<u>CS-457 HW No. 5</u> <u>Owais Waheed (ow07611)</u> <u>October 16, 2023</u>

Task 1:

Create a SQL database and separate tables for both datasets EmployeeAttrition1.csv and EmployeeAttrition2.csv using an RDBMS (PostgreSQL preferred).

Creating table 1 named as EmployeeAttrition1 with the following query,

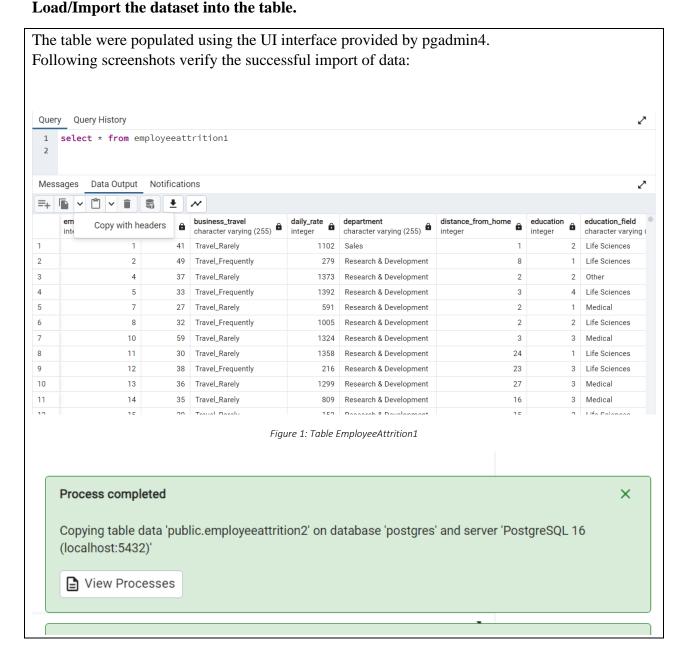
```
CREATE TABLE EmployeeAttrition1 (
Employee_Number INT,
Age INT,
Business_Travel varchar(255),
Daily_Rate INT,
Department varchar(255),
Distance_from_Home INT,
Education INT,
Education_Field varchar(255),
Environment_Satisfaction INT,
Gender varchar(255),
Hourly Rate INT,
Job_Involvement INT,
Job_Level INT,
Job_Role varchar(255),
Job Satisfaction INT,
Marital Status varchar(255),
Monthly Income INT,
Monthly_Rate INT,
Num_Companies_Worked INT,
Percent_Salary_Hike INT,
Performance_Rating INT,
Relationship_Satisfaction INT,
Standard Hours INT,
Stock Option Level INT,
Total Working Years INT,
Training Times Last Year INT,
Work_Life_Balance INT,
Years_At_Company INT,
Years_In_CurrentRole INT,
Years Since Last Promotion INT,
Years_With_Curr_Manager INT
);
```

Creating table 2 named as EmployeeAttrition2 with the following query,

```
CREATE TABLE EmployeeAttrition2 (
Employee_Number INT,
```

```
Over_18 varchar(4),
Over_Time varchar(4),
Attrition varchar(4)
);
```

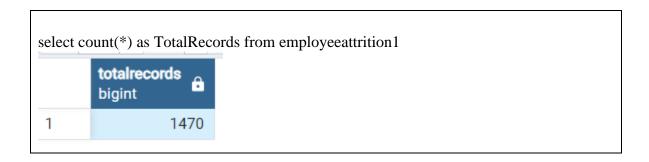
<u>Task 2:</u>



<u>Task 3:</u>

Query the database table for EmployeeAttrition1.csv and interpret the results, displaying:

1. the count of total number of records in the table

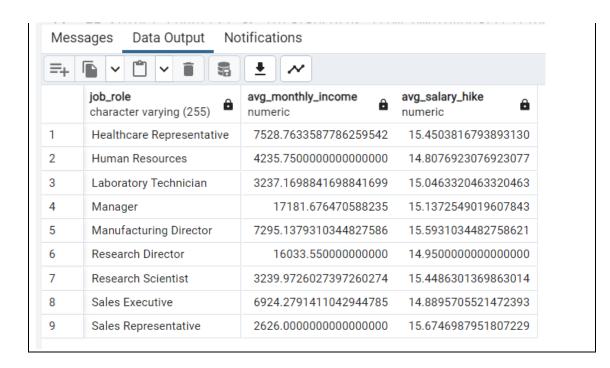


2. the count of records for each JobRole in descending order of count

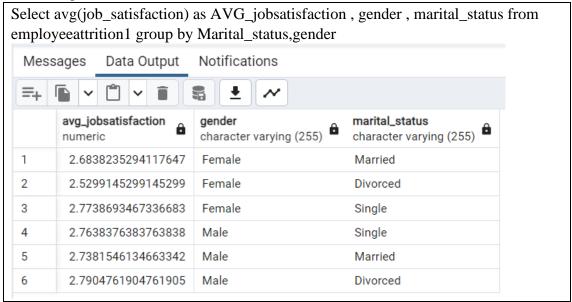


3. the average MonthlyIncome and PercentSalaryHike for each JobRole in ascending order of JobRole

SELECT Job_Role as JobRole, avg(Monthly_Income) as Average_Monthly_Income, avg(Percent_Salary_Hike) as Salary_Hike FROM EmployeeAttrition1 GROUP BY Job_Role ORDER BY Job_Role ASC;

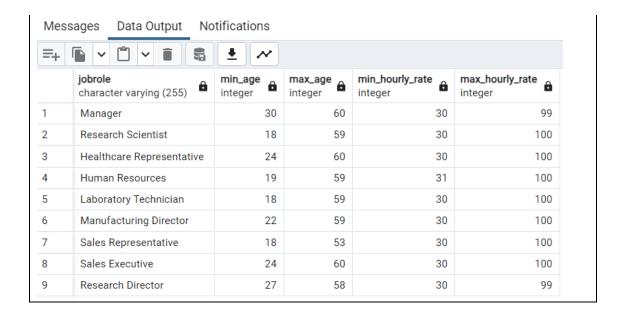


4. the average JobSatisfaction for each Gender and MaritalStatus



5. the range (Min and Max) of Age and HourlyRate for each JobRole

SELECT job_role as jobrole , min(Age) as min_age, Max(Age) as max_age, min(Hourly_Rate) as min_hourly_rate, Max(Hourly_Rate) as max_hourly_rate FROM EmployeeAttrition1
GROUP BY job_role;



- 6. Join two tables for EmployeeAttrition1.csv and EmployeeAttrition2.csv and display 20 records with the following columns
 - EmployeeNumber, Age, Gender, JobRole, OverTime and Attrition

