DATA AND
ARTIFICIAL INTELLIGENCE

SQL Training
Course-end Project

simplilearn

DATA AND
ARTIFICIAL INTELLIGENCE

ScienceQtech Employee Performance Mapping

Objectives

- To find the maximum salary of the employees and ensure that all jobs meet the organization's profile standard
- To calculate the bonuses to find the extra cost of the expenses

Prerequisites

- ER diagram
- Working of database
- Working of tables
- SQL functions
- SQL queries

Industry Relevance

- ER diagram: It is used to visualize the structure of a table as well as the relationships between logically related tables.
- Database: It is a collection of tables that store a specific set of structured data.
- Tables: It is a database object that contains all the data within it.
- SQL functions: Several built-in functions are available in SQL to calculate data.
- SQL queries: A query is a request for data or information from a database table or combination of tables. As a result of the structured query language (SQL), this data may be displayed as pictorials, graphs, or complex results, such as trend analyses from data mining tools.

Problem Statement

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 the project that the employees have undertaken, to analyze the employee database and extract specific data based on different requirements.

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.

Variable	-	Description
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

Dataset Description

Variable	-	Description
EXP	-	Years of experience the employee has
COUNTRY	-	The 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

Dataset Description

Proj_table: It contains information about the projects.

Variable	-	Description
PROJECT_ID	-	ID for the project
PROJ_Name	-	Name of the project
DOMAIN	-	Field of the project
START_DATE	-	The day the project began
CLOSURE_DATE	-	Day the project was or will be completed
DEV_QTR	-	Quarter in which the project was scheduled
STATUS	-	Current status of the project

Dataset Description

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

Variable	-	Description
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

Tasks to Perform

Perform the following tasks on the dataset provided using SQL:

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

Tasks to Perform

Perform the following tasks on the dataset provided using SQL:

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

Tasks to Perform

Perform the following tasks on the dataset provided using SQL:

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.

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.

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.

Tasks to Perform

Perform the following tasks on the dataset provided using SQL:

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.

Tasks to Perform

Perform the following tasks on the dataset provided using SQL:

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.

Tasks to Perform

Perform the following tasks on the dataset provided using SQL:

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 is given as follows:

    - Employee with experience less than or equal to 2 years, assign 'JUNIOR DATA SCIENTIST'
    - Employee with experience of 2 to 5 years, assign 'ASSOCIATE DATA SCIENTIST'
    - Employee with experience of 5 to 10 years, assign 'SENIOR DATA SCIENTIST'
    - Employee with experience of 10 to 12 years, assign 'LEAD DATA SCIENTIST',
    - Employee with experience of 12 to 16 years, assign 'MANAGER'

Tasks to Perform

Perform the following tasks on the dataset provided using SQL:

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.

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).

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

Project Outcome

- This project's goals are to determine the employees' maximum pay and confirm that every position satisfies the organization's profile criterion.
- This project involves calculating bonuses by determining the additional costs of the expenses.

DATA AND
ARTIFICIAL INTELLIGENCE

Submission Process

1. Complete the project in the Simplilearn lab
2. Complete each task listed in the problem statement
3. Take screenshots of the results for each question and the corresponding code
4. Save it as a document and submit using the assessment tab
5. Tap the "Submit" button (this will present you with three choices)
6. Attach three files and then click "Submit“

Note: Be sure to include screenshots of the output

DATA AND
ARTIFICIAL INTELLIGENCE

Thank You

simplilearn