SQL

Certification Project



**ScienceQtech Employee Performance Mapping.**

**Course-end Project 1**

**DESCRIPTION**

Science Qtech 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

**Steps to perform:**

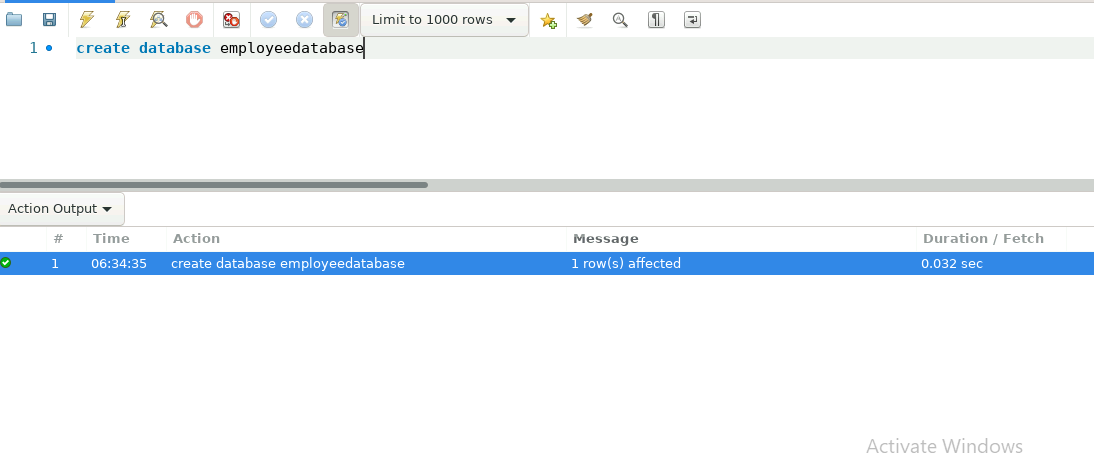
As the first step Load the my sql work bentch , and understand the data in tables from data dictionary .A sanity check, to ensure that you have clean records and the data types are correct and good to go ahead, is very important. And perform further analysis and provide insights where required.

 **SOLUTION**

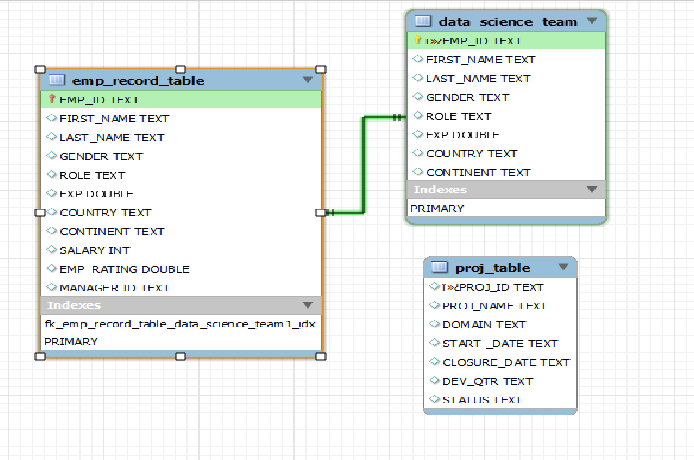
1. Create a database named employee, then import **data\_science\_team.csv** **proj\_table.csv** and **emp\_record\_table.csv** into the **employee** database from the given resources.

Here we created a database named as employee and imported the import data\_science\_team.csv proj\_table.csv and emp\_record\_table.csv into the employee database.

CREATE DATABASE employee;

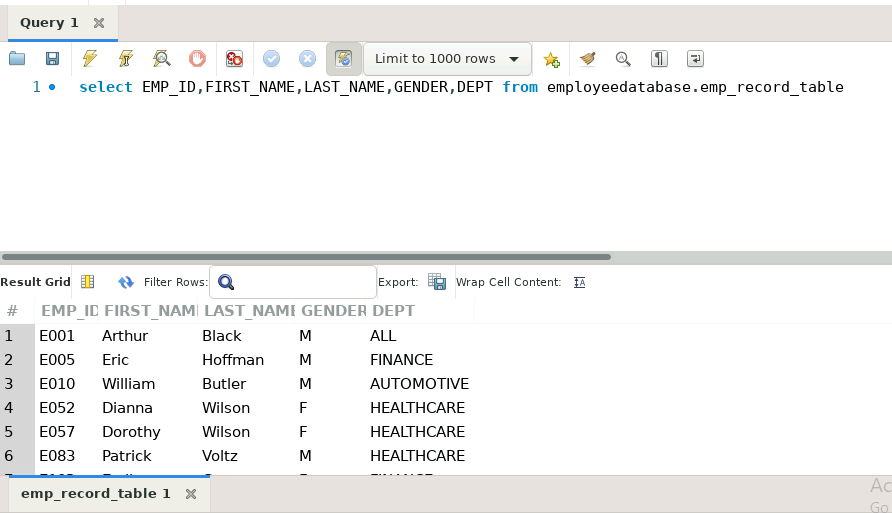


1. Create an ER diagram for the given **employee** database.



1. 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.

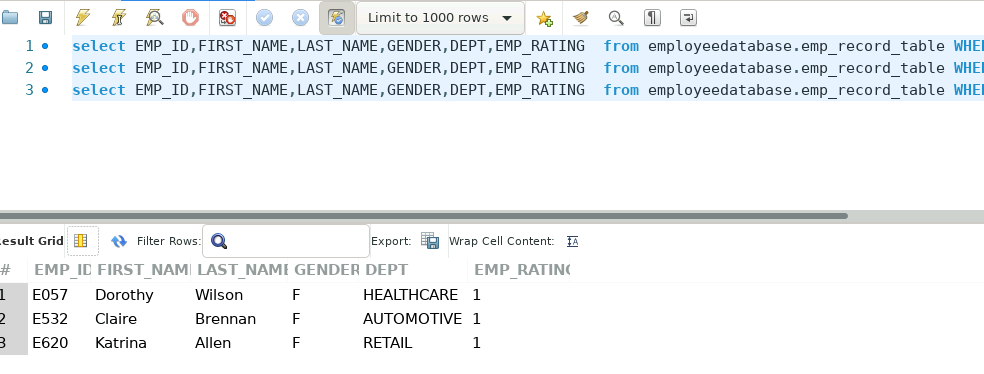
select EMP\_ID,FIRST\_NAME,LAST\_NAME,GENDER,DEPT from emp\_record\_table;



1. 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
* less than two

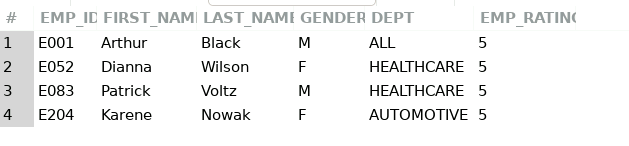
select EMP\_ID,FIRST\_NAME,LAST\_NAME,GENDER,DEPT,EMP\_RATING from emp\_record\_table where EMP\_RATING <2;



* greater than four

select EMP\_ID,FIRST\_NAME,LAST\_NAME,GENDER,DEPT,EMP\_RATING from emp\_record\_table

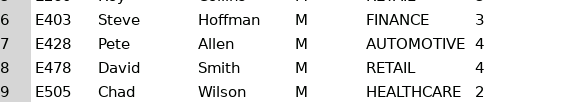
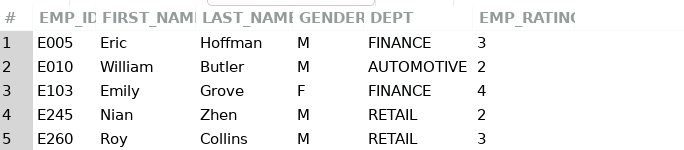
where EMP\_RATING >4;



* between two and four

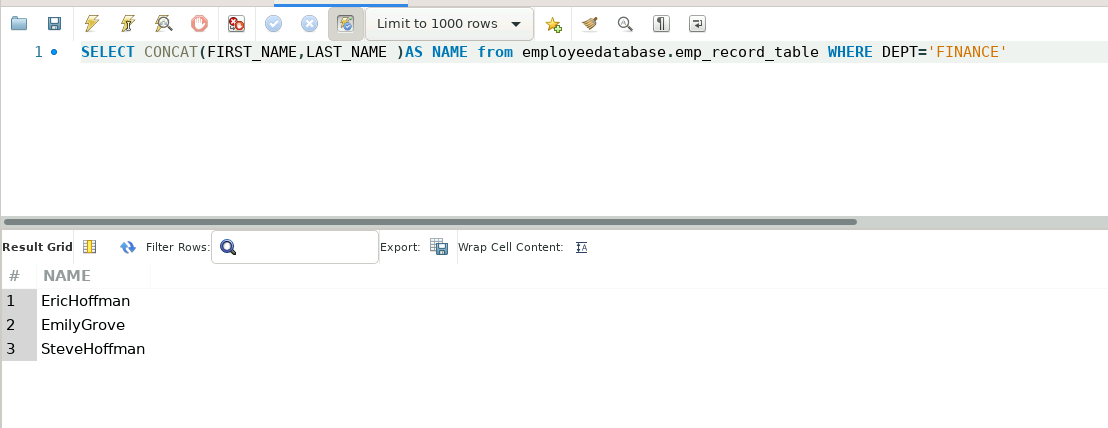
select EMP\_ID,FIRST\_NAME,LAST\_NAME,GENDER,DEPT,EMP\_RATING from emp\_record\_table

where EMP\_RATING between 2 and 4;



1. 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.

select CONCAT( FIRST\_NAME,LAST\_NAME ) AS NAME from emp\_record\_table where DEPT='FINANCE';



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

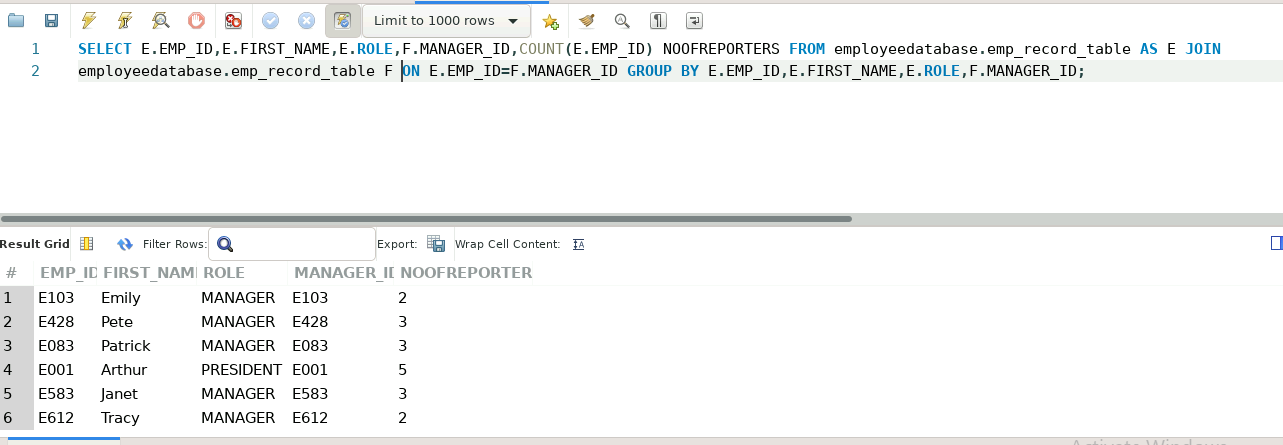
.

SELECT E.EMP\_ID,E.FIRST\_NAME,F.MANAGER\_ID,E.ROLE,count(E.EMP\_ID) AS NOOFREPORTERS FROM employee.dbo.emp\_record\_table AS E JOIN employee.dbo.emp\_record\_table AS F ON E.EMP\_ID=F.MANAGER\_ID group by E.EMP\_ID,E.FIRST\_NAME,F.MANAGER\_ID,E.ROLE;

( Or)

U can use this query

select coalesce(MANAGER\_ID,0),coalesce((count(MANAGER\_ID)),0)as count from [employee].[dbo].[emp\_record\_table] group by MANAGER\_ID;

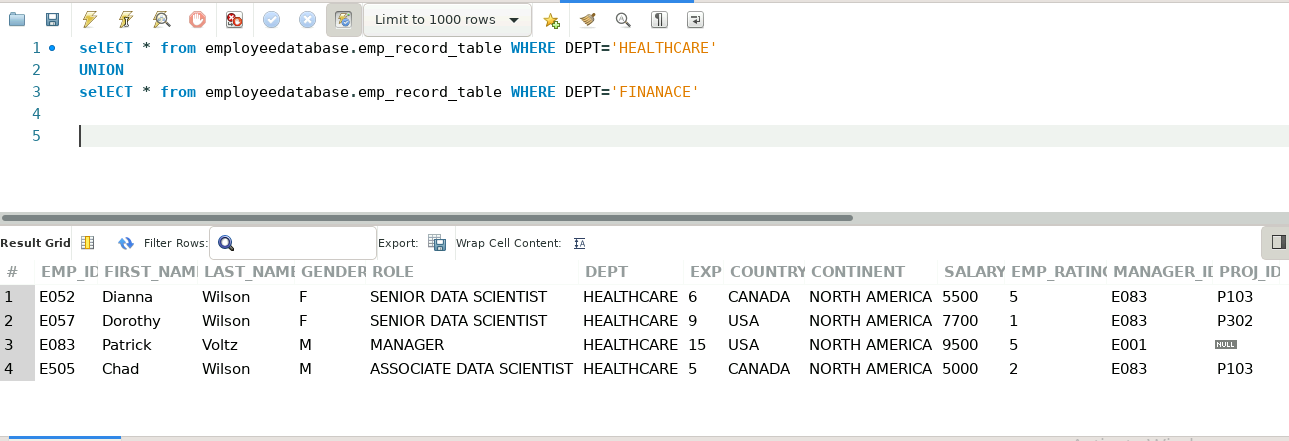


1. Write a query to list down all the employees from the healthcare and finance departments using union. Take data from the employee record table.

select \* FROM emp\_record\_table where DEPT='healthcare'

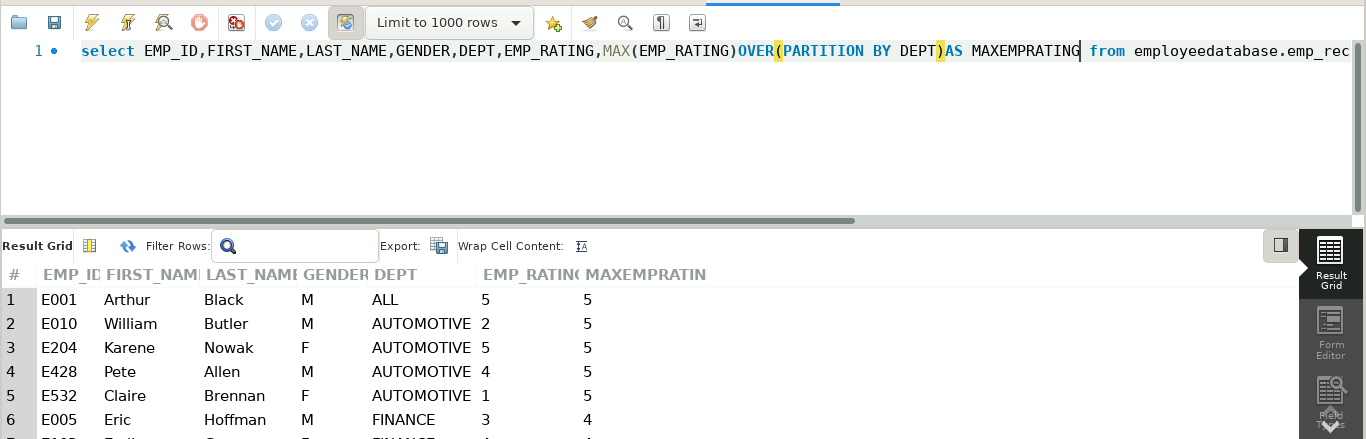
UNION

SELECT\* FROM emp\_record\_table where DEPT='finance'



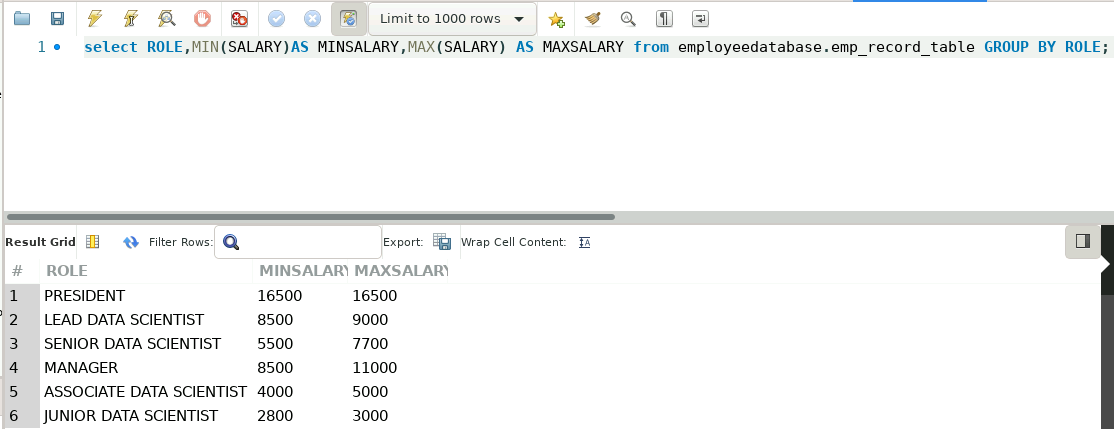
1. Write a query to list down employee details such as EMP\_ID, FIRST\_NAME, LAST\_NAME, ROLE, DEPARTMENT, and EMP\_RATING grouped by dept. Also include the respective employee rating along with the max emp rating for the department.

select EMP\_ID,FIRST\_NAME,LAST\_NAME,ROLE,DEPT ,EMP\_RATING,max(EMP\_RATING) over(partition by DEPT ) AS MAXRATINGBYDEPT from employee.dbo.emp\_record\_table



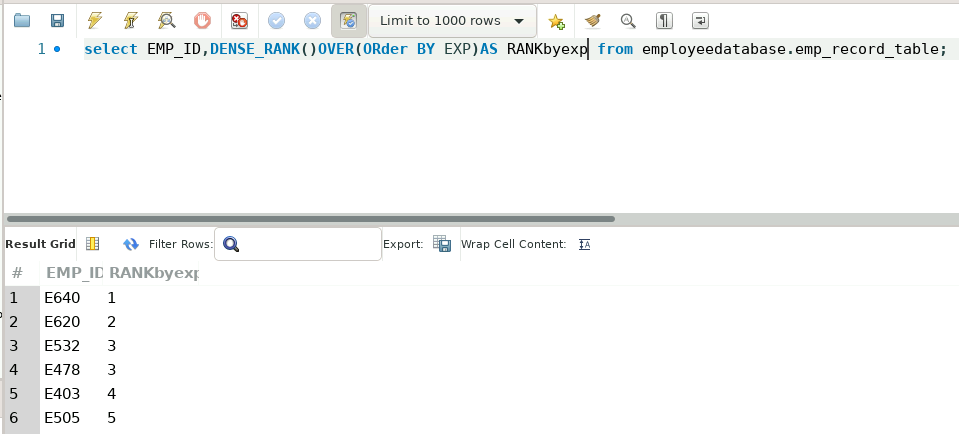
9 .Write a query to calculate the minimum and the maximum salary of the employees in each role. Take data from the employee record table.

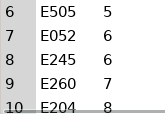
Select ROLE, min(SALARY) AS MINSALARY,max(SALARY)AS MAXSALARY FROM emp\_record\_table group by ROLE;

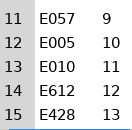


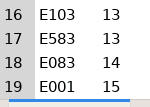
1. Write a query to assign ranks to each employee based on their experience. Take data from the employee record table.

SELECT EMP\_ID, DENSE\_RANK()OVER(ORDER BY EXP)AS RANKofexp FROM emp\_record\_table;









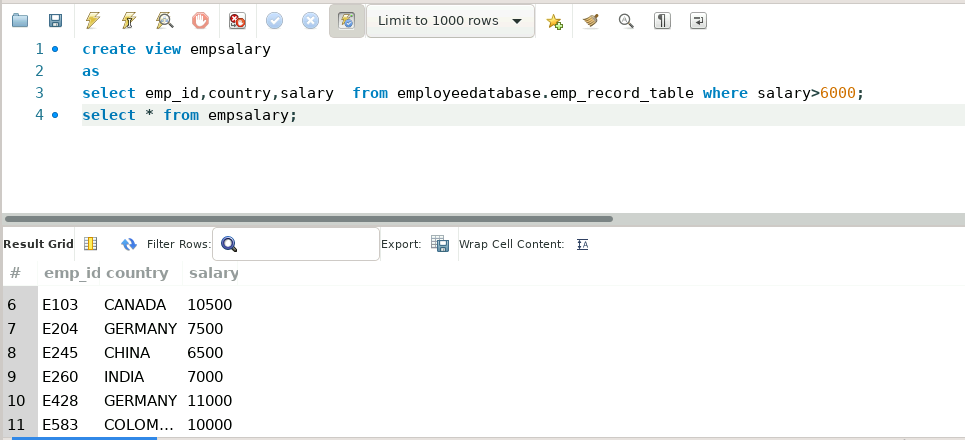
1. Write a query to create a view that displays employees in various countries whose salary is more than six thousand. Take data from the employee record table.

  CREATE VIEW EMPSALARY

AS

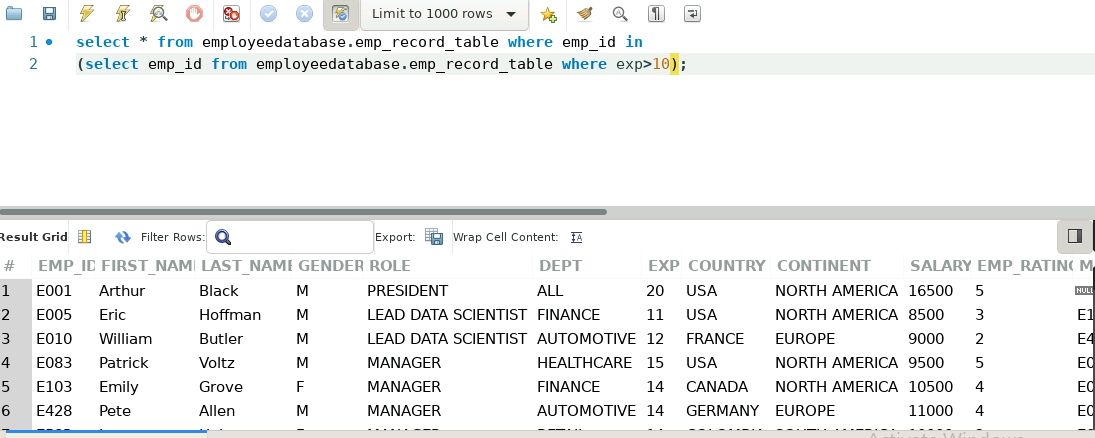
SELECT EMP\_ID,COUNTRY,SALARY FROM emp\_record\_table where SALARY>6000

SELECT\* FROM EMPSALARY;



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

SELECT \* FROM emp\_record\_table where EMP\_ID IN(SELECT EMP\_ID FROM EMP\_RECORD\_TABLE WHERE EXP>10)



1. 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.

CREATE PROCEDURE EXPMORETHAN3()

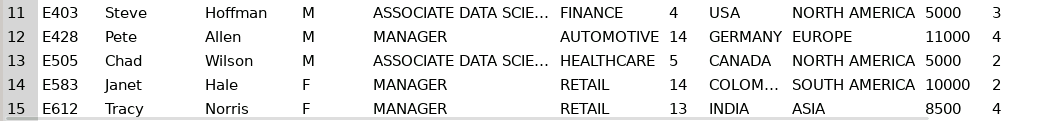
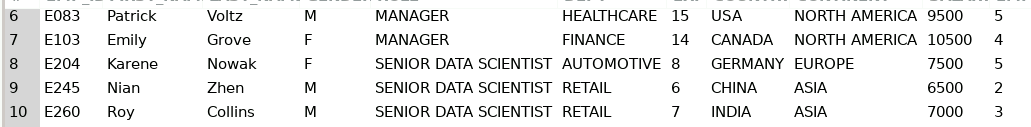
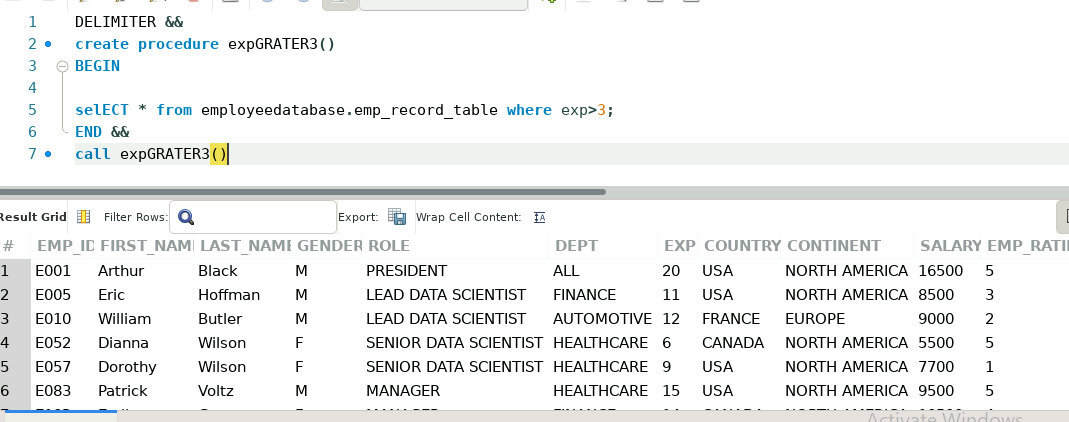
As

begin

SELECT \* FROM emp\_record\_table where EXP>3;

end

CALL EXPMORETHAN3



1. 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'.**

DELIMITER &&

CREATE FUNCTION Employee\_ROLE(EXP int)

RETURNS VARCHAR(40)

DETERMINISTIC

BEGIN

DECLARE Employee\_ROLE VARCHAR(40);

IF EXP BETWEEN 12 AND 16 THEN

SET Employee\_ROLE="MANAGER";

ELSEIF EXP BETWEEN 10 AND 12 THEN

SET Employee\_ROLE ="LEAD DATA SCIENTIST";

ELSEIF EXP BETWEEN 5 AND 10 THEN

SET Employee\_ROLE ="SENIOR DATA SCIENTIST";

ELSEIF EXP BETWEEN 2 AND 5 THEN

SET Employee\_ROLE ="ASSOCIATE DATA SCIENTIST";

ELSEIF EXP<=2 THEN

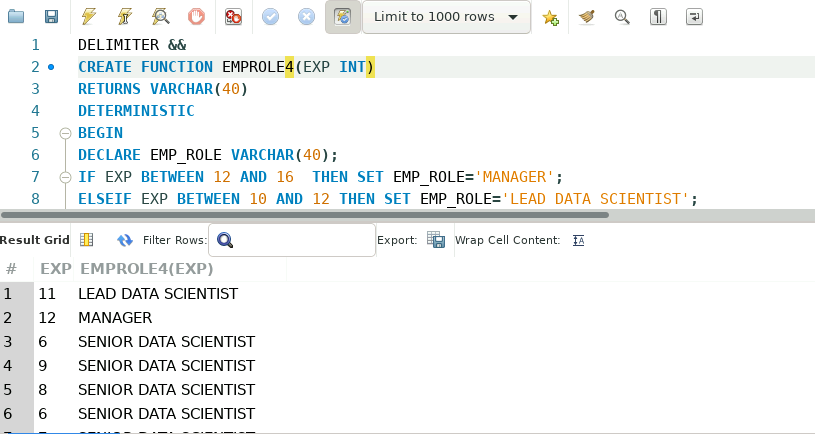
SET Employee\_ROLE ="JUNIOR DATA SCIENTIST";

END IF;

RETURN (Employee\_ROLE);

END &&

SELECT EXP,Employee\_ROLE(EXP) FROM data\_science\_team;



1. 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.

Create index idx\_emprecordtbale on employee.emp\_record\_table(EMP\_ID,FIRST\_NAME);

Select emp\_id from emp\_record\_table where first\_name=’Eric’

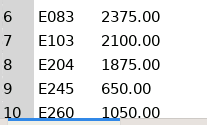
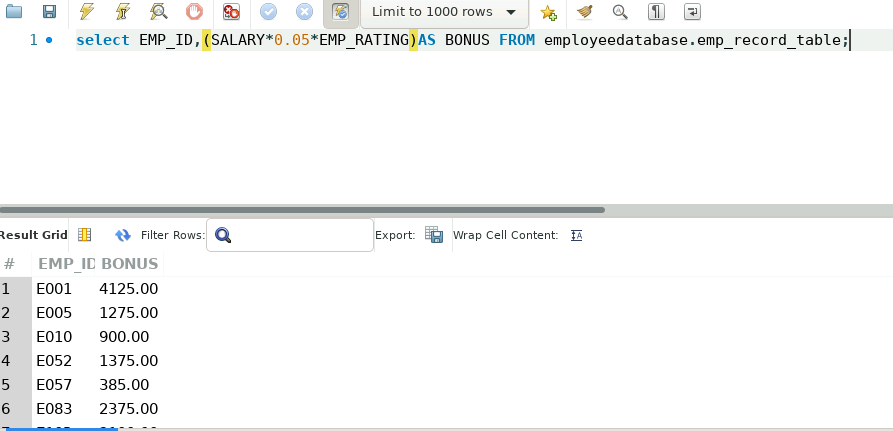
The index is created on the emp\_id,first\_name column to make the retrival of data fast.by this we can increase cost and performance of the query .

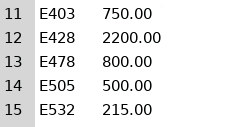


1. Write a query to calculate the bonus for all the employees, based on their ratings and salaries (Use the formula: 5% of salary \* employee rating).

select EMP\_ID,(SALARY\*0.05\*EMP\_RATING) AS BONUS FROM

employee.emp\_record\_table;





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

SELECT EMP\_ID,FIRST\_NAME,LAST\_NAME,SALARY,COUNTRY,CONTINENT,

AVG(salary)OVER(PARTITION BY COUNTRY)AVG\_salary\_IN\_COUNTRY,  
AVG(salary)OVER(PARTITION BY CONTINENT)AVG\_salary\_IN\_CONTINENT,  
COUNT(\*)OVER(PARTITION BY COUNTRY)COUNT\_IN\_COUNTRY,  
COUNT(\*)OVER(PARTITION BY CONTINENT)COUNT\_IN\_CONTINENT  
FROM emp\_record\_table;

