Data Acquisition and Manipulation using SQL

# Employee Performance Mapping

Overview

In this project, you will apply the fundamentals of SQL programming to practice, from data preparation to subqueries, and functions. The focus is on building an employee performance mapping solution for a data science startup. You will use SQL to accurately map employees and track their performance.

# Situation

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.

# Task

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 meet the organization's profile standard. You also need to calculate bonuses to find extra costs for expenses. This will improve the organization's overall performance by ensuring all the employees required receive training.

# Input Dataset

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 – The employee’s years of experience
* COUNTRY – The employee’s current country of residence
* CONTINENT – The employee’s continent of residence
* SALARY – Salary of the employee
* EMP\_RATING – Performance rating of the employee
* MANAGER\_ID – The manager mapped to the employee
* 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

# Action

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.
2. Create an ER diagram for the given employee database.
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.
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
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.
6. Write a SQL query to retrieve the employee ID, first name, role, and department of employees who hold leadership positions (Manager, President, or CEO).
7. Write a query to list all the employees from the healthcare and finance departments using the union. Take data from the employee record table.
8. Write a query to list 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.
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.
10. Write a query to assign ranks to each employee based on their experience. Take data from the employee record table.
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.
12. Write a nested query to find employees with experience of more than ten years. Take data from the employee record table.
13. 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'.

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.
2. 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).

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