
SQL Training

Course-End Project Problem Statement



Get Certified. Get Ahead.

ASSESSMENT – 1 - WRITEUP

Submitted By

Aruna Rajeswari Sivaperuman

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

DATASET OBSERVATION

Three tables were given.

emp_record_table

Stores all employee details including:

- Personal Info: EMP_ID, FIRST_NAME, LAST_NAME, GENDER, COUNTRY, CONTINENT
- Professional Info: ROLE, DEPT, EXP, SALARY, EMP_RATING, MANAGER_ID, PROJ_ID

Proj_table

Holds project information:

- PROJECT_ID, PROJ_NAME, DOMAIN, START_DATE, CLOSURE_DATE, DEV_QTR, STATUS

data_science_team

- Subset of emp_record_table for the Data Science department.

TASKS COMPLETED AND QUERIES EXECUTED

1. Database and Table Creation

Created a database named employee and imported emp_record_table.csv, proj_table.csv, and data_science_team.csv.

2. ER Diagram Creation

Developed an Entity Relationship (ER) diagram showing relationships between:

emp_record_table ↔ proj_table (via PROJ_ID)

emp_record_table ↔ itself (via MANAGER_ID)

data_science_team as a subset of emp_record_table

3. Basic Employee List by Department

Extracted employee ID, names, gender, and department from emp_record_table.

4. Employee Ratings – Filtered Results

Filtered records to show employees with ratings:

Less than 2

Between 2 and 4

Greater than 4

5. Concatenate Names for Finance Team

Merged FIRST_NAME and LAST_NAME for employees in the Finance department with alias NAME.

6. List of Managers and Number of Reportees

Identified all employees who have others reporting to them (including the President) and calculated number of direct reports.

7. Union Query for Healthcare and Finance Departments

Used UNION to combine employee details from Healthcare and Finance departments.

8. Group Employees by Department with Max Rating

Grouped employees by department and included both individual ratings and department's max rating.

9. Salary Range by Role

Used MIN() and MAX() functions to show salary distribution by role.

10. Rank Employees by Experience

Used RANK() or DENSE_RANK() to rank employees based on their years of experience.

11. View for High-Earning Employees by Country

Created a view for employees earning more than 6000, grouped by their respective countries.

12. Nested Query for Experienced Employees

Used a nested query to find employees with over 10 years of experience.

13. Stored Procedure: Experience > 3 Years

Created a stored procedure that returns details of employees with more than 3 years of experience.

14. Stored Function to Validate Job Role Standards

Implemented a stored function to map experience to expected roles:

≤ 2 yrs → Junior Data Scientist

2–5 yrs → Associate Data Scientist

5–10 yrs → Senior Data Scientist

10–12 yrs → Lead Data Scientist

12–16 yrs → Manager

Compared it with actual roles from the data_science_team.

15. Index Creation on FIRST_NAME

Added an index on the FIRST_NAME column to optimize search.

Resolved BLOB/TEXT error by specifying a prefix length (FIRST_NAME(20)).

16. Bonus Calculation Based on Salary and Rating

Calculated bonus using:

Bonus = 5% of Salary × Employee Rating

17. Average Salary by Continent and Country

Grouped data by CONTINENT and COUNTRY and calculated average salary per region.

INSIGHTS AND OBSERVATIONS

- Employee experience strongly influences roles and compensation.
- Performance ratings vary significantly across departments.
- Managers with many direct reports may require support or delegation.
- Some roles do not align with expected titles based on experience, suggesting a need for restructuring or training.
- Indexing on searchable fields like FIRST_NAME significantly boosts performance.

CONCLUSION

This project successfully utilized SQL to analyze employee data at ScienceQtech, delivering actionable insights for HR appraisal and organizational performance mapping. By combining analytical queries, procedural SQL, and performance optimization techniques, the task met all defined objectives.