Neha Petkar | SQL Practice Project 1

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.

Note: You must download the dataset from the course resource section in LMS and create a table to perform the above objective.

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.

Ans:

```
8
9  Select * from data_science_team;
10 Select * from emp_record_table;
11 Select * from proj_table;
```

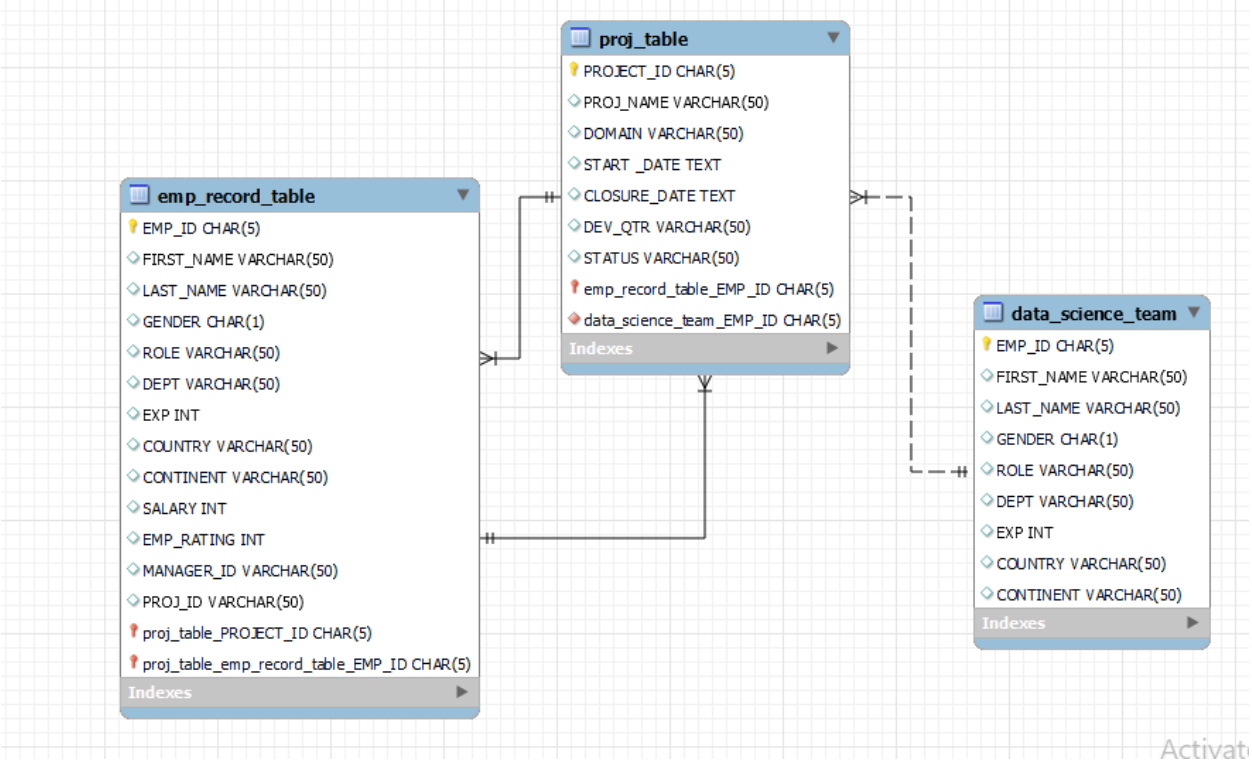
Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content:

EMP_ID	FIRST_NAME	LAST_NAME	GENDER	ROLE	DEPT	EXP	COUNTRY	CONTINENT
E005	Eric	Hoffman	M	LEAD DATA SCIENTIST	FINANCE	11	USA	NORTH AMERICA
E010	William	Butler	M	LEAD DATA SCIENTIST	AUTOMOTIVE	12	FRANCE	EUROPE
E052	Dianna	Wilson	F	SENIOR DATA SCIENTIST	HEALTHCARE	6	CANADA	NORTH AMERICA
E057	Dorothy	Wilson	F	SENIOR DATA SCIENTIST	HEALTHCARE	9	USA	NORTH AMERICA
E204	Karene	Nowak	F	SENIOR DATA SCIENTIST	AUTOMOTIVE	8	GERMANY	EUROPE
E245	Nian	Zhen	M	SENIOR DATA SCIENTIST	RETAIL	6	CHINA	ASIA
E260	Roy	Collins	M	SENIOR DATA SCIENTIST	RETAIL	7	INDIA	ASIA
E403	Steve	Hoffman	M	ASSOCIATE DATA SCIENTIST	FINANCE	4	USA	NORTH AMERICA
E478	David	Smith	M	ASSOCIATE DATA SCIENTIST	RETAIL	3	COLOMBIA	SOUTH AMERICA
E505	Chad	Wilson	M	ASSOCIATE DATA SCIENTIST	HEALTHCARE	5	CANADA	NORTH AMERICA
E532	Claire	Brennan	F	ASSOCIATE DATA SCIENTIST	AUTOMOTIVE	3	GERMANY	EUROPE
E620	Katrina	Allen	F	JUNIOR DATA SCIENTIST	RETAIL	2	INDIA	ASIA
E640	Jenifer	Jhones	F	JUNIOR DATA SCIENTIST	RETAIL	1	COLOMBIA	SOUTH AMERICA

data_science_team 51

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

Ans:



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.

Ans: select EMP_ID, FIRST_NAME, LAST_NAME, GENDER, DEPT as Department from emp_record_table;

```
15  select EMP_ID, FIRST_NAME, LAST_NAME, GENDER, DEPT as Department
16      from emp_record_table;
17
```

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

EMP_ID	FIRST_NAME	LAST_NAME	GENDER	Department
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

emp_record_table 4

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

Ans: select EMP_ID, FIRST_NAME, LAST_NAME, GENDER, DEPT as Department, emp_rating,
case
when emp_rating < 2 then 'Rating is below 2'
when emp_rating > 4 then 'Rating is above 4'
else 'between 2 and 4'
END Rating
from emp_record_table
Order by Rating;

```
36      -- Using Case
37          select EMP_ID, FIRST_NAME, LAST_NAME, GENDER, DEPT as Department, emp_rating,
38          case
39              when emp_rating < 2 then 'Rating is below 2'
40              when emp_rating > 4 then 'Rating is above 4'
41              else 'between 2 and 4'
42          END Rating
43          from emp_record_table
44          Order by Rating;
45
```

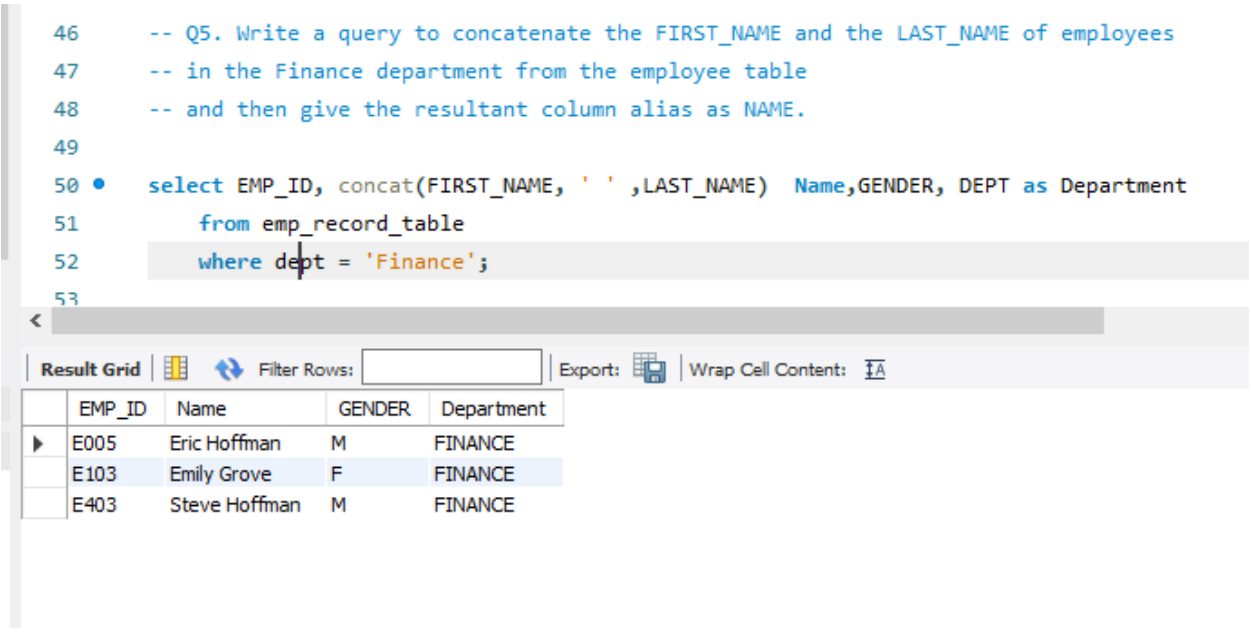
Result Grid | Filter Rows: | Export: | Wrap Cell Content:

EMP_ID	FIRST_NAME	LAST_NAME	GENDER	Department	emp_rating	Rating
E005	Eric	Hoffman	M	FINANCE	3	between 2 and 4
E010	William	Butler	M	AUTOMOTIVE	2	between 2 and 4
E103	Emily	Grove	F	FINANCE	4	between 2 and 4
E245	Nian	Zhen	M	RETAIL	2	between 2 and 4
E260	Roy	Collins	M	RETAIL	3	between 2 and 4
E403	Steve	Hoffman	M	FINANCE	3	between 2 and 4
E428	Pete	Allen	M	AUTOMOTIVE	4	between 2 and 4
E478	David	Smith	M	RETAIL	4	between 2 and 4
E505	Chad	Wilson	M	HEALTHCARE	2	between 2 and 4

Result 5

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.

Ans: select EMP_ID, concat (FIRST_NAME, ' ' , LAST_NAME) Name, GENDER, DEPT as Department from emp_record_table where dept = 'Finance';



```
46    -- Q5. Write a query to concatenate the FIRST_NAME and the LAST_NAME of employees
47    -- in the Finance department from the employee table
48    -- and then give the resultant column alias as NAME.
49
50    select EMP_ID, concat(FIRST_NAME, ' ' ,LAST_NAME)  Name,GENDER, DEPT as Department
51        from emp_record_table
52        where dept = 'Finance';
53
```

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

EMP_ID	Name	GENDER	Department
E005	Eric Hoffman	M	FINANCE
E103	Emily Grove	F	FINANCE
E403	Steve Hoffman	M	FINANCE

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

Ans:

```
select m.EMP_ID, m.FIRST_NAME as Manager, e.FIRST_NAME,
 Count (*) over (partition by m.FIRST_NAME) as Count_of_Reportee
 from emp_record_table e join emp_record_table m
 ON e.manager_id = m.emp_id
 order by m.emp_id;
```

```
60      -- With Window function
61      select m.EMP_ID, m.FIRST_NAME as Manager,e.FIRST_NAME,
62       Count(*) over(partition by m.FIRST_NAME ) as Count_of_Reportee
63          from emp_record_table e join emp_record_table m
64          ON e.manager_id = m.emp_id
65          order by m.emp_id;
66
```

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

EMP_ID	Manager	FIRST_NAME	Count_of_Reportee
E001	Arthur	Patrick	5
E001	Arthur	Emily	5
E001	Arthur	Pete	5
E001	Arthur	Janet	5
E001	Arthur	Tracy	5
E083	Patrick	Dianna	3
E083	Patrick	Dorothy	3
E083	Patrick	Chad	3
E103	Emily	Eric	2

esult 9

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.

Ans:

```
Select * from emp_record_table where dept = 'Healthcare'
    Union
    Select * from emp_record_table where dept = 'finance';
```

```
69 •   Select * from emp_record_table where dept = 'Healthcare'
70     Union
71     Select * from emp_record_table where dept = 'finance';
72
```

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

EMP_ID	FIRST_NAME	LAST_NAME	GENDER	ROLE	DEPT
E052	Dianna	Wilson	F	SENIOR DATA SCIENTIST	HEALTHCARE
E057	Dorothy	Wilson	F	SENIOR DATA SCIENTIST	HEALTHCARE
E083	Patrick	Voltz	M	MANAGER	HEALTHCARE
E505	Chad	Wilson	M	ASSOCIATE DATA SCIENTIST	HEALTHCARE
E005	Eric	Hoffman	M	LEAD DATA SCIENTIST	FINANCE
E103	Emily	Grove	F	MANAGER	FINANCE
E403	Steve	Hoffman	M	ASSOCIATE DATA SCIENTIST	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.

Ans:

```
select EMP_ID, FIRST_NAME, LAST_NAME, ROLE , DEPT as Department, emp_rating,
max (emp_rating) over (partition by dept) as MaxRating
from emp_record_table;
```

```
75
76 •     select EMP_ID, FIRST_NAME, LAST_NAME, ROLE , DEPT as Department, emp_rating,
77      max(emp_rating) over(partition by dept) as MaxRating
78      from emp_record_table;
79
```

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

EMP_ID	FIRST_NAME	LAST_NAME	ROLE	Department	emp_rating	MaxRating
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

Result 13

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.

Ans:

```
select EMP_ID, FIRST_NAME, LAST_NAME, DEPT, ROLE , Salary,
Max (salary) over (partition by Role) Maximum,
Min (salary) over (partition by Role) Minimum
from emp_record_table;
```

```
81
82 •  select EMP_ID, FIRST_NAME, LAST_NAME, DEPT,ROLE , Salary,
83     Max(salary) over(partition by Role) Maximum,
84     Min(salary) over(partition by Role) Minimum
85     from emp_record_table;
86
87     -- Q10. Write a query to assign ranks to each employee based on their experience. T
```

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

EMP_ID	FIRST_NAME	LAST_NAME	DEPT	ROLE	Salary	Maximum	Minimum
E403	Steve	Hoffman	FINANCE	ASSOCIATE DATA SCIENTIST	5000	5000	4000
E478	David	Smith	RETAIL	ASSOCIATE DATA SCIENTIST	4000	5000	4000
E505	Chad	Wilson	HEALTHCARE	ASSOCIATE DATA SCIENTIST	5000	5000	4000
E532	Claire	Brennan	AUTOMOTIVE	ASSOCIATE DATA SCIENTIST	4300	5000	4000
E620	Katrina	Allen	RETAIL	JUNIOR DATA SCIENTIST	3000	3000	2800
E640	Jenifer	Jhones	RETAIL	JUNIOR DATA SCIENTIST	2800	3000	2800
E005	Eric	Hoffman	FINANCE	LEAD DATA SCIENTIST	8500	9000	8500
E010	William	Butler	AUTOMOTIVE	LEAD DATA SCIENTIST	9000	9000	8500
E083	Patrick	Voltz	HEALTHCARE	MANAGER	9500	11000	8500

Result 14

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

Ans:
1. Dense_rank Function:
Select emp_id, first_name, last_name, exp,
dense_rank () over (order by exp desc) as RankbyExp from emp_record_table;

2. Rank Function:
Select emp_id, first_name, last_name, exp,
Rank () over (order by exp desc) from emp_record_table;

```
87      -- Q10. Write a query to assign ranks to each employee based on their exp
88      Select emp_id,first_name,last_name,exp,
89      dense_rank() over(order by exp desc) as RankbyExp from emp_record_table;
90
91      Select emp_id,first_name,last_name,exp,
92      rank() over(order by exp desc) from emp_record_table;
93
94
```

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

emp_id	first_name	last_name	exp	RankbyExp
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	4
E010	William	Butler	12	5
E005	Eric	Hoffman	11	6
E057	Dorothy	Wilson	9	7

Result 15

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.

Ans:

```
Create view V1_emp_6k
as
Select emp_id, first_name, last_name, Country, salary from emp_record_table where
salary > 6000;

select * from V1_emp_6k;
```

```
97
98  Create view V1_emp_6k
99  as
100 Select emp_id,first_name,last_name,Country,salary from emp_record_table where salary > 6000;
101
102 select * from V1_emp_6k;
103
104 -- Q12. Write a nested query to find employees with experience of more than ten years. Take data
```

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

emp_id	first_name	last_name	Country	salary
E001	Arthur	Black	USA	16500
E005	Eric	Hoffman	USA	8500
E010	William	Butler	FRANCE	9000
E057	Dorothy	Wilson	USA	7700
E083	Patrick	Voltz	USA	9500
E103	Emily	Grove	CANADA	10500
E204	Karene	Nowak	GERMANY	7500
E245	Nian	Zhen	CHINA	6500
E260	Roy	Collins	INDIA	7000

V1_emp_6k 16

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

Ans:
1. Using Nested Query:

```
Select EMP_ID, FIRST_NAME, LAST_NAME, EXP from emp_record_table where emp_id in
        (Select emp_id from emp_record_table where exp > 10);
```

2. Using CTE method

```
With CTE
as
(
Select * from emp_record_table
) Select EMP_ID, FIRST_NAME, LAST_NAME, EXP from cte where exp > 10;
```

```
106
107 •  Select EMP_ID, FIRST_NAME, LAST_NAME, EXP from emp_record_table where emp_id in
108        (Select emp_id from emp_record_table where exp > 10);
109
```

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

EMP_ID	FIRST_NAME	LAST_NAME	EXP
E001	Arthur	Black	20
E005	Eric	Hoffman	11
E010	William	Butler	12
E083	Patrick	Voltz	15
E103	Emily	Grove	14
E428	Pete	Allen	14
E583	Janet	Hale	14
E612	Tracy	Norris	13
NULL	NULL	NULL	NULL

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.

Ans:

```
call emp_above3year;

CREATE DEFINER=`root`@`localhost` PROCEDURE `emp_above3year`()
BEGIN
        Select * from emp_record_table where exp > 3;
END;
```

```
120 •  call emp_above3year;
121
122 •  CREATE DEFINER=`root`@`localhost` PROCEDURE `emp_above3year`()
123    BEGIN
124        Select * from emp_record_table where exp > 3;
125    END;
126
```

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
E052	Dianna	Wilson	F	SENIOR DATA SCIENTIST	HEALTHCARE	6	CANADA	NORTH AMERICA	5500	5
E057	Dorothy	Wilson	F	SENIOR DATA SCIENTIST	HEALTHCARE	9	USA	NORTH AMERICA	7700	1
E083	Patrick	Voltz	M	MANAGER	HEALTHCARE	15	USA	NORTH AMERICA	9500	5
E103	Emily	Grove	F	MANAGER	FINANCE	14	CANADA	NORTH AMERICA	10500	4
E204	Karene	Nowak	F	SENIOR DATA SCIENTIST	AUTOMOTIVE	8	GERMANY	EUROPE	7500	5
E245	Nian	Zhen	M	SENIOR DATA SCIENTIST	RETAIL	6	CHINA	ASIA	6500	2

Result 20

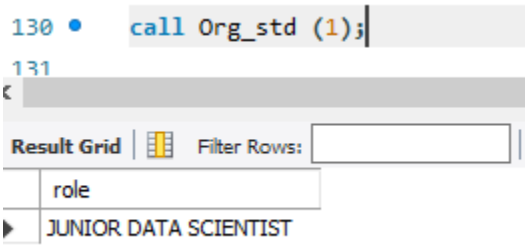
14. Write a query using stored functions in the project table to check whether the job profile assigned to each employee in the data science team matches the organization's set standard.

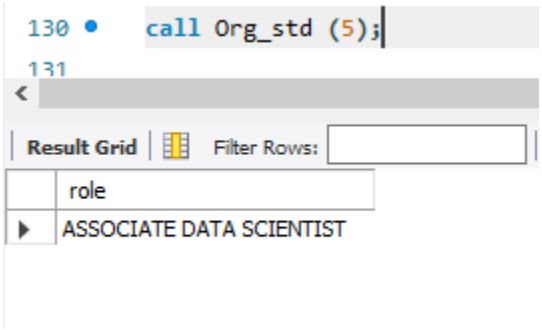
    The standard being:
    For an employee with experience less than or equal to 2 years assign 'JUNIOR DATA SCIENTIST',
    For an employee with the experience of 2 to 5 years assign 'ASSOCIATE DATA SCIENTIST',
    For an employee with the experience of 5 to 10 years assign 'SENIOR DATA SCIENTIST',
  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

Ans:

```
call Org_std (1);

CREATE DEFINER=`root`@`localhost` PROCEDURE `Org_std`(IN exp INT)
BEGIN
DECLARE role VARCHAR (50);
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;
SELECT role;
END
```





```
130 •   call Org_std (11);
131
```

Result Grid | Filter Rows:

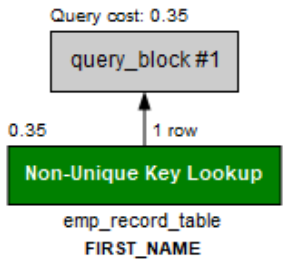
role
LEAD DATA SCIENTIST

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.

Ans:
Select * from emp_record_table where FIRST_NAME = 'Eric';

```
135
136 •   Select * from emp_record_table where FIRST_NAME = 'Eric';
137
```

Visual Explain | Display Info: Read + Eval cost | Overview: | View Source:



emp_record_table 31 ×

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).

Ans: Select first_name, last_name, Salary, ((salary* 0.5)*EMP_RATING) as Bonus from emp_record_table;

```
136 •  Select first_name,last_name,Salary, ((salary* 0.5)*EMP_RATING)
137    as Bonus from emp_record_table;
```

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

first_name	last_name	Salary	Bonus
Arthur	Black	16500	41250.0
Eric	Hoffman	8500	12750.0
William	Butler	9000	9000.0
Dianna	Wilson	5500	13750.0
Dorothy	Wilson	7700	3850.0
Patrick	Voltz	9500	23750.0
Emily	Grove	10500	21000.0
Karene	Nowak	7500	18750.0
Nian	Zhen	6500	6500.0
Roy	Collins	7000	10500.0
Steve	Hoffman	5000	7500.0
Pete	Allen	11000	22000.0
David	Smith	4000	8000.0
Chad	Wilson	5000	5000.0
Claire	Brennan	4300	2150.0
Janet	Hale	10000	10000.0

Result 47

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

Ans:
```
select CONTINENT, COUNTRY,
    avg (salary) over (partition by CONTINENT)
    from emp_record_table
    order by COUNTRY;
```

```
143 •   select CONTINENT,COUNTRY,
144        avg(salary) over(partition by CONTINENT) as AvgSal
145        from emp_record_table
146        order by COUNTRY;
```

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

CONTINENT	COUNTRY	AvgSal
NORTH AMERICA	CANADA	8525.0000
NORTH AMERICA	CANADA	8525.0000
NORTH AMERICA	CANADA	8525.0000
ASIA	CHINA	6250.0000
SOUTH AMERICA	COLOMBIA	5600.0000
SOUTH AMERICA	COLOMBIA	5600.0000
SOUTH AMERICA	COLOMBIA	5600.0000
EUROPE	FRANCE	7950.0000
EUROPE	GERMANY	7950.0000
EUROPE	GERMANY	7950.0000
EUROPE	GERMANY	7950.0000
ASIA	INDIA	6250.0000
ASIA	INDIA	6250.0000
ASIA	INDIA	6250.0000
NORTH AMERICA	USA	8525.0000
NORTH AMERICA	USA	8525.0000

Result 46