# $SQL\_EmployeeRecords$

### Jas Ahuja

### 2024-05-14

### Contents

| Course-end Project 1 - ScienceQtech Employee Performance Mapping   | 2           |
|--|-------------|
| Description  | 2           |
| Objective:   | 2           |
| Dataset description:   | 2           |
| emp_record_table: It contains the information of all the employees   | 2           |
| Proj_table: It contains information about the projects   | 2<br>3<br>3 |
| Data_science_team: It contains information about all the employees in the Data Science team.   | 3           |
| The tasks to be performed:   | 3           |
| 1. Create a database named employee, then import tables into the employee database from  | 0           |
| the given resources.   | 3<br>7      |
| 2. Create an ER diagram for the given employee database  | 1           |
| 3. Write a query to fetch EMP_ID, FIRST_NAME, LAST_NAME, GENDER, and DE-PARTMENT from the employee record table, and make a list of employees and details                    |             |
| of their department  | 7           |
| 4. Write a query to fetch EMP_ID, FIRST_NAME, LAST_NAME, GENDER, DEPART-   |             |
| MENT, and EMP_RATING if the EMP_RATING is:   | 8           |
| 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  | 0           |
| alias as NAME  | 9           |
| the number of reporters (including the President)  | 9           |
| 7. Write a query to list down all the employees from the healthcare and finance departments  | Э           |
| using union. Take data from the employee record table  | 11          |
| 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.   | 12          |
| 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   | 12          |
| 10. Write a query to assign ranks to each employee based on their experience. Take data from   |             |
| the employee record table  | 12          |
| 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   | 13          |
| 12. Write a nested query to find employees with experience of more than ten years. Take data   |             |
| from the employee record table   | 14          |
| 13. Write a query to create a <i>stored procedure</i> to retrieve the details of the employees whose   | 1.4         |
| experience is more than three years. Take data from the employee record table 14. Write a query using stored functions in the project table to check whether the job profile | 14          |
| assigned to each employee in the data science team matches the organization's set  |             |
| standard   | 15          |
| DUGITARIA I I I I I I I I I I I I I I I I I I  | 10          |

```
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.
18
```

css chunk to make tables fit

Hidden code chunk for mypassword

Make a MySQL connection from Rmarkdown

# Course-end Project 1 - ScienceQtech Employee Performance Mapping...

#### 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 tasks to be performed:

- 1. Create a database named employee, then import tables into the employee database from the given resources.
- data\_science\_team.csv
- proj\_table.csv and
- emp\_record\_table.csv

#### **Answer-1:** Done: Using the import function of MySQLWorkbench.

set nocount on

- ! This works with SQL database to run multiple queries from a code chunk. This does not work with MySQL and therefore I need to use one query per code chunk
- 1. Initiate use of our database.

#### use employee

2. Look at tables available in employee:

#### SHOW TABLES;

Table 1: 4 records

Tables\_in\_employee
data\_science\_team
emp\_record\_table
proj\_table
vw\_country

- 3. Fix the proj\_table
- Alter table as text feild  $project\_id$  cannot be PRIMARY KEY

ALTER TABLE proj\_table

MODIFY COLUMN project\_id VARCHAR(10)

ALTER TABLE proj\_table
ADD PRIMARY KEY(project\_id)

• Our altered table

SELECT \* FROM proj\_table

Table 2: 6 records

|                              |                                   |          | START         |               |         |                    |
|------------------------------|-----------------------------------|----------|---------------|---------------|---------|--------------------|
| $\operatorname{project\_id}$ | PROJ_NAME                         | DOMAIN   | _DATE         | $CLOSURE_{-}$ | _DADEV_ | _QT <b>B</b> TATUS |
| P103                         | Drug Discovery                    | HEALTHCA | A ROE-06-2021 | 6/20/2021     | Q1      | DONE               |
| P105                         | Fraud Detection                   | FINANCE  | 04-11-2021    | 6/25/2021     | Q1      | DONE               |
| P109                         | Market Basket Analysis            | RETAIL   | 04-12-2021    | 6/30/2021     | Q1      | DELAYED            |
| P204                         | Supply Chain Management           | AUTOMOT  | IWE/15/2021   | 9/28/2021     | Q2      | WIP                |
| P302                         | Early Detection of Lung<br>Cancer | HEALTHCA | ARE-08-2021   | 12/18/2021    | Q3      | YTS                |
| P406                         | Customer Sentiment Analysis       | RETAIL   | 07-09-2021    | 9/24/2021     | Q2      | WIP                |

- 4. Fix emp\_record\_table
- Lets look at the table

SELECT \* FROM emp\_record\_table

Table 3: Displaying records 1 - 10

| emp_ | _ifilrstna | unhAST_ | <u> SAM</u> | <b>DRXQ</b> LE              | DEPT  | EXF          | COUI           | N <b>CBN</b> TIN      | ISMTA] | RWP_ | RMAIN | QQEBj_ | _ <b>I</b> designation |
|------|------------|---------|-------------|-----------------------------|-------|--------------|----------------|-----------------------|--------|------|-------|--------|------------------------|
| E001 | Arthur     | Black   | M           | PRESIDENT                   | ALL   | 20           | USA            | NORTH<br>AMER-<br>ICA | 16500  | 5    | NA    | NA     | NA                     |
| E005 | Eric       | Hoffma  | ınM         | LEAD<br>DATA SCI-<br>ENTIST | FINAN | Œ            | USA            | NORTH<br>AMER-<br>ICA | 8500   | 3    | E103  | P105   | NA                     |
| E010 | William    | Butler  | M           | LEAD<br>DATA SCI-<br>ENTIST | AUTON | <b>/10</b> T | I <b>∀R</b> AN | N <b>ŒE</b> ROPE      | Е9000  | 2    | E428  | P204   | NA                     |

| $\overline{\mathrm{emp}}_{-}$ | _ifilrstna | and AST_ | SAM | <b>DIRO</b> LE | DEPT   | EXP           | COUN           | NCBNTIN               | USMTARWII | <u> </u> | RMAINA | GERj_ | <b>Ild</b> esignation |
|-------------------------------|------------|----------|-----|----------------|--------|---------------|----------------|-----------------------|-----------|----------|--------|-------|-----------------------|
| E052                          | Dianna     | Wilson   | F   | SENIOR         | HEALT  | HCA           | <b>REA</b> NA  | ADORTH                | 5500      | 5        | E083   | P103  | NA                    |
|                               |            |          |     | DATA SCI-      |        |               |                | AMER-                 |           |          |        |       |                       |
|                               |            |          |     | ENTIST         |        |               |                | ICA                   |           |          |        |       |                       |
| E057                          | Dorothy    | Wilson   | F   | SENIOR         | HEALT  | 'H <b>©</b> A | RISA           | NORTH                 | 7700      | 1        | E083   | P302  | NA                    |
|                               |            |          |     | DATA SCI-      |        |               |                | AMER-                 |           |          |        |       |                       |
|                               |            |          |     | ENTIST         |        |               | A              | ICA                   |           | _        | 7004   |       | 27.4                  |
| E083                          | Patrick    | Voltz    | Μ   | MANAGER        | HEALT  | H6'A          | KISA           | NORTH                 | 9500      | 5        | E001   | NA    | NA                    |
|                               |            |          |     |                |        |               |                | AMER-                 |           |          |        |       |                       |
| E102                          | Emile.     | Charra   | E   | MANAGER        | EIN AN | CITA          | CANIA          | ICA<br>N <b>D</b> RTH | 10500     | 4        | E001   | NA    | NA                    |
| E109                          | Emily      | Grove    | F   | MANAGER        | FINAIN | CIE           | CANE           | AMER-                 | 10000     | 4        | E001   | INA   | NA                    |
|                               |            |          |     |                |        |               |                | ICA                   |           |          |        |       |                       |
| E204                          | Karene     | Nowak    | F   | SENIOR         | AUTON  | лоті          | <b>KE</b> RN   | AEANRYOPE             | ₹7500     | 5        | E428   | P204  | NA                    |
| <b>L2</b> 01                  | riarono    | TYOWAIL  | -   | DATA SCI-      | 110101 | .1011         | <b>Q12</b> 101 | 111011011             | 21000     |          | 2120   | 1 201 | 1111                  |
|                               |            |          |     | ENTIST         |        |               |                |                       |           |          |        |       |                       |
| E245                          | Nian       | Zhen     | M   | SENIOR         | RETAI  | L 6           | CHIN           | ASIA                  | 6500      | 2        | E583   | P109  | NA                    |
|                               |            |          |     | DATA SCI-      |        |               |                |                       |           |          |        |       |                       |
|                               |            |          |     | ENTIST         |        |               |                |                       |           |          |        |       |                       |
| E260                          | Roy        | Collins  | M   | SENIOR         | RETAI  | L 7           | INDIA          | AASIA                 | 7000      | 3        | E583   | NA    | NA                    |
|                               |            |          |     | DATA SCI-      |        |               |                |                       |           |          |        |       |                       |
|                               |            |          |     | ENTIST         |        |               |                |                       |           |          |        |       |                       |

- Alter the table and put  $emp\_id$  as PRIMARY\_KEY

```
ALTER TABLE emp_record_table MODIFY COLUMN emp_id VARCHAR(10)
```

ALTER TABLE emp\_record\_table ADD PRIMARY KEY(emp\_id)

- $proj\_id$  needs to be set as **FOREIGN\_KEY**
- $\bullet\,$  it cannot be NA or Null! |
- feild type also needs to match in the two tables

```
SELECT * FROM emp_record_table
WHERE proj_id = 'NA' OR proj_id IS NULL
```

Table 4: 7 records

| $\overline{\mathrm{emp}}$ | _ifilrstna | and AST_ | _XAMI | <b>DRXQ</b> LE | DEPT  | EXI  | PCOUN | NTRONTIN                | ISMTARY | ИР_ | RMAINA | C[ERj_ | _ <b>I</b> designation |
|---------------------------|------------|----------|-------|----------------|-------|------|-------|-------------------------|---------|-----|--------|--------|------------------------|
| E001                      | Arthur     | Black    | M     | PRESIDENT      | ALL   | 20   | USA   | NORTH<br>AMER-          | 16500   | 5   | NA     | NA     | NA                     |
| E083                      | Patrick    | Voltz    | M     | MANAGER        | HEALT | ТНСА | RISA  | ICA<br>NORTH<br>AMER-   | 9500    | 5   | E001   | NA     | NA                     |
| E103                      | Emily      | Grove    | F     | MANAGER        | FINAN | CIE  | CANA  | ICA<br>ADWORTH<br>AMER- | 10500   | 4   | E001   | NA     | NA                     |
|                           |            |          |       |                |       |      |       | ICA                     |         |     |        |        |                        |

| emp_ifilrst_n | amAST_NAN | NIBIROBLE                        | DEPT EX  | PCOUN <b>TRA</b> NTI          | N <b>ISM</b> TA1 | RWP_ | RMAINA | GGERj_ | _ <b>I</b> designation |
|---------------|-----------|----------------------------------|----------|-------------------------------|------------------|------|--------|--------|------------------------|
| E260 Roy      | Collins M | SENIOR<br>DATA<br>SCIEN-<br>TIST | RETAIL 7 | INDIA ASIA                    | 7000             | 3    | E583   | NA     | NA                     |
| E428 Pete     | Allen M   | MANAGER                          | AUTOM@7  | TI <b>GE</b> RM <b>ÆNK</b> OF | PE11000          | 4    | E001   | NA     | NA                     |
| E583 Janet    | Hale F    | MANAGER                          | RETAIL14 | COLOM <b>BL</b> ATH<br>AMER-  | 0000             | 2    | E001   | NA     | NA                     |
|               |           |                                  |          | ICA                           |                  |      |        |        |                        |
| E612 Tracy    | Norris F  | MANAGER                          | RETAIL13 | INDIA ASIA                    | 8500             | 4    | E001   | NA     | NA                     |

- Fix the NA values in  $proj\_id$ 

In case of Error: 1175- go to menu "MySQLWorkbench" > "Settings" > "SQL Editor" > uncheck "Safe Updates" \_ on PC it is under "edit"

```
UPDATE emp_record_table
    SET proj_id=NULL
    WHERE proj_id='NA'
```

 $\bullet\,$  Now change type to match the proj\_table.

```
ALTER TABLE emp_record_table

MODIFY COLUMN proj_id VARCHAR(10)

ALTER TABLE emp_record_table

ADD CONSTRAINT fk_proj

FOREIGN KEY(proj_id) REFERENCES proj_table(PROJECT_ID)
```

• Let's see the altered table:

DESCRIBE emp\_record\_table

Table 5: Displaying records 1 - 10

| Field      | Type        | Null | Key | Default | Extra |
|------------|-------------|------|-----|---------|-------|
| emp_id     | varchar(10) | NO   | PRI | NA      |       |
| first_name | varchar(20) | YES  | MUL | NA      |       |
| LAST_NAME  | text        | YES  |     | NA      |       |
| GENDER     | text        | YES  |     | NA      |       |
| ROLE       | text        | YES  |     | NA      |       |
| DEPT       | text        | YES  |     | NA      |       |
| EXP        | int         | YES  |     | NA      |       |
| COUNTRY    | text        | YES  |     | NA      |       |
| CONTINENT  | text        | YES  |     | NA      |       |
| SALARY     | int         | YES  |     | NA      |       |

5. Similarly fix the data\_science\_team.

SELECT \* FROM data\_science\_team

Table 6: Displaying records 1 - 10

| emp_id | FIRST_N | AMAST_N | AMENDI | EROLE                       | DEPT    | EXP         | COUNTR | CONTINENT          |
|--------|---------|---------|--------|-----------------------------|---------|-------------|--------|--------------------|
| E005   | Eric    | Hoffman | M      | LEAD DATA<br>SCIENTIST      | FINANCE | 11          | USA    | NORTH<br>AMERICA   |
| E010   | William | Butler  | M      | LEAD DATA<br>SCIENTIST      | AUTOMO  | LIME        | FRANCE | EUROPE             |
| E052   | Dianna  | Wilson  | F      | SENIOR DATA<br>SCIENTIST    | HEALTHC | AR <b>E</b> | CANADA | NORTH<br>AMERICA   |
| E057   | Dorothy | Wilson  | F      | SENIOR DATA<br>SCIENTIST    | HEALTHC | ARÐ         | USA    | NORTH<br>AMERICA   |
| E204   | Karene  | Nowak   | F      | SENIOR DATA<br>SCIENTIST    | AUTOMO  | ΓIVE        | GERMAN | <b>E</b> UROPE     |
| E245   | Nian    | Zhen    | Μ      | SENIOR DATA<br>SCIENTIST    | RETAIL  | 6           | CHINA  | ASIA               |
| E260   | Roy     | Collins | M      | SENIOR DATA<br>SCIENTIST    | RETAIL  | 7           | INDIA  | ASIA               |
| E403   | Steve   | Hoffman | M      | ASSOCIATE DATA<br>SCIENTIST | FINANCE | 4           | USA    | NORTH<br>AMERICA   |
| E478   | David   | Smith   | M      | ASSOCIATE DATA<br>SCIENTIST | RETAIL  | 3           | COLOME | BISOUTH<br>AMERICA |
| E505   | Chad    | Wilson  | M      | ASSOCIATE DATA<br>SCIENTIST | HEALTHC | AR <b>5</b> | CANADA |                    |

• match emp\_id to emp\_record\_table

```
ALTER TABLE data_science_team
MODIFY COLUMN emp_id VARCHAR(10)
```

```
ALTER TABLE data_science_team

ADD CONSTRAINT fk_emp_record_table_emp_id

FOREIGN KEY(emp_id) REFERENCES emp_record_table(emp_id)
```

2. Create an ER diagram for the given employee database.

Answer-2: goto "database" > "Reverse Engineer" and follow prompts to get Figure-1

3. Write a query to fetch EMP\_ID, FIRST\_NAME, LAST\_NAME, GENDER, and DE-PARTMENT 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
```

Table 7: Displaying records 1 - 10

| $\overline{\mathrm{emp\_id}}$ | first_name | last_name | gender   | dept       |
|-------------------------------|------------|-----------|----------|------------|
| E001                          | Arthur     | Black     | M        | ALL        |
| E005                          | Eric       | Hoffman   | M        | FINANCE    |
| E010                          | William    | Butler    | ${ m M}$ | AUTOMOTIVE |
| E052                          | Dianna     | Wilson    | F        | HEALTHCARE |
| E057                          | Dorothy    | Wilson    | F        | HEALTHCARE |
| E083                          | Patrick    | Voltz     | M        | HEALTHCARE |

| $\overline{\mathrm{emp\_id}}$ | first_name | last_name | gender       | dept       |
|-------------------------------|------------|-----------|--------------|------------|
| E103                          | Emily      | Grove     | F            | FINANCE    |
| E204                          | Karene     | Nowak     | F            | AUTOMOTIVE |
| E245                          | Nian       | Zhen      | $\mathbf{M}$ | RETAIL     |
| E260                          | Roy        | Collins   | M            | RETAIL     |

### 4. Write a query to fetch EMP\_ID, FIRST\_NAME, LAST\_NAME, GENDER, DEPART-MENT, and EMP\_RATING if the EMP\_RATING is:

#### -4.1 *emp\_\_rating* less than two

```
SELECT emp_id, first_name, last_name, gender, dept
FROM emp_record_table
WHERE emp_rating < 2</pre>
```

Table 8: 3 records

| emp_id | first_name | last_name | gender | dept       |
|--------|------------|-----------|--------|------------|
| E057   | Dorothy    | Wilson    | F      | HEALTHCARE |
| E532   | Claire     | Brennan   | F      | AUTOMOTIVE |
| E620   | Katrina    | Allen     | F      | RETAIL     |

#### -4.2 emp\_rating greater than four

```
SELECT emp_id, first_name, last_name, gender, dept
   FROM emp_record_table
   WHERE emp_rating > 4
```

Table 9: 4 records

| emp_id | first_name | last_name | gender | dept       |
|--------|------------|-----------|--------|------------|
| E001   | Arthur     | Black     | M      | ALL        |
| E052   | Dianna     | Wilson    | F      | HEALTHCARE |
| E083   | Patrick    | Voltz     | M      | HEALTHCARE |
| E204   | Karene     | Nowak     | F      | AUTOMOTIVE |

#### -4.3 *emp\_rating* Between two and four

```
SELECT emp_id, first_name, last_name, gender, dept
FROM emp_record_table
WHERE emp_rating BETWEEN 2 AND 4
```

Table 10: Displaying records 1 - 10

| $\overline{\mathrm{emp\_id}}$ | first_name | last_name | gender    | dept       |
|-------------------------------|------------|-----------|-----------|------------|
| E005                          | Eric       | Hoffman   | M         | FINANCE    |
| E010                          | William    | Butler    | ${\bf M}$ | AUTOMOTIVE |
| E103                          | Emily      | Grove     | F         | FINANCE    |
| E245                          | Nian       | Zhen      | M         | RETAIL     |
| E260                          | Roy        | Collins   | M         | RETAIL     |

| emp_id | $first\_name$ | $last\_name$           | gender   | dept       |
|--------|---------------|------------------------|----------|------------|
| E403   | Steve         | Hoffman                | M        | FINANCE    |
| E428   | Pete          | Allen                  | M        | AUTOMOTIVE |
| E478   | David         | $\operatorname{Smith}$ | ${ m M}$ | RETAIL     |
| E505   | Chad          | Wilson                 | ${ m M}$ | HEALTHCARE |
| E583   | Janet         | Hale                   | F        | RETAIL     |
|        |               |                        |          |            |

#### -4.4 emp\_rating bins

```
SELECT emp_id, first_name, last_name, gender, dept
FROM emp_record_table
WHERE emp_rating < 2 OR emp_rating BETWEEN 2 AND 4 OR emp_rating > 4
```

Table 11: Displaying records 1 - 10

| 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    | $\mathbf{F}$ | HEALTHCARE |
| E057   | Dorothy    | Wilson    | $\mathbf{F}$ | HEALTHCARE |
| E083   | Patrick    | Voltz     | $\mathbf{M}$ | HEALTHCARE |
| E103   | Emily      | Grove     | $\mathbf{F}$ | FINANCE    |
| E204   | Karene     | Nowak     | F            | AUTOMOTIVE |
| E245   | Nian       | Zhen      | M            | RETAIL     |
| E260   | Roy        | Collins   | M            | RETAIL     |

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.

```
SELECT CONCAT(first_name, ",", last_name) AS NAME FROM emp_record_table
WHERE dept='FINANCE'
```

Table 12: 3 records

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).

```
SELECT * FROM emp_record_table
WHERE emp_id IN (SELECT DISTINCT manager_id from emp_record_table)
```

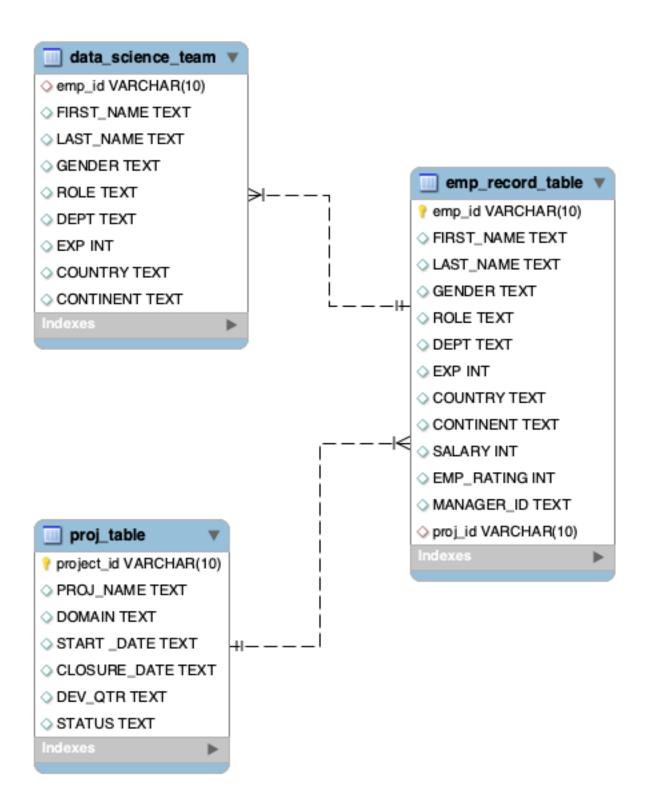


Figure 1: Figure-1: The Entity Relation diagram generated from MySQLWorkbench

Table 13: 6 records

| emp_ | i <b>d</b> irst_na | nLeAST_ | _N(#FM)      | D <b>RQ</b> LE | DEPT       | EXI   | PCOUN | NTCPOYNTINI      | ESVAIL A BE | MP_F | RAIANØ | €GβΩRj_ | <u>I</u> Iddesignati |
|------|--------------------|---------|--------------|----------------|------------|-------|-------|------------------|-------------|------|--------|---------|----------------------|
| E103 | Emily              | Grove   | F            | MANA           | GEIRIAN    | ICE4  | CANA  | ADMORTH          | 10500       | 4    | E001   | NA      | NA                   |
|      |                    |         |              |                |            |       |       | AMER-            |             |      |        |         |                      |
|      |                    |         |              |                |            |       |       | ICA              |             |      |        |         |                      |
| E428 | Pete               | Allen   | $\mathbf{M}$ | MANA           | CAEURTO:   | MOTI  | VERN  | <b>AFTUR</b> OPE | 11000       | 4    | E001   | NA      | NA                   |
| E083 | Patrick            | Voltz   | $\mathbf{M}$ | MANA           | CHETERAL Z | ГН(БА | RUESA | NORTH            | 9500        | 5    | E001   | NA      | NA                   |
|      |                    |         |              |                |            |       |       | AMER-            |             |      |        |         |                      |
|      |                    |         |              |                |            |       |       | ICA              |             |      |        |         |                      |
| E001 | Arthur             | Black   | $\mathbf{M}$ | PRESI          | DANIT      | 20    | USA   | NORTH            | 16500       | 5    | NA     | NA      | NA                   |
|      |                    |         |              |                |            |       |       | AMER-            |             |      |        |         |                      |
|      |                    |         |              |                |            |       |       | ICA              |             |      |        |         |                      |
| E583 | Janet              | Hale    | $\mathbf{F}$ | MANA           | CHETETA 1  | L14   | COLO  | MSKNATH          | 10000       | 2    | E001   | NA      | NA                   |
|      |                    |         |              |                |            |       |       | AMER-            |             |      |        |         |                      |
|      |                    |         |              |                |            |       |       | ICA              |             |      |        |         |                      |
| E612 | Tracy              | Norris  | $\mathbf{F}$ | MANA           | CHETETA 1  | L13   | INDIA | ASIA             | 8500        | 4    | E001   | NA      | NA                   |

# 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 * FROM emp_record_table
    WHERE dept = 'HEALTHCARE'
UNION
SELECT * FROM emp_record_table
    WHERE dept = 'FINANCE'
```

Table 14: 7 records

| emp_ifdrst_nathAST_INEN | <b>DERYC</b> LE                  | DEPT EX   | PCOU           | NCONTIN                | I <b>SMT</b> A <b>RM</b> P | RMAIN  | WAGEERIJ_ | _IDesignation |
|-------------------------|----------------------------------|-----------|----------------|------------------------|----------------------------|--------|-----------|---------------|
| E052 Dianna Wilson F    | SENIOR<br>DATA SCI-              | HEALTH(C  | ARAN.          | ADORTH<br>AMER-        | 5500 5                     | E083   | P103      | NA            |
| E057 DorothyWilson F    | ENTIST<br>SENIOR                 | HEALTH9C  | ARBA           | ICA<br>NORTH           | 7700                       | E083   | P302      | NA            |
|                         | DATA SCI-<br>ENTIST              |           |                | AMER-ICA               |                            |        | 27.1      |               |
| E083 Patrick Voltz M    | MANAGER                          | HEALTHSC. | ARBA           | NORTH<br>AMER-<br>ICA  | 9500 5                     | 6 E001 | NA        | NA            |
| E505 Chad Wilson M      | ASSOCIATE<br>DATA SCI-<br>ENTIST | HEALTHIC  | A <b>RA</b> N. |                        | 5000                       | E083   | P103      | NA            |
| E005 Eric HoffmanM      | LEAD<br>DATA SCI-<br>ENTIST      | FINANCE   | USA            |                        | 8500                       | E103   | P105      | NA            |
| E103 Emily Grove F      | MANAGER                          | FINANCE   | CAN            | ANORTH<br>AMER-<br>ICA | 10500                      | E001   | NA        | NA            |
| E403 Steve HoffmanM     | ASSOCIATE<br>DATA SCI-<br>ENTIST | FINANC    | USA            | NORTH<br>AMER-<br>ICA  | 5000 3                     | E103   | P105      | NA            |

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, last_name, role, dept, MAX(emp_rating) AS max_emp_rating
   FROM emp_record_table
-- this GROUP BY is used for max_emp_rating
   GROUP BY emp_id, first_name, last_name, role, dept
```

Table 15: Displaying records 1 - 10

| $\overline{\mathrm{emp\_id}}$ | $first\_name$ | $last\_name$ | role                | dept       | $\max_{\text{emp\_rating}}$ |
|-------------------------------|---------------|--------------|---------------------|------------|-----------------------------|
| E001                          | Arthur        | Black        | PRESIDENT           | ALL        | 5                           |
| E005                          | Eric          | Hoffman      | LEAD DATA SCIENTIST | FINANCE    | 3                           |
| E010                          | William       | Butler       | LEAD DATA SCIENTIST | AUTOMOTIVE | $^{2}$                      |
| E052                          | Dianna        | Wilson       | SENIOR DATA         | HEALTHCARE | 5                           |
|                               |               |              | SCIENTIST           |            |                             |
| E057                          | Dorothy       | Wilson       | SENIOR DATA         | HEALTHCARE | 1                           |
|                               |               |              | SCIENTIST           |            |                             |
| E083                          | Patrick       | Voltz        | MANAGER             | HEALTHCARE | 5                           |
| E103                          | Emily         | Grove        | MANAGER             | FINANCE    | 4                           |
| E204                          | Karene        | Nowak        | SENIOR DATA         | AUTOMOTIVE | 5                           |
|                               |               |              | SCIENTIST           |            |                             |
| E245                          | Nian          | Zhen         | SENIOR DATA         | RETAIL     | 2                           |
|                               |               |              | SCIENTIST           |            |                             |
| E260                          | Roy           | Collins      | SENIOR DATA         | RETAIL     | 3                           |
|                               |               |              | SCIENTIST           |            |                             |

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), MAX(salary)
FROM emp_record_table
GROUP BY ROLE
```

Table 16: 6 records

| 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 *, RANK() OVER (ORDER by exp DESC)
FROM emp_record_table
```

Table 17: Displaying records 1 - 10

RANK() OVER (ORDER by emp\_fidest\_naLnAcST\_CNEANDARCH\_E DEPT EXPCOUNTENTINSANTARMP\_RAMINAGER\_idesignation DESC) PRESIDENATLL 20 USA NORTH16500 NANA NA 1 E001 Arthur Black M AMER-**ICA** E083 PatrickVoltz M MANAGERHEALTIBCARR NORTH9500 E001 NA NA2 AMER-**ICA** E103 Emily Grove F MANAGERFINANCÆ CANANCÆRTH10500 E001 NA NA 3 AMER-ICA E428 Pete Allen M MANAGERAUTOMOTCFERMEANXOPE1000 E001 NA NA 3 E583 Janet Hale F MANAGERRETAIL4 COLOSOBIATH 10000 E001 NA NA 3 AMER-ICA MANAGERRETAIL3 INDIAASIA 8500 E612 Tracy Norris F 6 E001 NA NA E010 WilliamButler M LEAD AUTOMOTHRANCEROPE000 E428 P204 NA DATA SCIEN-TIST E005 Eric HoffmaM FINANCE USA NORTH8500 E103 P105 NA LEAD 8 DATA AMER-SCIEN-ICA TIST E057 Doroth Wilson F HEALT9HCAURE NORTH7700 SENIOR E083 P302 NA 9 DATA AMER-SCIEN-**ICA** TIST E204 KareneNowakF AUTOMOTGERMANNOPE500 E428 P204 NA 10 SENIOR DATA SCIEN-TIST

NB: Lets understand this one-

### 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.

Can't have two views by the same name!

DROP VIEW IF EXISTS vw\_country

CREATE VIEW vw\_country

AS

<sup>-</sup> \_SELECT \*: This selects all columns from the emp\_record\_table.\_

<sup>-</sup> RANK() OVER (ORDER by exp DESC): This is a window function that calculates the rank of each record based on the exp column in descending order.

<sup>-</sup> The ORDER BY exp DESC specifies that the records should be ranked based on the exp column in descending order

```
SELECT first_name, last_name, country FROM emp_record_table
WHERE salary > 6000
```

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

```
SELECT * FROM
(
   SELECT * FROM emp_record_table
        WHERE exp > 10
) AS T
```

Table 18: 8 records

| emp_iffirst_naihAST_KAN | IDIROLE DEPT                         | EXPCOUN'                   | TRANTIN                | ISMTARWP. | RMAING | GÆRj_ | <b>IB</b> esignation |
|-------------------------|--------------------------------------|----------------------------|------------------------|-----------|--------|-------|----------------------|
| E001 Arthur Black M     | PRESIDENTALL                         |                            | NORTH<br>AMER-<br>ICA  | 16500 5   | NA     | NA    | NA                   |
| E005 Eric HoffmanM      | LEAD FINAN<br>DATA<br>SCIEN-<br>TIST | NOTE USA                   |                        | 8500 3    | E103   | P105  | NA                   |
| E010 William Butler M   | LEAD AUTO DATA SCIEN- TIST           | M <b>Q</b> TI <b>VR</b> AN | ŒUROPE                 | 29000 2   | E428   | P204  | NA                   |
| E083 Patrick Voltz M    | MANAGER HEAL                         |                            | NORTH<br>AMER-<br>ICA  | 9500 5    | E001   | NA    | NA                   |
| E103 Emily Grove F      | MANAGER FINAN                        |                            | DWORTH<br>AMER-<br>ICA | 10500 4   | E001   | NA    | NA                   |
| E428 Pete Allen M       | MANAGER AUTO                         | MQTI <b>GE</b> RM          | <b>ANK</b> OPE         | 211000 4  | E001   | NA    | NA                   |
| E583 Janet Hale F       | MANAGER RETA                         |                            | MBLATH<br>AMER-<br>ICA | 10000 2   | E001   | NA    | NA                   |
| E612 Tracy Norris F     | MANAGER RETA                         | IL13 INDIA                 | ASIA                   | 8500 4    | E001   | NA    | NA                   |

- 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.
  - Need to drop procedure if exists

```
DROP PROCEDURE IF EXISTS sp_exp;

CREATE PROCEDURE sp_exp()
BEGIN
         SELECT * FROM emp_record_table WHERE exp > 3;
END;
```

In MySQLWorkbench we need a delimiter change. The code would therefore be-

```
DELIMITER //
CREATE PROCEDURE sp_exp()
BEGIN
SELECT * FROM emp_record_table WHERE exp > 3;
END//
DELIMITER ;
```

Here, // is used as the delimiter. It's changed back to; after the stored procedure definition.

This ensures that the SQL interpreter doesn't interpret the semicolons within the stored procedure as statement terminators until the entire procedure is defined.

### 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'.

```
SELECT * ,

CASE WHEN exp <= 2 THEN 'JUNIOR DATA SCIENTIST'

WHEN exp > 2 AND EXP <= 5 THEN 'ASSOCIATE DATA SCIENTIST'

WHEN exp > 5 AND EXP <= 10 THEN 'SENIOR DATA SCIENTIST'

WHEN exp > 10 AND EXP <= 12 THEN 'LEAD DATA SCIENTIST'

WHEN exp > 12 AND EXP <= 16 THEN 'MANAGER'

END AS designation

FROM emp_record_table
```

Table 19: Displaying records 1 - 10

| emp_first_nalmAeST_CAEAN  |                                  | DEPT EX          | PCOU     | NCIONYTINEANIARMI              | P_ <b>B</b> |      | GER_ | _i <b>de</b> signa | at <b>den</b> ignation           |
|---------------------------|----------------------------------|------------------|----------|--------------------------------|-------------|------|------|--------------------|----------------------------------|
| E001 Arthur Black M       | PRESIDEN                         | VALL 20          | USA      | NORTH 16500 S<br>AMER-<br>ICA  | 5 I         | NA   | NA   | NA                 | NA                               |
| E005 Eric Hoffma <b>M</b> | LEAD<br>DATA<br>SCIEN-<br>TIST   | FINANCE          | USA      | NORTH 8500 S<br>AMER-<br>ICA   | 3 1         | E103 | P105 | NA                 | LEAD<br>DATA<br>SCIEN-<br>TIST   |
| E010 WilliamButler M      | LEAD<br>DATA<br>SCIEN-<br>TIST   | AUTOM            | TFRA     | NEUEROP19000 2                 | 2 ]         | E428 | P204 | NA                 | LEAD<br>DATA<br>SCIEN-<br>TIST   |
| E052 Dianna Wilson F      | SENIOR<br>DATA<br>SCIEN-<br>TIST | HEALT <b>H</b> ( | CACBAEN. | ANDARTH 5500 E<br>AMER-<br>ICA | 5 1         | E083 | P103 | NA                 | SENIOR<br>DATA<br>SCIEN-<br>TIST |
| E057 DorothyWilson F      | SENIOR<br>DATA<br>SCIEN-<br>TIST | HEALTH           | CAUSIA   | NORTH 7700<br>AMER-<br>ICA     | l ]         | E083 | P302 | NA                 | SENIOR<br>DATA<br>SCIEN-<br>TIST |

| emp_first_  | naImAeST | CNEAN |                                  | DEPT EX            | PCOU.            | NIORYTI               | VSEAN TAIRIN | IP_ | BAATNIN | <b>G</b> ER | <b>ide</b> signa | at <b>den</b> ignation           |
|-------------|----------|-------|----------------------------------|--------------------|------------------|-----------------------|--------------|-----|---------|-------------|------------------|----------------------------------|
| E083 Patrio | k Voltz  | M     | MANAGE                           | ERHEALT <b>I</b> C | AUSIA            | NORTH<br>AMER-<br>ICA | 9500         | 5   | E001    | NA          | NA               | MANAGER                          |
| E103 Emily  | Grove    | F     | MANAGE                           | ERFINAN <b>CÆ</b>  | CAN.             |                       | 10500        | 4   | E001    | NA          | NA               | MANAGER                          |
| E204 Karen  | ie Nowak | F     | SENIOR<br>DATA<br>SCIEN-<br>TIST | AUTOM807           | T <b>(XEE</b> R) |                       | <b>₹</b> 500 | 5   | E428    | P204        | NA               | SENIOR<br>DATA<br>SCIEN-<br>TIST |
| E245 Nian   | Zhen     | M     | SENIOR<br>DATA<br>SCIEN-<br>TIST | RETAII6            | CHIN             | J <b>A</b> SIA        | 6500         | 2   | E583    | P109        | NA               | SENIOR<br>DATA<br>SCIEN-<br>TIST |
| E260 Roy    | Collins  | sM    | SENIOR<br>DATA<br>SCIEN-<br>TIST | RETAII7            | INDI             | AASIA                 | 7000         | 3   | E583    | NA          | NA               | SENIOR<br>DATA<br>SCIEN-<br>TIST |

• If we were to add a column like this -

```
ALTER TABLE emp_record_table
ADD designation VARCHAR(50)

UPDATE emp_record_table
SET designation=CASE WHEN exp <= 2 THEN 'JUNIOR DATA SCIENTIST'
WHEN exp > 2 AND EXP <= 5 THEN 'ASSOCIATE DATA SCIENTIST'
WHEN exp > 5 AND EXP <= 10 THEN 'SENIOR DATA SCIENTIST'
WHEN exp > 10 AND EXP <= 12 THEN 'LEAD DATA SCIENTIST'
WHEN exp > 12 AND EXP <= 16 THEN 'MANAGER'
END
```

lets remove this designation column before moving on

```
ALTER TABLE emp_record_table
DROP COLUMN designation
```

### 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.

- index helps in speeding up queries, its performance is quite observable in large datasets.
- index actually creates a backend table that would keep a ~binary record of these values.
- text was not compatible with indexing so we need to change to  $\mathbf{VARCHAR}$

```
ALTER TABLE emp_record_table
MODIFY COLUMN first_name VARCHAR(20)

CREATE INDEX ix_firstname ON emp_record_table(FIRST_NAME)
```

• This table is too short to tell the difference in query time

```
SELECT * FROM emp_record_table WHERE first_name = "eric"
```

Table 20: 1 records

| emp_iflrst_ | namheAST_NGAEN | <b>IDRO</b> LE                 | DEPT EXPCOU  | N <b>CRW</b> TIN      | ESMTLA RWI | · | R <b>MIAIN</b> | GprRj_1 | <b>ib</b> lesignation |
|-------------|----------------|--------------------------------|--------------|-----------------------|------------|---|----------------|---------|-----------------------|
| E005 Eric   | HoffmanM       | LEAD<br>DATA<br>SCIEN-<br>TIST | FINANCIE USA | NORTH<br>AMER-<br>ICA | 8500       | 3 | E103           | P105    | NA                    |

# 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 * , salary*0.05*emp_rating as bonus
FROM emp_record_table
```

Table 21: Displaying records 1 - 10

| $\overline{\mathrm{emp}}$ | ifdrst_n | a <b>h</b> AST_ | ØAN          | DECRLE                           | DEPT  | EXF            | COU.  | NOTORNITIN                       | NEAHARM | IP_ | RMANN | GER_ | <b>Id</b> esigna | t <b>iloo</b> mus |
|---------------------------|----------|-----------------|--------------|----------------------------------|-------|----------------|-------|----------------------------------|---------|-----|-------|------|------------------|-------------------|
| E001                      | Arthur   | Black           | M            | PRESIDENT                        | TALL  | 20             | USA   | NORTH<br>AMER-<br>ICA            | 16500   | 5   | NA    | NA   | NA               | 4125              |
| E005                      | Eric     | Hoffma          | a <b>i</b> M | LEAD<br>DATA<br>SCIEN-<br>TIST   | FINAN | ŒE             | USA   | NORTH<br>AMER-<br>ICA            | 8500    | 3   | E103  | P105 | NA               | 1275              |
| E010                      | Willian  | mButler         | M            | LEAD<br>DATA<br>SCIEN-<br>TIST   | AUTO  | M <b>0</b> T   | IFA:  | NEEROP                           | E9000   | 2   | E428  | P204 | NA               | 900               |
| E052                      | Dianna   | wilson          | ı F          | SENIOR<br>DATA<br>SCIEN-<br>TIST | HEALT | TH6CA          | ARÆN. | AIN/ARTH<br>AMER-<br>ICA         | 5500    | 5   | E083  | P103 | NA               | 1375              |
| E057                      | Doroth   | yWilson         | ı F          | SENIOR<br>DATA<br>SCIEN-<br>TIST | HEALT | ` <b>Н9</b> СА | AR\$A | NORTH<br>AMER-<br>ICA            | 7700    | 1   | E083  | P302 | NA               | 385               |
| E083                      | Patrick  | Voltz           | M            | MANAGER                          | HEALT | <b>Ъ</b> СА    | AR\$A | NORTH<br>AMER-<br>ICA            | 9500    | 5   | E001  | NA   | NA               | 2375              |
| E103                      | Emily    | Grove           | F            | MANAGER                          | FINAN | Œ              | CAN.  | A <b>N⁄A</b> RTH<br>AMER-<br>ICA | 10500   | 4   | E001  | NA   | NA               | 2100              |
| E204                      | Karene   | e Nowak         | F            | SENIOR<br>DATA<br>SCIEN-<br>TIST | AUTO  | M <b>®</b> T   | KYER. | MEAURYOP1                        | E7500   | 5   | E428  | P204 | NA               | 1875              |
| E245                      | Nian     | Zhen            | M            | SENIOR<br>DATA<br>SCIEN-<br>TIST | RETAI | L6             | CHIN  | V <b>A</b> SIA                   | 6500    | 2   | E583  | P109 | NA               | 650               |

| emp_fdrst_nafn&ST_ <b>CIANIFR</b> PLE    | DEPT EXPCOUN <b>CON</b> T | IN <b>EAT</b> TARMP_ | _RMATNACGER_ | <b>idD</b> esignat <b>iloo</b> nus |
|--|---------------------------|----------------------|--------------|------------------------------------|
| E260 Roy Collins M SENIOR DATA SCIENTIST | RETAIL7 INDIAASIA         | 7000 3               | E583 NA      | NA 1050                            |

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

```
SELECT CONTINENT, COUNTRY, AVG(SALARY) AS average_salary
FROM
    emp_record_table
GROUP BY
    continent, country;
```

Table 22: 7 records

| CONTINENT     | COUNTRY  | average_salary |
|---------------|----------|----------------|
| NORTH AMERICA | USA      | 9440.000       |
| EUROPE        | FRANCE   | 9000.000       |
| NORTH AMERICA | CANADA   | 7000.000       |
| EUROPE        | GERMANY  | 7600.000       |
| ASIA          | CHINA    | 6500.000       |
| ASIA          | INDIA    | 6166.667       |
| SOUTH AMERICA | COLOMBIA | 5600.000       |
|               |          |                |