

# CS 457 - Homework Assignment 5: SQL

rk06451

CS 457 Data Science  
Habib University  
Spring 2023

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

## Solution:

### Table Creation for EmployeeAttrition1

```
1 Create Table EmployeeAttrition1
2 (EmployeeNumber INT, Age INT,
3 BusinessTravel varchar(255), DailyRate varchar(255), Department varchar(255),
4 Education INT, EducationField varchar(255), EnvironmentSatisfaction INT,
5 Gender varchar(255), HourlyRate INT, JobInvolvement INT, JobLevel INT, JobRole
6 JobSatisfaction INT, MaritalStatus varchar(255), MonthlyIncome INT, MonthlyRate
7 NumCompaniesWorked INT, PercentSalaryHike varchar(255), PerformanceRating varc
8 RelationshipsSatisfaction varchar(255), StandardHours varchar(255), StockOption
9 TotalWorkingYears varchar(255), TrainingTimesLastYear varchar(255), WorkLifeBa
10 YearsAtCompany varchar(255), YearsInCurrentRole varchar(255), YearsSinceLastPro
11 YearsWithCurrManager varchar(255));
```

Data Output Messages Notifications

CREATE TABLE

Query returned successfully in 65 msec.

### Table Creation for EmployeeAttrition2

```
13
14 -- Creating Second Table
15 Create Table EmployeeAttrition2(
16 EmployeeNumber INT,
17 Over18 varchar(64),
18 OverTime varchar(64),
19 Attrition varchar(64)
20 );
```

Data Output Messages Notifications

CREATE TABLE

Query returned successfully in 46 msec.

2. Query the database table for EmployeeAttrition1.csv and interpret the results, displaying:
  - (a) the count of total number of records in the table

## Solution:

SQL Query:

```
select count(EmployeeNumber) from employeeattrition1;
```

Data Output		Messages	Notifications
	count bigint		
1	1470		
Total rows: 1 of 1		Query complete 00:00:00.324	

- (b) the count of records for each JobRole in descending order of count

**Solution:**

```
select JobRole, count(JobRole) from employeeattrition1
group by JobRole order by count(JobRole) DESC;
```

Data Output		Messages	Notifications
	jobrole character varying (255)		count bigint
1	Sales Executive		326
2	Research Scientist		292
3	Laboratory Technician		259
4	Manufacturing Director		145
5	Healthcare Representative		131
6	Manager		102
7	Sales Representative		83
8	Research Director		80
9	Human Resources		52

- (c) the average MonthlyIncome and PercentSalaryHike for each JobRole in ascending order of JobRole

**Solution:** select JobRole, avg(MonthlyIncome), PercentSalaryHike from employeeattrition1  
group by JobRole, PercentSalaryHike Order by JobRole ASC;

Data Output Messages Notifications			
	jobrole character varying (255)	avg numeric	percentsalaryhike character varying (255)
1	Healthcare Representative	7635.4285714285714286	11
2	Healthcare Representative	7157.4285714285714286	16
3	Healthcare Representative	7406.2352941176470588	14
4	Healthcare Representative	11214.5000000000000000	22
5	Healthcare Representative	7963.8000000000000000	15
6	Healthcare Representative	6940.0000000000000000	12
7	Healthcare Representative	8471.7500000000000000	23
8	Healthcare Representative	7604.5000000000000000	24
9	Healthcare Representative	7602.6666666666666667	21
10	Healthcare Representative	7602.7142857142857143	20
11	Healthcare Representative	6997.4705882352941176	13
12	Healthcare Representative	7476.0000000000000000	25
13	Healthcare Representative	8018.5000000000000000	19
14	Healthcare Representative	8280.3333333333333333	18
15	Healthcare Representative	7195.1250000000000000	17
16	Human Resources	3182.0000000000000000	17
17	Human Resources	5504.6666666666666667	19
18	Human Resources	4938.0000000000000000	14
Total rows: 131 of 131		Query complete 00:00:01.097	

- (d) the average JobSatisfaction for each Gender and MaritalStatus

**Solution:**

**select avg(JobSatisfaction), Gender, MaritalStatus from employeeattrition1  
group by Gender, MaritalStatus;**

Data Output Messages Notifications			
	avg numeric	gender character varying (255)	maritalstatus character varying (255)
1	2.5299145299145299	Female	Divorced
2	2.6838235294117647	Female	Married
3	2.7638376383763838	Male	Single
4	2.7904761904761905	Male	Divorced
5	2.7381546134663342	Male	Married
6	2.7738693467336683	Female	Single
Total rows: 6 of 6		Query complete 00:00:00.441	

- (e) the range (Min and Max) of Age and HourlyRate for each JobRole

**Solution:**

**select JobRole, min(Age) as  $Age_{Min}$ , Max(Age) as  $Age_{Max}$ , min(HourlyRate) as  $HourlyRate_{Min}$ , Max(HourlyRate) as  $HourlyRate_{Max}$  from employeeattrition1 group by JobRole;**

Data Output Messages Notifications						
	jobrole character varying (255)	age_min integer	age_max integer	hourlyrate_min integer	hourlyrate_max integer	
1	Manager	30	60	30	99	
2	Research Scientist	18	59	30	100	
3	Healthcare Representative	24	60	30	100	
4	Human Resources	19	59	31	100	
5	Laboratory Technician	18	59	30	100	
6	Manufacturing Director	22	59	30	100	
7	Sales Representative	18	53	30	100	
8	Sales Executive	24	60	30	100	
9	Research Director	27	58	30	99	
Total rows: 9 of 9 Query complete 00:00:00.981						

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

**Solution:**

```
select employeeattrition1.EmployeeNumber, Age, Gender, JobRole,
employeeattrition2.OverTime, employeeattrition2.Attrition from employeeattrition1
join employeeattrition2 on
employeeattrition1.EmployeeNumber = employeeattrition2.EmployeeNumber;
```

Data Output Messages Notifications						
	employeenumber integer	age integer	gender character varying (255)	jobrole character varying (255)	overtime character varying (64)	attrition character varying (64)
1	1	41	Female	Sales Executive	Yes	Yes
2	2	49	Male	Research Scientist	No	No
3	4	37	Male	Laboratory Technician	Yes	Yes
4	5	33	Female	Research Scientist	Yes	No
5	7	27	Male	Laboratory Technician	No	No
6	8	32	Male	Laboratory Technician	No	No
7	10	59	Female	Laboratory Technician	Yes	No
8	11	30	Male	Laboratory Technician	No	No
9	12	38	Male	Manufacturing Director	No	No
10	13	36	Male	Healthcare Representative	No	No
11	14	25	Male	Laboratory Technician	No	No
Total rows: 1000 of 1470 Query complete 00:00:01.182 Ln 31, Col 98						