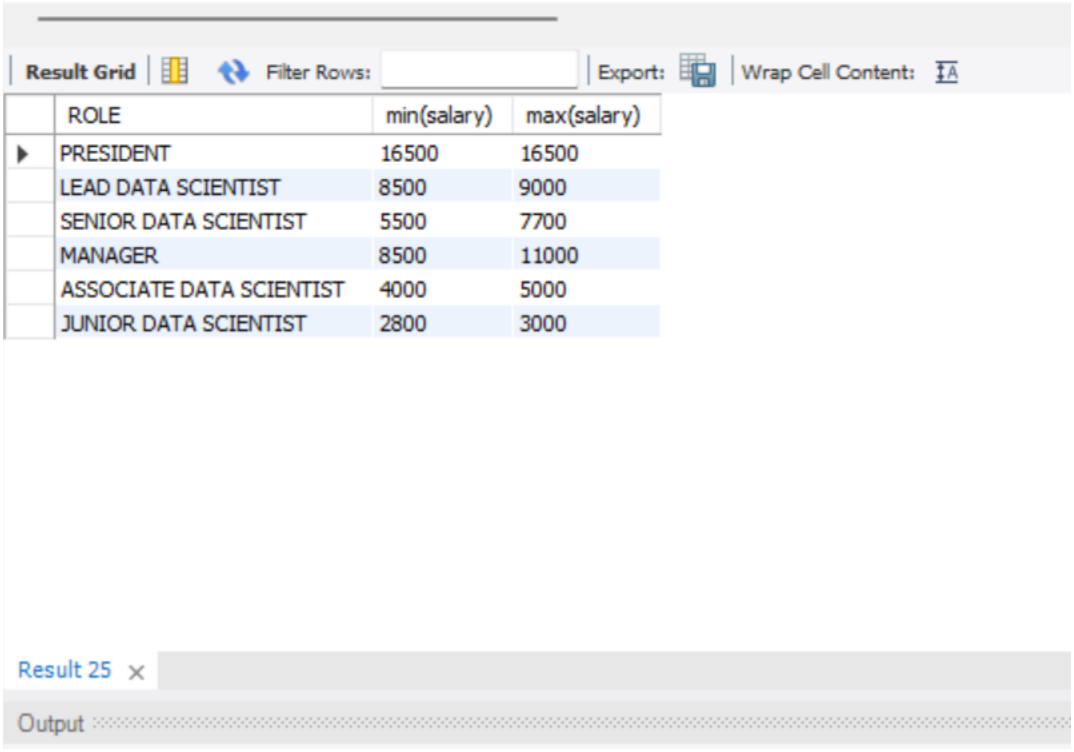
Output

9. Write a query to calculate the minimum and the maximum salary of the employees in each role. Take data from the employee record table.

- **Sql code**

```
select ROLE, min(salary), max(salary)
from  emp_record_table
group by ROLE;
```

output



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

ROLE	min(salary)	max(salary)
PRESIDENT	16500	16500
LEAD DATA SCIENTIST	8500	9000
SENIOR DATA SCIENTIST	5500	7700
MANAGER	8500	11000
ASSOCIATE DATA SCIENTIST	4000	5000
JUNIOR DATA SCIENTIST	2800	3000

Result 25

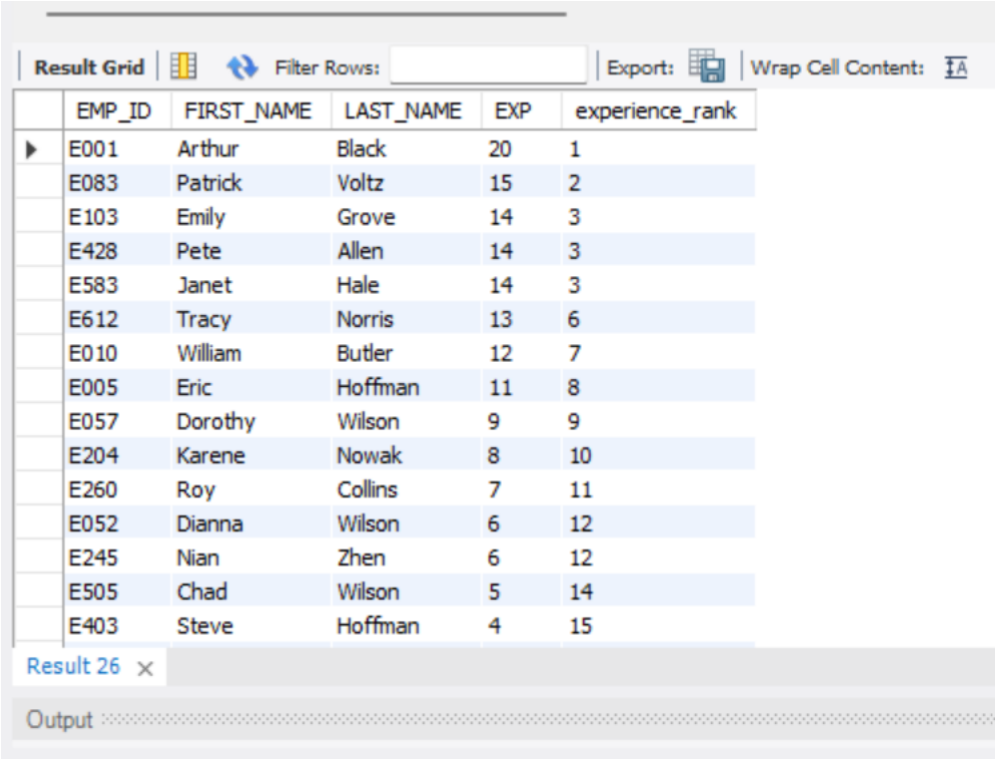
Output

10. Write a query to assign ranks to each employee based on their experience. Take data from the employee record table.

- **Sql code:**

```
select  EMP_ID,FIRST_NAME,LAST_NAME,EXP,
rank() over (order by EXP desc) as experience_rank
from emp_record_table
order by experience_rank;
```

output



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

EMP_ID	FIRST_NAME	LAST_NAME	EXP	experience_rank
E001	Arthur	Black	20	1
E083	Patrick	Voltz	15	2
E103	Emily	Grove	14	3
E428	Pete	Allen	14	3
E583	Janet	Hale	14	3
E612	Tracy	Norris	13	6
E010	William	Butler	12	7
E005	Eric	Hoffman	11	8
E057	Dorothy	Wilson	9	9
E204	Karene	Nowak	8	10
E260	Roy	Collins	7	11
E052	Dianna	Wilson	6	12
E245	Nian	Zhen	6	12
E505	Chad	Wilson	5	14
E403	Steve	Hoffman	4	15

Result 26 ×

Output

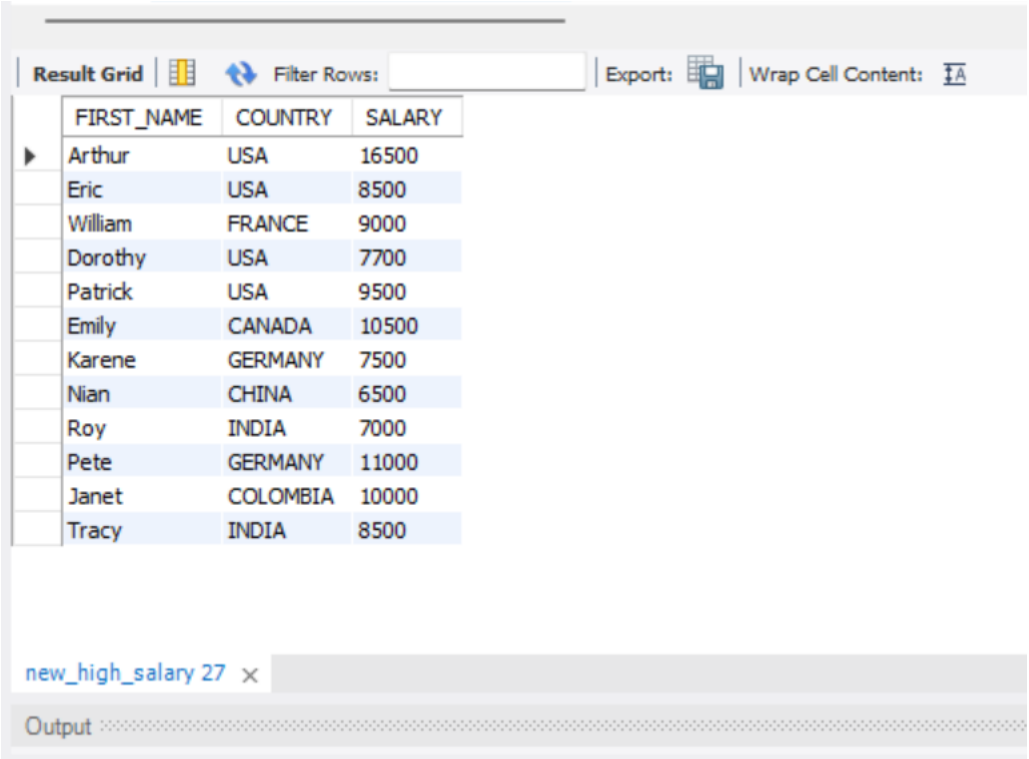
11. Write a query to create a view that displays employees in various countries whose salary is more than six thousand**.** Take data from the employee record table.

- **Sql code**

```
create view new_high_salary as
select FIRST_NAME,COUNTRY,SALARY
from emp_record_table
where salary >6000;

select * from new_high_salary;
```

output



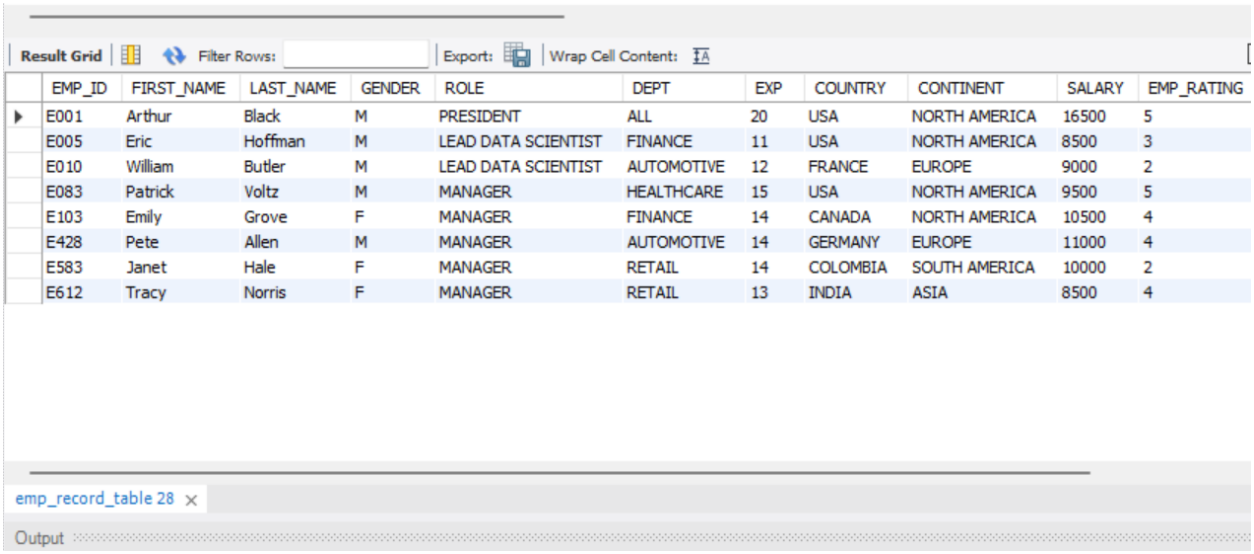
FIRST_NAME	COUNTRY	SALARY
Arthur	USA	16500
Eric	USA	8500
William	FRANCE	9000
Dorothy	USA	7700
Patrick	USA	9500
Emily	CANADA	10500
Karene	GERMANY	7500
Nian	CHINA	6500
Roy	INDIA	7000
Pete	GERMANY	11000
Janet	COLOMBIA	10000
Tracy	INDIA	8500

12. Write a nested query to find employees with experience of more than ten years. Take data from the employee record table.

- **Sql code**

```
select *
from emp_record_table
where EXP >(
        select 10);
```

output



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

EMP_ID	FIRST_NAME	LAST_NAME	GENDER	ROLE	DEPT	EXP	COUNTRY	CONTINENT	SALARY	EMP_RATING
E001	Arthur	Black	M	PRESIDENT	ALL	20	USA	NORTH AMERICA	16500	5
E005	Eric	Hoffman	M	LEAD DATA SCIENTIST	FINANCE	11	USA	NORTH AMERICA	8500	3
E010	William	Butler	M	LEAD DATA SCIENTIST	AUTOMOTIVE	12	FRANCE	EUROPE	9000	2
E083	Patrick	Voltz	M	MANAGER	HEALTHCARE	15	USA	NORTH AMERICA	9500	5
E103	Emily	Grove	F	MANAGER	FINANCE	14	CANADA	NORTH AMERICA	10500	4
E428	Pete	Allen	M	MANAGER	AUTOMOTIVE	14	GERMANY	EUROPE	11000	4
E583	Janet	Hale	F	MANAGER	RETAIL	14	COLOMBIA	SOUTH AMERICA	10000	2
E612	Tracy	Norris	F	MANAGER	RETAIL	13	INDIA	ASIA	8500	4

emp_record_table 28 ×

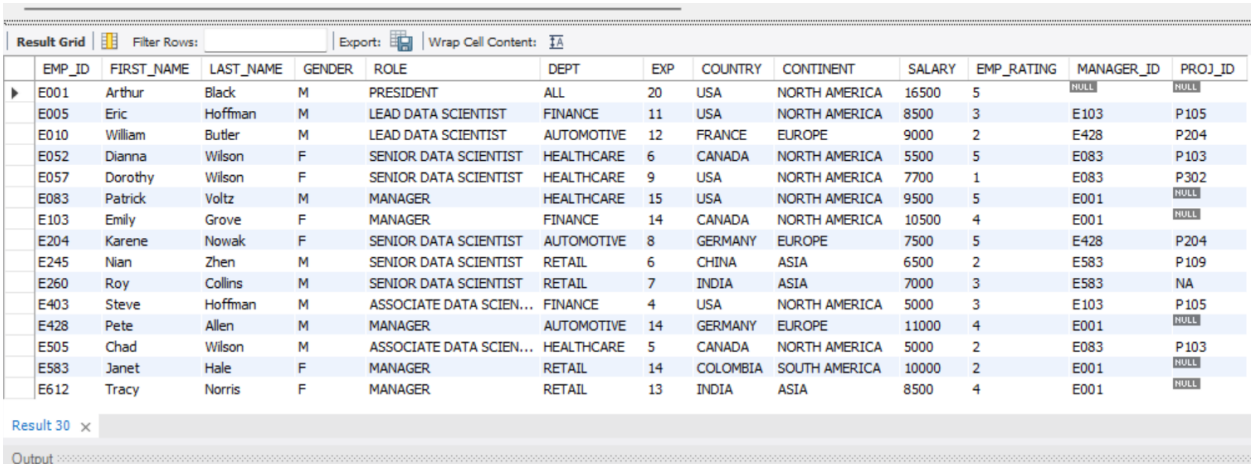
Output

13. Write a query to create a stored procedure to retrieve the details of the employees whose experience is more than three years. Take data from the employee record table.

- **Sql code**

```
DELIMITER //
create procedure getExperience_employee()
begin
        select *
   from emp_record_table
   where EXP >3;
end//
DELIMITER ;
call getExperience_employee;
```

output



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

EMP_ID	FIRST_NAME	LAST_NAME	GENDER	ROLE	DEPT	EXP	COUNTRY	CONTINENT	SALARY	EMP_RATING	MANAGER_ID	PROJ_ID
E001	Arthur	Black	M	PRESIDENT	ALL	20	USA	NORTH AMERICA	16500	5	NULL	NULL
E005	Eric	Hoffman	M	LEAD DATA SCIENTIST	FINANCE	11	USA	NORTH AMERICA	8500	3	E103	P105
E010	William	Butler	M	LEAD DATA SCIENTIST	AUTOMOTIVE	12	FRANCE	EUROPE	9000	2	E428	P204
E052	Dianna	Wilson	F	SENIOR DATA SCIENTIST	HEALTHCARE	6	CANADA	NORTH AMERICA	5500	5	E083	P103
E057	Dorothy	Wilson	F	SENIOR DATA SCIENTIST	HEALTHCARE	9	USA	NORTH AMERICA	7700	1	E083	P302
E083	Patrick	Voltz	M	MANAGER	HEALTHCARE	15	USA	NORTH AMERICA	9500	5	E001	NULL
E103	Emily	Grove	F	MANAGER	FINANCE	14	CANADA	NORTH AMERICA	10500	4	E001	NULL
E204	Karene	Nowak	F	SENIOR DATA SCIENTIST	AUTOMOTIVE	8	GERMANY	EUROPE	7500	5	E428	P204
E245	Nian	Zhen	M	SENIOR DATA SCIENTIST	RETAIL	6	CHINA	ASIA	6500	2	E583	P109
E260	Roy	Collins	M	SENIOR DATA SCIENTIST	RETAIL	7	INDIA	ASIA	7000	3	E583	NA
E403	Steve	Hoffman	M	ASSOCIATE DATA SCIEN...	FINANCE	4	USA	NORTH AMERICA	5000	3	E103	P105
E428	Pete	Allen	M	MANAGER	AUTOMOTIVE	14	GERMANY	EUROPE	11000	4	E001	NULL
E505	Chad	Wilson	M	ASSOCIATE DATA SCIEN...	HEALTHCARE	5	CANADA	NORTH AMERICA	5000	2	E083	P103
E583	Janet	Hale	F	MANAGER	RETAIL	14	COLOMBIA	SOUTH AMERICA	10000	2	E001	NULL
E612	Tracy	Norris	F	MANAGER	RETAIL	13	INDIA	ASIA	8500	4	E001	NULL

Result 30 ×

Output

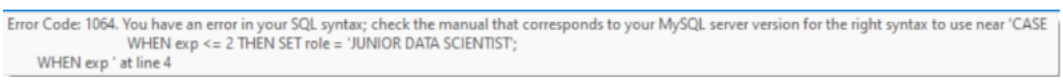
14. Write a query using stored functions in the project table to check whether the job profile assigned to each employee in the data science team matches the organization's set standard.

    The standard being:
    For an employee with experience less than or equal to 2 years assign 'JUNIOR DATA SCIENTIST',
    For an employee with the experience of 2 to 5 years assign 'ASSOCIATE DATA SCIENTIST',
    For an employee with the experience of 5 to 10 years assign 'SENIOR DATA SCIENTIST',
    For an employee with the experience of 10 to 12 years assign 'LEAD DATA SCIENTIST',
    For an employee with the experience of 12 to 16 years assign 'MANAGER'.

- **Sql code**

```
DELIMITER //
CREATE PROCEDURE check_role()
BEGIN
        SELECT * FROM emp_record_table
                CASE
                        WHEN exp <= 2 THEN SET role = 'JUNIOR DATA
SCIENTIST';
                        WHEN exp BETWEEN 3 AND 5 THEN SET role =
'ASSOCIATE DATA SCIENTIST';
                        WHEN exp BETWEEN 6 AND 10 THEN SET role = 'SENIOR
DATA SCIENTIST';
                        WHEN exp BETWEEN 11 AND 12 THEN SET role = 'LEAD
DATA SCIENTIST';
                        WHEN exp BETWEEN 13 AND 16 THEN SET role =
'MANAGER';
                        ELSE SET role = 'all good';
        END CASE;
END //
DELIMITER ;
```

output

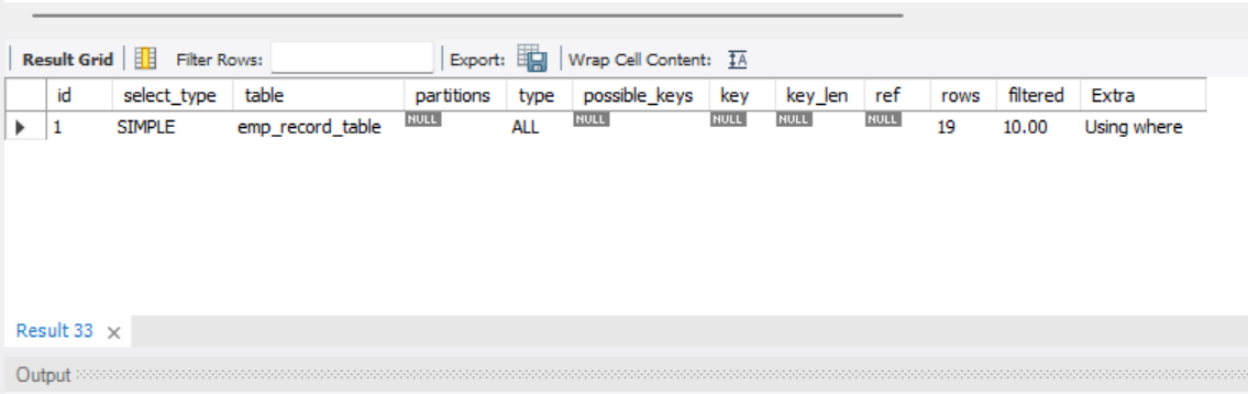


15. Create an index to improve the cost and performance of the query to find the employee whose FIRST_NAME is ‘Eric’ in the employee table after checking the execution plan.

- **Sql code**

```
EXPLAIN SELECT * FROM emp_record_table WHERE FIRST_NAME = 'Eric';
CREATE INDEX idx_first_name ON emp_record_table(FIRST_NAME);
EXPLAIN SELECT * FROM emp_record_table WHERE FIRST_NAME = 'Eric';
```

Output

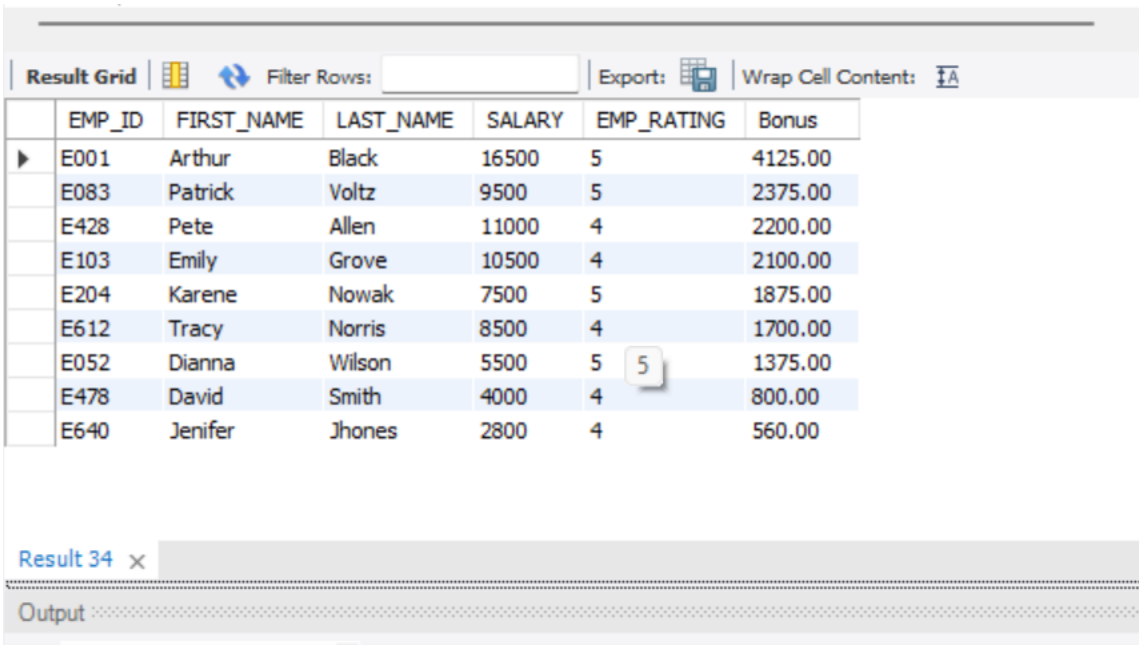


16. Write a query to calculate the bonus for all the employees, based on their ratings and salaries (Use the formula: 5% of salary * employee rating).

- **Sql code**

```
SELECT
   EMP_ID,
   FIRST_NAME,
   LAST_NAME,
   SALARY,
   EMP_RATING,
   (0.05 * SALARY * EMP_RATING) AS Bonus
FROM
   emp_record_table
WHERE EMP_RATING > 3
ORDER BY Bonus DESC;
```

Output



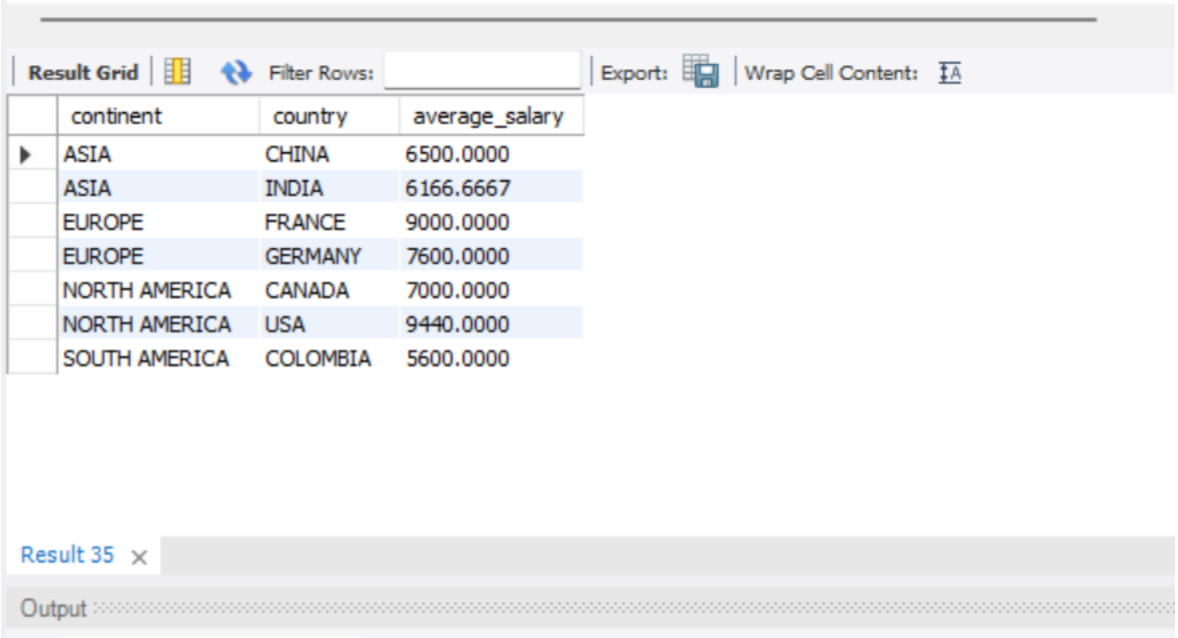
EMP_ID	FIRST_NAME	LAST_NAME	SALARY	EMP_RATING	Bonus
E001	Arthur	Black	16500	5	4125.00
E083	Patrick	Voltz	9500	5	2375.00
E428	Pete	Allen	11000	4	2200.00
E103	Emily	Grove	10500	4	2100.00
E204	Karene	Nowak	7500	5	1875.00
E612	Tracy	Norris	8500	4	1700.00
E052	Dianna	Wilson	5500	5	1375.00
E478	David	Smith	4000	4	800.00
E640	Jenifer	Jhones	2800	4	560.00

17. Write a query to calculate the average salary distribution based on the continent and country. Take data from the employee record table.

- **Sql code**

```
SELECT
    continent,
    country,
    AVG(Salary) AS average_salary
FROM
    emp_record_table
GROUP BY
    continent,
    country
ORDER BY
    continent,
    country;
```

output



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

continent	country	average_salary
ASIA	CHINA	6500.0000
ASIA	INDIA	6166.6667
EUROPE	FRANCE	9000.0000
EUROPE	GERMANY	7600.0000
NORTH AMERICA	CANADA	7000.0000
NORTH AMERICA	USA	9440.0000
SOUTH AMERICA	COLOMBIA	5600.0000

Result 35

Output