

95 – 703 B: Database Management

DB Implementation Project

This project is to be completed in Oracle Database 11g Express by groups of two students. Pairing will be provided by the instructor. This document is broken down into the following parts: description of the model (Part I), the specific tasks you need to do (Part II), and what is required for submission (Part III). You have about six weeks to complete the project. Submit your project (hardcopy) on December 1st, by class time.

Part I. Description of the model

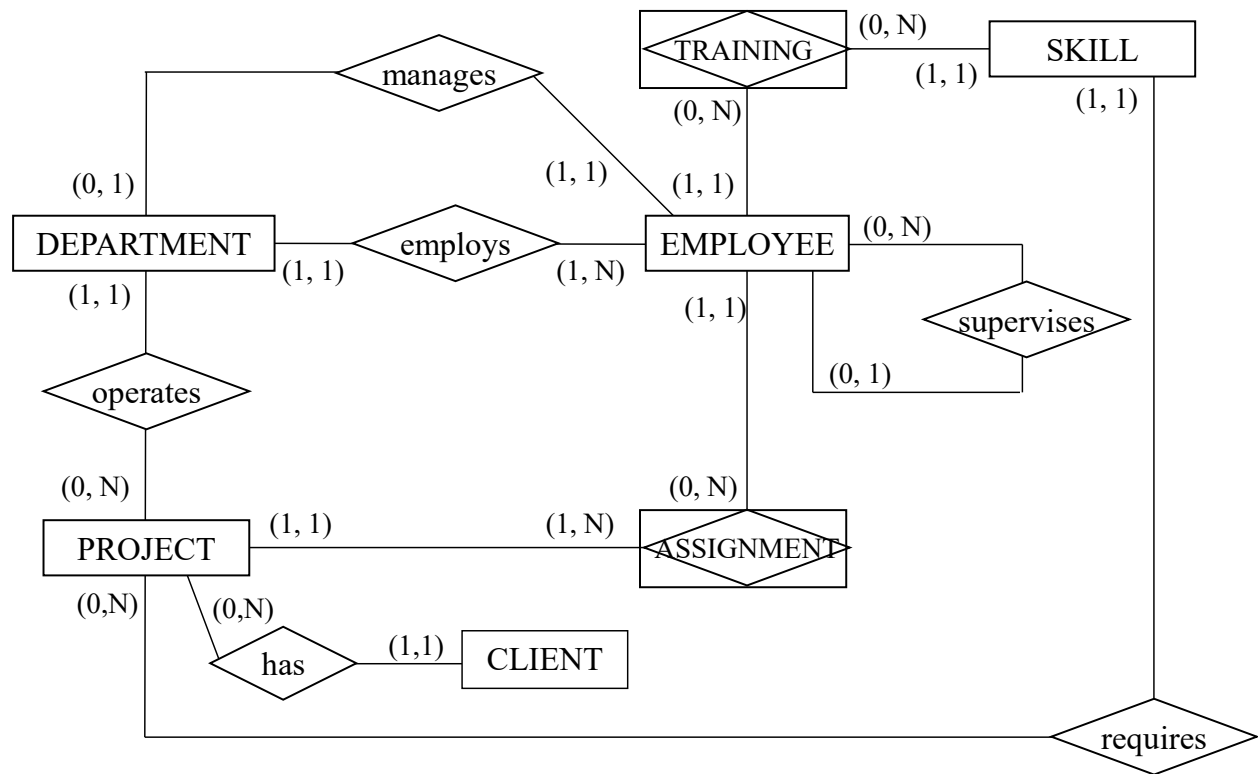
In *Homer Consulting* each employee is employed by one department. One of the employees from a department manages the department. Some employees (especially in larger departments) supervise other employees in their department. Thus, one employee may have a supervisor and may also supervise other employees as well. The CEO supervises the top-level employee of each department.

The company runs projects for external clients. Information on both current clients (those that have had projects done in the past) as well as potential clients (those who have not requested projects yet) is kept. For billing purposes, each project is contained within one department. Administration department does not run any projects. For each project start date and the total cost is recorded. The total cost of a project can only be estimated (and recorded) after the project is completed. For an ongoing project, the total cost will have a NULL value.

Each project has several employees working on (assigned to) the project. Most (but not necessarily all) of employees assigned to a project come from one department. Newly hired employees are not assigned to any project for the first few months and may be trained internally or by attending a formal workshop. Each assignment is recorded on month-by-month base. Within each month the date employee was assigned to that project, and the end date of the assignment in that month (not necessarily the last day of the month) is recorded. In addition, the total number of hours spent by the employee on the project in that month is kept. If an employee worked on a project for two months, then two records would be created, one record for each month. Number of hours used within each month is known and recorded after the monthly assignment is completed. For an assignment in a current month, both the end date and hours spent are NULL.

Homer Consulting maintains a list of desired skills. For each desired skill, code, name and category is kept. Not all of these skills might be represented among the employees. Whenever an employee acquires a skill, through some form of training (i.e. completing a course at a college, completing a technical training program, or a workshop) approved by the company, the date when it was acquired and a comment regarding the training is recorded. Each employee may take the same training more than once since some of the skills might need to be updated or renewed. Each project requires one “major” skill.

ER Diagram:



Schema:

SKILL (Code, Name, Category)

TRAINING (Train_Num, Code[@], Emp_Num[@], Name, Date_Acquired, Comments)

DEPARTMENT (Dept_Code, Name, Location, Phone, Manager_ID[@])

EMPLOYEE (Emp_Num, LName, FName, DOB, Hire_Date, Super_ID[@], Dept_Code[@])

CLIENT (Client_ID, Name, Street, City, State, Zip_Code, Industry, Web_Address, Phone, Contact_LName, Contact_FName)

PROJECT (Proj_Number, Name, Start_Date, Total_Cost, Dept_Code[@], Client_ID[@], Code[@])

ASSIGNMENT (Assign_Num, Proj_Number[@], Emp_Num[@], Date_Assigned, Date_Ended, Hours_Used)

Part II.

A. Create two separate SQL scripts for creating tables, data integrity constraints, and for inserting data into the tables to implement the database as defined above.

1. Script #1: Write all CREATE TABLE statements that have the necessary integrity constraints including primary keys, foreign keys, check constraints, etc. Start the script with a set of DROP TABLE statements that will allow you to clean up the database before creating it (useful when you recreate the database using the script). Use the names of tables as well as the names of attributes exactly as provided in the model above. *All constraints, except one, must be created and enforced within the “Create Table” statements. All constraints must be named using the convention discussed in class (lecture 12).* Run the script to create the DB objects and to create the missing constraint (as a last statement in your script). Your CREATE TABLE statements must also include the constraints defined below:
 - a. The employee’s *Hire_Date*, the project *Start_Date*, and the *Date_Assigned* should have a default value that is current date,
 - b. The *Date_Ended* cannot be earlier than the *Date_Assigned*, and
 - c. Three more appropriate and useful check constraints (but not the *Not Null* constraints) of your choice. Make sure that you clearly identify them.
2. Script #2: After the DB tables and all data integrity constraints are created, insert about ten rows of data into each table. Provide data that would allow you to test all queries as defined in section B below (each query must give you results). It may be required to add more than 10 rows in some tables in order to show that your queries work as intended. The *Department* and *Skill* tables could have a few rows. Run the script to perform the task of populating the tables in your database. (*Note: All constraints must be enabled and enforced – not deferred – before you populate your tables*)
3. After all tables are populated, list the complete content of each table. Format content of each table to make it readable and understandable.

B. Create SQL queries to answer the following questions:

1. The company’s human resources department needs information on the current training levels of the employees. They need information that shows, for each employee, and his/her acquired skills, the number of times he or she received training for that skill, the earliest date of the training, and the number of months (full months only) that have passed since the most recent training for the skill.
2. Due to a recent re-organization, the company needs a list that shows who supervises who. List all the employees in the company clearly indicating the organizational hierarchy. Include the “level” of the hierarchy each employee is at and the department name of the employee.
3. For every ongoing project provide the project name, project start date. In addition, for each month of the project list the number of employees assigned that month and hours spent on the project that month. At the end provide an overall total of the number of employees assigned to the project and the total hours spent on the project.

4. The company has decided to adopt a bonus policy for their employees. Each employee will receive a bonus of \$200 for every project started first quarter of this calendar year, on which they have charged at least 150 hours. As a result, management has asked you to update the employee records to include their current bonus information. In order to accomplish this, you will have to add another column to the EMPLOYEE table (i.e., column BONUS_AMT), before you update the data. Use a single Update statement to add values to the new column, and then list the contents of the Employee table.
5. For each employee hired in the second quarter of this calendar year, list ID & name, the hire date, name of any training received (if true), date of the training, and the number of days between the hire date and the training. Also include number of projects that employee worked on so far (if any).
6. Some projects are discontinued for certain period of time and resumed later. Find out project name and start date of those projects that have discontinuous activity. Indicate whether the “discontinued” projects have been completed (i.e., your output should clearly indicate this by “completed” or “on-going” value in a column).
7. We need a summary of project information for each quarter of this year. For each quarter include: number of projects started in the quarter, number of employees working on those projects in the quarter, number of hours spent on average per project.
8. For each employee and each skill, list the number of times the employee had training and the latest date (most recent) when the employee acquired that particular skill. Also, provide the number of trainings provided/attended for each skill, and the number of skills acquired by each employee. The output should look like this:

ID	Employee Name	ETL	Latest Date Acquired	Spark	Latest Date Acquired	Python	Latest Date Acquired	Tableau	Latest Date Acquired	Number of Skills:
100	Mary Smith	1	10/20/21	0	-----	0	-----	0	-----	1
110	Alice Rodriguez	0	-----	1	11/02/21	2	6/24/21	2	9/07/21	3
120	Kunal Shah	1	9/25/21	2	6/02/21	2	11/23/20	1	4/07/21	4
130	Tao Li	1	6/23/21	1	12/21/20	1	11/23/20	1	10/17/20	4
140	Pat Gomez	0	-----	1	11/02/21	1	6/24/21	1	10/17/20	3
---	Number of Trainings:	3	-----	5	-----	6	-----	5	-----	-----

(Note: the values in the table are made up; use the names/values from your own database):

9. For each department and each skill combination, list the number of trainings completed within the department that was associated with the particular skill. Also provide a rank of each skill within each department. The rank should be based on the number of trainings completed for that skill. Same rank should be given when the number of trainings is the same.
10. List project(s) with at least five monthly assignment records and with no less than 60 days in the top three records/months. Also list the total number of days employees worked on those projects.
11. For the 4 most senior employees provide, the employee ID, last name, hire date, the name of department managed by the employee (if any), and the number of employees he or she supervises directly.

12. We are interested in what types of clients we have. Based on the “web address” information, group the client by the type of client (e.g. “edu” indicates an educational institution, “gov” a government agency, “org” a non-for-profit organization, and “com” a for-profit company)¹ and provide how many clients we have of each type/category and how many project from each type of client. If the web address does not match any of the four major types, then use “Other” for the client type. If we do not have a web address for a client, then the “client type” should be “Not Available”.
13. In order to schedule training for current employees the company requires information on employees who were not assigned to any project after July of this year. List the names of employees and their departments. Include the last project the employee was assigned to. Order the result by department name and then by employee last name.
14. For each skill category, list number of trainings employees have had and the number of projects that required skills of that category. Include the grand total as last row.
15. Create a single query that accesses the system catalog to retrieve information about your database tables and their columns, as well as any constraints on those columns. Include the following:
 - a. Table name
 - b. Column name
 - c. Constraint name
 - d. Constraint type. Instead of the Oracle’s one letter codes, use the two letter indicators as defined in lecture #11
 - e. Search condition (for any check constraints), and
 - f. Table name and column name that each FK references

Sort the result by table name. List column in the same order as they were created in their tables. Include commands used for formatting the results (e.g., use BREAK command to suppress repeating names of tables and columns and COLUMN command on individual columns as needed).

Part III.

Create a title page that includes your names, email addresses, and the group number as it was assigned to your group. Clearly identify each part of your submitted project and each answer. Submit the following:

1. Printout of the two separate SQL script files (*Script#1* & *Script#2*) as you created in Part II.A. as well as the content of the tables. (*Note: Do not print your “spool file” of executing the scripts*)
2. The queries created in Part II.B. Include the SQL statements as you execute them as well as the results. Make sure you appropriately format both the statements and results (refer to the “Formatting SQL Output” document provided for your SQL Assignments).

¹ Example: <http://www.heinz.cmu.edu/>

Grading

Projects will be graded based on the following criteria:

- Correctness of SQL based on requirements given in the assignment. Although there may be more than one way to approach each question, the SQL query/results must match the requirements as defined in each question.
- Readability of your SQL statements as they are executed and well formatted results.
- Punctuality. The deadline for the submission is class time on December 1st.
 - Projects submitted on December 1rd but later than the class time will receive a 20% penalty
 - Projects will not be accepted after December 1st.