MYSQL

Course-end Project 1

BY: Manar Khamis

ScienceQtech Employee Performance Mapping

Course-end Project 1

Description

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.

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.

- 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

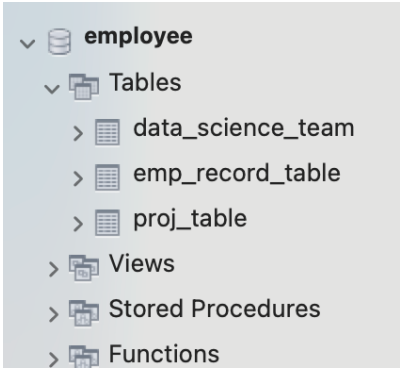
Data_science_team: It contains information about all the employees in the Data Science team.

- 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

The task to be performed:

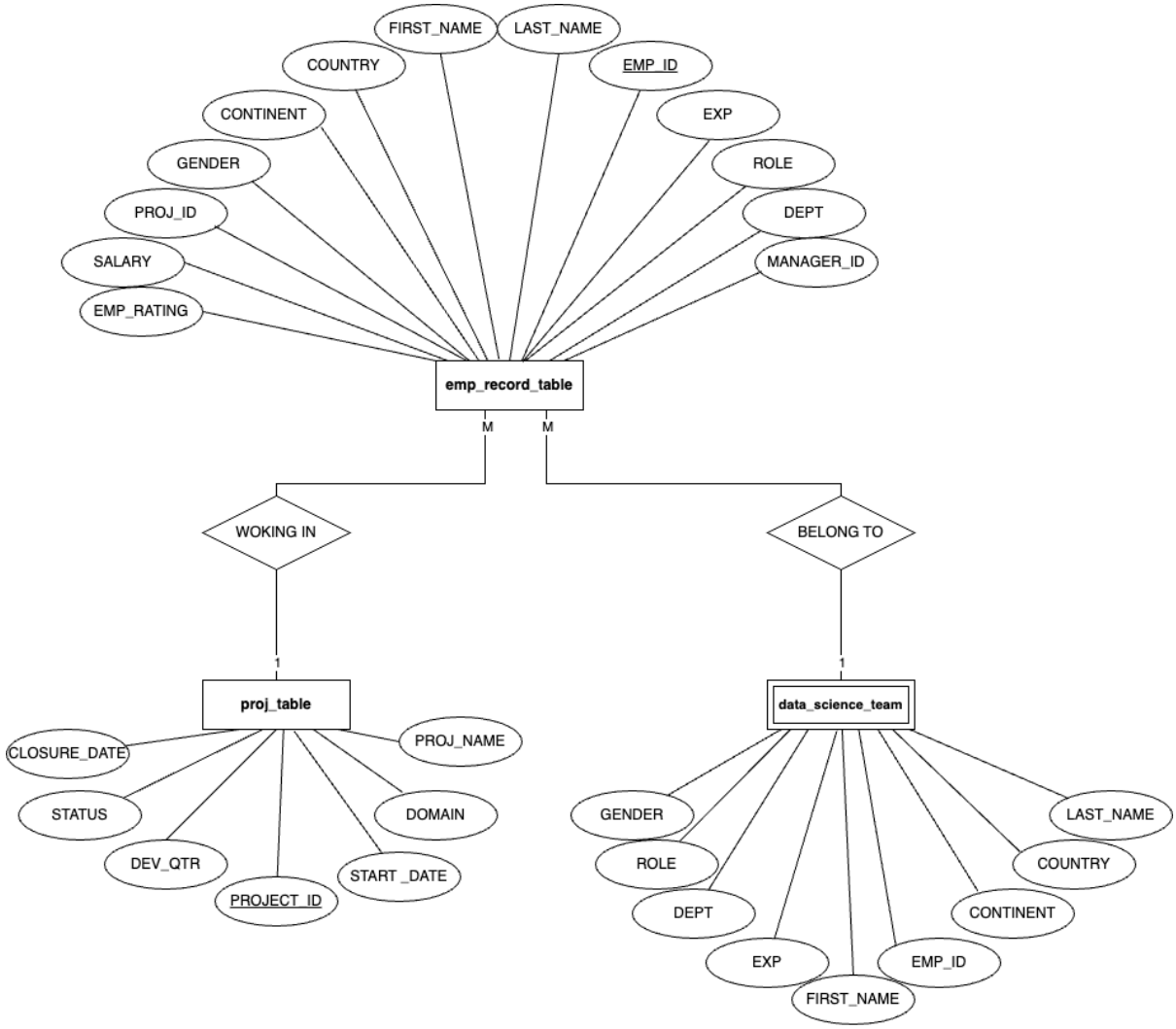
1. Create a database named employee, then import **data_science_team.csv proj_table.csv** and **emp_record_tabl e.csv** into the **employee** database from the given resources.

```
/* 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.
 */

CREATE DATABASE employee;

USE employee;
```



---

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.

```
SELECT EMP_ID, FIRST_NAME, LAST_NAME, GENDER, DEPT
FROM emp_record_table;
```

EMP_ID	FIRST_NAME	LAST_NAME	GENDER	DEPT
E001	Arthur	Black	M	ALL
E005	Eric	Hoffman	M	FINANCE
E010	William	Butler	M	AUTOMOTIVE
E052	Dianna	Wilson	F	HEALTHCARE
E057	Dorothy	Wilson	F	HEALTHCARE
E083	Patrick	Voltz	M	HEALTHCARE
E103	Emily	Grove	F	FINANCE
E204	Karene	Nowak	F	AUTOMOTIVE
E245	Nian	Zhen	M	RETAIL
E260	Roy	Collins	M	RETAIL
E403	Steve	Hoffman	M	FINANCE
E428	Pete	Allen	M	AUTOMOTIVE
E478	David	Smith	M	RETAIL
E505	Chad	Wilson	M	HEALTHCARE
E532	Claire	Brennan	F	AUTOMOTIVE
E583	Janet	Hale	F	RETAIL
E612	Tracy	Norris	F	RETAIL
E620	Katrina	Allen	F	RETAIL
E640	Jenifer	Jhones	F	RETAIL

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

```
-- less than two
SELECT EMP_ID, FIRST_NAME, LAST_NAME, GENDER, DEPT, EMP_RATING
FROM emp_record_table
WHERE EMP_RATING < 2;
```

EMP_ID	FIRST_NAME	LAST_NAME	GENDER	DEPT	EMP_RATING
E057	Dorothy	Wilson	F	HEALTHCARE	1
E532	Claire	Brennan	F	AUTOMOTIVE	1
E620	Katrina	Allen	F	RETAIL	1

- greater than four

```
-- greater than four
SELECT EMP_ID, FIRST_NAME, LAST_NAME, GENDER, DEPT, EMP_RATING
FROM emp_record_table
WHERE EMP_RATING > 4;
```

EMP_ID	FIRST_NAME	LAST_NAME	GENDER	DEPT	EMP_RATING
E001	Arthur	Black	M	ALL	5
E052	Dianna	Wilson	F	HEALTHCARE	5
E083	Patrick	Voltz	M	HEALTHCARE	5
E204	Karene	Nowak	F	AUTOMOTIVE	5

- between two and four

```
-- 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;
```

EMP_ID	FIRST_NAME	LAST_NAME	GENDER	DEPT	EMP_RATING
E005	Eric	Hoffman	M	FINANCE	3
E010	William	Butler	M	AUTOMOTIVE	2
E103	Emily	Grove	F	FINANCE	4
E245	Nian	Zhen	M	RETAIL	2
E260	Roy	Collins	M	RETAIL	3
E403	Steve	Hoffman	M	FINANCE	3
E428	Pete	Allen	M	AUTOMOTIVE	4
E478	David	Smith	M	RETAIL	4
E505	Chad	Wilson	M	HEALTHCARE	2
E583	Janet	Hale	F	RETAIL	2
E612	Tracy	Norris	F	RETAIL	4
E640	Jenifer	Jhones	F	RETAIL	4

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.

```sql
SELECT CONCAT(FIRST_NAME, ' ', LAST_NAME) AS NAME
FROM emp_record_table
WHERE DEPT = 'Finance';
```

NAME
Eric Hoffman
Emily Grove
Steve Hoffman

---

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

```sql
SELECT  MANAGER_ID, COUNT(EMP_ID) AS reporters_count
FROM emp_record_table
GROUP BY MANAGER_ID;
```

MANAGER_ID	reporters_count
NULL	1
E103	2
E428	3
E083	3
E001	5
E583	3
E612	2

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.

```
SELECT EMP_ID,FIRST_NAME, LAST_NAME, ROLE,DEPT  FROM emp_record_table
WHERE DEPT = 'Finance'
UNION
SELECT EMP_ID,FIRST_NAME, LAST_NAME, ROLE,DEPT FROM emp_record_table
WHERE DEPT = 'HEALTHCARE';
```

EMP_ID	FIRST_NAME	LAST_NAME	ROLE	DEPT
E005	Eric	Hoffman	LEAD DATA SCIENTIST	FINANCE
E103	Emily	Grove	MANAGER	FINANCE
E403	Steve	Hoffman	ASSOCIATE DATA SCIENTIST	FINANCE
E052	Dianna	Wilson	SENIOR DATA SCIENTIST	HEALTHCARE
E057	Dorothy	Wilson	SENIOR DATA SCIENTIST	HEALTHCARE
E083	Patrick	Voltz	MANAGER	HEALTHCARE
E505	Chad	Wilson	ASSOCIATE DATA SCIENTIST	HEALTHCARE

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.

```
SELECT EMP_ID,FIRST_NAME,ROLE,DEPT,EMP_RATING,MAX(EMP_RATING)
OVER(PARTITION BY DEPT) AS MAX_EMP_RATING FROM emp_record_table;
```

EMP_ID	FIRST_NA...	ROLE	DEPT	EMP_RATING	MAX_EMP_RATING
E001	Arthur	PRESIDENT	ALL	5	5
E010	William	LEAD DATA SCIENTIST	AUTOMOTIVE	2	5
E204	Karene	SENIOR DATA SCIENTIST	AUTOMOTIVE	5	5
E428	Pete	MANAGER	AUTOMOTIVE	4	5
E532	Claire	ASSOCIATE DATA SCIENTIST	AUTOMOTIVE	1	5
E005	Eric	LEAD DATA SCIENTIST	FINANCE	3	4
E103	Emily	MANAGER	FINANCE	4	4
E403	Steve	ASSOCIATE DATA SCIENTIST	FINANCE	3	4
E052	Dianna	SENIOR DATA SCIENTIST	HEALTHCARE	5	5
E057	Dorothy	SENIOR DATA SCIENTIST	HEALTHCARE	1	5
E083	Patrick	MANAGER	HEALTHCARE	5	5
E505	Chad	ASSOCIATE DATA SCIENTIST	HEALTHCARE	2	5
E245	Nian	SENIOR DATA SCIENTIST	RETAIL	2	4
E260	Roy	SENIOR DATA SCIENTIST	RETAIL	3	4
E478	David	ASSOCIATE DATA SCIENTIST	RETAIL	4	4
E583	Janet	MANAGER	RETAIL	2	4
E612	Tracy	MANAGER	RETAIL	4	4
E620	Katrina	JUNIOR DATA SCIENTIST	RETAIL	1	4
E640	Jenifer	JUNIOR DATA SCIENTIST	RETAIL	4	4

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 `Min Salary` , MAX(SALARY) `Max Salary`
FROM emp_record_table
GROUP BY ROLE;
```

ROLE	Min Salary	Max Salary
PRESIDENT	16500	16500
LEAD DATA SCIENTIST	8500	9000
SENIOR DATA SCIENTIST	5500	7700
MANAGER	8500	11000
ASSOCIATE DATA SCIENTIST	4000	5000
JUNIOR DATA SCIENTIST	2800	3000

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

```
SELECT EMP_ID, EXP,
RANK() OVER (ORDER BY EXP DESC) AS ranks
FROM emp_record_table;
```

EMP_ID	EXP	ranks
E001	20	1
E083	15	2
E103	14	3
E428	14	3
E583	14	3
E612	13	6
E010	12	7
E005	11	8
E057	9	9
E204	8	10
E260	7	11
E052	6	12
E245	6	12
E505	5	14
E403	4	15
E478	3	16
E532	3	16
E620	2	18
E640	1	19

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.

```
CREATE OR REPLACE VIEW employees_in_countries AS
SELECT EMP_ID, COUNTRY, SALARY
FROM emp_record_table
WHERE SALARY > 6000;

SELECT * FROM employees_in_countries;
```

EMP_ID	COUNTRY	SALARY
E001	USA	16500
E005	USA	8500
E010	FRANCE	9000
E057	USA	7700
E083	USA	9500
E103	CANADA	10500
E204	GERMANY	7500
E245	CHINA	6500
E260	INDIA	7000
E428	GERMANY	11000
E583	COLOMBIA	10000
E612	INDIA	8500

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

```
SELECT EMP_ID, FIRST_NAME, LAST_NAME, EXP
FROM emp_record_table
WHERE EXP IN
(
 SELECT EXP FROM emp_record_table
 WHERE EXP > 10
);
```

EMP_ID	FIRST_NAME	LAST_NAME	EXP
E001	Arthur	Black	20
E005	Eric	Hoffman	11
E010	William	Butler	12
E083	Patrick	Voltz	15
E103	Emily	Grove	14
E428	Pete	Allen	14
E583	Janet	Hale	14
E612	Tracy	Norris	13

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.

```
DELIMITER %%
CREATE PROCEDURE NUM_OF_EXP_detail()
 BEGIN
  SELECT EMP_ID, FIRST_NAME, LAST_NAME, EXP
  FROM emp_record_table
  WHERE  EXP > 3;
 END %%
DELIMITER ;

CALL NUM_OF_EXP_detail();
```

EMP_ID	FIRST_NAME	LAST_NAME	EXP
E001	Arthur	Black	20
E005	Eric	Hoffman	11
E010	William	Butler	12
E052	Dianna	Wilson	6
E057	Dorothy	Wilson	9
E083	Patrick	Voltz	15
E103	Emily	Grove	14
E204	Karene	Nowak	8
E245	Nian	Zhen	6
E260	Roy	Collins	7
E403	Steve	Hoffman	4
E428	Pete	Allen	14
E505	Chad	Wilson	5
E583	Janet	Hale	14
E612	Tracy	Norris	13

---

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 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>12 AND 16 THEN
SET Employee_ROLE="MANAGER";
ELSEIF EXP>10 AND 12 THEN
SET Employee_ROLE ="LEAD DATA SCIENTIST";
ELSEIF EXP>5 AND 10 THEN
SET Employee_ROLE ="SENIOR DATA SCIENTIST";
ELSEIF EXP>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 &&
DELIMITER ;
```

EXP	Employee_ROLE(EXP)
11	LEAD DATA SCIENTIST
12	LEAD DATA SCIENTIST
6	SENIOR DATA SCIENTIST
9	SENIOR DATA SCIENTIST
8	SENIOR DATA SCIENTIST
6	SENIOR DATA SCIENTIST
7	SENIOR DATA SCIENTIST
4	ASSOCIATE DATA SCIENTIST
3	ASSOCIATE DATA SCIENTIST
5	ASSOCIATE DATA SCIENTIST
3	ASSOCIATE DATA SCIENTIST
2	JUNIOR DATA SCIENTIST
1	JUNIOR DATA SCIENTIST

```
SELECT EXP,Employee_ROLE(EXP)
FROM data_science_team;
```

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.

```
CREATE INDEX idx_first_name
ON emp_record_table(FIRST_NAME(20));
```

```
EXPLAIN SELECT * FROM emp_record_table
WHERE FIRST_NAME='Eric';
```

id	select_ty...	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
1	SIMPLE	emp_record_table	NULL	ref	idx_first_name	idx_first_name	83	const	1	100.00	Using where

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).

```
SELECT EMP_ID, FIRST_NAME, LAST_NAME, EMP_RATING, SALARY,
( SELECT SALARY * 0.05 * EMP_RATING) AS Bouns
FROM emp_record_table;
```

EMP_ID	FIRST_NA...	LAST_NAME	EMP_RATING	SALARY	Bouns
E001	Arthur	Black	5	16500	4125.00
E005	Eric	Hoffman	3	8500	1275.00
E010	William	Butler	2	9000	900.00
E052	Dianna	Wilson	5	5500	1375.00
E057	Dorothy	Wilson	1	7700	385.00
E083	Patrick	Voltz	5	9500	2375.00
E103	Emily	Grove	4	10500	2100.00
E204	Karene	Nowak	5	7500	1875.00
E245	Nian	Zhen	2	6500	650.00
E260	Roy	Collins	3	7000	1050.00
E403	Steve	Hoffman	3	5000	750.00
E428	Pete	Allen	4	11000	2200.00
E478	David	Smith	4	4000	800.00
E505	Chad	Wilson	2	5000	500.00
E532	Claire	Brennan	1	4300	215.00
E583	Janet	Hale	2	10000	1000.00
E612	Tracy	Norris	4	8500	1700.00
E620	Katrina	Allen	1	3000	150.00
E640	Jenifer	Jhones	4	2800	560.00

17. 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
FROM emp_record_table;
```

EMP_ID	FIRST_NA...	LAST_NAME	SALARY	COUNTRY	CONTINENT	AVG_salary_IN_COUNTRY	AVG_salary_IN_CONTINENT
E245	Nian	Zhen	6500	CHINA	ASIA	6500.0000	6250.0000
E260	Roy	Collins	7000	INDIA	ASIA	6166.6667	6250.0000
E612	Tracy	Norris	8500	INDIA	ASIA	6166.6667	6250.0000
E620	Katrina	Allen	3000	INDIA	ASIA	6166.6667	6250.0000
E010	William	Butler	9000	FRANCE	EUROPE	9000.0000	7950.0000
E204	Karene	Nowak	7500	GERMANY	EUROPE	7600.0000	7950.0000
E428	Pete	Allen	11000	GERMANY	EUROPE	7600.0000	7950.0000
E532	Claire	Brennan	4300	GERMANY	EUROPE	7600.0000	7950.0000
E052	Dianna	Wilson	5500	CANADA	NORTH AMERICA	7000.0000	8525.0000
E103	Emily	Grove	10500	CANADA	NORTH AMERICA	7000.0000	8525.0000
E505	Chad	Wilson	5000	CANADA	NORTH AMERICA	7000.0000	8525.0000
E001	Arthur	Black	16500	USA	NORTH AMERICA	9440.0000	8525.0000
E005	Eric	Hoffman	8500	USA	NORTH AMERICA	9440.0000	8525.0000
E057	Dorothy	Wilson	7700	USA	NORTH AMERICA	9440.0000	8525.0000
E083	Patrick	Voltz	9500	USA	NORTH AMERICA	9440.0000	8525.0000
E403	Steve	Hoffman	5000	USA	NORTH AMERICA	9440.0000	8525.0000
E478	David	Smith	4000	COLOMBIA	SOUTH AMERICA	5600.0000	5600.0000
E583	Janet	Hale	10000	COLOMBIA	SOUTH AMERICA	5600.0000	5600.0000
E640	Jenifer	Jhones	2800	COLOMBIA	SOUTH AMERICA	5600.0000	5600.0000